Handbook of Workability and Process Design

Edited by

George E. Dieter
Howard A. Kuhn
S. Lee Semiatin

The Materials Information Society
www.asminternational.org

First printing, September 2003

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

ASM International staff who worked on this project include Steve Lampman, Acquisitions Editor; Bonnie Sanders, Manager of Production; Carol Terman, Jill Kinson, and Nancy Hrivnak, Production Editors; and Scott Henry, Assistant Director of Reference Publications.

Library of Congress Cataloging-in-Publication Data

Handbook of workability and process design / edited by George E. Dieter, Howard A.
Kuhn, Lee Semiatin.
p. cm.
Updated and expanded ed. of: Workability testing techniques. 1984.
ISBN 0-87170-778-0
1. Metals—Formability—Testing. I. Dieter, George Ellwood. II. Kuhn, Howard A. III.
Semiatin, S. L. IV. Workability testing techniques.

TA460.H3195 2003
671.3—dc21

2003052197

ISBN: 0-87170-778-0
SAN: 204-7586

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

Printed in the United States of America

Multiple copy reprints of individual articles are available from ASM International, Technical Department.

Contents

Preface

Workability is a vital aspect of the processing of materials, having roots in both material behavior and process design. Whether a part can be produced by plastic deformation without cracking or the generation of other defects is of important economic consequence. Because of the complex nature of the workability of metals, there is no single test that can be used to evaluate it. Several laboratory tests have been developed that are useful in screening materials for workability, but in other instances, very specialized tests that are specific to the process are commonly used.

The *Handbook of Workability and Process Design* is an update and expansion in scope of *Workability Testing Techniques* that was published by the American Society for Metals in 1984. This original work was developed by the Metal Working Group of ASM to provide a readily available description and interpretation of the most common workability tests in the deformation processing of metals. Prior to its introduction, this information was widely scattered in the literature. The nearly 20 year life of this book bears witness to the value and acceptance of the concept behind this project.

At the time of the formulation of *Workability Testing Techniques,* the use of finite element methods (FEMs) for the modeling and simulation of metal deformation processes was in its infancy. In the ensuing 20 years, the use of FEM analysis for process design has become rather commonplace. Therefore, in contemplating this revision and update, the editors decided to expand the scope to incorporate process design, especially as influenced by FEM analysis. By doing this, the *Handbook of Workability and Process Design* takes on a more mathematical flavor than its predecessor while still retaining a balance with its original intent. Thus, the chapters that describe the various workability tests continue true to the original intent of providing practical workability testing techniques that can be used by the inexperienced practitioner.

We appreciate the contributions from the many experts who have contributed to this *Handbook.* Also, special thanks go to Steve Lampman, of the ASM staff, who not only provided editorial guidance throughout this project but also expertly provided the chapters that describe the basics of forging, rolling, extrusion, and wiredrawing.

George E. Dieter
College Park, MD
May 1, 2003

Howard A. Kuhn
Johnstown, PA
May 1, 2003

S. Lee Semiatin
Dayton, OH
May 1, 2003

Publication title	Product code
Handbook of Workability and Process Design	#06701G

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p3-21
DOI:10.1361/hwpd2003p003

Chapter 1

Workability and Process Design—An Introduction

WROUGHT FORMS are produced by a wide variety of metalworking operations that can be classified either as bulk working or sheet forming. Bulk working operations (Fig. 1) (Ref 1) include (a) primary operations where cast ingots, continuously cast bars, or consolidated powder billets are worked into mill shapes (such as bar, plate, tube, sheet, wire) and (b) secondary operations where mill products are further formed into finished products by hot forging, cold forging, drawing, extrusion, straightening, sizing, and so forth. From a geometric viewpoint, bulk forming operations are distinguished by large changes in cross-sectional area (e.g., round bar extrusion or flat rolling) and may be accompanied by large changes in shape (e.g., impression die forging or shape rolling). In contrast, sheet forming operations (Fig. 2) (Ref 1) typically involve large changes in shape (e.g., cup forming from a flat blank) without a significant change in the sheet thickness. Sheet forming has several characteristics that distinguish it from bulk working; for example, sheet formability includes different criteria such as springback and the resistance of a sheet material to thinning. Sheet formability and process design are not addressed in this Handbook and are left as topics for another publication. Nevertheless, many of the same concepts and methods described in the present Handbook can be applied to sheet forming processes.

While the major role of bulk forming operations is to produce the desired shape, in doing so they also modify the material structure and surface. Generally, the effects of bulk working processes are beneficial, leading to improved internal quality (closure of casting porosity, refinement of grain structure, and grain alignment) as well as improved surface quality (burnished surfaces and worked surface material). However, the large amount of metal movement during bulk forming operations also can introduce material discontinuities that are potential defects (i.e., imperfections that degrade intended function).

Some general types of surface and internal discontinuities of wrought products are illustrated in Fig. 3 for the example of rolled bar. Imperfections such as segregation, porosity, or seams can influence the potential or likelihood of a defective part, and their prevention is one basic objective of process design and control. Material control is also important, because many problems can be traced back to the process of melting and solidification. For example, porosity and shrinkage cavities (pipe) in an ingot can be passed on to the wrought form (Fig. 4).

The demands of high-performance products and rapid product development also can be major concerns in the manufacturing of new products. For example, metallurgical features (such as dispersoids and reinforcing particles) that lead to desirable properties in modern high-performance materials may also render them hard-to-work. At the same time, rapid product and process development have achieved a high level of sophistication through the use of modern design methods and tools, enhanced by computerization. These innovations can be classified as materials testing and data acquisition, process and product modeling and simulation, and sensors and model based process control. Effective application of such methods and tools leads to process design and controls that prevent defects and reach the full advantages offered by bulk-forming operations in the production of high-performance components.

This Handbook focuses on bulk-forming processes, the defects that frequently occur in such processes, and the roles of materials testing, process design, and process control in avoiding defects. The types of workability problems that may occur are introduced first, and the general concepts of process modeling for designing and controlling bulk working processes are described. This introductory chapter also includes a brief overview on modeling of bulk forming processes by numerical techniques such as finite element analysis (FEA). Finite element

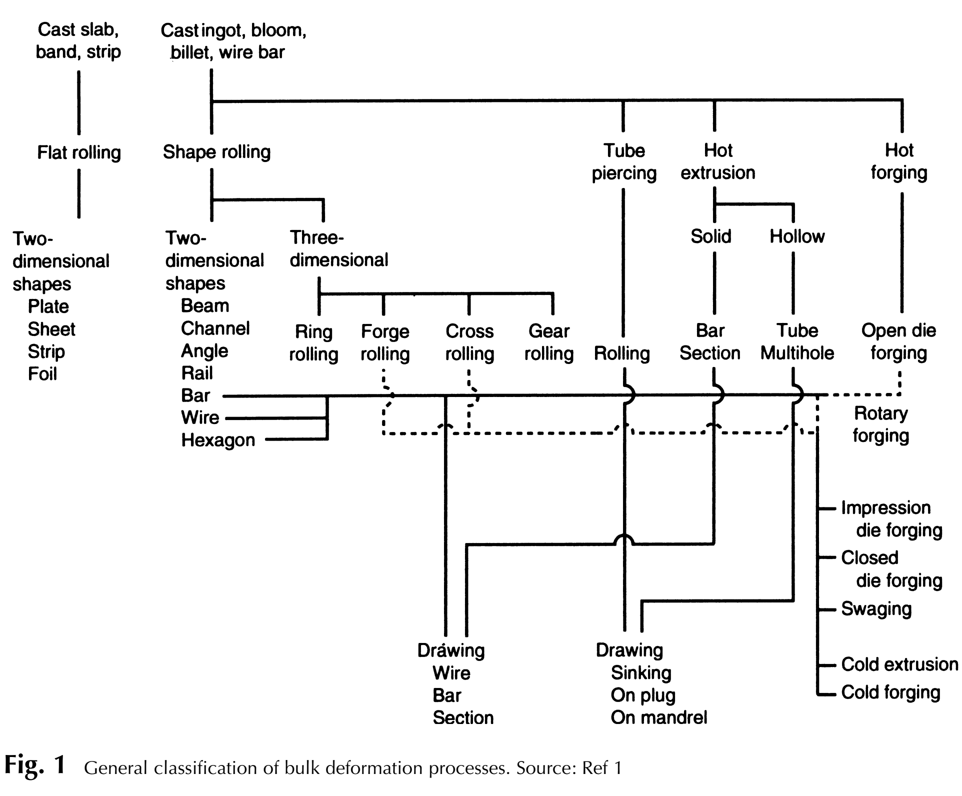


Fig. 1 General classification of bulk deformation processes. Source: Ref 1

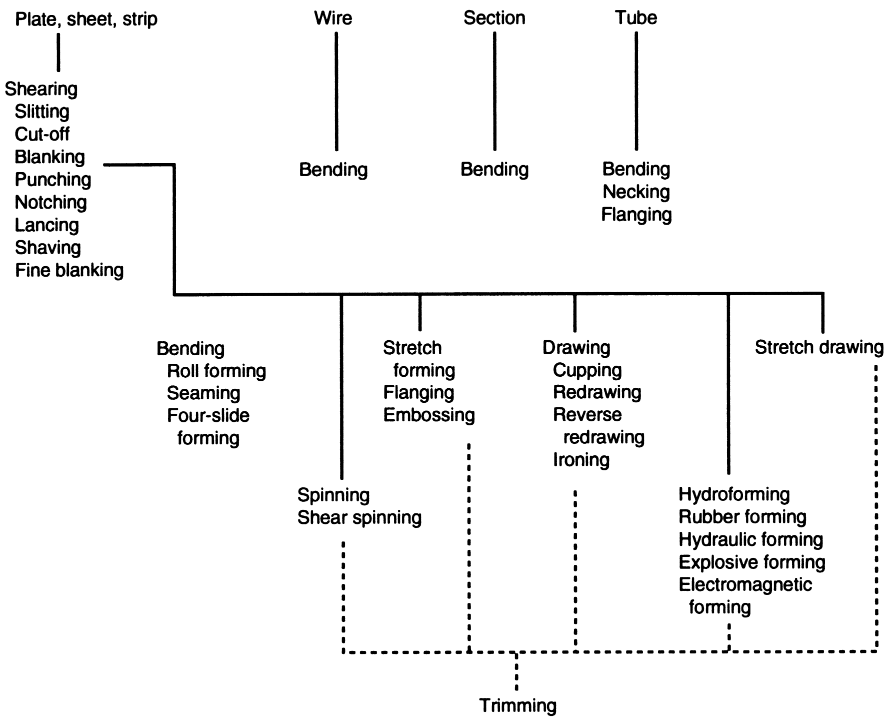


Fig. 2 General classification of sheet forming processes. Source: Ref 1

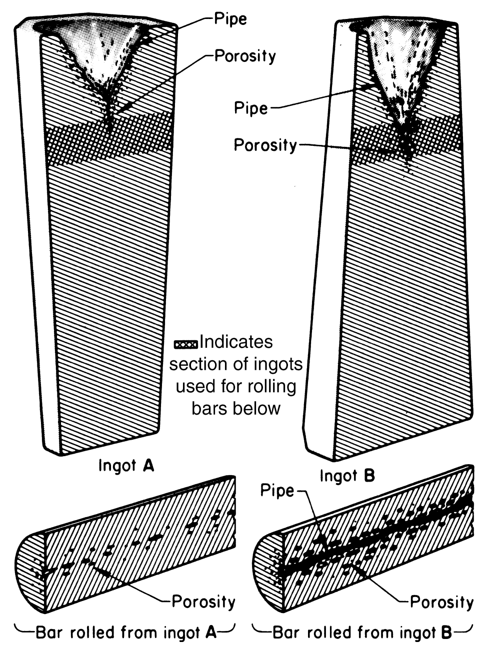


Fig. 4 Longitudinal sections of two types of ingots showing typical pipe and porosity. Pipe is a condition that develops in the nominal top centerline of the ingot because of metal shrinkage during solidification. When the ingots are rolled into bars, pipe and porosity become elongated throughout the center of the bars.

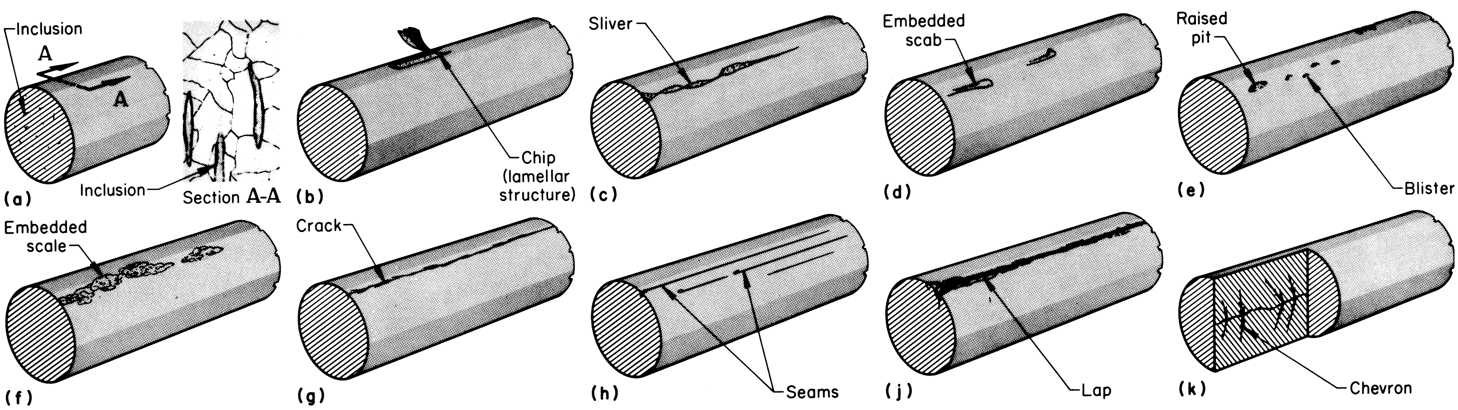


Fig. 3 Ten different types of imperfections that might be found in rolled bar

analysis is an effective method for solving a wide variety of engineering problems and has been useful in the design and analysis of both bulk and sheet forming processes (Ref 2).

Workability Problems

Workability refers to the relative ease with which a material can be shaped through plastic deformation. Workability is usually thought of as being limited by the onset of fracture. Greater workability of a material allows greater deformation and/or a more complex shape that can be produced before fracture occurs. In general, however, a workability problem occurs when the part produced by the bulk working process is unacceptable and must be scrapped or reworked. From this practical point of view, workability also may be defined by other factors such as the generation of a rough surface finish or the inability to achieve a required tolerance on a critical dimension.

Workability also is a complex technological concept that is related to both material and process characteristics. Workability depends not only on the fracture resistance (ductility) of the material but also on the specific details (stress state) of the process as influenced by die geometry, workpiece geometry, and lubrication conditions. Ease of manufacture is aided when the material has a low flow stress (yield strength) so that the force that must be applied by the processing equipment and the stresses on the dies are lower. On the other hand, a poorly designed or controlled process can produce defects in an easy-to-work material, leading to a scrap part. By way of example, lead is a very ductile and workable material that can be formed readily by compressive operations such as forging and rolling; yet, lead fails to form easily in tensile operations such as drawing.

A hard-to-work alloy presents other challenges to provide a deformation process environment that prevents defect formation. The evaluation of a material by the various workability testing and analysis methods described in this Handbook provide a framework for intelligently choosing materials for best workability or for changing the design of the process to enhance the workability of a given material. As a first step toward devising a solution to a workability problem, it is useful to categorize workability problems in terms of their general source:

- *Fracture-related problems:* for example, internal bursts or chevron cracks; cracks on free surfaces; cracks on die-contacted surfaces
- *Metal-flow-related problems:* for example, end grain and poor surface performance; inhomogeneous grain size; shear bands and locally weakened structures; cold shuts, folds, and laps; flow-through defects
- *Control, material selection, and utilization problems:* for example, underfill, part distortion, and poor dimensional control; tool overload and breakage; excessive tool wear; high initial investment due to equipment cost; poor material utilization and high scrap loss

These types of problems are introduced briefly in the following sections of this chapter. Each type of problem may involve different kinds of methods or solutions in the design of the process and/or product. For example, workability can be improved by changes in die geometry, workpiece geometry, lubrication conditions, or processing temperature. Much plastic deformation of metals is carried out at elevated temperature (hot working) because flow stress decreases with increasing temperature.

The same general approach applies to sheet forming operations in that the major emphasis of formability evaluation also is on measuring and predicting the limits of deformation before fracture. Sometimes the term *formability* is used in conjunction with either the sheet formability or bulk workability. However, the term *formability* is limited more properly to sheet forming operations because there are major distinctions in the conditions of sheet forming and bulk forming processes. In sheet forming, metal is deformed plastically by tensile loads, often without significant changes in sheet thickness or surface characteristics and with the possibility of significant elastic recovery or springback because the magnitudes of plastic and elastic deformation may be similar. In contrast, metal is deformed plastically by compressive loads during bulk forming, and the plastic deformation is proportionally much more prevalent than elastic deformation.

Fracture-Related Problems

The general types of fracture in bulk working are:

- Free surface fracture
- Die contact surface fracture
- Internal fracture

The most common types of fracture failures in bulk working are free surface fracture (at hot or cold processing temperatures) and internal fracture. Internal fracture occurs by mechanisms such as triple-point cracking/cavitation at hot working temperatures or inhomogeneous deformation that cause internal defects such as central bursts or chevron cracking. Internal fracture can be an extremely dangerous type of defect because it cannot be detected visually.

Free Surface Cracking (Adapted from Ref 3). A free surface, by definition, has neither a stress normal to it nor a shear stress on it. Free surface fractures occur on surfaces undergoing free expansion due to compressive loads on contact surfaces between the tools (rolls or dies) and the workpiece. The tensile stresses leading to free surface fracture are often referred to as secondary stresses since they are not applied directly by metalworking equipment. Edge cracking in rolling of plates, strip, or rings is a primary example of free surface cracking. Another is the surface cracking occurring on exposed expanding surfaces during upsetting or on the leading edges of localized areas of extrusion in forgings.

One of the most successful and useful design tools to come from bulk workability research is the workability diagram for free surface fracture during the cold working of wrought and powder metals. An example of a workability diagram of this type is shown in Fig. 5 from Chapter 3 of Ref 4, "Cold Upset Testing." The graph indicates the locus of free surface normal strains (one tensile and one compressive) that cause fracture. The workability diagrams are used during process design by plotting calculated or estimated surface strain paths that are to be imposed during forming on the fracture locus diagram (Fig. 5). If the final strains lie above the locus, part failure is likely, and changes are necessary in preform design, lubrication, and/or material. The fracture locus concept has been used to prevent free surface cracking in forging and to prevent edge cracking in rolling. With modifications, the fracture locus approach has also provided insight into such failure modes as center bursting in extrusion and forging and die-workpiece contact fractures in forging. These limits change with chemistry, grain size, temperature, second-phase content, and possibly with strain rate.

The concept of a working limit for free surface fracture is important because the workability of a metal may be characterized for a particular set of process conditions. In general, as the working temperature is increased, the location of the fracture line will move upward, indicating that higher deformation can be accommodated before fracture. However, higher temperatures are not always beneficial, as in the case of IN718 nickel-base superalloy, which has a temperature limit of about 1120 °C (2050 °F) for hot working (Ref 5 and Chapter 12, "Workability Theory and Application in Bulk Forming Processes" in this Handbook). Experimental observations have shown that the slope of the line increases with strain rate for some metals, most notably some brasses and austenitic stainless steels (Ref 6). The position of the line drops (lower workability) as the second-phase content increases, much like the tensile ductility decreases with second phase. The fracture line location is also sensitive to the microstructure. For example, a spheroidized structure for a high-carbon steel has a higher fracture-line position than a pearlitic microstructure has. More details and references on workability in terms of free surface cracking are in Chapter 11, "Design for Deformation Processes."

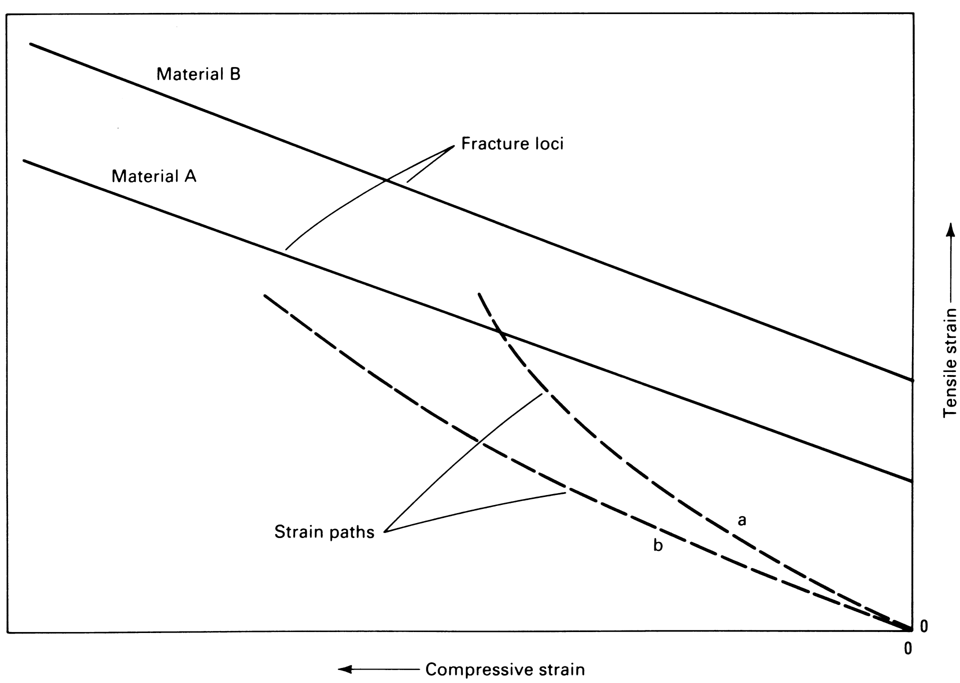


Fig. 5 Schematic workability diagrams for bulk forming processes. Strain path a would lead to failure for material A. Both strain paths (a and b) can be used for the successful forming of material B. Source: Ref 4

Die-Contact Surface Cracking. Cracking on surfaces in contact with a die is a common problem. Frequently, cracks occur during forging on surfaces that are in contact with the dies. One common location of such defects is in the vicinity of a die or punch corner. A combination of shear deformation with tension or low values of hydrostatic pressure in the vicinity of a die corner are responsible for surface cracking (Ref 7).

Observation of a variety of such defects shows that an apparently common characteristic is an abrupt change in frictional shear traction distribution in the region of the crack. High friction to retard metal flow in advance of the crack location is one method for preventing such defects. These cracks usually do not propagate deeply into the workpiece but instead result in unacceptable surface quality or unacceptable machining depths if that surface is to be finish machined. As the forging community has moved closer to net-shape forming, this type of defect has become an increasing problem. Causes of this problem include nonuniform lubrication, flow around die corners, and an improper starting workpiece surface.

Central or Internal Bursts (Chevron Cracking). Central bursts are internal fractures caused by high hydrostatic tension in combination with internal material weaknesses. Chevrons are internal flaws named for their shape (Fig. 3k). A central burst, or chevron crack, is associated most commonly with extrusion and drawing operations, although it can be generated by forging and rolling processes as well. Internal bursts in rolled and forged metals result from the use of equipment that has insufficient capacity to work the metal throughout its cross section. If the working force is not sufficient, the outer layers of the metal will be deformed more than the inside metal, sometimes causing wholly internal, intergranular fissures that can act as initiation sites for further crack propagation during service loads. In forward cold extrusion, the occurrence of central bursts or chevrons is nearly always restricted to isolated lots of material and usually to only a small percentage of the pieces extruded in any particular production run.

A change in deformation zone geometry is usually sufficient to eliminate the problem. The conservative design approach is to ensure that no hydrostatic tension develops. Often, however, the part or tooling design cannot be changed sufficiently to eliminate hydrostatic tension. If the level of hydrostatic tension can be kept below a critical level, bursting can likely be avoided. This may be accomplished by a change in lubricant, die profile, temperature, deformation level, or process rate.

The probability of internal burst is enhanced in areas of material weakness if they are acted on by hydrostatic tension stress states during the deformation process. For example, with ingot imperfections (such as pipe, porosity, segregation, or inclusions), tensile stresses can be sufficiently high to tear the material apart internally, particularly if the forging temperature is too high (Ref 8). Similarly, if the metal contains low-melting phases resulting from segregation, these phases may rupture during forging. Ingot pipe, unhealed center conditions, or voids associated with melt-related discontinuities may also induce center bursts if reduction rates are too severe or temperatures are incorrect during working. The conversion practice to bar or billet 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. 6.

It also is useful to point out that, if the stress state is compressive in areas where material weaknesses occur, the compressive stress state will close up any porosity or pipe and will retard any inclusions or segregation from initiating cracks. The stress state, as controlled by the process parameters, thus has an important role in the degree of soundness in the worked material. The classic work by Lou Coffin and Harry Rogers (Ref 9) showed very clear relationships between damage generation and tensile stress states (as well as damage healing and compressive stress states) during deformation processing.

Macroetching and ultrasonic inspection methods are the most widely used for identifying regions of unsoundness. Bursts usually display a distinct pattern of cracks and do not show spongy areas, thus distinguishing them from pipes. Bursts are readily detected by macroetching. Figure 7 shows a large burst that occurred during the forging of an electroslag-remelted (ESR) ingot. The cause was traced to a weak solidification plane near the bottom of the ingot

Fig. 6 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.

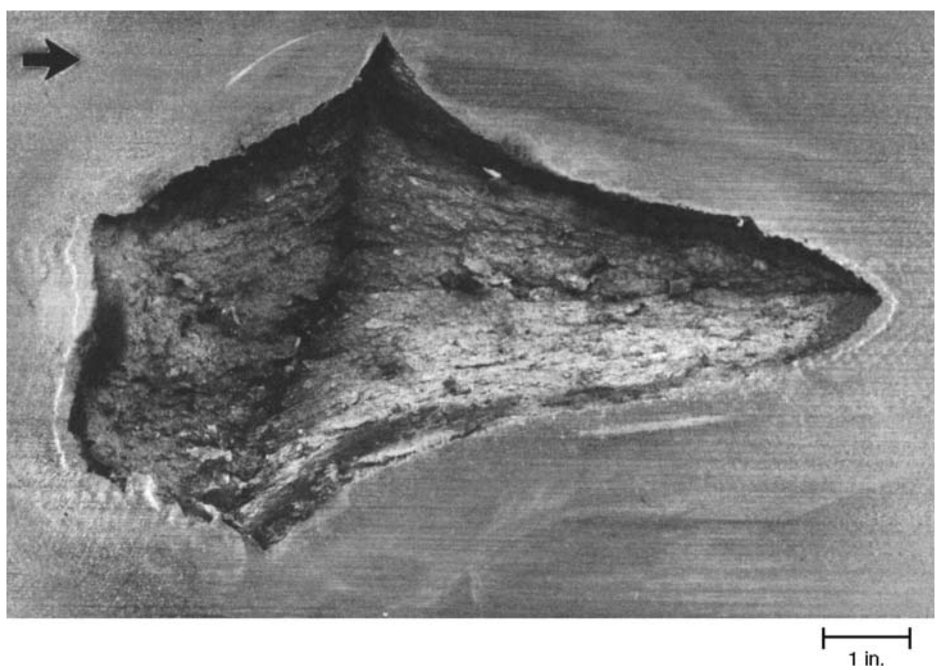


Fig. 7 Cross section of a forged bar showing a forging burst. The burst is located approximately at the centerline of the workpiece. Arrow indicates the direction of working.

combined with higher than normal forging temperatures.

Flow-Related Defects

The defects described in this section are related to the distribution of metal. They can be avoided by proper die design, preform design, and choice of lubrication system. Strictly speaking, these defects are not fundamental to the workability of the material but are related to the process details. However, knowledge of these common forging defects is necessary for a practical understanding of workability. These are the defects that commonly limit deformation in secondary operations.

Most types of flow-related defects occur in hot forging, which is described in more detail in Chapter 13, "Workability in Forging." However, the following provides a general introduction to typical types of defects that may occur from bulk working. They are common for impression-die forging, in which case defect formation may also involve entrapment of oxides and lubricant. When this occurs, the metal is incapable of rewelding under high forging pressures; the term *cold shut* is frequently applied in conjunction with laps, flow-through defects, and so on to describe the flaws generated.

Underfill may not seem like a flow-related defect, but aside from simple insufficient starting mass, the reasons for underfill are flow related. These include improper fill sequence, insufficient forging pressure, insufficient preheat temperature, lubricant buildup in die corners, poor or uneven lubrication, and excessive die chill. An improper fill sequence may result in excessive flash loss, or it may be the result of extraordinary pressure requirements to fill a particular section. Sometimes, venting may eliminate the problem; more often than not, a change in the incoming workpiece shape or a change in the deformation sequence is required.

Laps and Folds. Laps are surface irregularities that appear as linear defects and are caused by the folding over of hot metal at the surface. These folds are worked into the surface but are not metallurgically bonded (welded) because of the oxide present between the surfaces (Fig. 8). Thus, a discontinuity with a sharp notch is created.

In rolling, laps are most often caused by excessive material in a given hot roll pass being squeezed out into the area of the roll collar. When turned for the following pass, the material is rolled back into the bar and appears as a lap on the surface (Fig. 3). A lap or fold occurs in die forging from an improper progression in fill sequence. Normally, a lap or fold is associated with flow around a die corner, as in the case of an upper rib or lower rib, or with a reversal in metal-flow direction.

In die forging, a general rule of thumb is to keep metal moving in the same direction. The die corner radius is a critical tool dimension, and it should be as generous as possible. In progressing through a forging sequence, the die corners should become tighter so that the workpiece fillets are initially large and progressively become smaller as the forging steps are completed. Figure 9 (Ref 10) shows schematically a lap forming as metal flows around a die corner.

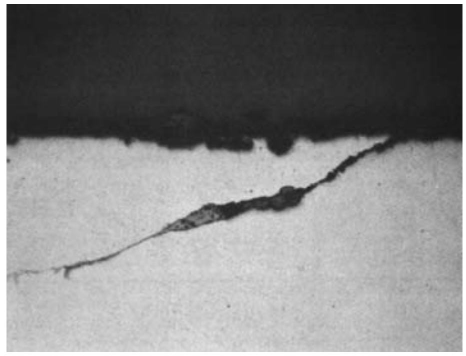


Fig. 8 Micrograph of a forging lap. Note the included oxide material in the lap. 20×

Extrusion-Type Defects. The tail of an extrusion is unusable because of nonuniform flow through the extrusion die. This results in a center-to-surface velocity gradient, with metal from the workpiece interior moving through the die at a slightly higher velocity than the outer material. The result shows up at the tail of the extrusion as a suck-in or pipe, and, for extrusions, the tail is simply cut off and discarded. Alternatively, a follower block of cheaper material may be added so that most of the defect falls in the cheaper material, and less length of the extruded workpiece is lost.

For forgings that involve forward or backward extrusion to fill a part section, the same situation can develop. Metal flow into a rib or hub can result in a suck-in defect, which in a worst-case scenario would show up as a fold on the face opposite to the rib. A best case would be a depression on what otherwise should be a flat surface. One method of eliminating this type of defect is to position more material on the back face initially. Another method is to change the rib geometry (aspect ratio and/or angles). If neither of these changes can be accomplished, an extra forging step may be needed to limit the amount of extrusion that is done in any one step.

Extrusion-type defects are formed when centrally located ribs formed by extrusion-type flow draw too much metal from the main body or web of the forging. A defect similar to a pipe cavity is thus formed (Fig. 10). Methods of minimizing the occurrence of these defects include increasing the thickness of the web or designing the forging with a small rib opposite the larger rib, as shown in Fig. 10.

Shear-Related or Flow-Through Defects. Shearing defects are also known as *flow-through defects* because they result from excessive metal flow past a filled detail of the part. Flow-through defects are formed when metal is forced to flow past a recess after the recess has filled or when material in the recess has ceased to deform because of chilling. An example of this is shown in Fig. 11 (Ref 7) for a trapped-die forging that has a rib on the top surface. The rib denoted by "2" is filled early in the forging sequence, and significant mass must flow past the rib in order to fill the inner hub, zone "4." The result can be a complete shearing-off of the rib in the worst case, with a lesser case being the formation of a shear-type crack.

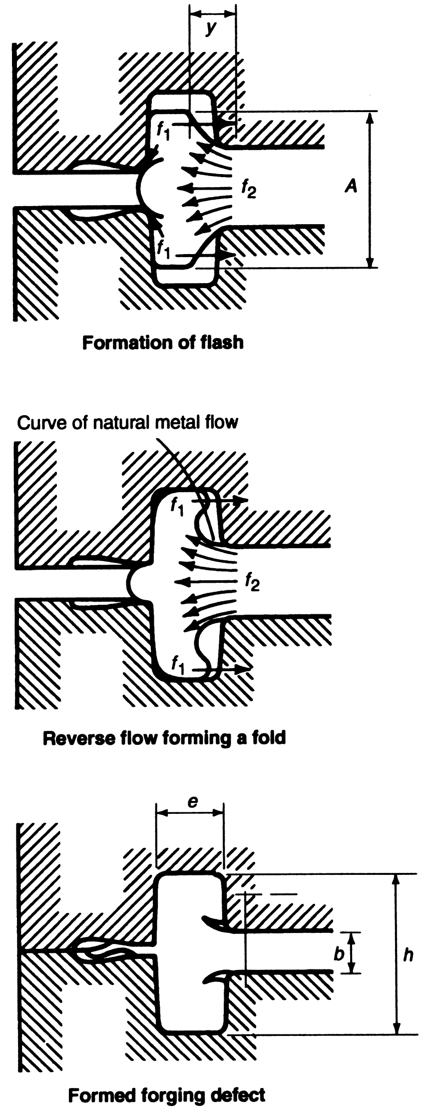


Fig. 9 Lap formation in a rib-web forging caused by improper radius in the preform die. Source: Ref 10

Similar to laps in appearance, flow-through defects can be shallow, but they are indicative of an undesirable grain flow pattern or shear band that extends much deeper into the forging. An example is shown in (Fig. 12) (Ref 12). Flow-through defects can also occur when trapped lubricant forces metal to flow past an impression.

Seams are crevices in the surface of the metal that have been closed, but not welded, by working the metal. Seams result from elongated trapped-gas pockets or from cracks. Even

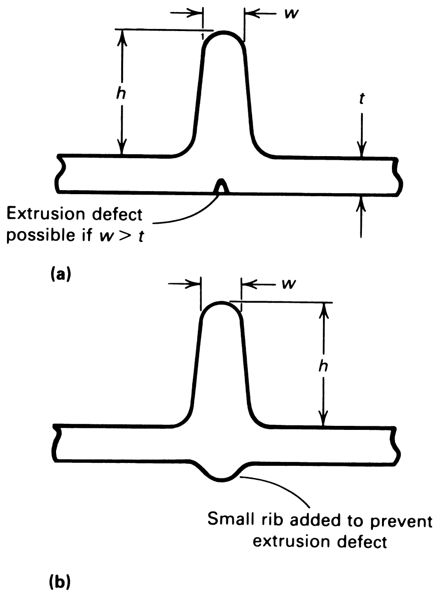


Fig. 10 Extrusion-type defect (a) in centrally located rib and (b) die-design modification used to avoid defect. Source: Ref 11

though seams can occur from cracks, the two can be distinguished from the presence of oxides. In mill processing, for example, 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. In contrast, cracks

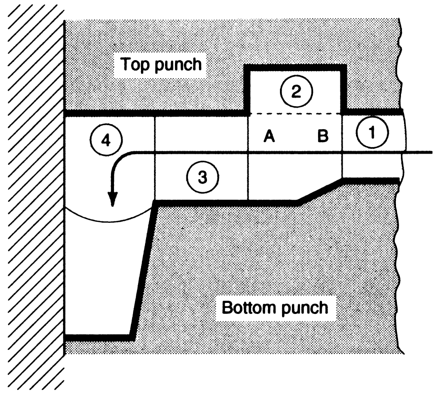


Fig. 11 Schematic of a flow-through crack at the base of a rib in a trapped-die forging. Excessive metal flow past region 2 causes a shear crack to form at A and propagate toward B. Source: Ref 7

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. Seams are generally heavily oxidized and decarburized in steels (Fig. 13).

Seams have a large number of possible origins, some mechanical and some metallurgical. Seams can develop from cracks caused by working or from an imperfection in the ingot surface, such as a hole, that becomes oxidized and is prevented from healing during working. In this case, the hole simply stretches out during forging or rolling, producing a linear seam in the workpiece surface. Seams also result from trapped-gas pockets, cracks, a heavy cluster of nonmetallic inclusions, or a deep lap.

Seams may be continuous or intermittent, as indicated in Fig. 3(h). Depth of seams varies widely, and surface areas sometimes may be welded together in spots. Seams seldom penetrate to the core of bar stock. Seams can be difficult to detect because they may appear as scratches or because a machining process may obliterate them. Seams are normally closed tight enough that no actual opening can be detected visually without some nondestructive inspection techniques such as magnetic particle inspection. Figure 14 is an example of a seam detected by routine magnetic particle inspection of a hot-rolled 4130 steel bar. No stringer type inclusions were observed in the region of the flaw, but it did contain a substantial amount of oxide (Fig. 14b).

Seams may not become evident until the part has been subjected to installation and service stresses. For example, seams are sometimes difficult to detect in an unused fastener, but they are readily apparent after a fastener has been subjected to installation and service stresses. Seams also may not become evident until the constraint exerted by the bulk of material is removed from the neighborhood of a seam. The incomplete removal of seams from forging stock can cause additional cracking in hot forging and quench cracking during heat treatment.

Microstructure and Properties

A product with poor properties is another potential problem, and it usually arises from an inadequate microstructure such as grain flow and grain size. A major advantage of shaping metal parts by rolling, forging, or extrusion stems from the opportunities such processes offer the designer with respect to the control of grain flow. The strength of these and similar wrought prod-

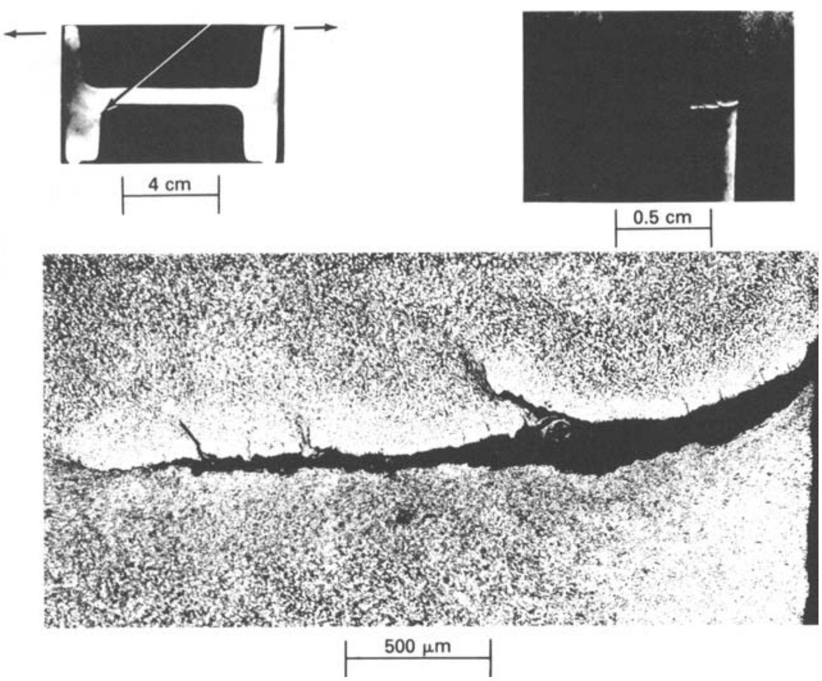


Fig. 12 Flow-through defect in Ti-6Al-4V rib-web structural part. Source: Ref 12

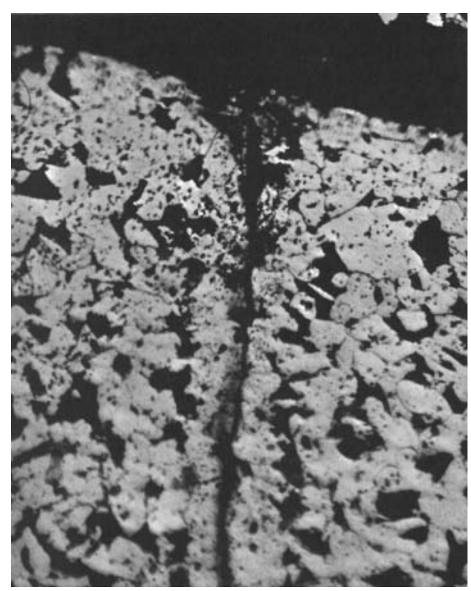


Fig. 13 Micrograph of a seam in a cross section of a 19 mm ($^3/_4$ in.) diameter medium-carbon steel bar showing oxide and decarburization in the seam. 350×

ucts is almost always greatest in the longitudinal (or equivalent) direction of grain flow. The maximum load-carrying ability in the finished part is attained by providing a grain flow pattern parallel to the direction of the major applied service loads when, in addition, sound, dense, good-quality metal of satisfactorily fine grain size has been produced throughout.

Grain Flow and Anisotropy. Metal that is rolled, forged, or extruded develops and retains a fiberlike grain structure that is aligned in the principal direction of working. This characteristic becomes visible on external and sectional surfaces of wrought products when the surfaces are suitably prepared and etched (Fig 15). The "fibers" are the result of elongation of the microstructural constituents of the metal in the direction of working. Thus the phrase *direction of grain flow* is commonly used to describe the dominant direction of these fibers within wrought metal products from the crystallographic reorientation of the grains during deformation and/or the alignment of nonmetallic inclusions, voids, and chemical segregation. The occurrence and severity of fibering varies with such factors as composition, extent of chemical segregation, and the amount of work or reduction the workpiece receives.

In wrought metal, the direction of grain flow is also evidenced by measurements of mechanical properties. Strength and ductility are almost always greater in the direction parallel to that of working. The characteristic of exhibiting different strength and ductility values with respect to the direction of working is referred to as "mechanical anisotropy" and is exploited in the design of wrought products. Although best properties in wrought metals are most frequently the longitudinal (or equivalent), properties in other directions may yet be superior to those in products not wrought—that is, in cast ingots or in forging stock taken from ingot only lightly worked.

Although all wrought metals are mechanically anisotropic, the effects of anisotropy on mechanical properties vary among different metals and alloys. For example, a vacuum-melted steel of a given composition is generally less mechanically anisotropic than a conventionally killed, air-melted steel of the same composition. Response to etching to reveal the grain flow characteristic of anisotropy also varies. Metals with poor corrosion resistance are readily etched, whereas those with good corrosion resistance require more corrosive etchants and extended etching times to reveal grain flow. Because grain flow can be a vital factor in the service performance of a part, it is useful to use arrows on forging drawings to show optimum grain flow direction that corresponds to the direction of principal service stress. Effects of grain flow are described in more detail in Chapter 11 "Design for Deformation Processes."

Grain Size. The influence of grain size on properties such as strength and ductility is generally well understood, and the effects of ther-

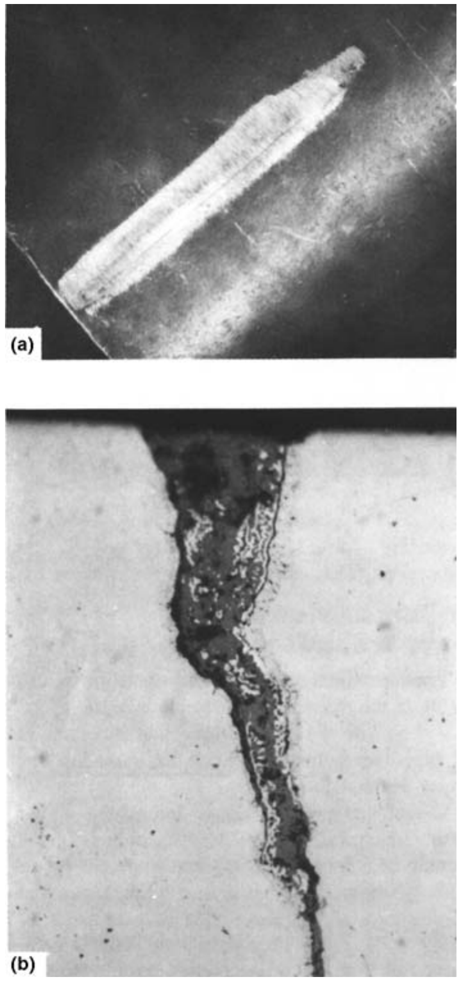


Fig. 14 Seam in rolled 4130 steel bar. (a) Close-up of seam. Note the linear characteristics of this flaw. (b) Micrograph showing cross section of the bar. Seam is normal to the surface and filled with oxide. 30×

momechanical processing on grain size are critical in obtaining products with satisfactory properties. This requires a more extensive evaluation of the dynamic and thermophysical conditions that influence metallurgical processes at a microscopic level. Thus, the effects of thermomechanical processing on grain size require more complex models based on the principles of physical metallurgy and the continuum mechanics of mechanical flow.

Grain size and grain structure also influence workability. Workability typically decreases with increasing grain size, because cracks may initiate and propagate easily along the grain boundaries. When the grain size is large relative to the overall size of the workpiece, as in conventionally cast ingot structures, hot working is required because of the low workability of the material. In general, the workability of metals increases with temperature. However, as temperature increases, grain growth also occurs. Thus, the design of thermomechanical processes may involve a complex set of factors such as materials control (preforms with suitably fine grain

Fig. 15 Section through a forged head on a threaded fastener showing uniform grain flow that minimizes stress raiser and unfavorable shear planes

size for good workability) and an optimal working temperature for adequate workability without excessive growth of grain size.

Control of grain size during thermomechanical processing is a topic of significant practical importance, and extensive efforts have been employed in this regard. For example, the concept of grain size control has been used for many years in the production of flat-rolled products. Small niobium additions increase the austenite recrystallization temperature, and controlled rolling is used to refine the relatively coarse austenite structure by a series of high-temperature rolling and recrystallization steps. It combines repeated deformation and recrystallization steps with the addition of austenite grain-growth inhibitors such as titanium nitride to refine the starting austenitic grain size and to restrict grain growth after recrystallization. This is a topic of ongoing interest, as described in more detail in the section "Microstructural Modeling" in this chapter.

Similar metallurgical effects on grain size apply to other materials and hot working operations. Minor variations in composition may also cause large variations in workability, grain size, and final properties. In one case, for example, wide heat-to-heat variations in grain size occurred in parts forged from nickel–base alloy 901 (UNS N09901) in the same sets of dies. For some parts, optimal forging temperatures had to be determined for each incoming heat of material by making sample forgings and examining them after heat treatment for variations in grain size and other properties. This illustrates the complexity of process design and modeling of thermomechanical forming.

Material Control, Selection, and Equipment Utilization

The efficiency and productivity of a forming operation are influenced by a number of factors such as material selection and control, equipment cost, tool wear, part distortion, poor dimensional control, poor material utilization, and high scrap loss. Material control is important, as material variation can have significant effects on properties and characteristics of the product such as grain size and mechanical properties. The responsibility of assuring and verifying the properties and characteristics of the product is vested in material control, which controls all processes employed in production, from selection of raw material to final inspection. It establishes manufacturing standards to ensure reproducibility in processing and product uniformity. Material control depends on the proper application of drawings, specifications, manufacturing process controls, and quality-assurance programs to satisfy all requirements for metallurgical integrity, mechanical properties, and dimensional accuracy. It also provides for identification and certification, so that a product history can be traced.

Effective utilization of materials and equipment also depends on the type of workpiece material and forming operation. Other chapters describe briefly equipment and workability for specific types of bulk-working operations. Chapter 11 "Design for Deformation Processes" also describes briefly tool materials. A wide range of materials for tools and dies is available to designers. Among the important attributes are hardenability; machinability; and resistance to wear, plastic deformation, shock loading, and heat checking. The needed levels of resistance to wear, plastic deformation, and so forth, are determined by factors such as type of equipment used, workpiece temperature, expected die temperature, and number of parts to be fabricated. Low-alloy steels and hot-work die steels are often suitable for conventional metalworking. On the other hand, high-temperature die materials are required for special applications such as isothermal forging of titanium and nickel-base alloys. These die materials include various superalloys and TZM molybdenum (Mo-0.5Ti-0.1Zr). Recommendations on the selection of these materials are made in Chapter 11 "Design for Deformation Processes." The approach used in making these recommendations and the tool materials are discussed in detail for hot forging tooling.

Process Design and Modeling

A considerable array of mechanical tests and modeling techniques has been developed to prevent defects and/or optimize results in bulk forming operations such as forging, extrusion, and rolling. In general, the tools and techniques for process design have one or more of the following objectives:

- Prevent improper part shape or final geometry that differs from expectations
- Prevent defects or cracking that occur during shaping
- Prevent poor properties from the development of inadequate microstructure
- Optimize results in terms of desired process characteristics (such as energy consumption and maximum productivity) and/or product properties (such as microstructural homogeneity, grain flow characteristics, and grain size)

Each objective requires different types of analytical tools or models. For example, models based on continuum mechanics can address flow-related problems such as insufficient die fill, poor shape control, and fracture conditions.

The four major design considerations in a bulk forming process are material flow, workability, resultant properties (microstructure) of the product, and utilization factors (economics, efficiency, productivity, etc.). The first consideration is the workpiece material and its flow stress behavior. *Flow stress* is the stress needed to cause plastic deformation and is affected by temperature, rate of deformation, and amount of previous plastic deformation. Flow stress behavior is based on mechanical testing and can be modeled by constitutive equations that describe mathematically the relationship between stress and strain during plastic deformation.

The second consideration is the fracture behavior of the material and the effects of temperature, stress state, and strain rate on fracture; this combined view of ductility and stress state is termed workability for bulk forming processes. Even if the desired shape is obtained, cracking or some other cracklike discontinuity may occur. For example, edge cracking may occur during rolling, or central bursts may occur during extrusion. Cracking of the tool itself may even occur. Many methods have been developed to evaluate workability, as described in more detail in subsequent chapters, with damage criteria detailed in Chapter 12 "Workability Theory and Application in Bulk Forming Processes".

The third major consideration is a determination of the desired final microstructure needed to produce an acceptable product. Microstructural optimization or prediction requires additional models besides just continuum mechanics. Microstructural modeling also requires quantifying the thermal field in the materials and the associated metallurgical phenomena. The practical application of such physical metallurgy-based models continues to advance further as numerical techniques such as finite element analysis (FEA) become more sophisticated with improvements in computer hardware. The ability to accurately predict thermomechanical histories by finite element modeling is being used to predict the evolution of microstructure, thereby replacing data-intensive empirical methods with more knowledge-based analytical/numerical methods using the fundamental principles of transport phenomena, continuum mechanics, and physical metallurgy.

The fourth consideration involves added constraints of available equipment and economics in addition to flow stress, forming, and part performance considerations. The fourth consideration usually dominates the other considerations, sometimes to the detriment of the material being worked. Material utilization may also include factors of economic productivity, efficiency, tool wear, and scrap loss.

Concepts of Process Modeling

Engineering models are often used to determine the response of structure, component, process, or system to a set of conditions. The development of any model requires the definition of appropriate boundary conditions and the application of appropriate mathematical equations (which are typically differential equations for many physical situations). The complexity of the boundary conditions and the governing differential equations determine the possible methods of solution. In some cases, solutions may be adequately achieved by analytical (closed-form) equations, while in other cases numerical techniques (such as FEA) may be required.

In the case of deformation processes like a forg-

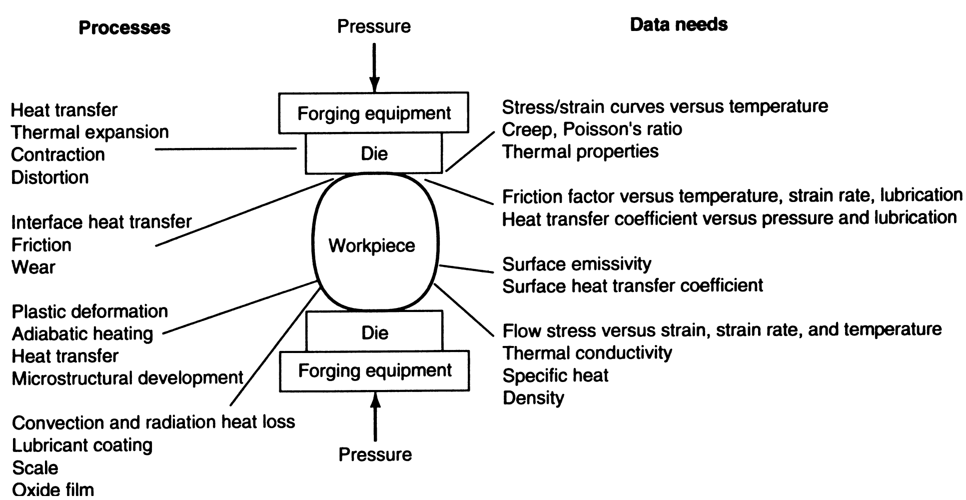


Fig. 16 Typical physical phenomena occurring during a forging operation

ing operation (Fig. 16), the major process variables and interactions are shown in Fig. 17. It can be seen in Fig. 17 that for a metal forming analysis, one needs to satisfy the equilibrium conditions, compatibility equations/strain-displacement relations, constitutive equations, and, in some instances, the heat balance equation. Modeling the microstructural effects of these variables could, in principle, require a more robust set of appropriate mathematical equations for the process phenomena (such as convection, radiation, chemical reaction, and diffusion, in addition to deformation).

Historically, deformation processing has fallen in the gap between the traditional disciplines of metallurgy and mechanics, and, as a result, this area has often been neglected in an academic sense. A main reason for this is the difference in length scales by which mechanics and materials science view a material. The length scale of deformation is at the atomic level, and it is at this level that materials science addresses deformation. In manufacturing, deformation effects are related or measured at a macroscopic level, and continuum mechanics is applied to analyze and explain plasticity quantitatively. However, to understand the macroscopic response of materials to temperatures and rates of deformation, consideration must be at a lower length scale level, at least to qualitatively explain material behavior. The empirical equations used to address plasticity do not generally capture the microscopic aspects of deformation.

Fortunately, the metals that are commonly processed by bulk deformation methods have many grains per unit volume, and microscopic events are suitably averaged at the macroscopic level. The macroscopic or continuum mechanics approach begins to break down when the grain size approaches the physical size of the workpiece or when a dominant crystallographic texture is present in a workpiece. Examples of the former include fine wire drawing, bending of fine wire, and sheet forming processes. Primary examples of the latter are sheet-metal forming processes or bending of heavily drawn wire where crystallographic texture plays a dominant role. An additional complication is the fact that most metals have more than one phase present in their microstructure. The second phase may be present due to alloying (e.g., cementite in iron), or it may be an unwanted phase (e.g., sulfide or silicate inclusions in steel). The effects of these second phases are again averaged at the macroscopic level, and the material has not been adequately described as having separate phases from a mathematical sense. The materials science community and the mechanics community are trying to bridge this length-scale problem, but for now the most useful analysis tools are combinations of continuum mechanics and empirical results. This is changing, however, as computer-based numerical techniques allow more effective modeling from first principles.

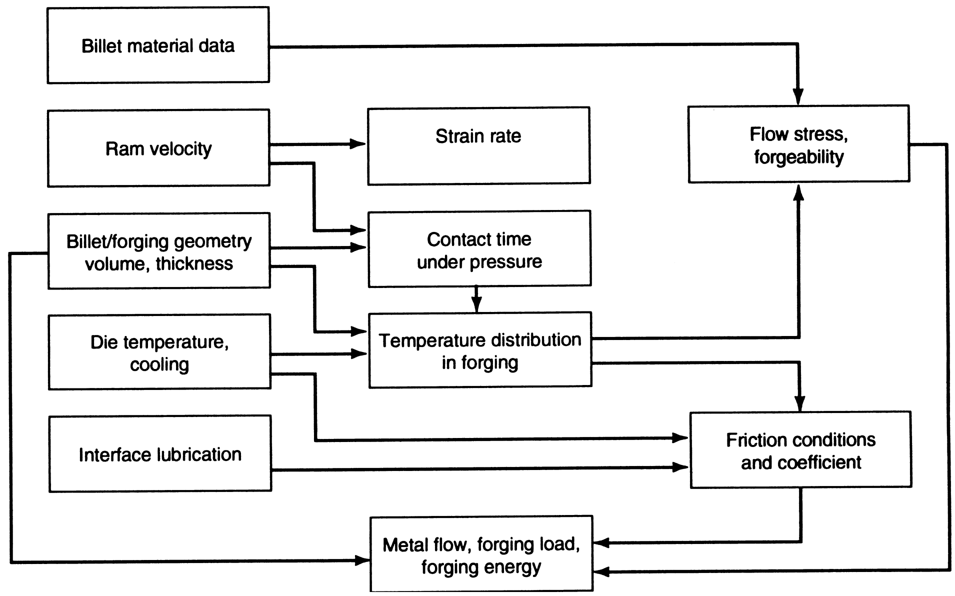


Fig. 17 Interaction among major process variables during forming

Types of Differential Equations. The physical response or behavior of a system, such as the plastic deformation of material during bulk working processes, can be described by differential equations. Partial differential equations are required when the behavior is a function of time and space or of more than one space variable. However, if the problem can be simplified to one independent variable (time or one space variable), ordinary differential equations can be used in understanding the effect of certain parameters on the response or behavior of the process or object.

Partial differential equations can be divided into three categories: hyperbolic, parabolic, and elliptic. Standard hyperbolic equations include the wave equation:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

where c is the wave speed. Parabolic partial differential equations include the diffusion equation:

$$D\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

where D is the diffusivity.

Elliptic equations are usually used to model steady-state phenomena. When hyperbolic or parabolic equations are assumed to be invariant with time, then they reduce to elliptic equations. For example, when the time dependence is removed from the wave equation or the diffusion equation, they reduce to the Laplace or Poisson equations for steady-state heat conduction in solids with constant properties:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \qquad \text{(Laplace equation)}$$

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + \frac{g(x,y,z)}{k} = 0 \qquad \text{(Poisson equation)}$$

where the Possion equation includes a heat-source function, $g(x, y, z)$, with k as the thermal conductivity of the material.

Continuum Mechanics Equations. Continuum mechanic models are based on steady-state equations of mechanical equilibrium and constitutive equations for mechanical flow. The complete set of equations serves as the foundation for continuum mechanics models of bulk working processes. A model for each case is developed by imposing appropriate boundary conditions and initial conditions (tool and workpiece geometry, temperature, heat flow, etc.) on the solutions for the set of equations. In addition, models of thermomechanical processes may also require description of thermophysical behavior and the contact interface between the tool and worked material.

Equilibrium and Compatibility (Strain-Displacement) Equations. Steady-state equations include the forces acting on an element and the compatibility of strain displacements. The steady-state equations in describing the various forces acting on an element in mechanical equilibrium are:

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} = -F_x$$
$$\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} = -F_y$$
$$\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_z}{\partial z} = -F_z \qquad \text{(Eq 1)}$$

where σ is the normal stress component, τ is the shear stress component, and F is the body force/unit volume component. Similarly, the strain-displacement relationships are given as:

$$\varepsilon_x = \frac{\partial u}{\partial x} \quad \gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}$$
$$\varepsilon_y = \frac{\partial v}{\partial y} \quad \gamma_{yz} = \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}$$
$$\varepsilon_z = \frac{\partial w}{\partial z} \quad \gamma_{zx} = \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \qquad \text{(Eq 2)}$$

where ε is the normal strain, γ is the shear strain, and u, v, and w are the displacements in the x, y, and z directions, respectively. In addition to boundary conditions, the solution of these equations may be complicated further by time dependence of the force functions, the nonlinear stress-strain behavior of plastic deformation (i.e., constitutive equations), and the flow rules of plastic deformation (such as the Von Mises yield citerion) under combined stresses.

Constitutive Equations. For a given material, the relations between stress components in Eq 1 and strain components in Eq 2 are given by the constitutive equations representing the behavior of that material. The simplest example of a constitutive equation is the well-known Hooke's law in the elastic regime:

$$\sigma = \varepsilon E$$

where E is the elastic modulus of the material, which is measured in a simple tension test or by ultrasonic means.

During plastic deformation of most metallic materials, the stress-strain curve becomes nonlinear, because hardening (or less frequently softening) of the material can occur when continuing plastic strain is built up. The strain rate can also influence the hardening or softening of a material. The general form of the constitutive equation for deformation processing is:

$$\bar{\sigma} = f(\bar{\varepsilon}, \dot{\bar{\varepsilon}}, T)$$

where $\bar{\sigma}$ is the equivalent (or effective) combined stress, $\bar{\varepsilon}$ is the equivalent true strain, $\dot{\bar{\varepsilon}}$ is the equivalent true strain rate, and T is the processing temperature. Most software packages for bulk forming modeling have options to input the testing data in a tabular form or as a constitutive equation. The tabular form is easy to use but is not based on metallurgical principles, as with some constitutive equations.

The most frequently used constitutive equation is:

$$\bar{\sigma} = K\,\bar{\varepsilon}^n\,\dot{\bar{\varepsilon}}^m + Y$$

where n is the strain hardening exponent, m is the strain-rate sensitivity, and Y and K are coefficients. Strain-rate sensitivity is important at elevated temperatures, while it has little influence at room temperature for most metallic materials. In contrast, the importance of the strain-hardening exponent becomes more significant with decreasing temperature.

The most immediately preceding equation does not reflect the influence of temperature. For each temperature, there is a set of equations. A more fundamentally sound equation has been proposed by Sellars and Tegart (Ref 13) by assuming materials flow during deformation as a thermally activated process:

$$\dot{\bar{\varepsilon}} = A\,[\sinh(\alpha\bar{\sigma})]^{n'} \exp\left(-\frac{Q}{RT}\right)$$

where A, α, and n' are constants determined by fitting empirical data, and Q is the apparent activation energy. At low stresses ($\alpha\bar{\sigma} < 0.8$), the equation reduces to a power law:

$$\dot{\bar{\varepsilon}} = A_1 \bar{\sigma}^{n'} \exp\left(-\frac{Q}{RT}\right)$$

At high stresses ($\alpha\bar{\sigma} > 1.2$), the equation reduces to an exponential form:

$$\dot{\bar{\varepsilon}} = A_2 \exp(\beta\bar{\sigma}) \exp\left(-\frac{Q}{RT}\right)$$

where $\beta = \alpha n'$.

Other constitutive relations have been proposed to describe dynamic recovery and dynamic recrystallization, such as the Laasoui-Jonas model (Ref 14) and the internal variable model (Ref 15). In all cases, constitutive equations are empirical-based relations derived from the reduction of test data.

Yield Criteria (Flow Rules). The theory of continuum plasticity involves the definition of yield criterion for when a material yields plastically or flows. In structural analysis, yield criterion may be characterized as a "failure theory," because plastic deformation is an undesired outcome in structural design. In bulk deformation, yielding of the workpiece is intended, and yield criteria are used in the modeling of flow under combined stresses.

The continuum mechanics of metallic materials includes several theories for yielding, as described in a historical sketch of continuum plasticity theory with an introduction on computational methods in solid mechanics (Ref 16). The first yield criterion for metals was proposed by Henri Tresca in the 1860s. The Tresca criterion is based on the premise that yielding is dependent on just shear stresses, whereby plastic flow begins when the shear stresses exceed the shear yield strength of the metallic material. Although the Tresca yield criterion is adequate, it neglects the intermediate principal stress, σ_2.

The Levy-von Mises yield criterion is considered to be a more complete and generally applicable yielding criteria. It is based on the theory of Richard von Mises that incorporated a proposal by M. Levy, which stated that the tensor components of plastic-strain increments are in proportion to each other just as are the tensor components for deviatoric stress. It is based on the second tensor invariant of the deviatoric stresses (that is, of the total stresses minus those of a hydrostatic state with pressure equal to the average normal stress over all planes) (Ref 16). The von Mises yield criterion is thus expressed as:

$$2\sigma_0^2 = (\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2$$

where σ_0 is the uniaxial yield (flow) stress measured in tension or compression, and σ_1, σ_2, σ_3 are the three principal stresses.

Fracture Criteria. Continuum mechanics can also be used in conjunction with multiaxial fracture criteria to gain insights and solutions on the prevention of flow related cracks. Various fracture (or microstructural damage) criteria have been developed to evaluate workability (e.g., Ref 17–21), but not all are easily incorporated into continuum mechanics models of processes. One criterion that gives a very accurate description of workability and is easily implemented in the models is that due to Crockcroft and Latham (Ref 17):

$$\int_0^{\bar{\varepsilon}_f} \sigma_1\, d\bar{\varepsilon} \geq C$$

where σ_1 is the maximum principal stress, $\bar{\varepsilon}_f$ is the equivalent strain at fracture, and C is a constant representing the workability of the material. If $\sigma_1 < 0$, then there are only compressive stresses and no fracture occurs. The Crockcroft-Latham criterion is phenomenological based rather than mechanistic, but it captures the physical concept and intuitive understanding that damage accumulation, or workability, is dependent both on the degree of plastic deformation (represented by the integral over effective strain) and tensile stress (represented by the maximum principal stress). Furthermore, both of these physical quantities are readily available outputs of continuum mechanics models and can be evaluated at every localized region or point throughout the material to determine potential sites of crack initiation. As with any constitutive relation, the value of incorporating this workability criterion into process analyses depends on accurate representation of the actual material behavior, represented by the coefficient C in this case.

A more recent fracture criterion is theorized as (Ref 20):

$$\sigma_{1f} \cdot \frac{\bar{\varepsilon}}{\bar{\sigma}} \geq C'$$

where σ_{1f} is the maximum principal stress at fracture, and C' is a workability constant. This criteria can be related directly to bulk workability tests. For incremental loading, Bandstra (Ref 21) proved that the criterion can be expressed as:

$$\int_0^{\bar{\varepsilon}_f} \frac{\sigma_1}{\bar{\sigma}}\, d\bar{\varepsilon} \geq C'$$

A review of fracture criteria can be found in Ref 20 and some examples are given in Chapter 12, "Workability Theory and Application in Bulk Forming Processes."

Methods of Solution

The general methods of solving the underlying equations of a model depend on the complexity of boundary conditions and process variables such as time, space, and internal states. Developing an analytical or a closed-form solution model may be advantageous in many instances, but numerical techniques may be required. In some circumstances analytic solutions of partial differential equations can be obtained. However, this is only true of simple forms of the equations and in simple geometric regions. For most practical problems, computational or numerical solutions are needed.

There are several techniques for solving differential equations. Numerical algorithms to solve differential equations consist of *lumped-parameter methods* or the so-called *meshed-solution methods* (such as finite element analysis). If the problem can be simplified for the use of ordinary differential equations, then a lumped-parameter model might be used. A lumped-parameter model may help in understanding the effect of certain parameters on the process. However, these methods do not model spatial variation directly, and the parameters do not necessarily always have direct physical significance.

Typically, partial differential equations are required to describe the process in terms of time, space, field variables, and internal states. Several methods have been developed for the numerical solution of partial differential equations. This typically involves meshed-solution methods such as the finite element method, the finite difference method (*FDM*), and the boundary element method (*BEM*). Each has different suitability for different types of problems. For example, the finite difference method is often adopted in fluid mechanics but is seldom used in solid and structural mechanics. Finite element methods are the most common for linear and nonlinear continuum mechanics, although the boundary element method (BEM) has advantages in some applications of continuum mechanics.

These numerical methods provide approximate solutions by converting a complex continuum model into a discrete set of smaller problems with a finite number of degrees of freedom. The meshes are created by using structured elements like rectilinear blocks or unstructured meshes with variable-shaped elements (e.g., tetrahedra, bricks, hexahedral, prisms, and so forth) for better fidelity to the macroscopic conditions or boundaries. Once a discrete-element model has been created, mathematical techniques are used to obtain a set of equilibrium equations for each element and the entire model. By applying various boundary conditions and loads to the model, the solution of the simultaneous set of equations provides the resulting response anywhere in the model while still providing continuity and equilibrium. The process of solution is essentially a computer-based numerical method, where interpolation functions (polynomials) are used to reduce the behavior at an infinite field of points to a finite number of points.

Meshed-solution models have advantages over the typical closed-form solutions, as they more readily give solutions to irregular shapes, variable material properties, and irregular boundary conditions. Construction of a discrete meshed model for numerical solution may be necessary if the modeled volume:

- Has a complex shape (as is common in bulk forming)
- Contains different phases and grains
- Contains discontinuous behavior such as a phase change
- Has a nonlinear physical process such as when the heat transfer coefficient is a nonlinear function of the temperature

In many instances, meshed models are supplemented by some non-meshed symbolic or analytical modeling. This is done in order to decide on appropriate boundary conditions for the meshed part of the problem, because it is the boundary conditions that effectively model the physical problem and control the form of the final solution. Analytic models are always useful for distinguishing between mechanisms that have to be modeled separately or modeled as a coupled set.

Boundary Conditions. Application of appropriate boundary conditions is a major part of the activity of process modeling. Boundary and initial conditions represent geometric factors (e.g., symmetry, tool shape, and workpiece shape), thermal factors (e.g., heat flux and temperature) and loads (e.g., tool pressure and tool friction) pertinent to the particular problem being modeled. These conditions or constraints influence the complexity of the problem and the most appropriate method or algorithm for solving the equations of the model.

For example, consider the simple example of a uniformly loaded cantilever beam (Fig. 18a). In this case (assuming uniform loading and a rigid connection), then closed-formed equations can provide analytical solutions for the deflection and bending stress over the length of the beam (Fig. 18a). However, if the boundary conditions are altered by the addition of simple supports (Fig. 18b), then the system becomes statically indeterminate (i.e., there is no longer a closed-formed solution that specifies deflection over the length of the beam). In this case, numerical techniques (such as finite element analysis) are required to find approximate solutions of deflection and bending stresses. Likewise, the solution of dynamic problems may also require numerical techniques, depending on the complexity of the equations in the models and the appropriate boundary conditions.

In analytical solutions models, boundary conditions need to be set at a very early stage. In meshed-solution models, boundary conditions are typically represented separate from the main equations and decoupled to some extent from the model itself. Therefore, sensitivity analysis can be done much easier using meshed methods.

Material Properties. All models of bulk-working processes require input of accurate material properties so that the fundamental materials behavior can be represented faithfully by the constitutive equations. Acquiring these properties can be difficult and expensive. Sensitivity analysis of the model with respect to variations in property data should be done. In many instances, it may be possible to use models with inexact material property information in order to predict trends, as opposed to determining actual values. Problems may arise if the material properties are improperly extrapolated beyond their range of validity. In deformation modeling, Young's modulus, Poisson's ratio, anisotropic behavior, and flow stress (as functions of stress, strain, strain rate, and temperature) are needed.

Closed-Form and Numerical Methods of Solution. Several types of theoretical methods are available for metal forming analysis, as

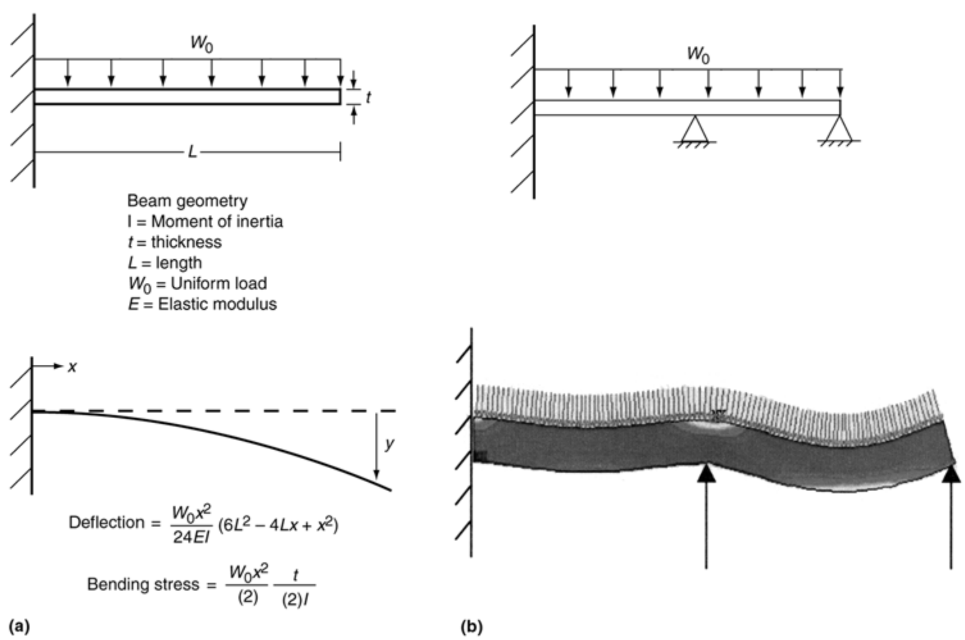


Fig. 18 Effect of boundary conditions on the solution of a cantilever problem. (a) The beam deflection and bending stresses for a uniformly loaded cantilever can be solved by a closed-form equation as shown. (b) A supported cantilever beam is statically indeterminate, and numerical methods are required to approximate defection and bending-stress conditions that are consistent with the boundary conditions established by the additional supports.

described in Ref 22 and 23. Generally, the methods of deformation modeling fall into one of two categories:

- Closed form analytical approximations
- Numerical methods using discrete elements

However, this division is not entirely distinct, as closed-form methods are often solved by discrete-element numerical methods, as well. The general distinction between the two sets of analytical methods is that the first set is based on simplifying assumptions that permit closed form solution as well as rapid numerical and graphical methods. Solving the partial differential equations by numerical methods, in contrast, reduces the need for simplifying assumptions. Regardless of the solution method, continuum mechanics is the starting point. Each problem starts with the continuum equations and then, through various simplifying assumptions, leads to a solution method that may be closed form or numerical. The next two sections of this chapter briefly introduce some typical closed formed methods and numerical techniques.

Approximate and Closed Form Solutions

The equations of continuum mechanics are the basic starting point of a model, and then various simplifying assumptions lead to a solution method that may be closed form or numerical. In general, the boundary conditions in metal forming operations are too complicated for analytical solution of plasticity in the forming of parts with even relatively simple geometric features. However, the need to obtain at least approximate solutions may be satisfied by simplifying assumptions, but each of the analytical approaches has its limitations. The most commonly used approximate methods include:

- The slab method, which restricts the change of stress to only one direction
- The uniform deformation energy method, which neglects redundant work involved in internal shearing due to nonuniform deformation
- The slip-line field solution, which is limited to rigid-plastic materials under plane-strain conditions
- The bounding methods, which can provide fairly good estimates of upper and lower limits of the deformation force but cannot provide details of local stress and strain distributions

These methods are discussed only briefly without attempting a rigorous description of the equations that are solved. More details are available in several other sources (Ref 24–30). These methods are described in detail with examples in the cited references and are presented here in increasing order of complexity.

The Sachs (slab) method is applicable to problems in which one dimension is greater than the others, such as strip rolling, which has a large length of contact zone compared to the small strip thickness, or wire drawing, which has a large length of contact in the die compared to the wire diameter. Under these conditions, the normal stresses can be assumed to vary only in the direction of the large dimension. This reduces the complex set of continuum equations to one ordinary differential equation of equilibrium and one strain equation. The ordinary differential equation can be solved in closed form, or, for complex geometries, it can be solved by numerical methods.

In this approach, the deformation is assumed to be homogeneous, and the force equilibrium equations are set up and solved using an appropriate yield criterion. The slab method is a quick way of obtaining approximate load and strain estimates in axisymmetric and plane-strain problems and is therefore widely used. This method does not account for the contribution of *redundant work,* which is the additional work expended due to inhomogeneous deformation. In this method, stress and strain gradients are accounted for in only one direction and are assumed to be uniform in the perpendicular direction.

Slip Line Method. The slip line field method (Ref 31) utilizes the wave nature of the hyperbolic differential equations that describe plastic deformation to produce a graphical method of solution of the stress field in a two-dimensional plane strain or axisymmetric field. These constructions can be carried manually for problems having simple geometry or by numerical techniques for complex geometries. (Interestingly, the equations describing hypersonic fluid flow are also hyperbolic, and methods similar to slip line fields are used to design rocket nozzles and expanders. The equivalent of slip lines in plastic flow are the shock waves occurring in hypersonic fluid flow.)

The slip line field approach was developed for plane-strain problems. It assumes that the material is rigid and ideally plastic (that is, the material does not strain harden). The theory is based on the fact that any state of stress in plane strain can be represented as the sum of a hydrostatic stress and a pure shear stress. Given the force and velocity boundary conditions, this slip line field is constructed. The main advantage of this method over the slab method is that it can provide local stress calculations even when the deformation is not homogeneous. The major limitations of the slip-line field approach are:

- It is usable only for plane-strain problems.
- It assumes rigid and ideally plastic materials.
- The method is tedious, and solutions are difficult to verify.

This technique has been applied to forging, extrusion, and other processes.

The upper bound method breaks up the deformation volume into a number of simple triangular or rectangular sections and applies the principle of compatibility of deformation along with the limit theorem on power dissipation. Rearrangement or refinement of the deformation sections leads to progressively smaller estimates of the power dissipated, hence the term "upper bound." The equilibrium equations are not solved in this case, but as the upper limit is minimized, satisfaction of equilibrium is approached. With the upper bound method, again, closed form solutions can be developed for simple geometries, while numerical methods are used for complex geometries.

The upper bound method is based on the limit theorem stating that the power dissipated by the boundary forces at their prescribed velocities is always less than or equal to the power dissipated by the same forces under any other kinematically admissible velocity field. A *kinematically admissible velocity field* is one that satisfies the velocity boundary conditions and material incompressibility. This method allows kinematically admissible velocity fields to be set up as a function of an unknown parameter. Power dissipation is then minimized with respect to the unknown parameter to yield a reasonable estimate of load.

The main disadvantage of this method is that the choice of velocity field is rather arbitrary, and the poorer the selection, the more the estimated load will exceed the true load. Another limitation is that no local stress field is compared. The upper bound method does offer a relatively simple way of calculating the major force requirements (Ref. 27).

The lower bound method is not of great practical significance, because forming loads are underestimated. However, it does provide an indication of how conservative the upper bound solution is if the lower bound solution is known. The lower bound approach is based on the limit theorem stating that the power dissipation of the actual surface forces at their prescribed velocities is always greater than the power dissipation of the surface tractions corresponding to any other statically admissible stress field. A statically admissible stress field must satisfy force equilibrium and not violate the yield criterion.

Numerical Techniques in Process Modeling

Numerical solution of the continuum equations is required when the simplifying assumptions of methods described in the preceding are not justified. The general types of numerical techniques used in the solution of partial differential equations includes:

- The finite element method
- The boundary element method
- The finite difference method

With the continued improvement in computer capabilities, these numerical techniques have become very effective in the engineering analysis of static problems, dynamic conditions (where the calculation of inertial and/or damping forces involves derivatives with respect to time), or quasi-static conditions where rate-dependent plasticity may require a realistic estimation of time (but where inertial forces may still be neglected.). The general theory and practical use of the methods are described in Ref 32.

In applications involving continuum mechanics, the finite element method is the most common technique, although the boundary element method (BEM) has advantages in some applications of continuum mechanics. The finite difference method has proven to be useful in fluid and thermal problems but is seldom used in solid and structural mechanics. Nonetheless, the finite difference method is a simple and efficient method for solving ordinary differential equations in problem regions with simple boundaries. For each node of the mesh, the unknown function values are found, replacing the relevant differential equation, (i.e., $dy = f[x,y]dx$) by a difference equation:

$$\Delta y = f(x + \Delta x/2, y + \Delta y/2)\Delta x$$

where Δx Δy are steps in an iterative procedure. Hyperbolic and parabolic partial differential equations are often solved using a hybrid of the FEM and FDM. The spatial variables are modelled using the FEM, and their variation with time is modelled by the FDM.

Other numerical techniques include:

- Finite-volume methods, which are important in highly nonlinear problems of fluid mechanics
- Spectral methods, which are based on transforms that map space and/or time dimensions to spaces where the problem is easier to solve
- Mesh-free methods, which are a recent development of finite difference methods with arbitrary grids

These techniques and the finite difference method are only mentioned for reference without further discussion. The finite element and boundary element methods are more common in the modeling of bulk deformation processes, as described later in this chapter in more detail. Application of these computer models has become an essential tool in meeting product requirements of dimensional tolerances, surface finish, and consistency of properties.

Boundary Element Method

The boundary element method (BEM) is a technique for representing a complex structure or component as a computer model in order to determine its response to a set of given conditions. Like the finite element method (FEM), the model is formed by subdividing the structure into small elements to form the overall model. However, unlike the FEM, only the surface (or boundary) of the problem requires subdivision, thereby reducing the dimensionality of the problem and thus dramatically reducing the computational effort in obtaining a solution.

The BEM has a more restricted range of application than FEM has. In general it is applicable mainly to linear elliptic partial differential equations. It also requires reformulation of the governing partial differential equations into a Fredholm integral equation, which applies to a range of physical problems. For example, elliptic partial differential equations like the Laplace or Helmholtz equations can be reformulated as Fredholm integral equations and then be solved by the BEM. The advantage is that the mesh need only cover the boundaries of the domain.

The BEM can be an effective tool in the analysis of various metal forming problems in rolling and extrusion. For example, Ref 33 demonstrates that the BEM can be used to efficiently and accurately analyze planar and axisymmetric forming problems involving both material and geometric nonlinearities, along with complicated interface conditions. Like FEM, the use of BEM in the modeling of metal forming operations may require consideration of elastic-plastic behavior and elastic-viscoplastic problems involving large strains. Elastic strains are assumed to be small, while nonelastic strains (plastic or viscoplastic) are presumed to be large. When strains become large, the original mesh may become so distorted that the interpolation polynomials are incapable of modeling the geometry of the elements and their relevant state variables. This requires a process of remeshing.

Finite-Element Analysis

Finite element analysis (FEA) is a computer-based analytical technique that allows numerical solutions to be obtained for complex mathematical and engineering problems by creating a discrete or finite number of individual nodes and elements. Discrete elements fill the appropriate geometry, and the method enables the systematic solution of equilibrium equations for each element and the entire model with as much fidelity to geometry as needed.

Finite element analysis is a powerful analysis tool that is flexible for solving problems with irregular shapes, variable material properties, and irregular boundary conditions. With advancements in computer technology, the use of numerical methods such as FEA has grown. The method was originally developed for structural problems (stress displacement of complex geometries), but the same concepts and principles apply to other kinds of engineering problems such as those listed in Table 1. In deformation processes, FEA is a useful tool in die design and process analysis. Common problems solved by FEA include insufficient die filling, poor shape control, poor flow of material, cracks and voids that lead to fracture, and inadequate properties from microstructural variations (grain size).

The application of finite element modeling (FEM) in metal forming (e.g., Ref 2, 23, 34) has brought great changes to design methodologies that were formerly based on trial and error approaches. For example, shape changes during forging are easily predicted by FEM. The detailed temperature, strain, and strain-rate histories at each individual material point in a workpiece are also available from FEM simulations. Simulation of thermomechanical processes can be based on models of continuum mechanics, transport phenomena (heat flow), and metallurgical phenomena (e.g., grain growth and recrystallization).

Table 1 Engineering problems addressed by the finite element method

Discipline	Typical unknown	Possible boundary conditions
Structural	Displacement	Stress or displacement
Thermal	Temperature	Heat flux or convective term or radiative term
Electrical	Voltage	Current source
Magnetic	Electromotive force	Magnetic field source or intensity
Fluid flow	Pressure, velocity	Velocity
Diffusion (Fickian)	Mass concentration	Flux of species
Diffusion (porous media)	Flow velocity	Boundary flow
Corrosion	Anode consumption rate	Electropotential
Crack propagation	Strain energy release rate	Stress
Acoustic noise	Sound pressure level	Velocity

The range of FEA applications in the area of material processing is extremely wide, and a brief review of the finite element techniques applied to metal forming, nonmetal forming, and powder metallurgy are briefly discussed in Ref 35 with an encyclopedic view of the different possibilities in these various fields of application. Many texts (e.g., Ref 2, 23, 34) also describe finite element theory and how it is used in forming analyses. Models may be based on continuum mechanics in the evaluation of flow and/or fracture problems supplemented by more sophisticated thermomechanical/thermophysical models for the simulation of microstructural evolution.

Finite element models of flow during deformation processes are based on the equations of mechanical equilibrium and flow behavior, as previously described in the section "Continuum Mechanics Equations" in this chapter. As noted, constitutive equations are mathematical expressions that describe stress-strain curve in the region of nonelastic (plastic or viscoplastic) flow. Constitutive equations for plastic deformation are typically nonlinear, as plastic deformation of materials is generally a function of the strain hardening/softening and the strain-rate hardening/softening response of a material for different conditions of temperature, stress, and microstructure. Constitutive equations are required for realistic modeling, and they are unique for each material under each processing condition. Constitutive equations may be developed from data obtained under simplified experimental conditions or from thermomechanical testing. Constitutive equations are then extended to more complex situations of combined stresses by well-known hypotheses of flow rules (e.g., von Mises yield criterion) from the theory of continuum plasticity.

Finite element methods, as they apply to metal forming analysis, can be classified into either elastic-plastic or rigid-viscoplastic methods, depending on the assumptions made with regard to the material flow behavior. The elastic-plastic method assumes that the material deformation includes a small, recoverable elastic part and a much larger, nonrecoverable plastic part that is time independent. In contrast, the rigid-viscoplastic method assumes time-dependent deformation behavior (i.e., creep or viscous behavior).

Elastic-plastic FEA can give details regarding deformation loads, stresses and strains, and residual stresses. Reference 23 describes in detail elastic-plastic FEA and how it is used in forming analyses. The elastic-plastic FEM has been applied to a large variety of problems, including upsetting (Ref 36 and 37), indentation (Ref 38), rolling (Ref 11, 39), extrusion (Ref 40), and expansion of a hole in a plate (Ref 41). However, because of the large change in the material flow behavior between elastic and plastic deformation and the need to check the status of each element, the deformation steps must be small, and this makes the method uneconomical.

The rigid-viscoplastic FEA method assumes that the deformation stresses are primarily dependent on deformation (strain) rates. An early example of a rigid-viscoplastic FEM is the ALPID (Analysis of Large Plastic Incremental Deformation computer program) (Ref 42). This program uses the approach of Kobayashi et al. (Ref 43) and includes the incorporation of convenient features including capabilities for handling arbitrary die geometries and remeshing.

Although predictions regarding residual stresses cannot be made with the rigid-viscoplastic FEM, the larger steps that can be used in modeling metal forming procedures make the method very economical, especially for modeling hot deformation. Another example is a rigid-viscoplastic model used to predict surface defects from deformation processes (Ref 44) where defect initiation and development are predicted and some critical processing parameters are identified. Other application examples of rigid-viscoplastic models are described in Ref 45–49.

Regardless of the flow model, the modeling of deformation processes by FEM (and BEM) typically requires techniques for the remeshing of elements. When the original mesh becomes highly distorted, the interpolation functions are incapable of modeling the geometry and state variables. Large strains in finite element simulations of metal forming require several remeshing procedures during the complete simulation cycle. These remeshings may be performed automatically both for saving computational time and for convenience. The methods of remeshing involve the generation of new elements by using the interpolation functions, as briefly described in Ref 23. Automatic remeshing algorithms have been well established for two-dimensional (2D) applications but have also been established for some three-dimensional (3D) cases (e.g., Ref 50). Other examples of automatic remeshing techniques are given in Ref 51–53.

Remeshing algorithms, whether in 2D or 3D, offer the possibility of using a finer grid in certain areas of the finite element mesh. These remeshing/refinement algorithms can play an important role in the simulation of microstructure. In general, it is preferable to define loads and boundary conditions relative to the geometry rather than to the finite element mesh. This permits remeshing of the geometry without the need to redefine loads and boundary conditions. It is also desirable to specify hard points, lines, or surfaces where parts will interact with other parts, so that the interaction can be properly represented.

Example 1: Use of FEA to Study Ductile Fracture during Forging (Ref 54). Large nickel-copper alloy K-500 (K-Monel) shafts sometimes revealed fractures after forging and heat treating (Ref 55). Inspection of failed fracture surfaces revealed that the cracks were intergranular and occurred along a carbon film that had apparently developed during a slow cooling cycle (such as during casting of the ingot stock). Cracks were observed in both longitudinal and transverse directions and arrested just short of the outer bar diameter. Finite element analysis was used to determine the cause and remedy for this cracking. In addition, new processing parameters were determined to eliminate cracking.

A 2D model of the bar was developed and discretized into a finite element mesh. A public-domain, nonlinear FEA code from Lawrence Livermore National Laboratory (Ref 56), NIKE2D, was used to calculate the stress plots under forging conditions. The stress/strain histories resulting from forging were required for ductile fracture analyses. In this particular case, the Oyane et al. ductile fracture criteria (Ref 18) were used (Ref 57). These criteria attempt to predict ductile fracture based on a porous plasticity model.

Figure 19 (Ref 55) is a plot of in-process hydrostatic stress history at the center of a bar during forging with two rigid, parallel flat dies. (Hydrostatic stress is the mean of any three mutually perpendicular normal stresses at a given point. It is invariant with direction.) Various strain rate and temperature conditions are shown having similar trends. At very low reductions (1.25%), the center of the bar is still within the elastic range and, therefore, does not experience large stresses. At 2.5% reduction, the outward movement of the free sides of the bar produces a high tensile hydrostatic stress at the center of the bar. As reduction continues, the compressive stresses acting in the vertical direction begin to grow. This acts to reduce the hydrostatic stress at the center. When the bar has been reduced by 11%, the hydrostatic stress is compressive in all cases. Beyond this reduction, the hydrostatic stress at the center of the bar continues to grow in a compressive manner.

Figure 20(b) (Ref 55) shows a map of ductile fracture accumulation after one forging pass for air-melted material at 930 °C (1706 °F) and a strain rate of 10 s^{-1}. Notice that the maximum damage is located at the center of the bar. After this pass, the bar would be rotated and further reductions would be taken. Damage would accu-

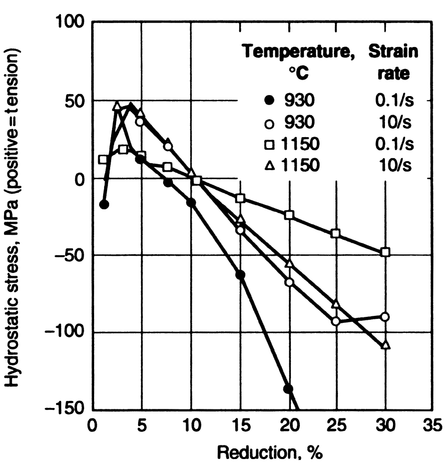


Fig. 19 Hydrostatic stress at the bar center during forging. Source: Ref 55

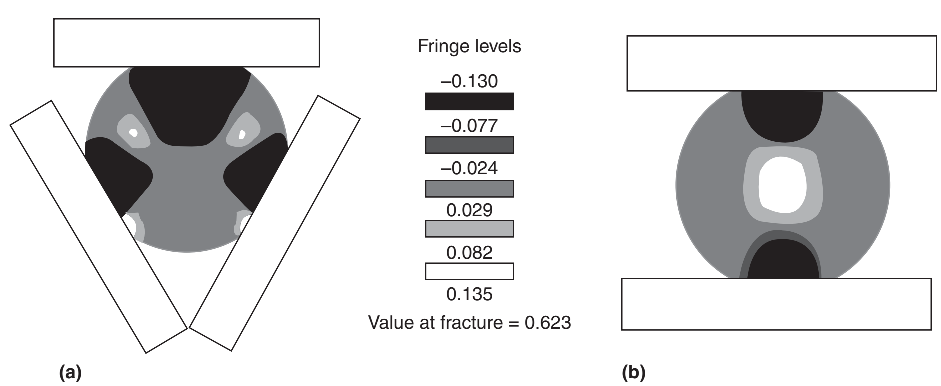


Fig. 20 Ductile fracture map of a nickel-copper alloy K-500 bar at 10% reduction, 930 °C, and a strain rate of 10.0 s^{-1} using the Oyane et al. ductile fracture criteria. (a) Three forging dies. (b) Two forging dies. Source: Ref 55

mulate during these subsequent passes. Ductile fracture was predicted to occur at the bar center from the FEA results after six reducing passes under the specified conditions. One possible way of eliminating this ductile fracture is to use a V-shaped die assembly as shown in Fig. 20(a). In this case, the fracture does not occur at the center but at the edges and gets an opportunity to heal as the bar is rotated during subsequent passes.

Microstructural Modeling

Microstructural modeling is an active area of practical importance in the design and analysis of thermomechanical processing. One early example is the pioneering work of Sellars and coworkers, who developed methods for the microstructural modeling of carbon-manganese steels in hot rolling (Ref 58–61). Similar methods of microstructural modeling also have been developed for hot rolling operations by others (Ref 62). Though the details of the models differ from group to group, these models emphasize the importance of quantifying the thermal and metallurgical phenomena. Each basically involves the discrete characterization of dynamic structural changes (e.g., dynamic recrystallization, static recrystallization, and grain growth) by using empirical equations based on fundamental physical metallurgy principles. Such physical metallurgy-based models grasp the essence of the microstructural changes.

Models for process design continue to advance from data-intensive empirical methods to more knowledge-based analytical/numerical methods using the fundamental principles of transport phenomena, continuum mechanics, and physical metallurgy. By this means, expensive trials can be minimized, and robust processes to make high-integrity parts within tight tolerances can be readily designed. The ability to accurately predict thermomechanical histories is thus being used to predict the evolution of microstructure by coupling microstructure models to FEM results. For example, Ref 63 describes carefully designed workability tests and FEM simulations to develop a general model for microstructural evolution during thermomechanical processing of a nickel-base superalloy. This process-modeling tool was then applied to predict the percentage of recrystallization and grain size in actual Waspaloy forgings.

Quantification of the evolution of microstructure during industrial processing (as described in more detail in Chapter 3, "Evolution of Microstructure during Hot Working") is increasingly recognized as an essential tool for metal producers and fabricators. One major challenge is the need to understand the microstructural consequences of processing variables in different elements of a mesh model. This is discussed in Ref 64 for different types of models applied to hot rolling, where the success and limitation of predicting variations in local microstructures are described for microstructural models based on FEMs. Understanding the microstructural consequences of different strain paths, thermal histories, and strain rates remains a major challenge for physical metallurgy if the full predictive capabilities of finite element models are to be realized.

Processing Maps

Workability is also evaluated by the use of processing maps that show the occurrence of damage as a function of processing variables such as temperature and strain-rate sensitivity. This approach, however, involves extensive testing, and it is difficult to account for all the variable factors in mechanistic models. For example, the location of the fracture boundaries in processing maps is very sensitive to microstructure, prior thermomechanical history, and numerous material parameters such as diffusivity. Nonetheless, processing maps, such as that shown in Fig. 21 (Ref 65), are a very useful guide for the selection of deformation processing conditions.

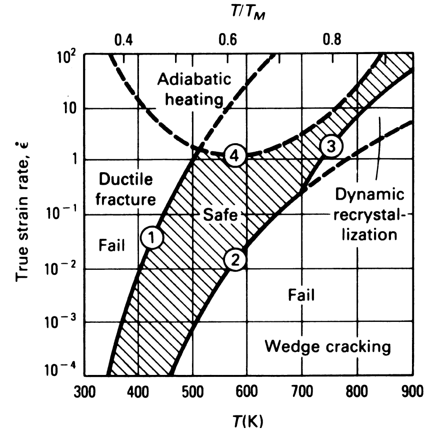


Fig. 21 Composite processing map for aluminum showing the safe region for forming. Boundaries shift with microstructure. Instabilities due to purely continuum effects, such as shear localization in sheet metal forming, are not considered. Source: Ref 65

Dynamic Material Modeling

Processing maps are also used in conjunction with a method known as dynamic material modeling (DMM), which is a top-down approach that begins with macroscopic determination of flow stress as a function of temperature, strain rate, and strain and ends with a microscopic evaluation of the microstructure and the final properties of the forging (Ref 66). This method maps the power efficiency of the deformation of the material in a strain rate/temperature space. The DMM methodology describes the dynamic path in response to an instantaneous change in strain rate at a given temperature and strain. As such, it describes how workpiece material dissipates applied power by various metallurgical processes, depending on the instantaneous change in strain rate at a given temperature and strain. By measuring flow stress as a function of strain, strain rate, and temperature and by calculating strain rate sensitivity at each value of temperature (T) and strain rate ($\dot{\varepsilon}$), it is then possible to identify boundaries between safe and unsafe regions on a processing map (Fig. 22) (Ref 67). The safe regions on the DMM processing map correspond to conditions of DMM stability (see the following paragraphs), which is determined by Liapunov stability analysis.

The DMM method allows modeling for when stable deformation and microstructure evolution can be expected. A collection of DMM processing maps is contained in Ref 68, and the application of DMM concepts is discussed in some chapters of the section "Multidisciplinary Process Design and Optimization" in this Handbook. In its general form, the dynamic material model is sometimes considered to be a generalized concept of workability that plays a role in unifying the relationships among constitutive behavior, hot workability, and microstructure development (Fig. 23) (Ref 69). Dynamic material modeling was developed in response to the need for a macroscopic description of flow, fracture, and workability of complex engineering materials under hot-working conditions. As noted, this approach makes use of the constitutive equations for plastic flow determined from hot deformation tests performed at different temperatures and strain rates. Parameters such as the strain-rate sensitivity and temperature sensitivity of the flow stress are related to the manner in which the workpiece dissipates energy instantaneously during hot deformation. The rate of change of stress with strain rate at constant levels of strain (ε) and temperature (T) is known as the strain-rate sensitivity parameter (m), which is defined as follows:

$$m = [\partial(\log \sigma)/\partial(\log \dot{\varepsilon})]_{\varepsilon,T} \quad \text{(Eq 3)}$$

The temperature sensitivity of the flow stress is analyzed in terms of a parameter (s):

$$s = \frac{1}{T}\left[\partial \ln \sigma/\partial(1/T)\right]_{\dot{\varepsilon},\varepsilon} \quad \text{(Eq 4)}$$

Furthermore, an apparent activation energy (Q) can be determined from computations of m and s using:

$$Q = \frac{sRT}{m} \quad \text{(Eq 5)}$$

where R is the universal gas constant.

Metallurgical Interpretation of DMM Stability (Ref 65). The regions of DMM stability correspond to areas on a DMM map where the energy dispersion processes of the material are in a steady state. The DMM stability criteria are:

$$0 < m \leq 1 \quad \text{(Eq 6)}$$

$$\partial m/\partial(\log \dot{\varepsilon}) < 0 \quad \text{(Eq 7)}$$

$$s \geq 1 \quad \text{(Eq 8)}$$

$$\partial s/\partial(\log \dot{\varepsilon}) < 0 \quad \text{(Eq 9)}$$

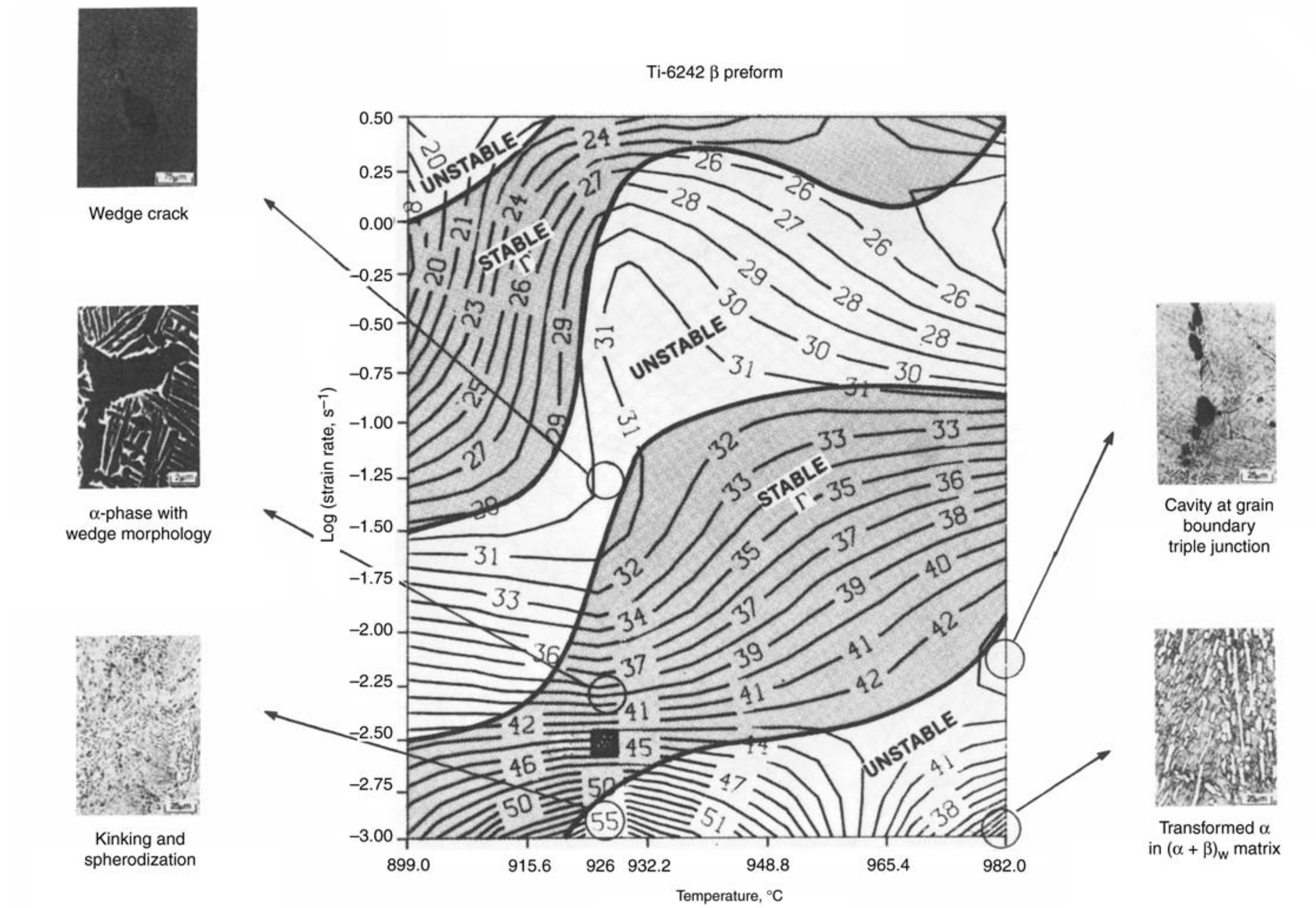


Fig. 22 Processing map for Ti-6242 β microstructure with stable regions identified. Source: Ref 67

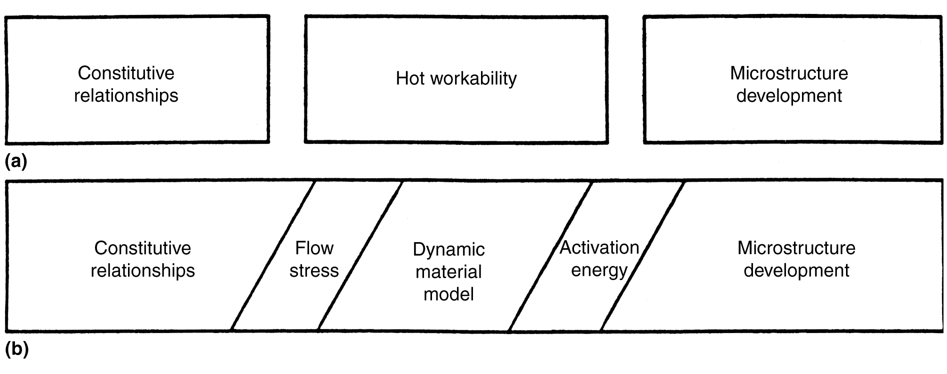


Fig. 23 Models of material behavior in hot deformation studies. (a) Discrete approach where flow and microstructure development are separate in the evaluation of workability. (b) The DMM approach that is linked with mechanical flow and the activation energy of microstructural development. Source: Ref 69

The range of strain-rate-sensitivity (m) values for stable material flow (Eq 6) is derived from theoretical considerations of the maximum rate of power dissipation by material systems and experimental observations. Metals and alloys generally satisfy this criterion under hot working conditions. Negative values of the strain-rate-sensitivity parameter are obtained only under conditions promoting dynamic strain aging (a product of the interaction of mobile dislocations and solute atoms). However, the imposed strain rates for hot working operations are, in general, sufficiently high that such interactions are avoided.

Another mechanism that can lead to negative m values is the dynamic propagation of preexisting or newly formed microcracks in the workpiece, which eventually lead to fracture of the workpiece. As the value of m increases, the tendency for localized deformation decreases, enabling extensive elongation of specimens under tensile loading without necking. Similarly, the occurrence of shear-band formation can be inhibited very effectively at high m values. For metals and alloys, $m = 1$ represents ideal superplastic behavior resulting from the Newtonian flow typical of a glass material.

The stability criterion relating to the variation of m with log $\dot{\varepsilon}$ (Eq 7) stems from the theoretical requirement for the material system to continuously lower its total energy. If fracture stress is assumed to be independent of strain rate, then an increasing $m(\dot{\varepsilon})$ will probably lead to catastrophic failure at high strain rates. In contrast, a decreasing $m(\dot{\varepsilon})$ has a lower probability of inducing fracture in the workpiece. Moreover, this stability criterion leads to more uniform stress fields across the workpiece and a decreased tendency for strain localization, which is clearly desirable.

The lower limit of s values (Eq 8) is derived from the premise that the net entropy-production rate associated with irreversible processes must be positive for stable conditions. It is well established that when dynamic recovery alone operates as a softening mechanism, the temperature dependence of flow stress is relatively weak. In contrast, when both dynamic recrystallization and dynamic recovery are in operation, the flow stress varies markedly with temperature. Thus, low values of s are indicative of the dynamic recovery processes, while high values of s are usually associated with dynamic recrystallization processes.

The stability criterion relating to the variation of s with log $\dot{\varepsilon}$ (Eq 9) arises from the necessity for the material system to continuously lower its total energy. As the strain rate increases, the extent of adiabatic heating in local regions increases significantly. Hence, in any particular region of the workpiece, if the local strain rate increases above the nominal value, then it is accompanied by a reduction in the local flow stress. However, if $s(\dot{\varepsilon})$ increases with $\dot{\varepsilon}$, then a very significant thermal softening will be encountered in the regime of high strain rates, which will produce severe strain localization and adiabatic shear bands. This autocatalytic process is likely to lead to severe cracking of the workpiece if the material has poor resistance to nucleation and growth of cracks. On the other hand, compliance with this stability criterion has a mitigating influence on this tendency for flow localization.

Thus, the four stability criteria of the DMM approach have a sound metallurgical basis. Since the material behavior is both dynamic and stochastic, these criteria should be regarded as probabilistic indicators of the behavioral trends exhibited by a material during hot working. The effects of these instabilities on the workability of a material can be alleviated substantially by proper control of the die design. For instance, die designs that result in a high ratio of mean-to-effective stress can, in part, offset the adverse effect of positive values of $\partial m/\partial(\log \dot{\varepsilon})$. This observation is exemplified by the remarkable successes associated with processes such as hydrostatic extrusion, pack rolling, and closed-die forging.

Summary

Workability refers to the relative ease with which a material can be shaped through plastic deformation. It is a complex technological concept that depends not only on the fracture resistance (ductility) of the material but also on the specific details of the stress state generated by the deformation process. The chief process variables that determine the stress state are the workpiece geometry, the die geometry, and lubrication conditions. The material fracture resistance is determined by its composition, microstructure, and freedom from inclusions, as well as important process parameters such as the working temperature and the speed of deformation.

Fractures found in deformation processing can be classified as free surface fracture, die contact surface fracture, and internal fracture such as center bursts. The mechanism by which the fracture originates depends on the temperature and strain rate of the deformation.

Other workability related defects are not due to fracture, but are related to the distribution of the metal in the deformation process. These are described by such terms as die underfill, laps and folds, extrusion-type defects, flow-through defects, and seams.

Deformation processing not only changes the shape of the material, but it also changes its mechanical properties. Ideally, these changes result in property improvement. Improvement results from the closure of pores and voids and the refinement of grain size. Metal that is rolled, forged, or extruded, however, develops and retains a fiberlike structure that is aligned in the principal direction of working. There also may be crystallographic reorientation of grains. These structural changes result in properties that vary with orientation to the fiber structure.

Proper design of process tooling and optimum selection of material and material process conditions can do much to minimize workability defects and optimize the properties resulting from the deformation. The merging of metallurgy with mechanics and the availability of powerful computer-based models has greatly enhanced our ability to achieve favorable workability, even for hard-to-work materials. This chapter has laid the foundation for understanding the concepts and mathematics of deformation modeling that will be useful as subsequent chapters discuss specific processes and examples. The development follows a historical order, starting with the slab method, advancing to the slip-line field theory, the upper and lower bound methods, the boundary element method, and finally to the widely used finite element method. It also reviews current work on the use of computer modeling to predict the microstructures produced from thermomechanical processing and to predict the regions of temperature and strain rate that are most favorable for achieving good workability.

ACKNOWLEDGMENTS

Portions of this chapter were adapted from the following *ASM Handbook* articles:

- J. Schey, Manufacturing Processes and Their Selection, *Materials Selection and Design,* Vol 20, *ASM Handbook,* ASM International, 1997, p 687–704
- L. Ferguson, Design for Deformation Process, *Materials Selection and Design,* Vol 20, *ASM Handbook,* ASM International, 1997, p 736–737
- H.L. Gegel et al., Modeling Techniques Used in Forging Process Design, *Forming and Forging,* Vol 14, *Metals Handbook,* 1988, p 425

The editors also thank Jim Malas for providing the text from Ref 65 on "Metallurgical Interpretation of DMM Stability."

REFERENCES

1. J. Schey, Manufacturing Processes and Their Selection, *Materials Selection and Design,* Vol 20, *ASM Handbook,* ASM International, 1997, p 687–704
2. R.H. Wagoner and J.-L. Chenot, *Metal Forming Analysis,* Cambridge University Press, 2001
3. L. Ferguson, Design for Deformation Process, *Materials Selection and Design,* Vol 20, *ASM Handbook,* ASM International, 1997, p 736–737
4. P.W. Lee and H.A. Kuhn, Cold Upset Testing, *Workability Testing Techniques,* G.E. Dieter, Ed., American Society for Metals, 1984, p 37–50
5. S.L. Semiatin, Workability in Forging, *Workability Testing Techniques,* G.E. Dieter, Ed., American Society for Metals, 1984, p 197–247
6. H.A. Kuhn and G.E. Dieter, Workability in Bulk Forming Processes, *Fracture 1977, Fourth International Conf. on Fracture,* Vol 1, (Waterloo, Canada), 1977, p 307–324
7. H.A. Kuhn and B.L. Ferguson, *Powder Forging,* Metal Powder Industries Federation, 1990
8. V.N. Whittacker, *Nondestructive Testing,* Oct 1971, p 320
9. L.F. Coffin Jr. and H.C. Rogers, Influence of Pressure on the Structural Damage in Metal Forming Processes, *ASM Trans. Quart.,* Vol 60 (No. 4), Dec 1967, p 672–686
10. A. Chamouard, *Closed Die Forging*: Part 1, Dunod, Paris, 1964
11. Z.-C. Lin and C.-C. Shen, Rolling Process Two-Dimensional Finite Element Model Analysis, *Finite Elem. Anal. Des.,* Vol 26 (No. 2), June 1997, p 143–160
12. F.N. Lake and D.J. Moracz, "Comparison of Major Forging Systems," Technical Report AFML-TR-71-112, TRW, Inc., May 1971
13. C.M. Sellars and W.J. McG. Tegart, *Mem. Sci. Rev. Metall.,* Vol 63, 1966, p 731
14. A. Laasroui and J.J. Jonas, Prediction of Steel Flow Stresses at High Temperatures and Strain Rates, *Metall. Trans. A,* Vol 22, 1991, p 1545–1558
15. S.B. Brown, K.H. Kim, and L.A. Anand, An Internal Variable Constitutive Model for Hot Working of Metals, *Int. J. Plast.,* Vol 5, 1989, p 95–130
16. J.R. Rice, Foundations of Solid Mechanics, *Mechanics and Materials: Fundamentals and Linkages,* M.A. Meyers, R.W. Armstrong, and H. Kichner, Ed., John Wiley & Sons, 1999
17. M.G. Cockcroft and D.J. Latham, Ductility and Workability of Metals, *J. Inst. Met.,* Vol 96, 1968, p 33–39
18. M. Oyane, Criteria of Ductile Fracture Strain, *Bull Jpn. Soc. Mech. Eng.,* Vol 15, 1972, p 1507–1513
19. A.M. Freudenthal, *The Inelastic Behaviour of Solids,* John Wiley & Sons, 1950

20. D. Zhao, J.P. Bandstra, and H.A. Kuhn, A New Fracture Criterion for Fracture Prediction in Metalworking Processes, *Concurrent Engineering Approach to Materials Processing,* S.N. Dwivedi, A.J. Paul, and F.R Dax, Ed., TMS, 1992, p 107–119
21. J.P. Bandstra, "3D Extension of Kuhn Surface Fracture Criterion," CTC Memorandum, CTC/JPB-M0469-95, Concurrent Technologies Corporation, 1995
22. H.L. Gegel et al., Modeling Techniques Used in Forging Process Design, *Forming and Forging,* Vol 14, *Metals Handbook,* 1988, p 425
23. G.W. Rowe, C.E.N. Sturgess, P. Hartley, and I. Pillingre, *Finite–Element Plasticity and Metalforming Analysis,* Cambridge University Press, 1991
24. T. Altan, S. Oh, and H. Gegel, *Metal Forming: Fundamentals and Applications,* American Society for Metals, 1983
25. O. Hoffman and G. Sachs, *Introduction to the Theory of Plasticity for Engineers,* McGraw-Hill, 1953
26. S. Kobayashi and E.G. Thomson, Approximate Solutions to a Problem of Press Forging, *Trans. ASME,* Series B, Vol 81, 1959, p 217–227
27. B. Avitzur, *Metal Forming: Processes and Analysis,* McGraw-Hill, 1968
28. W. Johnson and H. Kudo, *The Mechanics of the Metal Extrusion,* Manchester University Press, 1962
29. K. Lange, *Handbook of Metal Forming,* McGraw-Hill, 1985
30. R. Hill and S.J. Tupper, A New Theory of the Plastic Deformation in Wire Drawing, *J. Iron Steel Inst.,* Vol 159, 1948, p 353–359
31. W. Johnson, R. Sowerby, and R.D. Venter, *Plane Strain Slip Line Fields for Metal Deformation Processes: A Source Book and Bibliography,* Pergamon Press, 1982
32. J. Raamachandran, *Boundary and Finite Elements: Theory and Problems,* CRC Press, FL, 2000
33. A Chandra, Analyses of Metal Forming Problems by the Boundary Element Method, *Int. J. Solids Struct,* Vol 31 (No. 12–13), June–July 1994, p 1695–1736
34. S. Kobayashi, S.I. Oh, and T. Altan, *Metal Forming and the Finite Element Method,* Oxford University Press, New York, 1989, p 222–243
35. N. Brannberg and J. Mackerle, Finite Element Methods and Material Processing Technology, *Eng. Comput. (WALES),* Vol 11 (No. 5), Oct 1994, p 413–455
36. C.H. Lee and S. Kobayashi, Analysis of Axisymmetrical Upsetting and Plane-Strain Side-Pressing of Solid Cylinders by Finite Element Method, *J. Eng. Ind. (Trans. ASME),* Vol 93, 1971, p 445
37. I. Pillinger, M. Bakhshi-Jooybari, P. Hartley, and T. A. Dean, Finite Element Simulation and Experimental Study of Hot Closed-Die Upsetting, *Int. J. Mach. Tools Manuf.,* Vol 36 (No. 9), Sept 1996, p 1021–1032
38. C.H. Lee and S. Kobayashi, Elastoplastic Analysis of Plane-Strain and Axisymmetric Flat Punch Indentation by the Finite Element Method, *Int. J. Mech. Sci.,* Vol 12, 1970, p 349
39. Z.-C. Lin and C.-C. Shen, Three-Dimensional Asymmetrical Rolling of an Aluminium Flat Strip Using an Elastic-Plastic Finite Element Method, *Int. J. Comput. Appl. Technol.,* Vol 9 (No. 5–6), 1996, p 281–297
40. A. Mihelic and B. Stok, Tool Design Optimization in Extrusion Processes, *Comput. Struct.,* Vol 68 (No. 1–3), July–Aug 1998, p 283–293
41. J.S. Gunasekera and J.M. Alexander, Analysis of the Large Deformation of an Elastic-Plastic Axially-Symmetric Continuum, *International Symposium on Foundations of Plasticity,* Noordhoff International, 1972, p 25–146
42. S.I. Oh, Finite Element Analysis of Metal Forming Problems with Arbitrarily Shaped Dies, *Int. J. Mech. Sci.,* Vol 17, 1982, p 293
43. S. Kobayashi, Rigid Plastic Finite-Element Analysis of Axisymmetrical Metal Forming Processes, *Numerical Modeling of Manufacturing Processes ASME,* PVP-PB-025, 1977, p 49–68
44. Y. Peng, X. Ruan, and T. Zuo, Defect Prediction during Conform Process by FEM, *J. Mater. Process. Technol.,* Vol 45 (No. 1–4), Sept 1994, p 539–543
45. M.S. Joun, J.H. Chung, and R. Shivpuri, An Axisymmetric Forging Approach to Preform Design in Ring Rolling Using a Rigid-Viscoplastic Finite Element Method, *Int. J. Mach. Tools Manuf.,* Vol 38 (No. 10–11), Oct–Nov 1998, p 1183–1191
46. S. G. Xu, K. J. Weinmann, D. Y. Yang, and J. C. Lian, Simulation of the Hot Ring Rolling Process by Using a Thermo-Coupled Three-Dimensional Rigid-Viscoplastic Finite Element Method, *J. Manuf. Sci. Eng.,* Vol 119 (No. 4A), Nov 1997, p 542–549
47. D. Y. Yang, N. K. Lee, and J. H. Yoon, A Three-Dimensional Simulation of Isothermal Turbine Blade Forging by the Rigid-Viscoplastic Finite-Element Method, *J. Mater. Eng. Perform.,* Vol 2 (No. 1), Feb 1993, p 119–124
48. S. M. Hwang and M. S. Joun, Analysis of Hot-Strip Rolling by a Penalty Rigid-Viscoplastic Finite Element Method, *Int. J. Mech. Sci.,* Vol 34 (No. 12), Dec 1992, p 971–984
49. M. Fu and Z. Luo, The Prediction of Macro-Defects During the Isothermal Forging Process by the Rigid-Viscoplastic Finite-Element Method, *J. Mater. Process. Technol.,* Vol 32 (No. 3), Aug 1992, p 599–608
50. L.H. Tack, R. Schneiders, J. Debye, R. Kopp, and W. Oberschelp, Two- and Three-Dimensional Remeshing, Mesh Refinement and Application to Simulation of Micromechanical Processes, *Conference: Computational Modeling of the Mechanical Behavior of Materials,* 15–16 Nov 1993, published in *Comput. Mater. Sci.,* Vol 3 (No. 2), Nov 1994, p 241–246
51. J. Zhu, M. Gotoh, J. Shang, and Z.G. Sun, An Integrated FEM System Facilitated with Automatic Remeshing, *Metals and Mater.* Vol 4 (No. 4), 1998, p 657–661
52. R. Krishna Kumar and Ch. Pavana, Remeshing Issues in the Finite Element Analysis of Metal Forming Problems, *J. Mater. Process. Technol.,* Vol 75 (No. 1–3), March 1998, p 63–74
53. Y.Y. Zhu, T. Zacharia, and S. Cescotto, Application of Fully Automatic Remeshing to Complex Metal-Forming Analyses, *Comput. Struct.,* Vol 62 (No. 3), Feb 1997, p 417–427
54. A.J. Paul, Modeling of Manufacturing Processes *Materials Selection and Design,* Vol 20, *ASM Handbook,* 1997, p 708–709
55. M.L. Tims, J.D. Ryan, W.L. Otto, and M.E. Natishan, Crack Susceptibility of Nickel-Copper Alloy K-500 Bars During Forging and Quenching, *Proc. First International Conf. on Quenching and Control of Distortion,* ASM International, 1992, p 243–250
56. J.O. Hallquist, "NIKE2D—A Vectorized Implicit, Finite Deformation Finite Element Code for Analyzing the Static and Dynamic Response of 2-D Solids with Interactive Rezoning and Graphics," User's Manual, UCID-19677, Rev. 1, Lawrence Livermore National Laboratory, 1986
57. S.E. Clift, P. Hartley, C.E.N. Sturgess, and G.W. Rowe, Fracture Prediction in Plastic Deformation Processes, *Int. J. Mech. Sci.,* Vol 32 (No. 1), p 1–17
58. C.M. Sellars, *Hot Working and Forming Processes,* C.M. Sellars and G.J. Davies, Ed., TMS, London, 1979, p 3–15
59. C.M. Sellars, *Mater. Sci. Technol.,* 1990, Vol 6, p 1072–1081
60. C.M. Sellars, *International Conf. on Physical Metallurgy of Thermomechanical Processing of Steels and Other Metals, THERMEC 88,* 6–10 June 1988, (Tokyo), I. Tamura, Ed., Iron and Steel Institute of Japan, Tokyo, 1988, p 448–457
61. J.H. Beynon, P.R. Brown. S.I. Mizban, A.R.S. Ponter, and C.M. Sellars, *Proc. of NUMIFORM Conf.,* Aug 25–29 1986, (Gothenburg, Sweden), K. Mattiasson, A Samuelsson, R.D. Wood, and O.C. Zienkiewicz, Ed., A.A. Balkerna, Rotterdam, Holland, p 213–218
62. C. Devadas, I.V. Samarasekera, and E.B. Hawbolt, The Thermal and Metallurgical State of Steel Strip during Hot Rolling, Part III: Miscrostructural Evolution, *Metall. Trans. A,* 1991, Vol 22, p 335–349

63. G. Shen, S.L. Semiatin, and R. Shivpuri, Modeling Microstructural Development during the Forging of Waspaloy, *Metall. and Materials Trans.,* Vol 26A, July 1995, p 1795–1803
64. C. M. Sellars, The Role of Computer Modelling in Deformation Processing., Conference: Fifty Years of Metallurgy: Retrospect and Prospect, New Delhi, India, Nov 1996, *Trans. Indian Inst. Met.,* Vol 50 (No. 6), Dec 1997, p 563–571
65. R. Raj, Development of a Processing Map for Use in Warm-Forming and Hot-Forming Processes, *Metall. Trans. A,* Vol 12, 1981, p 1089–1097
66. Y.V.R.K. Prasad, H.L. Gegel, S.M. Doraivelu, J.C. Malas, J.T. Morgan, K.A. Lark, and D.R. Barker, Modeling of Dynamic Material Behavior in Hot Deformation: Forging of Ti-6242, *Metall. Trans. A,* Vol 15, 1984, p 1883–1892
67. H.L. Gegel, Synthesis of Atomistics and Continuum Modeling to Describe Microstructure, *Computer Simulation in Materials Processing, Proc. of Materials Science Seminar,* ASM International, 1987
68. Y.V.R. K. Prasad and S. Sasidhara, *Hot Working Guide: A Compendium of Processing Maps,* ASM International, 1997
69. J.C. Malas and V. Seetharaman, Using Material Behavior Models to Develop Process Control Strategies, *JOM,* Vol 44 (No. 6), June 1992, p 8–13

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p22-34
DOI:10.1361/hwpd2003p022

Chapter 2

Bulk Workability of Metals

George E. Dieter, University of Maryland

WORKABILITY refers to the relative ease with which a metal can be shaped through plastic deformation. In this chapter, consideration is restricted to the shaping of materials by such bulk deformation processes as forging, extrusion, and rolling. The evaluation of workability of a material involves both the measurement of the resistance to deformation (strength) and determination of the extent of possible plastic deformation before fracture (ductility). Therefore, a complete description of the workability of a material is specified by its flow stress dependence on processing variables (for example, strain, strain rate, preheat temperature, and die temperature), its failure behavior, and the metallurgical transformations that characterize the alloy system to which it belongs.

However, the major emphasis in workability is on measurement and prediction of the limits of deformation before fracture. The emphasis in this chapter is on understanding the factors that determine the extent of deformation a metal can withstand before cracking or fracture occurs. It is important, however, to allow for a more general definition in which workability is defined as the degree of deformation that can be achieved in a particular metalworking process without creating an undesirable condition. Generally, the undesirable condition is cracking or fracture, but it may be another condition, such as poor surface finish, buckling, or the formation of laps, which are defects created when metal folds over itself during forging. In addition, in the most general definition of workability, the creation by deformation of a metallurgical structure that results in unsatisfactory mechanical properties, such as poor fracture toughness or fatigue resistance, can be considered to be a limit on workability.

Generally, workability depends on the local conditions of stress, strain, strain rate, and temperature in combination with material factors, such as the resistance of a metal to ductile fracture. In addition to a review of the many process variables that influence the degree of workability, the mathematical relationships that describe the occurrence of room-temperature ductile fracture under workability conditions are summarized in this chapter. The most common testing techniques for workability prediction are discussed in Chapter 4, "Bulk Workability Testing" in this Handbook. Much greater detail on these tests and on workability tests specific to a particular bulk forming process, such as forging, are given in subsequent chapters of this Handbook.

Stress, Strain, and Stress-Strain Curves

Stress and Strain. This section provides the definitions of stress and strain for readers without a strong background in mechanics. This knowledge is needed to understand coverage in subsequent chapters.

In simplest terms, stress is force per unit area. Two fundamentally different types of stress can be readily identified (Fig. 1). The normal stress (σ) acting on plane *ab* (Fig. 1) is the force perpendicular to the plane divided by its area, A. Alternatively, the normal stress is the axial force (P) divided by the cross-sectional area of the plane perpendicular to the axis of the rod:

$$\sigma = \frac{P\cos\theta}{A_{ab}} = \frac{P}{A} \qquad \text{(Eq 1)}$$

where $A = A_{ab}/\cos\theta$.

Shear stresses are those stresses applied along a specific direction in a specific plane. To determine the shear stress, the applied load P is resolved along the prescribed shearing direction (for example, in the direction along *ab* lying at an angle λ in Fig. 1):

$$P_{\text{shear}} = P\cos\lambda \qquad \text{(Eq 2a)}$$

Hence, the shear stress, τ, is given by:

$$\tau = \frac{P_{\text{shear}}}{A_{ab}} = \frac{P}{A}\cos\lambda\cos\theta \qquad \text{(Eq 2b)}$$

When the plane and direction for which the shear stress is determined is a crystallographic slip system, and the shear stress is that which activates plastic deformation, then Eq 2(b) is known as Schmid's law.

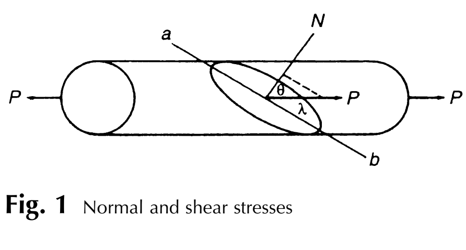


Fig. 1 Normal and shear stresses

The state of stress at a point is defined by three normal stresses and six shear stresses. To simplify, it is common practice to work with three principal stresses. The principal stresses lie parallel to three mutually perpendicular principal axes and are normal to the three principal planes. The defining feature is that no shear stress lies in any of the principal planes. Any state of stress can be resolved into the equivalent three principal stresses (Ref 1). It is important to note that both maximum and minimum values of normal stresses occur along the principal axes. The usual convention is to denote the principal stresses as σ_1, σ_2, σ_3, where algebraically $\sigma_1 > \sigma_2 > \sigma_3$.

Strain is defined in terms of changes in dimensions. If two gage marks were placed a distance L_0 apart on the specimen in Fig. 1, and if those marks were L_1 apart after a tensile load P_1 was applied, then the engineering normal strain (e) would be:

$$e = \frac{L_1 - L_0}{L_0} = \frac{\Delta L}{L_0} = \frac{1}{L_0}\int_{L_0}^{L_1} dL \qquad \text{(Eq 3)}$$

If the bar had been compressed from an initial length L_0 to a shorter length L_1, then the normal strain calculated from Eq 3 would have a negative sign, indicating that the strain was compressive.

Deformation of a body can result not only in the change of length of a line in the body, but it may also produce a change in the angle between two lines. The resulting angular change is known as *shear strain,* γ. Figure 2 illustrates the strain produced by the simple shear on one face

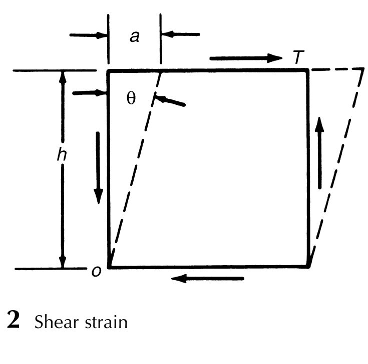


Fig. 2 Shear strain

of a cube. The angle at *O,* which originally was 90°, is decreased by a small amount, θ, by the application of a shear stress. The shear strain is equal to the displacement, *a,* divided by the distance between the shear planes, *h:*

$$\gamma = \frac{a}{h} = \tan\theta \cong \theta \text{ (at small values of } \theta\text{)} \quad \text{(Eq 4)}$$

As with stress, the complete description of the strain at a point requires the determination of three normal strains and six shear strains.

Engineering normal strain is based on a definition of strain with reference to an original unit length, L_0. This definition of strain is satisfactory for elastic strains where ΔL is very small. However, in metalworking processes where the strains often are large, it is no longer logical to base strain on the original gage length. In this case, the calculation of strain is based on the instantaneous gage length. This is called the *true strain,* ε:

$$\varepsilon = \sum\left(\frac{L_1 - L_0}{L_0} + \frac{L_2 - L_1}{L_1} + \frac{L_3 - L_2}{L_2} + \cdots\right) \quad \text{(Eq 5)}$$

or

$$\varepsilon = \int_{L_0}^{L} \frac{dL}{L} = \ln\frac{L}{L_0} \quad \text{(Eq 6)}$$

True strain and engineering normal strain are easily related. From Eq 3:

$$e = \frac{L - L_0}{L_0} = \frac{L}{L_0} - 1$$

and from Eq 6:

$$\varepsilon = \ln\frac{L}{L_0} = \ln(e+1) \quad \text{(Eq 7)}$$

Comparison values of true strain and engineering strain are given as follows:

True strain, ε	Engineering strain, *e*
0.01	0.01
0.10	0.105
0.20	0.22
0.50	0.65
1.0	1.72
4.0	53.6

Note that the normal stress given by Eq 1 is based on the actual area over which the force (*P*) acts. Therefore, it is a true normal stress, σ. If the normal stress were based on the original area (A_0) before the application of load, as is the case in the usual engineering stress-strain curve, then it would be defined as an engineering normal stress, *s*:

$$s = P/A_0 \quad \text{(Eq 8)}$$

Stress-Strain Curves. In the conventional engineering tensile test, a test specimen is gripped at opposite ends within the load frame of a testing machine, and the force and extension are recorded until the specimen fractures. The load is converted into engineering normal stress with Eq 8, and the extension between gage marks is converted to engineering strain with Eq 3. This results in the engineering stress-strain curve with a typical shape, as shown in Fig 3.

In the elastic region of the curve, stress is linearly related to strain, $\sigma = E \cdot e$, where *E* is the elastic (Young's) modulus. When the load exceeds a value corresponding to the yield stress, the specimen undergoes gross plastic deformation. If the specimen were loaded part way and then unloaded, it would be found to have been permanently deformed after the load returned to 0. The stress to produce permanent (plastic) deformation increases with increasing strain; the metal strain hardens. To a good engineering approximation, the volume remains constant as the specimen deforms plastically, $AL = A_0L_0$. Initially, the strain hardening more than compensates for the decrease in cross-sectional area of the specimen, and the engineering stress continues to rise with increasing strain. Eventually, a point is reached where the decrease in area is greater than the increase in load-carrying capacity arising from strain hardening. This condition is reached first at some point in the specimen that is weaker than the rest. All further plastic deformation is concentrated in this region, and the specimen begins to neck or thin down locally. Because the cross-sectional area is now decreasing far more rapidly than strain hardening is increasing the deformation load, the engineering stress continues to decrease until fracture occurs. The maximum in the engineering stress-strain curve is known as the *ultimate tensile strength,* s_u. The strain at maximum load, up to which point the cross-sectional area decreases uniformly along the gage length as the specimen elongates, is known as the *uniform elongation,* e_u.

The necking instability that occurs in the tension test makes interpretation of the curve beyond maximum load more difficult. Because this is the region that is often of importance in metalworking processes, a better interpretation needs to be used. The falloff in stress beyond P_{max} is artificial and occurs only because the stress is calculated on the basis of the original cross-sectional area, A_0, when in fact the area at the necked region is now much smaller than A_0. If the true stress based on the actual cross-sectional area of the specimen is used, the stress-strain curve increases continuously up to fracture (Fig 4). Then, if the strain is expressed as true strain, the true-stress/true-strain curve results. This often is called the flow curve because it gives the stress required to plastically deform (flow) the metal under conditions of no geometrical or frictional constraint.

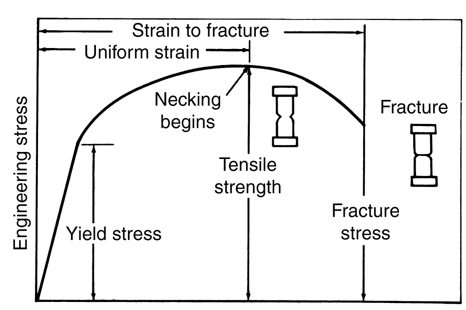


Fig. 3 Engineering stress-strain curve

The formation of a neck in the tensile specimen introduces a complex triaxial state of stress in that region. The necked region is, in effect, a mild notch. A notch under tension produces radial stress (σ_r) and transverse stress (σ_t) that raise the value of longitudinal stress required to cause the plastic flow. Therefore, the average true stress at the neck, which is determined by dividing the axial tensile load by the minimum cross-sectional area of the specimen at the neck, is higher than the stress that would be required to cause flow if simple tension prevailed.

Figure 5 illustrates the geometry at the necked region and the stresses developed by this localized deformation. The radius of curvature of the neck (*R*) can be measured either by projecting the contour of the necked region on a screen or by using a tapered, conical radius gage. A mathematical analysis has been made that provides a correction to the average axial stress to compen-

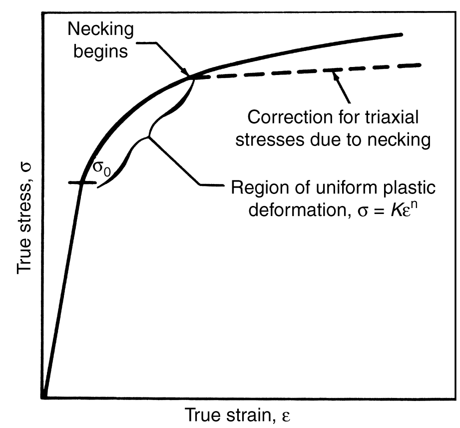


Fig. 4 True-stress/true-strain curve (flow curve)

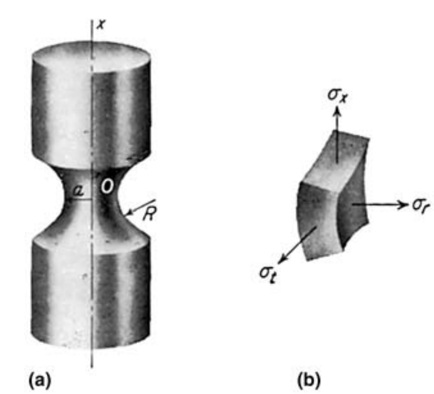


Fig. 5 Stress distribution at the neck of a tensile specimen. (a) Geometry of necked region. *R* is the radius of curvature of the neck; *a* is the minimum radius at the neck. (b) Stresses acting on element at point *O*. σ_x is the stress in the axial direction; σ_r is the radial stress; σ_t is the transverse stress.

sate for the introduction of transverse stresses (Ref 2). According to this analysis, known as the *Bridgman correction,* the uniaxial flow stress corresponding to that which would exist in the tension test if necking had not introduced triaxial stresses is:

$$\sigma = \frac{(\sigma_x)_{\text{avg}}}{\left(\dfrac{1+2R}{a}\right)\left[\ln\left(1+\dfrac{a}{2R}\right)\right]} \qquad \text{(Eq 9)}$$

where $(\sigma_x)_{\text{avg}}$ is the measured stress in the axial direction (load divided by minimum cross section), and a is the minimum radius at the neck. Figure 4 shows how the application of the Bridgman correction changes the true-stress/true-strain curve. The values of a/R needed for the analysis can be obtained either by straining a specimen a given amount beyond necking and unloading to measure a and R directly, or by measuring these parameters continuously past necking, using photography or videography.

To avoid these tedious measurements, Bridgman presented an empirical relation between a/R and the true strain in the neck. Figure 6 shows that this gives close agreement for steel specimens but not for other metals with widely different necking strains. It has been shown that the Bridgman correction factor, B, can be estimated from (Ref 3):

$$B = 0.83 - 0.186 \log \varepsilon \; (0.15 \,''\, \varepsilon \geq 3) \qquad \text{(Eq 10)}$$

where $B = \sigma/(\sigma_x)_{\text{avg}}$.

True-Stress/True-Strain Curve. The true-stress/true-strain curve obviously has many advantages over the engineering stress-strain curve for determining the flow and fracture characteristics of a material. This section develops many of the properties that can be obtained from this test. While the test is considered a valuable basic test of material mechanical behavior, its usefulness in workability studies is somewhat limited, due to the test being limited to relatively small strains, because fracture soon follows the onset of necking. Thus, it is not generally possible to achieve strains of the same magnitude as those found in the metal deformation process.

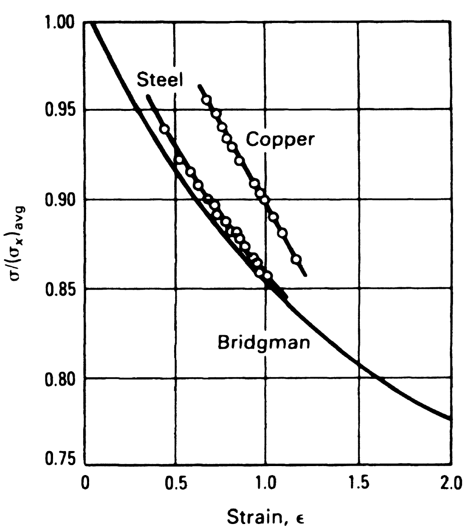


Fig. 6 Relationship between Bridgman correction factor, $\sigma/(\sigma_x)_{\text{avg}}$, and true tensile strain

As discussed previously, true stress in tension is the deformation load divided by the actual cross-sectional area:

$$\sigma = \frac{P}{A} \qquad \text{(Eq 11)}$$

The true strain, ε, may be determined from the engineering or conventional strain, e, by:

$$\varepsilon = \ln(e+1) = \ln\frac{L}{L_0} \qquad \text{(Eq 12)}$$

This equation is applicable only to the onset of necking, because it assumes homogeneity of deformation along the specimen gage length. Beyond maximum load, the true strain should be based on actual area or diameter, D, measurements:

$$\varepsilon = \ln\frac{A_0}{A} = \ln\frac{\left(\dfrac{\pi}{4}\right)D_0^2}{\left(\dfrac{\pi}{4}\right)D^2} = 2\ln\frac{D_0}{D} \qquad \text{(Eq 13)}$$

The true stress at maximum load corresponds to the true tensile strength. For most materials, necking begins at maximum load at a value of strain where the true stress equals the slope of the flow curve. Let σ_u and ε_u denote the true stress and true strain at maximum load when the cross-sectional area of the specimen is A_u. The ultimate tensile strength can be defined as:

$$s_u = \frac{P_{\max}}{A_0} \qquad \text{(Eq 14)}$$

and

$$\sigma_u = \frac{P_{\max}}{A_u} \qquad \text{(Eq 15)}$$

Eliminating $P_{\max}$ yields the true stress at maximum load:

$$\sigma_u = s_u \frac{A_0}{A_u} \qquad \text{(Eq 16)}$$

The true fracture stress is the load at fracture divided by the cross-sectional area at fracture. This stress should be corrected for the triaxial state of stress existing in the tensile specimen at fracture. Because the data required for this correction frequently are not available, true fracture stress values are frequently in error.

The true fracture strain, ε_f, is the true strain based on the original area, A_0, and the area after fracture, A_f:

$$\varepsilon_f = \ln\frac{A_0}{A_f} \qquad \text{(Eq 17)}$$

This parameter represents the maximum true strain that the material can withstand before fracture and is analogous to the total strain to fracture of the engineering stress-strain curve. Because Eq 12 is not valid beyond the onset of necking, it is not possible to calculate ε_f from measured values of engineering fracture strain, e_f. However, for cylindrical tensile specimens, the reduction in area, q, is related to the true fracture strain by:

$$\varepsilon_f = \ln\frac{1}{1-q} \qquad \text{(Eq 18)}$$

where

$$q = \frac{A_0 - A_f}{A_0} \qquad \text{(Eq 19)}$$

Ductility in the Tension Test. In addition to evaluating the strength of a material, the tension test also gives an indication of material ductility. Ductility is the ability of a material to deform without fracture. This is very similar to the concept of workability. The common distinction is that workability refers to situations of large deformation, as in metalworking processes, while the term ductility is usually used in the context of machines and structures, where it is thought of as the ability to accommodate to a small amount of plastic deformation without fracture.

The conventional measures of ductility that are obtained from the tension test are the engineering strain to fracture and the reduction of area at fracture. The percentage elongation, *%e,* is the elongation between the gage marks of the specimen divided by the original gage length, expressed as a percentage:

$$\%e = \left(\frac{L_f - L_0}{L_0}\right) \times 100 \qquad \text{(Eq 20)}$$

Because an appreciable fraction of the deformation is concentrated in the necked region of the tensile specimen, the percentage elongation for a material is higher if the gage length is short than if the gage length was longer. Thus, it is always important to use a standard gage length and to report the gage length over which the percentage elongation was measured.

Empirical studies have shown that elongation at fracture, e_f, correlates with specimen geometry, according to the Bertella-Oliver equation (Ref 4):

$$e_f = e_0\left(\frac{L}{\sqrt{A}}\right)^{-a} = e_0 K^{-a} \qquad \text{(Eq 21)}$$

where $K = L/\sqrt{A}$ is the slimness ratio of the specimen. If elongation is plotted against K on a log scale, a straight line results. The value of elongation at $K = 1$ is e_0, and a is the slope of the line. Equation 21 predicts that a given elongation will be produced in a material if

$$K = L/\sqrt{A}$$

is maintained constant. Thus, at a constant percentage elongation:

$$L_1/\sqrt{A_1} = L_2/\sqrt{A_2}$$

where 1 and 2 are two different specimens. To provide comparable test results using a gage length L_2 on a specimen with area A_2 by means of measurements in a specimen with area A_1, it would only be necessary to adjust the gage length to

$$L_1 = L_2\sqrt{A_1/A_2}$$

Reduction of area, q, as measured by Eq 19, is the other ductility measure provided by the tension test. The occurrence of necking in the tension test, however, makes any quantitative conversion between elongation and reduction in area impossible. Although elongation and reduction in area usually vary in the same way—for example, as a function of test temperature, tempering temperature, or alloy content—this is not always the case. Generally, elongation and reduction in area measure different types of material behavior. Provided the gage length is not too short, percent elongation is primarily influenced by uniform elongation, and thus it is dependent on the strain-hardening capacity of the material. Reduction in area is more a measure of the deformation required to produce fracture, and its chief contribution comes from the necking process. Because of the complicated stress state in the neck, values of reduction in area are dependent on specimen geometry and deformation behavior, and they should not be taken as true material properties. However, reduction in area is the most structure-sensitive ductility parameter, and as such, it is very useful in detecting the influence of structural changes in the material, such as second-phase particles, inclusions, and porosity.

Mathematical Expressions for the Flow Curve. A simple power curve relation can express the flow curve of many metals in the region of uniform plastic deformation, that is, from yielding up to maximum load:

$$\sigma = K\varepsilon^n \qquad \text{(Eq 22)}$$

where n is the strain-hardening exponent, and K is the strength coefficient. A log-log plot of true stress and true strain up to maximum load results in a straight line if Eq 22 is satisfied by the data (Fig. 7). The linear slope of this line is n, and K is the true stress at $\varepsilon = 1.0$ (corresponds to $q = 0.63$). As shown in Fig. 8, the strain-hardening exponent may have values from $n = 0$ (perfectly plastic solid) to $n = 1$ (elastic solid). For most metals, n has values between 0.10 and 0.50.

Note that the rate of strain hardening, $d\sigma/d\varepsilon$,

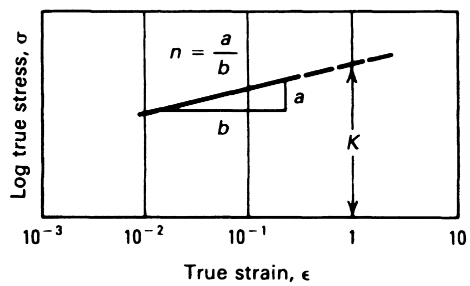


Fig. 7 Log-log plot of true-stress/true-strain curve. n is the strain-hardening exponent; K is the strength coefficient.

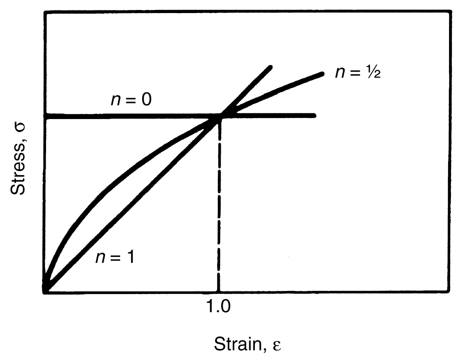


Fig. 8 Various forms of power curve $\sigma = K\varepsilon^n$

is not identical to the strain-hardening exponent. From the definition of n:

$$n = \frac{d(\log\sigma)}{d(\log\varepsilon)} = \frac{d(\ln\sigma)}{d(\ln\varepsilon)} = \frac{\varepsilon}{\sigma}\frac{d\sigma}{d\varepsilon} \qquad \text{(Eq 23)}$$

$$\frac{d\sigma}{d\varepsilon} = n\frac{\sigma}{\varepsilon} \qquad \text{(Eq 24)}$$

It can be readily shown that the strain-hardening exponent is equal to the true uniform strain, ε_u (Ref 1):

$$n = \varepsilon_u = \ln\frac{A_0}{A_u} \qquad \text{(Eq 25)}$$

where A_u is the area at maximum load.

Deviations from Eq 22 frequently are observed, often at low strains (10^{-3}) or high strains ($\varepsilon \approx 1.0$). One common type of deviation is for a log-log plot of Eq 22 to result in two straight lines with different slopes. Sometimes data that do not plot according to Eq 22 yield a straight line according to the relationship:

$$\sigma = K(\varepsilon_0 + \varepsilon)^n \qquad \text{(Eq 26)}$$

where ε_0 can be considered to be the amount of strain that the material received prior to the tension test (Ref 5).

Another common variation on Eq 22 is the Ludwik equation:

$$\sigma = \sigma_0 + K\varepsilon^n \qquad \text{(Eq 27)}$$

where σ_0 is the yield stress, and K and n are the same constants as in Eq 22. This equation may be more satisfying than Eq 22, because the latter implies that at zero true strain, the stress is zero. It has been shown that σ_0 can be obtained from the intercept of the strain-hardening portion of the stress-strain curve and the elastic modulus line (Ref 6):

$$\sigma_0 = \left(\frac{K}{E^n}\right)^{\frac{1}{1-n}} \qquad \text{(Eq 28)}$$

Effect of Strain Rate and Temperature. The rate at which strain is applied to the tension specimen has an important influence on the stress-strain curve. Strain rate is defined as $\dot{\varepsilon} = d\varepsilon/dt$. It is expressed in units of s^{-1}. Increasing strain rate increases the flow stress. Moreover, the strain-rate dependence of strength increases with increasing temperature. The yield stress and the flow stress at lower values of plastic strain are more affected by strain rate than the tensile strength.

If the crosshead velocity, v, of the testing machine is $v = dL/dt$, then the strain rate expressed in terms of conventional engineering strain, $\dot{e}$, is:

$$\dot{e} = \frac{de}{dt} = \frac{d(L - L_0)/L_0}{dt} = \frac{1}{L_0}\frac{dL}{dt} = \frac{v}{L_0} \qquad \text{(Eq 29)}$$

The engineering strain rate is proportional to the crosshead velocity. In a modern testing machine in which the crosshead velocity can be set accurately and controlled, it is a simple matter to carry out tension tests at a constant engineering strain rate. The true strain rate, $\dot{\varepsilon}$, is given by:

$$\dot{\varepsilon} = \frac{d\varepsilon}{dt} = \frac{d[\ln(L/L_0)]}{dt} = \frac{1}{L}\frac{dL}{dt} = \frac{v}{L} \qquad \text{(Eq 30)}$$

Equation 30 shows that for a constant crosshead velocity, the true strain rate decreases as the specimen elongates or cross-sectional area shrinks. To run tension tests at a constant true strain rate requires monitoring the instantaneous cross section of the deforming region with closed-loop control feedback to increase the crosshead velocity as the area decreases (Ref 7).

The strain-rate dependence of flow stress at constant strain and temperature is given by:

$$\sigma = C(\dot{\varepsilon})^m\big|_{\varepsilon,T} \qquad \text{(Eq 31)}$$

The exponent m in Eq 31 is known as the strain-rate sensitivity, and C is the strain-hardening coefficient. The exponent m can be obtained from the slope of a plot of log σ versus log $\dot{\varepsilon}$. However, a more sensitive way to determine m is with a rate-change test (Fig. 9). A tensile test is carried out at strain rate $\dot{\varepsilon}_1$, and at a certain flow stress, σ_1, the strain rate is suddenly increased to $\dot{\varepsilon}_2$. The flow stress quickly increases to σ_2. The strain-rate sensitivity, at constant strain and temperature, can be determined from Eq 32:

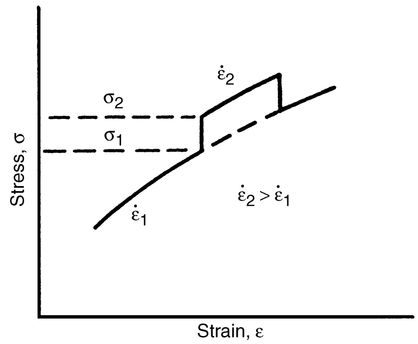


Fig. 9 Strain-rate change test, used to determine strain-rate sensitivity, *m*

$$m = \left(\frac{\partial \ln \sigma}{\partial \ln \dot{\varepsilon}}\right)_{\varepsilon,T} = \frac{\dot{\varepsilon}}{\sigma}\left(\frac{\partial \sigma}{\partial \dot{\varepsilon}}\right) = \frac{\Delta \log \sigma}{\Delta \log \dot{\varepsilon}}$$
$$= \frac{\log \sigma_2 - \log \sigma_1}{\log \dot{\varepsilon}_2 - \log \dot{\varepsilon}_1} = \frac{\log(\sigma_2/\sigma_1)}{\log(\dot{\varepsilon}_2/\dot{\varepsilon}_1)} \quad \text{(Eq. 32)}$$

The strain-rate sensitivity of metals is quite low (<0.1) at room temperature, but m increases with temperature. At hot working temperatures ($T/T_M > 0.5$), m values of 0.1 to 0.2 are common in metals. Polymers have much higher values of m and may approach $m = 1$ in room-temperature tests for some polymers.

The temperature dependence of flow stress can be represented by:

$$\sigma = C_2 e^{Q/RT}\big|_{\varepsilon,\,\dot{\varepsilon}} \quad \text{(Eq. 33)}$$

where Q is an activation energy for plastic flow, cal/g·mol; R is the universal gas constant, 1.987 cal/ K·mol; and T is testing temperature in degree Kelvin. From Eq 33, a plot of ln σ versus $1/T$ gives a straight line with a slope Q/R.

Constitutive equations are mathematical expressions that relate the stress in terms of the variable strain, strain rate, and temperature. For example, Eq 22, 31, and 33 are simple constitutive equations. Such expressions are necessary for the computer modeling of deformation of materials, but it must be noted that no universally accepted equations have been developed.

One of the oldest and most useful equations of this type is:

$$Z = A[\sinh(\alpha\sigma)]^{1/m} \quad \text{(Eq 34)}$$

where Z, the Zener-Holloman parameter, is:

$$Z = \dot{\varepsilon} \exp\left(\frac{Q}{RT}\right)$$

The evaluation of the parameter α is given in Ref 8.

The Johnson-Cook equation is widely used in computer codes that handle large plastic deformations:

$$\sigma = (\sigma_0 + K\varepsilon^n)\left(1 + C \ln \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\right)\left[1 - \left(\frac{T - T_0}{T_M - T_0}\right)^m\right] \quad \text{(Eq 35)}$$

In this equation, K, C, m, and n are material parameters, and ε_0 and T_0 are evaluated at reference conditions (Ref 9).

Multiaxial Stress States

In design and metal processing situations, one expects to find stresses acting in more than just a single direction. This section discusses some of the most common situations that are encountered.

Plane Stress and Plane Strain. Because of different geometries of the tooling and the workpiece and the different ways that forces are applied in metal deformation processes, different states of stress can be developed in the workpiece. It has been seen that the general state of stress can be defined by three principal stresses, $\sigma_1 > \sigma_2 > \sigma_3$. Two states of stress that frequently occur in metalworking applications—plane stress and plane strain—deserve special mention. The use of the adjective *plane* implies that the stresses are confined to a two-dimensional (plane) situation.

Plane stress occurs when the stress state lies in the plane of the member. This typically occurs when one dimension of the member is very small compared to the other two, and the member is loaded by a force lying in the plane of symmetry of the body so that the stress normal to the plane surface is zero. Examples are thin, plate-type structures, such as thin-wall pressure vessels.

Plane strain occurs when the strain in one of the three principal directions is zero (for example, $\varepsilon_3 = 0$), as in the rolling of a wide sheet. In this case, the sheet does not change in width while it is reduced in thickness and increases in length. Although $\varepsilon_3 = 0$ for plane strain, there is a stress acting in that direction. For plastic deformation, the stress in the principal direction for which strain is zero is the average of the other two principal stresses, that is:

$$\sigma_3 = \frac{\sigma_1 + \sigma_2}{2}$$

Effective Stress and Strain. Stress states in metalworking processes are often complex. It is convenient to be able to express this situation by a single expression, the effective stress, $\bar{\sigma}$:

$$\bar{\sigma} = \frac{1}{\sqrt{2}}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]^{\frac{1}{2}} \quad \text{(Eq 36)}$$

Tensile stresses are positive, and compressive stresses are negative. Note that for uniaxial tension, the effective stress reduces to $\bar{\sigma} = \sigma_1$.

The effective strain, $\bar{\varepsilon}$, in terms of total plastic strain, is:

$$\bar{\varepsilon} = \left[\frac{2}{3}(\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2)\right]^{\frac{1}{2}} \quad \text{(Eq 37)}$$

Hydrostatic Stress. In the theory of plasticity, it can be shown that the state of stress can be divided into a hydrostatic or mean stress and a deviator stress that represents the shear stresses in the total state of stress. Experience shows that the greater the compressive mean stress, the better the workability of a material. For example, it is known that most materials can be deformed more easily in an extrusion press and even more so with hydrostatic extrusion (Ref 10). The mean or hydrostatic component of the stress state is given by σ_m:

$$\sigma_m = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3} \quad \text{(Eq 38)}$$

In general, the greater the level of tensile stress, the more severe the stress system is with regard to workability. For a given material, temperature, and strain rate of deformation, the workability is much improved if the stress state is highly compressive. A general workability parameter, β, has been proposed that allows for the stress state (Ref 11):

$$\beta = \frac{3\sigma_m}{\bar{\sigma}} \quad \text{(Eq 39)}$$

Figure 10 shows the workability parameter plotted for various mechanical tests and metalworking processes. The strain to fracture is the ordinate. The curve is evaluated with three basic tests: tension, torsion, and compression. Other common metal deformation processes are superimposed. This figure emphasizes the critical role in workability that is played by the state of stress developed in the workpiece.

Yielding Criteria. The ease with which a metal yields plastically or flows is an important factor in workability. If a metal can be deformed at low stress, as in superplastic deformation, then the stress levels throughout the deforming workpiece are low, and fracture is less likely. The dominant metallurgical conditions and temperature are important variables, as is the stress state. Plastic flow is produced by slip within the individual grains, and slip is induced by a high resolved shear stress. Therefore, the beginning of plastic flow can be predicted by a maximum shear stress, or Tresca, criterion:

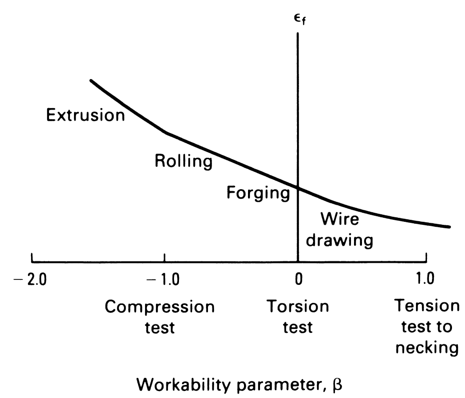


Fig. 10 Influence of the stress state on the strain to fracture, ε_f

$$\tau_{max} = \frac{1}{2}(\sigma_1 - \sigma_3) = \frac{\sigma_0}{2} \quad \text{(Eq 40)}$$

where τ_{max} is the maximum shear stress, and σ_0 is the yield (flow) stress measured in either a uniaxial tension or uniaxial compression test. Although adequate, this yield criterion neglects the intermediate principal stress, σ_2.

A more complete and generally applicable yielding criterion is that proposed by von Mises:

$$2\sigma_0^2 = (\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 \quad \text{(Eq 41)}$$

where $\sigma_1 > \sigma_2 > \sigma_3$ are the three principal stresses, and σ_0 is the uniaxial flow stress of the material. This is the same equation given earlier for effective stress, Eq 36.

The significance of yield criteria is best illustrated by examining a simplified stress state in which $\sigma_3 = 0$ (plane stress). The Tresca yield criterion then defines a hexagon, and the von Mises criterion an ellipse (Fig. 11).

Figure 11 illustrates how the stress required to produce plastic deformation varies significantly with the stress state and how it can be related to the basic uniaxial flow stress of the material through a yield criterion. Yielding (plastic flow) can be initiated in several modes. In pure tension, flow occurs at the flow stress, σ_0 (point 1 in Fig. 11). In pure compression, the material yields at the compressive flow stress, which, in ductile materials, is usually equal to the tensile flow stress (point 2). When a sheet is bulged by a punch or a pressurized medium, the two principal stresses in the surface of the sheet are equal (balanced biaxial tension) and must reach σ_0 (point 3) for yielding.

An important condition is reached when deformation of the workpiece is prevented in one of the principal directions (plane strain). This occurs because a die element keeps one dimension constant, or only one part of the workpiece is deformed, and adjacent nondeforming portions exert a restraining influence. In either case, the restraint creates a stress in that principal direction. The stress is the average of the two other principal stresses (corresponding to point 4). The stress required for deformation is still σ_0 according to Tresca but is 1.15 σ_0 according to von Mises. The latter is usually regarded as the plane-strain flow stress of the material. It is sometimes called the *constrained flow stress.*

An important stress state is pure shear, in which two principal stresses are of equal magnitude but opposite sign (point 5, Fig. 11), $\tau = \sigma_1 = -\sigma_2$ and $\sigma_3 = 0$. When these stresses are substituted into the von Mises yield criterion, Eq 41, it is seen that plastic deformation occurs at the shear flow stress:

$$\tau_0 = \sigma_0 / \sqrt{3}$$

In metalworking theory, the shear flow stress is often denoted by the symbol *k*.

Material Factors Affecting Workability

Fracture Mechanisms. Fracture in bulk deformation processing usually occurs as ductile fracture, rarely as brittle fracture. However, depending on temperature and strain rate, the details of the ductile fracture mechanism vary. Figure 12 illustrates in a schematic way the different modes of ductile fracture obtained in a tension test over a wide range of strain rates and temperatures. At temperatures below approximately one-half the melting point of a given material (below the hot working region), a typical dimpled rupture type of ductile fracture usually occurs. At very high temperatures, a rupture type of fracture occurs in which the material recrystallizes rapidly and pulls down to a point, with nearly 100% reduction in area. A transgranular creep type of failure occurs at temperatures less than those causing rupture. Intragranular voids form, grow, and coalesce into internal cavities, which result in a fracture with a finite reduction in area.

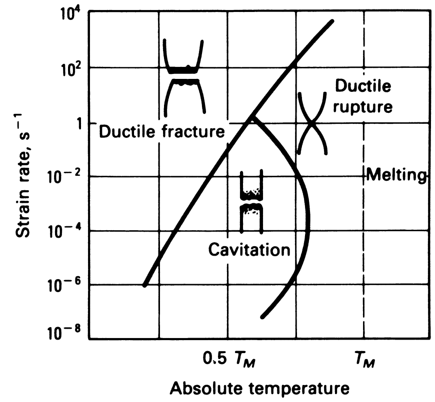


Fig. 12 Tensile fracture modes as a function of temperature (measured on the absolute temperature scale) and strain rate. T_M, melting temperature

A more commonly found representation of possible fracture mechanisms is the fracture mechanism (Ashby) map (Fig. 13). Such a map shows the area of dominance in terms of normalized stress versus normalized temperature for the dominant fracture mechanisms. The maps are constructed chiefly by using the best mechanistic models of each fracture process. For more on high-temperature fracture mechanisms, see Ref (12).

Failure in deformation processing at below one-half the melting temperature, 0.5 T_M, occurs by ductile fracture. The three stages of ductile fracture are shown schematically in Fig. 14. The first stage is void initiation, which usually occurs at second-phase particles or inclusions. Voids are initiated because particles do not deform, and this forces the ductile matrix around the particle to deform more than normal. This in turn produces more strain hardening, thus creat-

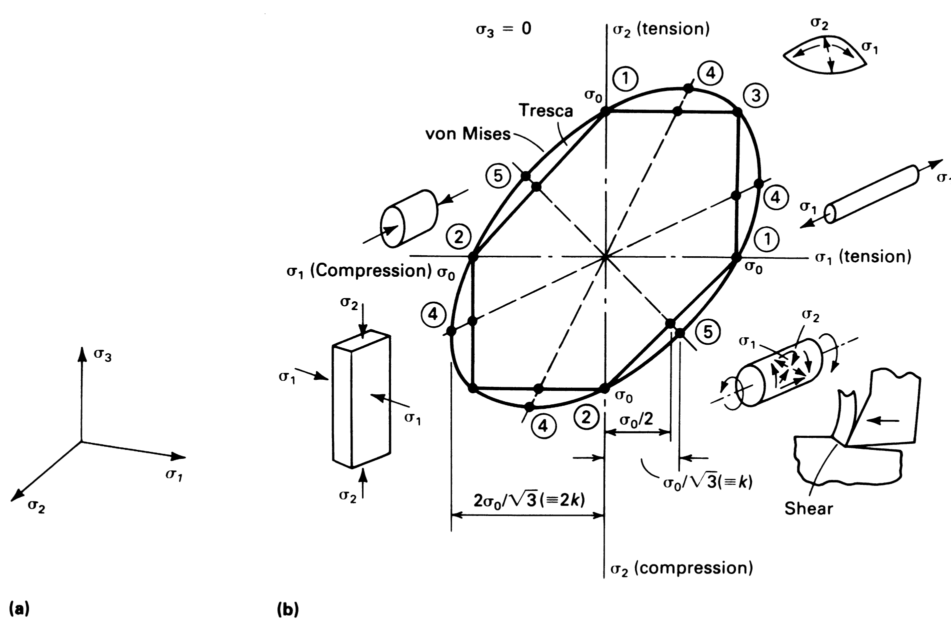


Fig. 11 (a) Directions of principal stresses and (b) yield criteria with some typical stress states

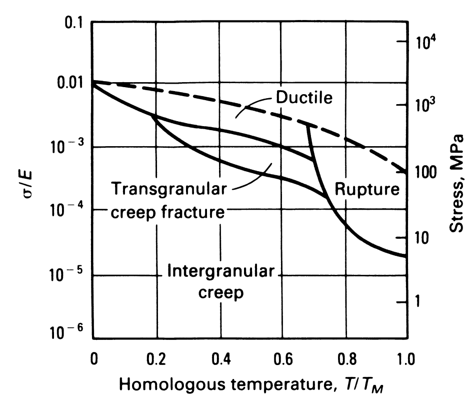


Fig. 13 Fracture mechanism map for nickel.

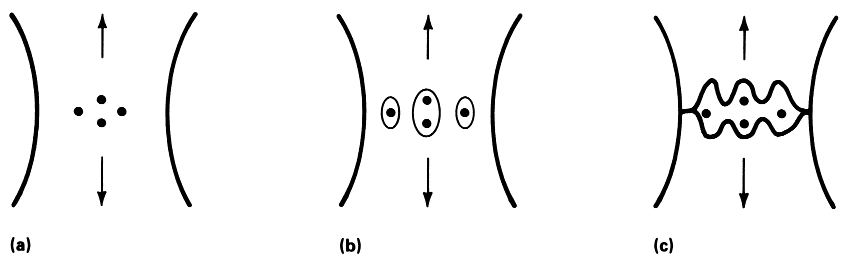


Fig. 14 Stages in the dimpled rupture mode of ductile fracture. (a) Void initiation at hard particles. (b) Void growth. (c) Void linking

ing a higher stress in the matrix near the particles. When the stress becomes sufficiently large, the interface may separate, or the particle may crack. As a result, ductility is strongly dependent on the size and density of the second-phase particles, as shown in Fig. 15.

The second stage of ductile fracture is void growth, which is a strain-controlled process. Voids elongate as they grow, and the ligaments of matrix material between the voids become thin. Therefore, the final stage of ductile fracture is hole coalescence through the separation of the ligaments that link the growing voids.

Ductile fracture by void growth and coalescence can occur by two modes. Fibrous tearing (mode I) occurs by void growth in the crack plane that is essentially normal to the tensile axis. In mode II void growth, voids grow in sheets at an oblique angle to the crack plane under the influence of shear strains. This type of shear band tearing is found on the surface of the cone in a ductile cup-and-cone tensile fracture. It commonly occurs in deformation processing in which friction and/or geometric conditions produce inhomogeneous deformation, leading to local shear bands. Localization of deformation in these shear bands leads to adiabatic temperature increases that produce local softening.

Increasing the temperature of deformation leads to significant changes in deformation behavior and fracture mode. At temperatures above one-half the melting point, particularly at low strain rates, grain-boundary sliding becomes prominent. This leads to wedge-shaped cracks that propagate along grain boundaries and result in low ductility. Such cracking is common at the low strain rates found in creep but is not a frequent occurrence at the faster strain rates in deformation processes unless there are brittle precipitates at the grain boundaries. This is because

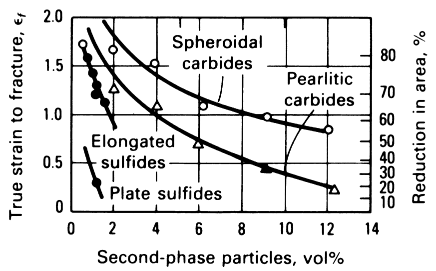


Fig. 15 Effect of volume fraction of second-phase particles on the tensile ductility of steel. Source: Ref 13

the probability of wedge cracking varies with the applied strain rate. If the strain rate is so high that the matrix deforms at a faster rate than the boundaries can slide, then grain-boundary sliding effects are negligible.

For high-temperature fracture initiated by grain-boundary sliding, the processes of void growth and coalescence, rather than void initiation, are the primary factors that control ductility. When voids initiated at the original grain boundaries have difficulty in linking because boundary migration is high as a result of dynamic recrystallization, hot ductility is high. In extreme cases, this can lead to highly ductile rupture, as shown in Fig. 12.

Compressive stresses superimposed on tensile or shear stresses by the deformation process can have a significant influence on closing small cavities and cracks or limiting their growth and thus enhancing workability. Because of this important role of the stress state, it is not possible to express workability in absolute terms. Workability depends not only on material characteristics but also on process variables, such as strain, strain rate, temperature, and stress state.

Hot Working. Many deformation processes for metals and alloys are performed at temperatures greater than 0.5 T_M, the hot working range. Three principal benefits accrue from hot working. First, because the flow stress is lower at higher temperatures, it offers an economical method for size reduction of large workpieces. Second, metals at high temperatures are generally capable of achieving larger deformation strains without fracture than at lower temperatures. However, for many alloy systems, the temperature and strain rate must be appropriately chosen. Lastly, high-temperature deformation assists in the homogenization of the ingot structure (chemical segregation) and in the closing of internal voids. These benefits come at the cost of surface oxidation, poorer dimensional accuracy, and the need to heat and cool the work appropriately.

The mechanisms of hot working are rather complex and vary considerably from alloy to alloy. First, one needs to understand some terminology. Those processes that occur during deformation are called *dynamic processes,* while those that occur between intervals of deformation or after deformation is completed are called *static processes.* The two dynamic processes involved in hot working are *dynamic recovery* and *dynamic recrystallization.*

Dynamic recovery results from the annihilation of dislocations due to ease of cross slip, climb, and dislocation unpinning at the hot working temperature. These deformation mechanisms produce a microstructure consisting of elongated grains, inside of which is a well-developed, fine subgrain structure, typically of the order of 1 to 10 μm. The stress-strain curve for a metal undergoing dynamic recovery shows an increase in flow stress up to a steady-state value that corresponds to the development of a steady-state substructure (Fig. 16, curve a). The level of the steady-state value increases with strain rate and a decrease in deformation temperature.

In dynamic recrystallization dislocation, annihilation only occurs when the dislocation density reaches such high levels that strain-free recrystallized grains are nucleated. Therefore, the rate of strain hardening is high until recrystallization occurs (Fig. 16, curve b). However, when it begins, the flow stress drops rapidly as recrystallization progresses.

Materials that experience rapid recovery and thus do not undergo dynamic recrystallization are body-centered cubic iron; beta-titanium alloys; hexagonal metals such as zirconium; and high stacking-fault energy, face-centered cubic metals such as aluminum. Face-centered cubic metals with lower stacking-fault energy, such as austenitic iron, copper, brass, and nickel, experience dynamic recrystallization in hot working. In these materials, dislocation climb is difficult. This leads to higher dislocation densities than in materials whose deformation is controlled by dynamic recovery. Additional discussion and references pertaining to the mechanisms of hot working can be found in Chapter 3, "Evolution of Microstructure during Hot Working," in this Handbook.

Flow Localization. Workability problems can arise when metal deformation is localized to a narrow zone. This results in a region of different structures and properties that can be the site of failure in service. Localization of deformation can also be so severe that it leads to cracking in the deformation process. In either mode, the presence of flow localization needs to be recognized and dealt with.

Flow localization is commonly caused by the formation of a dead-metal zone between the workpiece and the tooling. This can arise from

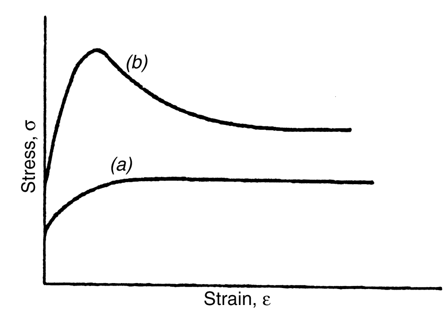


Fig. 16 (a) Stress-strain curve for a metal that undergoes dynamic recovery in the hot working region. (b) Stress-strain curve for a metal that undergoes dynamic recrystallization in the hot working region

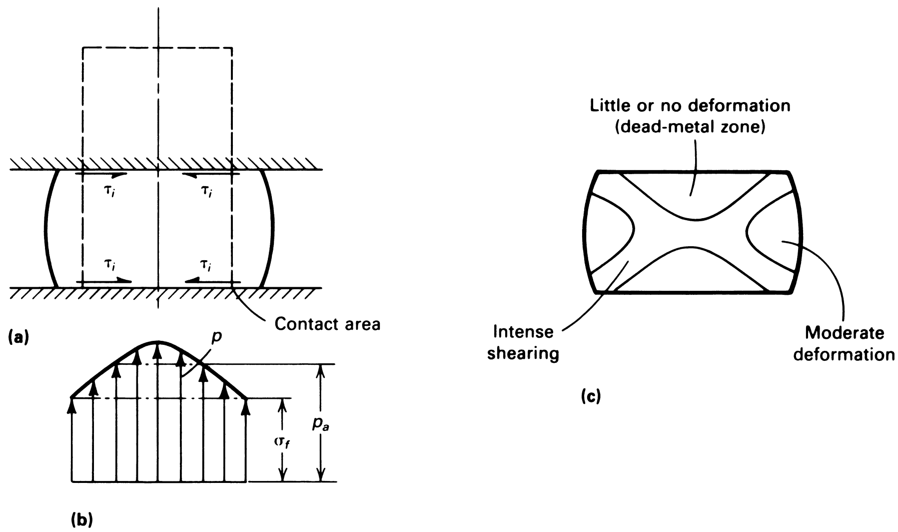


Fig. 17 Consequences of friction illustrated in the upsetting of a cylinder. (a) Direction of shear stresses. (b) Consequent rise in interface pressure. (c) Inhomogeneity of deformation. τ_i, average frictional shear stress; *p*, normal pressure; σ_f, flow stress; p_a, average die pressure

poor lubrication at the workpiece-tool interface. Figure 17 illustrates the upsetting of a cylinder with poorly lubricated platens. When the workpiece is constrained from sliding at the interface, it barrels, and the friction-hill pressure distribution is created over the interface. The inhomogeneity of deformation throughout the cross section leads to a dead zone at the tool interface and a region of intense shear deformation. A similar situation can arise when the processing tools are cooler than the workpiece; in this case, heat is extracted at the tools. Consequently, the flow stress of the metal near the interface is higher because of the lower temperature.

However, flow localization may occur during hot working in the absence of frictional or chilling effects. In this case, localization results from flow softening (negative strain hardening). Flow softening arises during hot working as a result of deformation heating or microstructural instabilities, such as the generation of a softer texture during deformation or dynamic spheroidization. Flow softening has been correlated with materials properties determined in uniaxial compression (Ref 14,15) by the parameter:

$$\alpha_c = \frac{\gamma' - 1}{m} \quad \text{(Eq 42)}$$

where γ', the normalized flow-softening rate, is given by $\gamma' = (1/\sigma)\, d\sigma/d\varepsilon$, and m is the strain-rate sensitivity. Nonuniform flow in compression is likely when $\alpha_c \geq 5$. For plane-strain compression, as in side pressing, flow localization is determined by $\alpha_p = \gamma'/m \geq 5$. Figure 18 shows a fracture that initiated at a shear band during high-speed forging of a complex austenitic stainless steel.

Metallurgical Considerations. Workability problems depend greatly on grain size and grain structure. When the grain size is large relative to the overall size of the workpiece, as in conventionally cast ingot structures, workability is lower, because cracks may initiate and propagate easily along the grain boundaries. Moreover, with cast structures, impurities are frequently segregated to the center and top or to the surface of the ingot, creating regions of low workability. Because chemical elements are not distributed uniformly on either a micro- or a macroscopic scale, the temperature range over which an ingot structure can be worked is rather limited.

Typically, cast structures must be hot worked. The melting point of an alloy in the as-cast condition is usually lower than that of the same alloy in the fine-grain, recrystallized condition because of chemical inhomogeneities and the presence of low-melting-point compounds that frequently occur at grain boundaries. Deformation at temperatures too close to the melting point of these compounds may lead to grain-boundary cracking when the heat developed by plastic deformation increases the workpiece temperature and produces local melting. This fracture mode is called *hot shortness.* It can be prevented by using a sufficiently low deformation rate that allows the heat developed by deformation to be dissipated by the tooling, by using lower working temperatures, or by subjecting the workpiece to a homogenization heat treatment prior to hot working. The relationship between the workability of cast and wrought structures and temperatures is shown in Fig. 19.

The intermediate temperature region of low ductility shown in Fig. 19 is found in many metallurgical systems (Ref 17). This occurs at a temperature that is sufficiently high for grain-boundary sliding to initiate grain-boundary cracking but not so high that the cracks are sealed off from propagation by a dynamic recrystallization process.

The relationship between workability and temperature for various metallurgical systems is summarized in Fig. 20. Generally, pure metals and single-phase alloys exhibit the best workability, except when grain growth occurs at high temperatures. Alloys that contain low-melting-point phases (such as gamma-prime-strengthened nickel-base superalloys) tend to be difficult to deform and have a limited range of working temperature. In general, as the solute content of the alloy increases, the possibility of forming low-melting-point phases increases, while the temperature for precipitation of second phases increases. The net result is a decreased region for good forgeability (Fig. 21).

During the breakdown of cast ingots and the subsequent working by forging nonuniformities

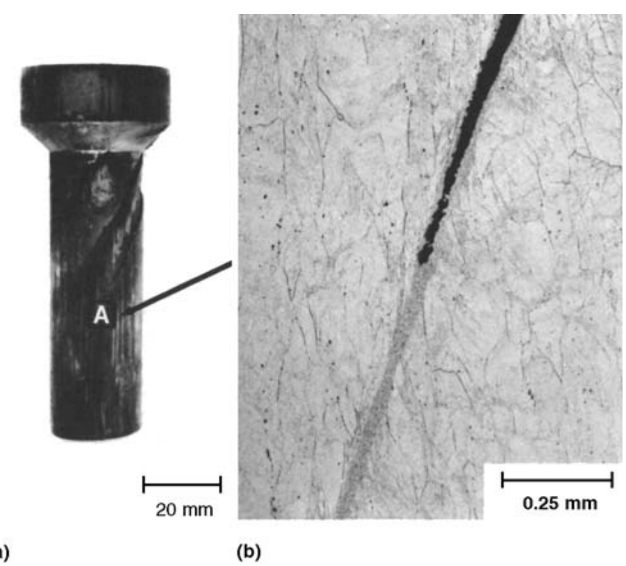


Fig. 18 Austenitic stainless steel high-energy-rate forged extrusion. Forging temperature: 815 °C (1500 °F); 65% reduction in area; $\dot{\varepsilon} = 1.4 \times 10^3\ s^{-1}$. (a) View of extrusion showing spiral cracks. (b) Microstructure at the tip of one of the cracks in area A of extrusion. Note that the crack initiated in a macroscopic shear band that formed at the bottom of the field of view. Source: Ref 16

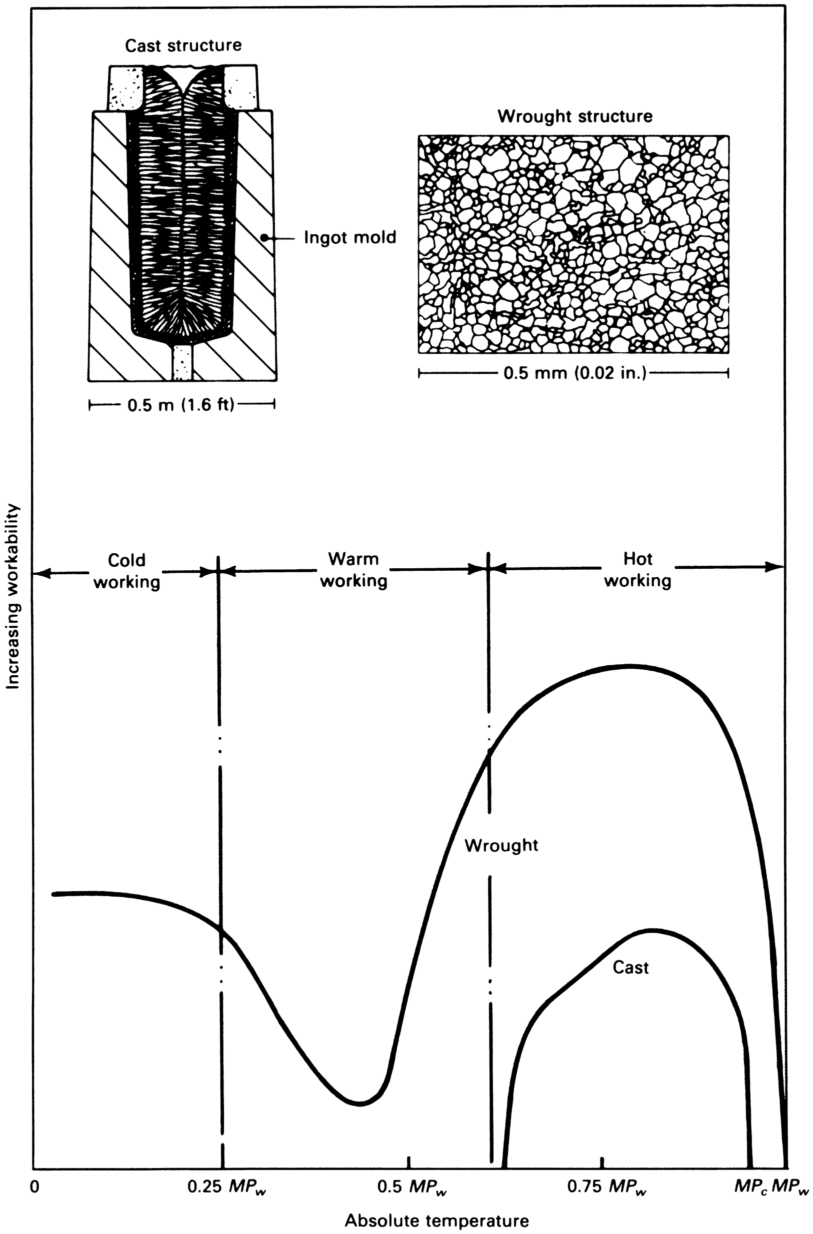


Fig. 19 Relative workabilities of cast metals and wrought recrystallized metals at cold, warm, and hot working temperatures. The melting point (or solidus temperature) is denoted as MP_c (cast metals) or MP_w (wrought and recrystallized metals).

in alloy chemistry, second-phase particles, inclusions, and the crystalline grains themselves are aligned in the direction of greatest metal flow. This directional pattern of crystals and second-phase particles is known as the grain flow pattern. This pattern is responsible for the familiar fiber structure of wrought metal products (Fig. 22). It also produces directional variation in such properties as strength, ductility, fracture toughness, and resistance to fatigue. This anisotropy in properties is greatest between the working (longitudinal) direction and the transverse direction (Fig. 23). In a properly designed forging, the largest stress should be in the direction of the forging fiber, and the parting line of the dies should be located so as to minimize disruption to the grain flow lines.

Process Variables Determining Workability

Strain. Stress and strain are defined in the section "Stress, Strain, and Stress-Strain Curves" earlier in this chapter. This section relates the subject more closely to deformation processing and workability. The principal objective in plastic deformation processes is to change the shape of the deformed product. A secondary objective is to improve or control the properties of the deformed product.

In dense metals, unlike porous powder compacts, the volume, V, of the workpiece remains constant in plastic deformation as it increases in cross-sectional area, A, and decreases in length, L:

$$V = A_0L_0 = A_1L_1 \quad \text{(Eq 43)}$$

If the plastic deformation is expressed as true strain, then the constancy-of-volume condition results in the following expression for the principal strains:

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0 \quad \text{(Eq 44)}$$

Deformation in metalworking is often expressed by the cross-sectional area reduction, R:

$$R = \frac{A_0 - A_1}{A_0} \quad \text{(Eq 45)}$$

From constancy of volume, Eq 43, and the definition of true strain, one can write:

$$\varepsilon = \ln\frac{L_1}{L_0} = \ln\frac{A_0}{A_1} = \ln\frac{1}{1-R} \quad \text{(Eq 46)}$$

Strain rate has three chief effects in metal deformation processes: increases in strain rate raise the flow stress, especially in strain-rate-sensitive materials with a high m; the temperature of the workpiece is increased by adiabatic heating, because there is little time for the heat to dissipate; and there is improved lubrication at the tool-metal interface, so long as the lubricant film can be maintained.

If one considers a cylinder of height h being upset in compression, then the strain rate is given by:

$$\dot{\varepsilon} = \frac{d\varepsilon}{dt} = \frac{dh}{h} \cdot \frac{1}{dt} = \frac{v}{h} \quad \text{(Eq 47)}$$

where dh/dt is the deformation velocity, v, and h is the instantaneous height of the cylinder. However, for most processes, the deformation velocity is not a constant value. It is usual to determine a time-averaged strain rate, $\bar{\dot{\varepsilon}}$. For example, for hot rolling:

$$\bar{\dot{\varepsilon}} = \frac{\varepsilon}{t} = \frac{\ln(h_0/h_1)}{L/v} = \ln\frac{h_0}{h_1}\left(\frac{2\pi rn}{r \times \Delta h}\right) \quad \text{(Eq 48)}$$

where L is the horizontal projection of the arc of contact, r is the roll diameter, and n is the angular speed in revolutions per second.

Temperature. Metalworking processes are commonly classified as hot working or cold working operations. Hot working refers to deformation under conditions of temperature and deformation velocity such that restorative processes (recovery and/or recrystallization) occur simultaneously with deformation (see the section "Hot Working" in this chapter). Cold working refers to deformation carried out under conditions for which restorative processes are not effective during the process. In hot working, the strain hardening and the distorted grain structure produced by deformation are eliminated rapidly by the restorative processes during or immediately after deformation.

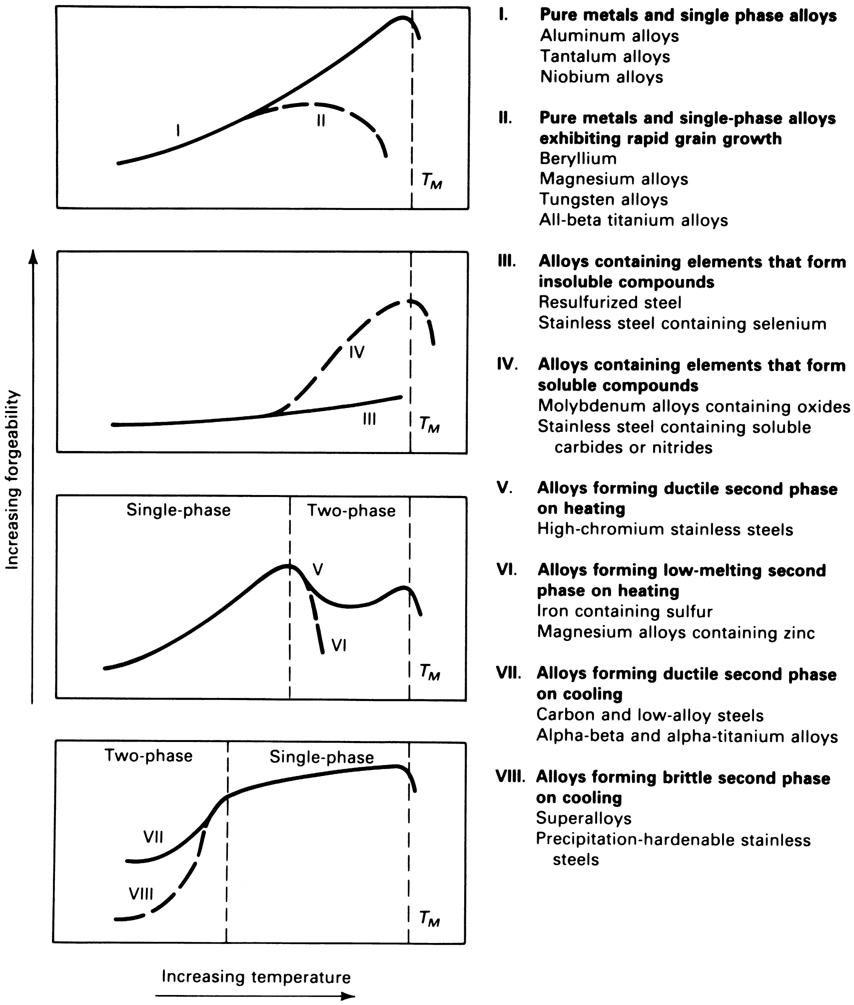


Fig. 20 Typical workability behavior exhibited by different alloy systems. T_M: absolute melting temperature. Source: Ref 18

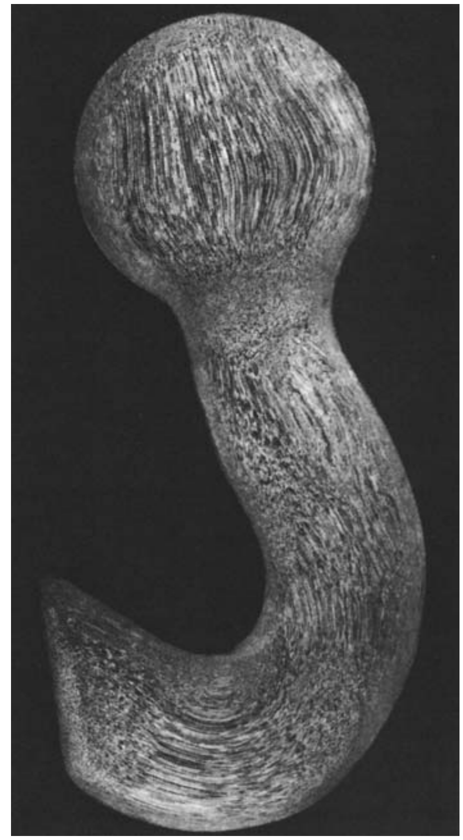


Fig. 22 Flow lines in a forged 4140 steel hook. Specimen was etched using 50% HCl. 0.5×

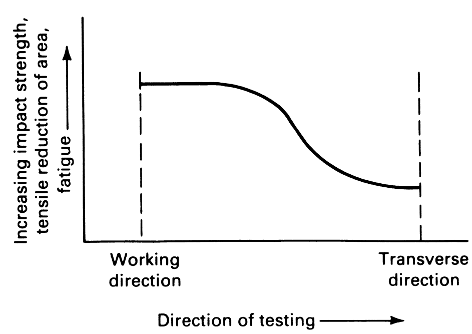


Fig. 23 Anisotropy in wrought alloys

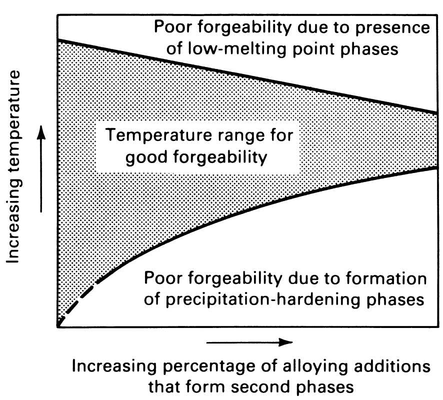


Fig. 21 Influence of solute content on melting and solution temperatures and therefore on forgeability

Very large deformations are possible in hot working, because the restorative processes keep pace with the deformation. Hot working occurs at essentially constant flow stress. Flow stress decreases with the increasing temperature of deformation. In cold working, strain hardening is not relieved, and the flow stress increases continuously with deformation. Therefore, the total deformation possible before fracture is less for cold working than for hot working, unless the effects of strain hardening are relieved by annealing.

Approximately 95% of the mechanical work expended in deformation is converted into heat. Some of this heat is conducted away by the tools or lost to the environment. However, a portion remains to increase the temperature of the workpiece. The faster the deformation process, the greater the percentage of heat energy that goes to increase the temperature of the workpiece. Flow localization defects are enhanced by temperature buildup in the workpiece (see the section "Flow Localization" in this chapter).

Friction. An important concern in all practical metalworking processes is the friction between the deforming workpiece and the tools and/or dies that apply the force and constrain the shape change. Friction occurs because metal surfaces, at least on a microscale, are never perfectly smooth and flat. Relative motion between such surfaces is impeded by contact under pressure.

The existence of friction increases the value of the deformation force and makes deformation more inhomogeneous (Fig. 17c), which in turn increases the propensity for fracture. If friction is high, seizing and galling of the workpiece surfaces occur, and surface damage results.

The mechanics of friction at the tool-workpiece interface are very complex; therefore, simplifying assumptions are usually used. One such assumption is that friction can be described by Coulomb's law of friction:

$$\mu = \frac{\tau_i}{p} \qquad \text{(Eq 49)}$$

where μ is the Coulomb coefficient of friction, τ_i is the shear stress at the interface, and p is the stress (pressure) normal to the interface.

Another simplification of friction is to assume that the shear stress at the interface is directly proportional to the flow stress, σ_0, of the material:

$$\tau_i = m\frac{\sigma_0}{\sqrt{3}} \quad \text{(Eq 50)}$$

where *m,* the constant of proportionality, is the interface friction factor. (The context of the situation usually allows one to differentiate whether *m* is friction factor or strain-rate sensitivity.) For given conditions of lubrication and temperature and for given die and workpiece materials, *m* is usually considered to have a constant value independent of the pressure at the interface. Values of *m* vary from 0 (perfect sliding) to 1 (no sliding). In the Coulomb model of friction, τ_i increases with *p* up to a limit at which interface shear stress equals the yield stress of the workpiece material.

Control of friction through lubrication is an important aspect of metalworking. High friction leads to various defects that limit workability. However, for most workability tests, conditions are selected under which friction is either absent or easily controlled. Most workability tests make no provision for reproducing the frictional conditions that exist in the production process; consequently, serious problems can result in the correlation of test results with actual production conditions.

Workability Fracture Criteria

Workability is not a unique property of a given material. It depends on such process variables as strain, strain rate, temperature, friction conditions, and the stress state imposed by the process. For example, metals can be deformed to a greater extent by extrusion than by drawing because of the compressive nature of the stresses in the extrusion process that makes fracture more difficult. It is useful to look at workability through the relationship:

$$\text{Workability} = f_1\ (\text{material}) \cdot f_2(\text{process}) \quad \text{(Eq 51)}$$

where f_1 is a function of the basic ductility of the material, and f_2 is a function of the stress and strain imposed by the process. Because f_1 depends on the material condition and the fracture mechanism, it is a function of temperature and strain rate. Similarly, f_2 depends on such process conditions as lubrication (friction) and die geometry. Therefore, to describe workability in a fundamental sense, a fracture criterion must be established that defines the limit of strain as a function of strain rate and temperature. Also required is a description of stress, strain, strain rate, and temperature history at potential fracture sites. The application of computer-aided finite-element analysis of plastic deformation and heat flow has facilitated this goal (see Chapters 10, "Thermomechanical Testing," and 11, "Design for Deformation Processes," in this Handbook). The forming-limit diagrams illustrated and discussed in Chapter 12, "Workability Theory and Application in Bulk Forming Processes," in this Handbook are a direct outgrowth of the previously mentioned concept of workability.

The simplest and most widely used fracture criterion (Ref 19) is not based on a micromechanical model of fracture but simply recognizes the joint roles of tensile stress and plastic strain in producing fracture:

$$\int_0^{\varepsilon_f} \bar{\sigma}\left(\frac{\sigma^*}{\bar{\sigma}}\right) d\bar{\varepsilon} = C \quad \text{(Eq 52)}$$

where $\bar{\sigma}$ is the effective stress, σ^* is the maximum tensile stress, and *C* is a material constant evaluated from the compression test. The effective strain is defined by:

$$\bar{\varepsilon} = \left[\frac{2}{3}(\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2)\right]^{\frac{1}{2}} \quad \text{(Eq 53)}$$

This fracture criterion indicates that fracture occurs when the tensile strain energy per unit volume reaches a critical value.

Use of this fracture criterion is shown in Fig. 24. The values of the reduction ratio at which centerburst fracture occurs in the cold extrusion of two aluminum alloys are illustrated. The energy conditions for different die angles are given by the three curves that reach a maximum. The fracture curves for the two materials slope down to the right. Centerburst occurs in the regions of reduction, for which the process energies exceed the material fracture curve. No centerburst occurs at small or large reduction ratios.

Process Maps

Mapping Based on Deformation Mechanisms. It has been seen that temperature and strain rate are critical parameters in ensuring that a deformation process achieves success, particularly at elevated temperature where many different mechanisms that lead to failure can occur. It would be advantageous if a processing map could be developed that considered all of the failure mechanisms that can operate in a material over a range of strain rates and temperatures. This kind of processing map was developed for aluminum and is based on theoretical models of fracture mechanisms (Ref 20); however, it agrees well with experimental work. As Fig. 25 shows, a safe region is indicated in which the material is free from cavity formation at hard particles, leading to ductile fracture, or wedge cracking at grain-boundary triple points. The processing map predicts that, at constant temperature, there should be a maximum in ductility with respect to strain rate. For example, at 500 K

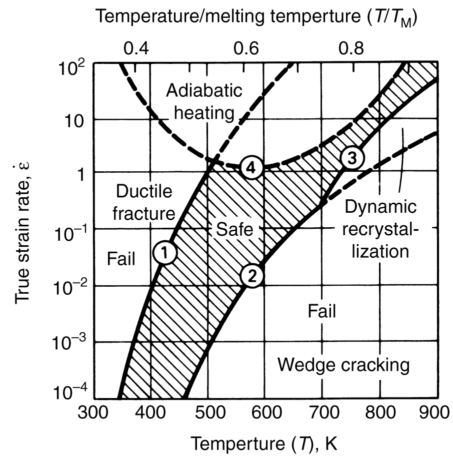


Fig. 25 Composite processing map for aluminum showing the safe region for forming. Boundaries shift with microstructure. Instabilities due to purely continuum effects, such as shear localization in sheet metal forming, are not considered. Source: Ref 20

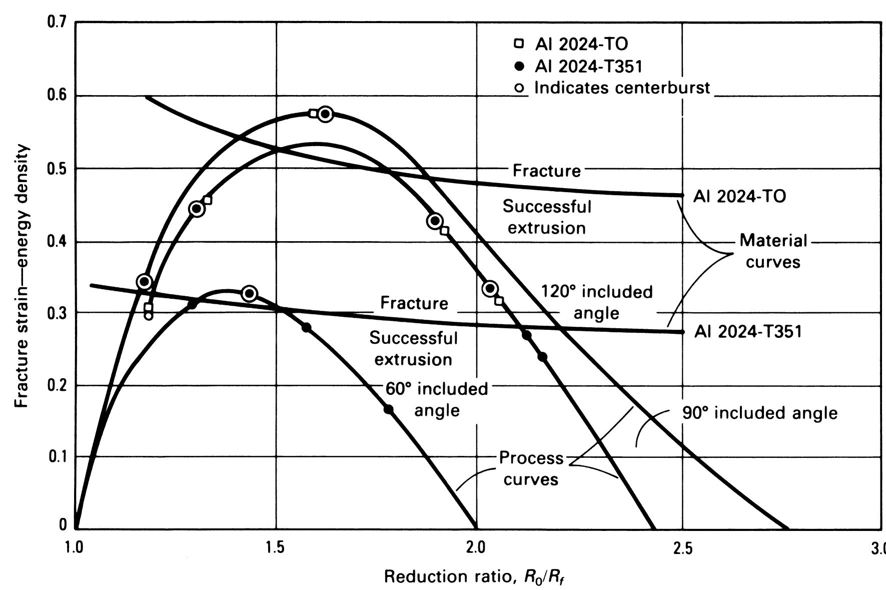


Fig. 24 Application of the Cockcroft-Latham fracture criterion (maximum tensile strain energy) to predict occurrence of centerburst fracture in aluminum alloy 2024 extrusions

(227 °C, or 440 °F), ductility should be at a maximum at a strain rate of 10^{-3} to 1 s^{-1}. Below the lower value, wedge cracking occurs; above this level, ductile fracture reduces ductility.

The safe region shown in Fig. 25 is sensitive to the microstructure of the metal. Decreasing the size and volume fraction of hard particles moves boundary 1 to the left. Increasing the size or fraction of hard particles in the grain boundary makes sliding more difficult and moves boundary 2 to the right.

While a processing map is a very useful guide for the selection of deformation processing conditions, maps that are constructed from mechanistic models of fracture are limited in practical application as a design tool. The analyses can be made only for pure metals and simple alloys, not for complex engineering materials in which strain-rate sensitivity is a function of temperature and strain rate. Moreover, the numerous material parameters, such as diffusivity, that must be introduced into the models are difficult to obtain for complex engineering alloys. Of more importance, the location of the boundaries in the processing map is very sensitive to microstructure and to prior thermomechanical history. It is difficult to account for these factors in the mechanistic models.

Dynamic Material Modeling. Many workability problems arise when deformation is localized into a narrow zone, resulting in a region of different structure and properties that can be a site of failure. In severe cases, this leads to failure during processing. When less severe, it serves as a potential site of failure in service. As has been seen, such flow localization can occur in hot working in the absence of die chilling or frictional effects. Flow softening causes localization of deformation due to deformation heating or microstructural instabilities, such as texture softening or dynamic spheroidization.

Because it is impossible to model from first principles so many structural phenomena, especially in complex alloys, a different approach is needed to create a processing map. The approach that has evolved uses continuum principles with macroscopic determination of flow stress as a function of temperature and strain rate and applies appropriate criteria of instability to identify the regions of T and $\dot{\varepsilon}$ that should be avoided in processing (Ref 21, 22). Then, microscopic studies identify the nature of the instability. Safe regions for processing are those that promote dynamic recrystallization, dynamic recovery, or spheroidization. Regions to avoid in processing are those that produce void formation at hard particles, wedge cracking, adiabatic shear band formation, and flow localization.

The technique known as *dynamic material modeling* (DMM) maps the power efficiency of the deformation of the material in a strain rate/temperature space (Fig. 26). At a hot working temperature, the power per unit volume, P, absorbed by the workpiece during plastic flow is:

$$P = \bar{\sigma}\dot{\bar{\varepsilon}} = \int_0^{\dot{\varepsilon}} \sigma d\dot{\varepsilon} + \int_0^{\sigma} \dot{\varepsilon} d\sigma \quad \text{(Eq 54)}$$

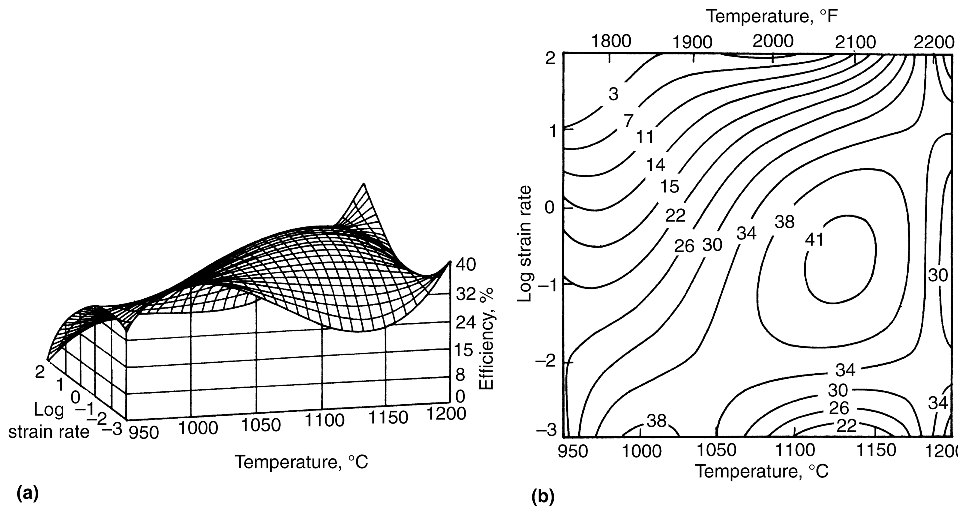


Fig. 26 Dynamic material modeling processing map for the nickel-base superalloy Nimonic AP1. (a) Three-dimensional plot of efficiency of power dissipation as a function of temperature and strain rate. (b) The corresponding contour map with numbers representing constant efficiency of power dissipation. Source: Ref 25

or

$$P = G + J \quad \text{(Eq 55)}$$

where G is the power dissipated by plastic work (most of it converted into heat), and J is the dissipator power co-content, which is related to the metallurgical mechanisms that occur dynamically to dissipate power. A strong theoretical basis for this position has been developed from continuum mechanics and irreversible thermodynamics (Ref 21, 22). Figure 27 illustrates the definitions of G and J. At a given deformation temperature and strain:

$$J = \int_0^{\bar{\sigma}} \dot{\bar{\varepsilon}} d\bar{\sigma} = \frac{\bar{\sigma}\dot{\bar{\varepsilon}}m}{m+1} \quad \text{(Eq 56)}$$

where the constitutive equation for the material is Eq 31, and m is the strain-rate sensitivity. The value of J reaches its maximum when $m = 1$. Therefore:

$$J_{max} = \frac{\bar{\sigma}\dot{\bar{\varepsilon}}}{2} \quad \text{(Eq 57)}$$

This leads to the chief measure of the power dissipation capacity of the material, the dimensionless parameter called the *efficiency of power dissipation,* η:

$$\eta = \frac{J}{J_{max}} = \frac{2m}{m+1} \quad \text{(Eq 58)}$$

Deformation processing should be focused on the regions of maximum efficiency of power dissipation, unless structural instabilities, for example, flow localization, intrude (Fig. 26). The location of regions of microstructural instability are found by mapping the instability parameter, $\xi(\dot{\varepsilon})$, where:

$$\xi(\dot{\varepsilon}) = \frac{\partial \ln[m/(m+1)]}{\partial \ln \dot{\varepsilon}} + m < 0 \quad \text{(Eq 59)}$$

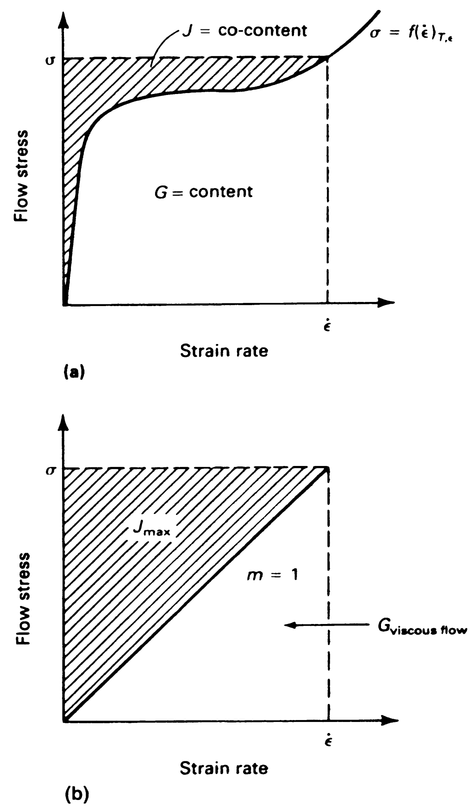


Fig. 27 Schematic of constitutive relation of material system as energy converter (dissipator). (a) Material system as nonlinear (general case) energy dissipator. (b) Material system as linear (special case) dissipator

The shaded region in Fig. 28 shows the processing conditions to be avoided because of flow localization. The instability criteria used in Eq 59 are based on the principle of maximum entropy production (Ref 23), that is, when the rate of entropy production by a microstructural change in the material is lower than the applied rate of entropy, the material exhibits instabilities during flow.

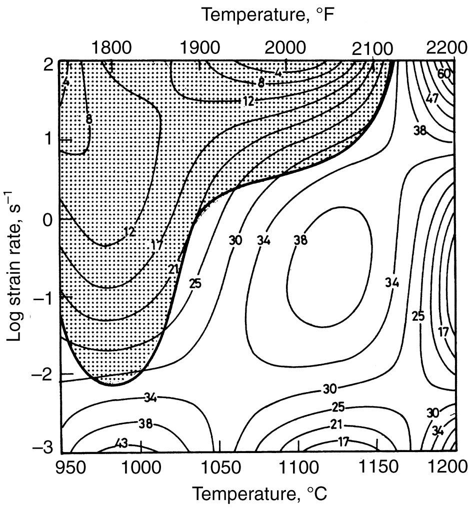


Fig. 28 Finished processing map for the nickel-base superalloy Nimonic AP1. Obtained by superposition of instability regions determined with Eq 59 with contours of percent efficiency of power dissipation. Shaded region corresponds to conditions of flow instability. Source: Ref 25

The DMM methodology describes the dynamic path a material element takes in response to an instantaneous change in $\dot{\varepsilon}$ at a given T and ε. As such, it is a map that graphically describes power dissipation by the material in stable and unstable ways. These boundaries correspond to safe and unsafe regions on a processing map. The use of the DMM methodology is in its infancy, but it appears to be a powerful tool for evaluating workability and controlling microstructure by thermomechanical processing in complex alloy systems (e.g., see Ref 24 and Chapter 22, "Multidisciplinary Process Design and Optimization: An Overview" in this Handbook). A major compendium of processing maps has been published (Ref 25).

Summary

The workability of metals is a complicated subject. Workability is both a material attribute and a process characteristic. The two aspects interact in complex ways. The process establishes the state of stress in the material as well as its temperature, strain, and strain profile. In addition, the process imposes frictional conditions that affect the stress state.

The material response to the process environment is expressed by the level of stress needed to deform the material, the flow stress, and by its propensity for fracture. The flow characteristics can be measured by methods described in Chapter 4, "Bulk Workability Testing," in this Handbook. Also discussed there are tests expressly designed to evaluate fracture.

Most commercial metals are complex alloys that undergo a wide range of structural behavior over the range of strain, strain rate, and temperature common to deformation processing. These mechanisms of fracture and instability must be understood and identified. Ideally, enough information is available to plot a process map that delineates the safe and unsafe regions for processing. One way of doing this is with the methodology called DMM. Another is to use the tests described in Chapter 4, "Bulk Workability Testing," in this Handbook and in more detail in other chapters of this Handbook to simulate the process of interest and to identify the safe processing regions.

REFERENCES

1. G.E. Dieter, *Mechanical Metallurgy*, 3rd ed., McGraw-Hill, 1986, p 18–36, 289–290
2. P.W. Bridgman, *Trans. ASM,* Vol 32, 1944, p 553
3. N.E. Dowling, *Mechanical Behavior of Materials,* Prentice-Hall, 1993, p 165
4. J.M. Holt, Uniaxial Tension Testing, *Mechanical Testing and Evaluation,* Vol 8, *ASM Handbook,* ASM International, 2000, p 131–132
5. J. Datsko, *Mechanical Properties and Manufacturing Processes,* John Wiley & Sons, 1966, p 18–20
6. W.B. Morrison, *Trans. ASM,* Vol 59, 1966, p 824
7. J.W. House and P.G. Gillis, Testing Machines and Strain Sensors, *Mechanical Testing and Evaluation,* Vol 8, *ASM Handbook,* ASM International, 2000, p 78–92
8. D. Zhao, Testing for Deformation Modeling, *Mechanical Testing and Evaluation,* Vol 8, *ASM Handbook,* ASM International, 2000, p 798–810
9. G.R. Johnson, J.M. Hoegfeldt, U.S. Lindholm, and A. Nagy, Response of Various Metals to Large Torsional Strains Over a Large Range of Strain Rates, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 105, 1983, p 42–53
10. H. Inoue and M. Nishihara, Ed., *Hydrostatic Extrusion,* Elsevier, 1985
11. V. Vujovic and A.H. Shabaik, A New Workability Criterion for Ductile Metals, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 108, 1986, p 245–249
12. T.H. Courtney, *Mechanical Behavior of Materials,* 2nd ed., McGraw-Hill, 2000, p 522–550
13. T. Gladman, B. Holmes, and L.D McIvor, *Effect of Second-Phase Particles on the Mechanical Properties of Steel,* Iron and Steel Institute, 1971, p 78
14. S.L. Semiatin and G.D. Lahoti, The Occurrence of Shear Bands in Isothermal Hot Forging, *Metall. Trans. A,* Vol 13, 1982, p 275–288
15. J.J. Jonas, R.A. Holt, and C.E. Coleman, Plastic Stability in Tension and Compression, *Acta Metall.,* Vol 24, 1976, p 911
16. M.C. Mataya and G. Krauss, A Test to Evaluate Flow Localization During Forging, *J. Appl. Metalwork.,* Vol 2, 1981, p 28–37
17. F.N. Rhines and P.J. Wray, Investigation of the Intermediate Temperature Ductility Minimum in Metals, *Trans. ASM,* Vol 54, 1961, p 117
18. A.M. Sabroff, F.W. Boulger, and H.J. Henning, *Forging Materials and Practices,* Reinhold, 1968
19. M.G. Cockcroft and K.J. Latham, Ductility and Workability of Metals, *J. Inst. Met.,* Vol 96, 1968, p 33–39
20. R. Raj, Development of a Processing Map for Use in Warm-Forming and Hot-Forming Processes, *Metall. Trans. A,* Vol 13, 1982, p 275–288
21. H.L. Gegal, Synthesis of Atomistics and Continuum Modeling to Describe Microstructure, *Computer Simulation in Materials Processing,* ASM International, 1987
22. S.V.S. Narayana Murty, B. Nageswara Rao, and B.P. Kashyap, Instability Criteria in Hot Deformation of Materials, *Int. Mater. Rev.,* Vol 45 (No.1), 2000, p 15–26
23. H. Ziegler, in *Progress in Solid Mechanics,* I.N. Sneddon and R. Hill, Ed., Vol 4, North-Holland, Amsterdam, 1963
24. J.C. Malas and V. Seetharaman, Using Material Behavior Models to Develop Process Control Strategies, *JOM,* Vol 44, (No. 6) 1992, p 8–13
25. Y.V.R.K. Prasad and S. Sasidhara, *Hot Working Guide: A Compendium of Processing Maps,* ASM International, 1997

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p35-44
DOI:10.1361/hwpd2003p035

Chapter 3

Evolution of Microstructure during Hot Working

S.L. Semiatin, Air Force Research Laboratory, Materials and Manufacturing Directorate

IN PROCESS DESIGN, a compromise is often necessary in the selection of conditions that give optimal workability versus those that yield a desired microstructure. The term *workability* is frequently applied in connection with the prevention of cracking or defects during working and the determination of flow stresses and total working loads in metal-forming practice. These factors determine the necessary conditions of plastic flow (such as temperature, strain, and strain rate) and the required press capacity, die materials and wear, and certain other die-design features. In addition, the development of microstructure during bulk working, especially at hot-working temperatures, is an equally important consideration when selecting processing conditions.

The development of desirable microstructure during hot working is a critical factor in process design, and the description of microstructure evolution (based on both phenomenological as well as mechanism-based approaches) is reaching a state of maturity. Thus, the design of thermomechanical processes to control and optimize microstructure for specific service applications is becoming increasingly widespread. In this chapter, the general aspects of microstructure evolution during thermomechanical processing are reviewed briefly. The effect of thermomechanical processing on microstructure evolution is briefly summarized to provide insight into this aspect of process design. Attention is focused on hot working and the key processes that control microstructure evolution—that is, dynamic and static recovery and recrystallization and grain growth. The discussion is divided into three main subsections dealing with the mechanisms, phenomenology, and modeling of microstructure evolution.

Mechanisms of Microstructure Evolution

The key mechanisms that control microstructure evolution during hot working and subsequent heat treatment are dynamic recovery, dynamic recrystallization, metadynamic recrystallization, static recovery, static recrystallization, and grain growth (Ref 1, 2).

Dynamic Recovery and Recrystallization. As its name implies, dynamic recovery and recrystallization occurs *during* hot working. As metals are worked, defects are generated in the crystal lattice. The most important defects are line defects known as dislocations. As deformation increases, the deformation resistance increases due to increasing dislocation content. However, the dislocation density does not increase without limit because of the occurrence of dynamic recovery and dynamic recrystallization.

In high stacking-fault-energy (SFE) metals (e.g., aluminum and its alloys, iron in the ferrite phase field, titanium alloys in the beta phase field), dynamic recovery (DRV) predominates. During such processes, individual dislocations or pairs of dislocations are annihilated because of the ease of climb (and the subsequent annihilation of dislocations of opposite sign) and the formation of cells and subgrains that act as sinks for moving (mobile) dislocations. Because subgrains are formed and destroyed continuously during hot working, the hot deformed material often contains a collection of *equiaxed* subgrains (with low misorientations across their boundaries) contained within *elongated* primary grains (Ref 2, 3). An example of such microstructural features for an aluminum-lithium alloy is shown in Fig. 1. Furthermore, the dynamic-recovery process leads to low stresses at high temperatures. Thus cavity nucleation and growth are retarded, and ductility is high. The evolution of microstructure in high-SFE (and some low-SFE) materials worked at lower temperatures, such as those characteristic of cold-working, is similar. At these temperatures, subgrains may also form and serve as sinks for dislocations. However, the subgrains are more stable. Thus, as more dislocations are absorbed into their boundaries, increasing misorientations are developed, eventually giving rise to an equiaxed structure of high-angle boundaries. Such a mechanism forms the basis for grain refinement in so-called severe plastic deformation (SPD) processes such as equal channel angular extrusion (ECAE) (Ref 4–6). This mechanism of grain refinement is sometimes called *continuous dynamic recrystallization* (CDRX) because of the gradual nature of the formation of high angle boundaries with increasing strain.

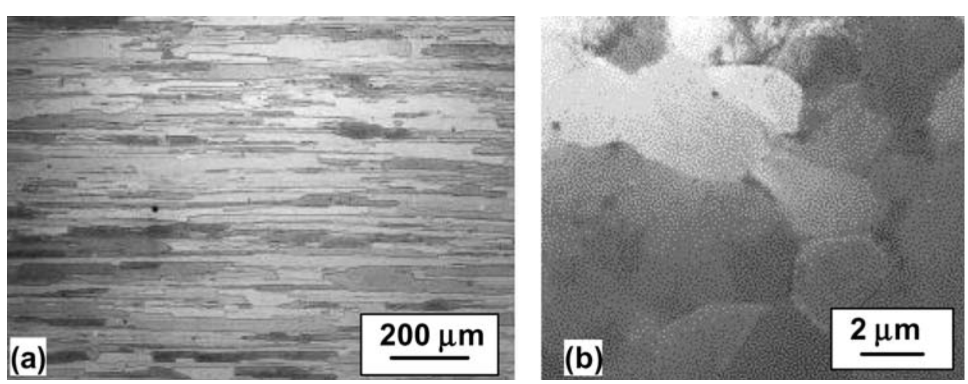


Fig. 1 Evolution of microstructure during hot-rolling of an aluminum lithium alloy undergoing dynamic recovery. (a) Optical micrograph showing heavily deformed elongated initial grains and (b) TEM micrograph showing equiaxed subgrains. Source: K.V. Jata, Air Force Research Laboratory

In low-SFE materials (e.g., iron and steel in the austenite phase field, copper, and nickel), dynamic recovery occurs at a lower rate under hot working conditions because mobile dislocations are dissociated. Therefore, climb is difficult. This leads to somewhat higher densities of dislocations than in materials whose deformation is controlled by dynamic recovery. Furthermore, as the temperature is increased, the mobility of grain boundaries increases rapidly. Differences in dislocation density across the grain boundaries, coupled with high mobility, lead to the nucleation and growth of new, strain-free grains via a *discontinuous dynamic recrystallization process* (DDRX) (Ref 2, 7). The evolution of a dynamically recrystallized microstructure is illustrated in Fig. 2. At large strains, a fully recrystallized structure is obtained. However, even at this stage, recrystallized grains are being further strained and thus undergo additional cycles of dynamic recrystallization. Nevertheless, a steady state is reached in which the rate of dislocation input due to the imposed deformation is balanced by dislocation annihilation due to the nucleation and growth of new grains (as well as some dynamic recovery). Hence, although a nominally equiaxed grain structure is obtained at large strains, the distribution of stored energy is not uniform.

The presence of second-phase particles may affect the evolution of microstructure during hot working of both high- and low-SFE materials. In high-SFE materials, particles may affect the homogeneity and magnitude of dislocation substructure that evolves. In low-SFE materials, particles may affect the evolution of substructure, serve as nucleation sites for dynamic recrystallization, and serve as obstacles to boundary migration during the recrystallization process.

Static Softening Processes and Microstructure Evolution. Residual dislocations from hot working play an important role in the evolution of microstructure during heat treatment following hot working. In high-SFE materials, static recovery or static recovery and recrystallization may occur, depending on the level of stored work, the rate at which the material is reheated, and the annealing temperature/boundary mobility. Static recovery is similar to dynamic recovery in that climb of dislocations and the absorption of dislocations into subboundaries occur. Furthermore, subgrain growth may occur, further reducing dislocation density. However, a totally strain-free material may not be obtained even after long annealing times in the absence of static recrystallization. In high-SFE metals such as aluminum that contain second-phase particles, static recrystallization may also occur as a result of particle-stimulated nucleation (PSN) (Ref 1). In these cases, subgrains in regions of locally higher deformation adjacent to the particles serve as nuclei for recrystallization. These subgrains grow rapidly, consuming the substructure of the surrounding material.

For low-SFE metals, a number of static softening processes may occur, depending on the level of deformation during hot working. These are summarized in Fig. 3 for pure nickel (Ref 8). For prestrains much less than those required to initiate DDRX, sufficient stored energy to nucleate static recrystallization is not available, and only static recovery occurs. Therefore, some dislocation substructure is retained, and full softening is not obtained. At somewhat higher prestrains just below those at which DDRX is initiated, static recovery is followed by static recrystallization, which involves a nucleation-and-growth process. In this case, a fully annealed/softened microstructure is obtained. For prestrains that exceed those at which DDRX was initiated during hot working, residual DDRX nuclei undergo very rapid growth without an incubation period in a process known as post- or meta-dynamic recrystallization (MDRX); this is followed by static recovery and recrystallization. Partially worked regions may also undergo metadynamic recovery, thus reducing the stored energy needed for nucleation of static recrystallization and thus the ability to obtain a fully softened condition.

Following static recovery and recrystallization, grain growth also occurs frequently (Ref 1). As in recovery and recrystallization, the driving force for grain growth is a reduction in stored energy. For grain growth, the stored energy is in the form of grain-boundary energy. Grain growth may lead to very large grain sizes (and sharp crystallographic textures) particularly in single-phase metals heat treated at high temperatures. In alloys containing second-phase particles, an equilibrium grain size may be reached due to pinning of the grain boundaries by the second phase.

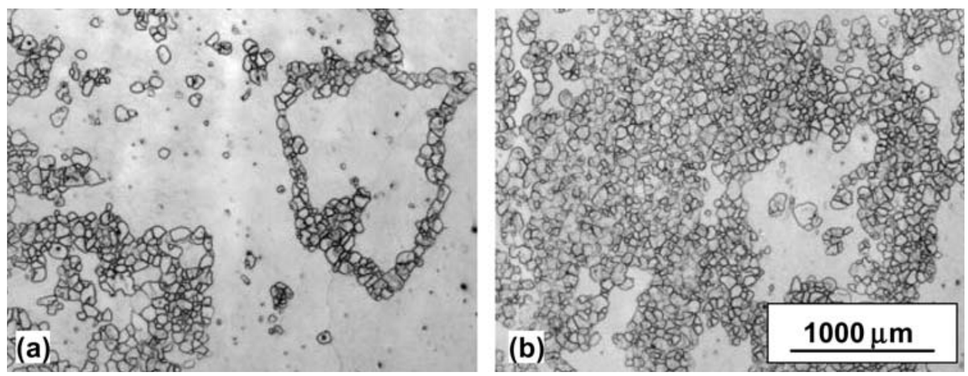


Fig. 2 Discontinuous dynamic recrystallization (DDRX) in an initially coarse-grained nickel-base superalloy. (a) Initial stage of DDRX and (b) nearly fully recrystallized microstructure

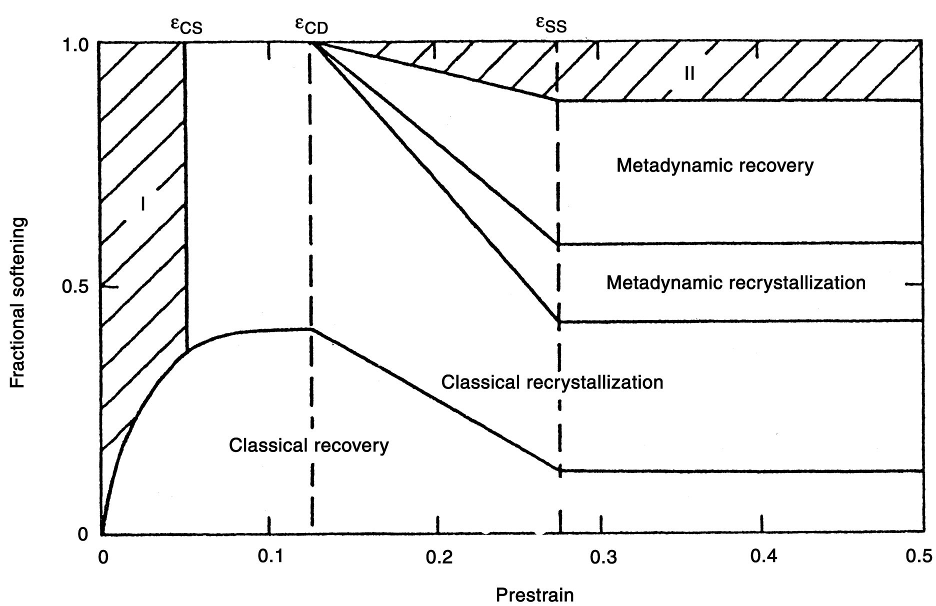


Fig. 3 Schematic illustration of the effect of hot-working prestrain on the subsequent static softening mechanisms and relative magnitudes of softening due to each for pure nickel. The cross-hatched areas indicate conditions under which incomplete softening occurs. Source: Ref 8

Phenomenology of Plastic Flow and Microstructure Evolution

Some of the key phenomenological descriptions of plastic flow and microstructure evolution are summarized in this section. The quantitative description of plastic flow (i.e., *stress-strain* or *flow curves*) and microstructure evolution depends on the specific mechanism controlling deformation (and heat-treatment) response.

Flow Curves

The stress-strain (flow) curves that are measured under hot working conditions are a function of the predominant dynamic softening mechanism.

Flow Curves for Dynamic Recovery. As mentioned earlier, the hot-working response of high-SFE metals is controlled by dynamic recovery. In such cases, dislocation generation is offset by dislocation annihilation due to recovery processes. The flow curve thus exhibits an initial stage of strain hardening followed by a steady-state (constant) flow stress. Typical flow curves for iron in the ferrite phase field are shown in Fig. 4 (Ref 9). The magnitude of the steady-state flow stress σ_{ss} decreases with increasing temperature, T, and decreasing strain rate, $\dot{\varepsilon}$, typically according to a relation as follows (Ref 1, 2):

$$Z \equiv \dot{\varepsilon}\exp(Q_{drv}/RT) = C_1 \sinh(C_2\sigma_{ss})^{n_{drv}} \quad \text{(Eq 1)}$$

in which Z denotes the Zener-Hollomon parameter, Q_{drv} is an apparent activation energy, R is the gas constant, and C_1, C_2, and n_{drv} are constants.

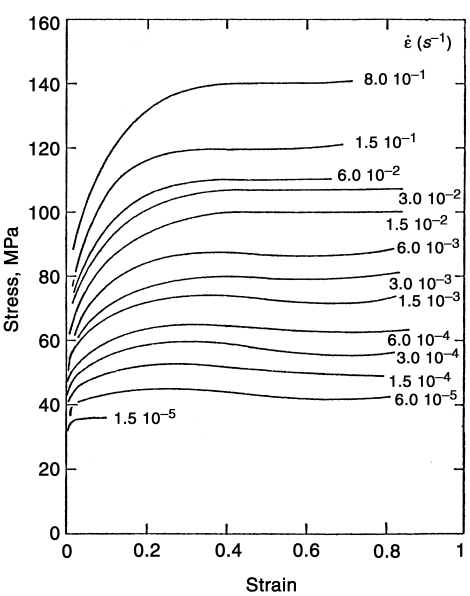


Fig. 4 Flow curves for Armco iron deformed under hot-working conditions in the ferrite-phase field. Source: Ref 9

A somewhat more quantitative insight into the shape of the flow curve for cases involving dynamic recovery may be obtained by an analysis of the overall rate of change of (mobile) dislocation density, ρ, with strain ε, $d\rho/d\varepsilon$, i.e.:

$$d\rho/d\varepsilon = d\rho/d\varepsilon\big|_{\text{storage}} - d\rho/d\varepsilon\big|_{\text{recovery}} \quad \text{(Eq 2)}$$

The specific functional form of the dislocation storage and annihilation terms in Eq 2 can be expressed in the following form (Ref 10, 11):

$$d\rho/d\varepsilon = U - \Omega\rho \quad \text{(Eq 3)}$$

In Eq 3, U denotes the rate of dislocation generation due to strain hardening, and Ω is a factor describing the rate of dynamic recovery. The rate of recovery is also directly proportional to the instantaneous level of dislocation density, ρ. At hot working temperatures, U is independent of strain rate and temperature to a first order (Ref 11). Thus, the strain rate and temperature dependence of the rate of dislocation multiplication is determined principally by $\Omega = \Omega(\dot{\varepsilon}, T)$. An example of such a dependence for a microalloyed (Nb-B) steel is shown in Fig. 5(a) (Ref 11).

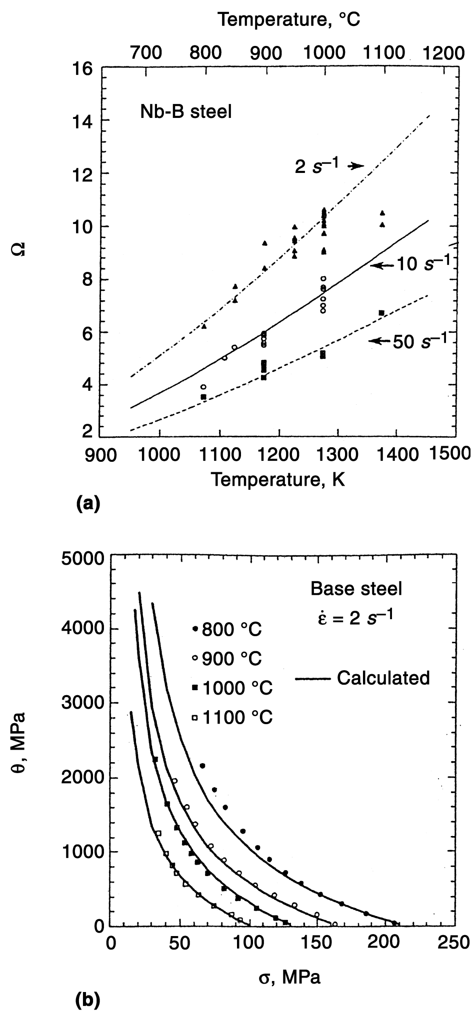


Fig. 5 Characterization of flow behavior under hot working conditions. (a) Ω as a function of temperature and strain rate for a Nb-B microalloyed steel and (b) the overall hardening rate, θ, = $d\sigma/d\varepsilon$ as a function of stress, σ, at a strain rate of 2 s^{-1} and various temperatures for a low carbon steel. Source: Ref 11

To a first order, the flow stress σ under working conditions is given by the following expression:

$$\sigma = \alpha G b\sqrt{\rho} \quad \text{(Eq 4)}$$

in which α denotes a constant whose magnitude is between 0.5 and 1.0, G is the shear modulus, and b is the length of the Burgers (slip) vector. Inspection of Eq 3 and 4 reveals that the strain-rate and temperature dependence of the overall rate of hardening in the flow curve is largely determined by Ω. The strain-hardening rate is frequently quantified in terms of plots of $d\sigma/d\varepsilon$ ($\equiv \theta$) as a function of σ. Typical plots for low-carbon steel are shown in Fig. 5(b) (Ref 11).

Equation 3 reveals that a steady-state dislocation density, ρ_{ss}, is reached when $U = \Omega\rho$, or:

$$\rho_{ss} = U/\Omega \quad \text{(Eq 5)}$$

The steady-state flow stress, σ_{ss}, is thus given by the following relation:

$$\sigma_{ss} = \alpha G b\sqrt{\rho_{ss}} = \alpha G \mathbf{b}\sqrt{U/\Omega} \quad \text{(Eq 6)}$$

There are a number of alternate approaches to the modeling of dynamic recovery under the broad framework of Eq 2. For example, Kocks (Ref 12) has shown that a linear dependence of strain-hardening rate ($d\sigma/d\varepsilon$) on stress (σ) is consistent with the following relation for $d\rho/d\varepsilon$:

$$d\rho/d\varepsilon = (k_1\sqrt{\rho} - k2\rho)/\mathrm{b} \quad \text{(Eq 7)}$$

in which k_1 and k_2 are constants. Similarly, for a strain-hardening rate that varies linearly with $1/\sigma$, Roberts (Ref 13) has shown that the following relation applies:

$$d\rho/d\varepsilon = k_3 - k_4\sqrt{\rho} \quad \text{(Eq 8)}$$

in which k_3 and k_4 are constants.

Flow Curves for Discontinuous Dynamic Recrystallization. Flow curves for materials undergoing discontinuous dynamic recrystallization (DDRX) have shapes that are distinctively different from those that characterize materials that soften solely by dynamic recovery. Those for DDRX exhibit an initial strain hardening transient, a peak stress, flow softening, and, finally, a period of steady-state flow. Typical curves for two austenitic stainless steels are shown in Fig. 6 (Ref 14).

Dynamic recrystallization typically initiates at a strain of approximately five-sixths of the strain corresponding to the peak stress. Because of this behavior, the peak stress for a material that undergoes dynamic recrystallization is *less* than that which would be developed if the material softened solely by dynamic recovery (Fig. 7a). The strain at which DDRX initiates as well as the steady-state flow stress that would be developed in the absence of DDRX are readily determined from a plot of $d\sigma/d\varepsilon$ as a function of σ (Fig. 7b).

From a phenomenological standpoint, the strain at the peak stress ε_p is usually found to depend on the initial grain size, d_o, and the Zener-Hollomon parameter, Z, per an expression of the following form (Ref 15–17):

$$\varepsilon_p = C_3\, d_o^{n_3} Z^{n_4} \qquad \text{(Eq 9)}$$

in which C_3, n_3, and n_4 are material-specific constants. The activation energy specific to DDRX is used in the determination of Z.

The regime of steady-state flow in stress-strain curves for materials that undergo dynamic recrystallization may be smooth or exhibit an oscillatory behavior that dampens with increasing strain. Temperature-strain rate conditions for which the dynamically recrystallized grain size is less than one-half of the initial grain size show the former behavior. By contrast, those that give rise to grain-size coarsening or a reduction of less than one-half of the starting grain size exhibit the oscillatory behavior (Ref 18).

Phenomenological Models for Microstructure Evolution

Dynamic Recovery. The key parameter that characterizes the microstructure that evolves in materials that dynamically recover is the steady-state subgrain size d_{ss}. Two of the common correlations of d_{ss} to process variables are those developed by McQueen and Hockett (Ref 3) and Derby (Ref 19). In the former work, d_{ss} was found to vary linearly with log Z (Fig. 8a). In the work of Derby, the normalized subgrain size (d_{ss}/b) of a number of materials was shown to vary inversely with the normalized steady-state flow stress (σ_{ss}/G) (Fig. 8b).

Discontinuous Dynamic Recrystallization (DDRX). The key parameters that characterize DDRX are the dynamically-recrystallized grain size d_{drx} and the kinetics of recrystallization as a function of strain ε. Devadas et al. (Ref 16) have summarized a number of such measurements. Typical relations are of the following form:

$$d_{drx} = C_4\, Z^{-n_5} \qquad \text{(Eq 10)}$$

$$X_{drx} = 1 - \exp\{-\ln 2[(\varepsilon - \varepsilon_c)/\varepsilon_{0.5}]^{n_6}\} \qquad \text{(Eq 11)}$$

In Eq 10 and 11, X_{drx} represents the fraction recrystallized; C_4, n_5, and n_6 are (positive) constants; and ε_c and $\varepsilon_{0.5}$ denote the critical strain for the onset of DDRX ($\approx 5\varepsilon_p/6$) and the strain for 50% recrystallization, respectively. Equation 11 is a classical Avrami-type relation (Ref 1, 20) for the nucleation-and-growth type processes that characterize DDRX. The Avrami exponent, denoted as n_6 above, typically has a value of approximately 2 for DDRX, but higher and lower values are not uncommon depending on the type (e.g., site saturation versus continuous) and sites (e.g., grain edges versus triple points) of nucleation and the dimensionality (e.g., two-dimensional or 2D versus three-dimensional or 3D) of growth (Ref 20–22). The strain $\varepsilon_{0.5}$ for 50% recrystallization is a weak function of Z, but depends strongly on the initial grain size d_o. Typically, $\varepsilon_{0.5} \sim d_o^{0.3}$. Such a dependence is not surprising in view of the relatively smaller volume fraction of nucleation sites (usually grain boundaries) in coarser-grain materials. Hence, greater amounts of hot work may be required to fully recrystallize a cast ingot material via DDRX as compared to a fine-grain wrought material unless nucleation occurs at sub-boundaries that develop within the coarse grains during hot working.

Metadynamic and Static Recrystallization. Although the nucleation-and-growth mechanisms are different for metadynamic (MDRX)

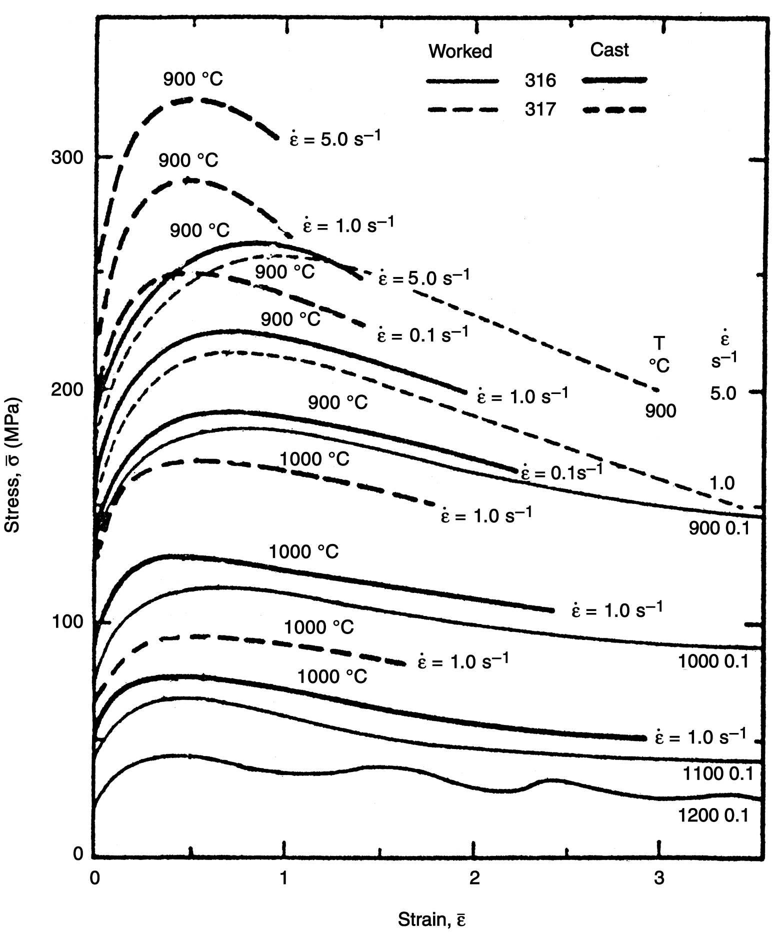


Fig. 6 Flow curves for 316 and 317 stainless steels deformed under hot-working conditions. Source: Ref 14

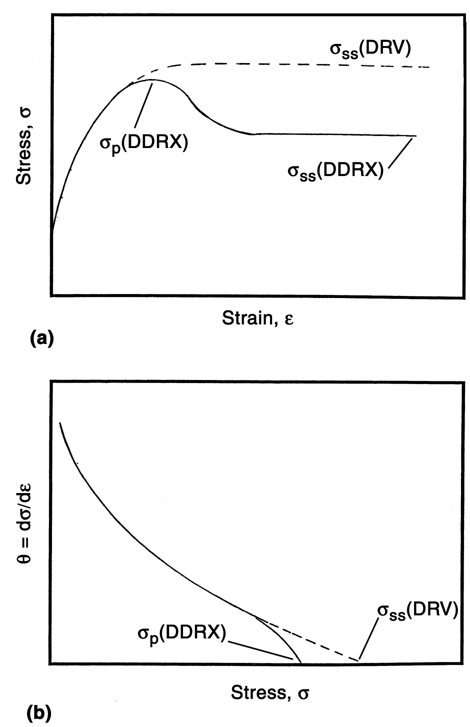


Fig. 7 Schematic illustration of work hardening behavior for a material undergoing dynamic recrystallization at hot-working temperatures. (a) Stress-strain curve and (b) corresponding plot of $d\sigma/d\varepsilon$ as a function of stress, σ

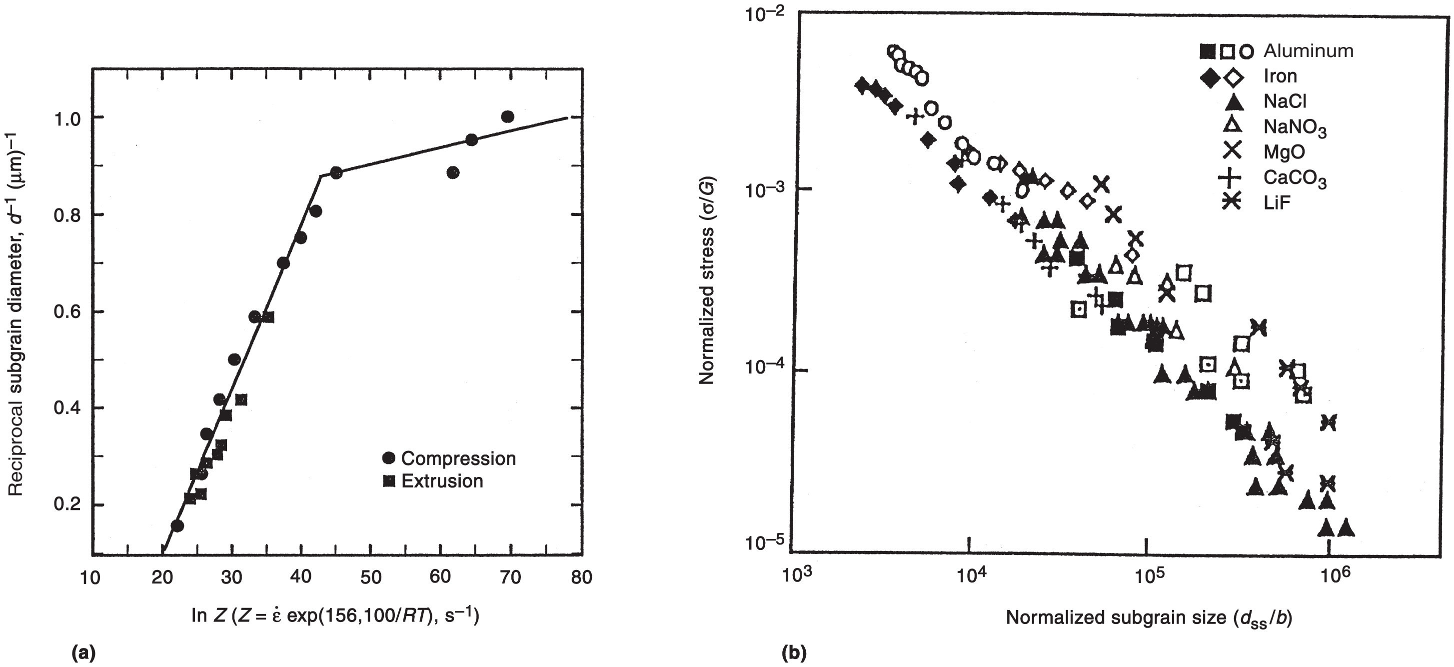


Fig. 8 Equilibrium subgrain sizes d_{ss} developed due to dynamic recovery during hot-working. (a) d_{ss} vs. Z for pure aluminum and (b) σ/G vs. d_{ss}/b for various materials. Source: Ref 3, 19

and static (SRX) recrystallization, as discussed earlier, the kinetics of both processes are also usually well fit by an Avrami relation of the following form (Ref 15, 16):

$$X_{\mathrm{mdrx,srx}} = 1 - \exp[-\ln 2(t/t_{0.5})^{n_7}] \quad \text{(Eq 12)}$$

in which $X_{\mathrm{mdrx,srx}}$ denotes the fraction transformed by metadynamic or static recrystallization, and t and $t_{0.5}$ are time and time for 50% recrystallization, respectively. The Avrami exponent n_7 takes a value typically of the order of 1 (MDRX) or 2 (SRX). For static recrystallization, the functional form for $t_{0.5}$ depends on whether the hot-working prestrain is greater or less than that required to initiate DDRX:

$$t_{0.5} = C_5\, \varepsilon^{-n_8}\, d_o^{\,n_9} \exp(Q_s/RT) \quad (\varepsilon < 5\varepsilon_p/6) \quad \text{(Eq 13a)}$$

$$t_{0.5} = C_6\, Z^{-n_{10}} \exp(Q_s/RT) \quad (\varepsilon > 5\varepsilon_p/6) \quad \text{(Eq 13b)}$$

In Eq 13a and 13b, C_5, C_6, n_8, n_9, and n_{10} are positive constants, and Q_s is the activation energy for static recrystallization. Similar expressions for $t_{0.5}$ apply for the static recrystallization of high-SFE energy metals that undergo only dynamic recovery during hot working (Ref 15); in these cases, different fits are used depending on whether the prestrain is greater or less than that at which the maximum (steady-state) flow stress is achieved.

The statically recrystallized grain size, d_{srx}, in metals that undergo DDRX is also a function of hot-working variables. Typical expressions are as follows (Ref 15, 16):

$$d_{\mathrm{srx}} = C_7 d_o^{\,n_{11}} \varepsilon^{-n_{12}} \quad (\varepsilon < \varepsilon^*) \quad \text{(Eq 14a)}$$

$$d_{\mathrm{srx}} = C_8 Z^{n_{13}} \quad (\varepsilon \geq \varepsilon^*) \quad \text{(Eq 14b)}$$

in which

$$\varepsilon^* = C_9 d_o^{\,n_{14}} \varepsilon_p \quad \text{(Eq 14c)}$$

and C_7, C_8, C_9, n_{11}, n_{12}, n_{13}, and n_{14} denote positive constants.

Grain Growth. Following recrystallization and the elimination of dislocation substructure, the principal driving force for microstructural changes is the reduction of grain-boundary energy. At elevated temperatures, classical isothermal grain growth in single-phase alloys is usually modeled with a phenomenological expression of the following form:

$$D^n - D_o^{\,n} = Kt \exp(Q_g/RT) \quad \text{(Eq 15)}$$

in which D and D_o denote the initial and final grain sizes, n is the grain-growth exponent (whose value typically ranges between 2 and 4), K is a constant, t is time, and Q_g is the activation energy for grain growth. Care must be exercised in applying this expression particularly for materials that contain or develop noticeable crystallographic textures in which the grain-growth exponent n may vary during growth (Ref 1). In addition, alloys containing second phase particles may exhibit parabolic growth kinetics initially, but eventually develop a stable grain size due to particle pinning. For volume fractions f_v of several percent or less of particles whose average diameter is d_p, the stable grain diameter D_s due to such "Zener pinning" is given by (Ref 1):

$$D_s \cong 2\alpha_p d_p / 3 f_v \quad \text{(Eq 16a)}$$

in which α_p is a constant between 0.25 and 1.0. Hellman and Hillert (Ref 23) and Hazzledine and Oldershaw (Ref 24) have developed analytical and computer models, respectively, for the stable, pinned grain size for higher volume fractions of particles (Fig. 9). For example, the former work leads to the following expression:

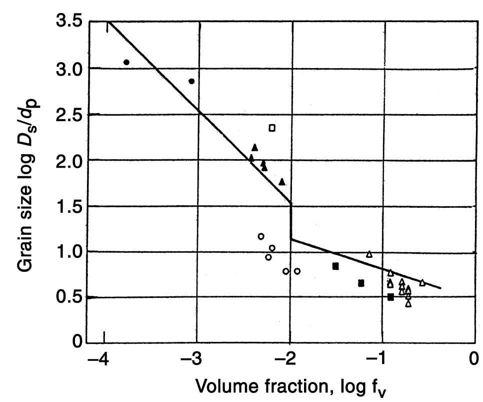


Fig. 9 Comparison of the measured dependence of the ratio of stable grain size (D_s) to particle size (d_p) on volume fraction of particles (f_v) (data points) and Zener-model predictions ($f_v \leq 0.01$) or computer simulation predictions of Hazzledine and Oldershaw ($f_v \geq 0.01$) (solid lines). Source: Ref 24

$$D_s \cong d_p (6/f_v)^{1/3} \quad \text{(Eq 16b)}$$

Mechanistic Models for Microstructure Evolution

A number of mechanism-based approaches have been developed to model microstructure evolution during hot working and annealing. These models incorporate deterministic and statistical aspects to varying degrees. In this section, several broad classes of models that are currently under development are discussed. These include cellular automata, Monte-Carlo techniques, analytical models of grain (and subgrain) growth, and texture evolution models. Each is briefly discussed subsequently.

Cellular Automata (CA)

Cellular automata (also sometimes referred to as cellular automaton) is a numerical procedure used to model a number of metallurgical processes based on nucleation and growth or growth alone. These include static and dynamic recrystallization, precipitation, and grain growth. Irrespective of the specific process under investigation, rules for nucleation and growth are specified and applied to a material domain composed of discrete cells. Each cell is surrounded by a *neighborhood* of adjacent cells of a given geometry that determines the shape of the transformed phase. The rules may be applied at random sites, thus introducing a stochastic element into the simulations as well. Because of the discrete nature of the calculations, spatial as well as average (temporal) predictions of structure evolution are obtained.

An example of the application of CA is static recrystallization (Ref 25–27). In the simulation, cells may be either unrecrystallized or recrystallized. A nucleation rule based on site saturation (all nuclei seeded at the beginning of the simulation) or continuous nucleation is most common. The specific sites chosen as nuclei are selected using a random number generator. For example, under continuous (constant rate N) nucleation conditions, the overall probability of nucleation is first calculated as $P_N = N dt / N_{CA}$, in which dt is the time step, and N_{CA} is the number of CA cells. A cell is then selected at random, and a random number between 0 and 1 is generated. If the random number is greater than P_N, the cell becomes a nucleus. The growth of a nucleus to consume cells in its neighborhood may be automatic (i.e., every cell with a previously recrystallized cell in its neighbor undergoes recrystallization) or specified based on some sort of probability-based rule as well. Typical predictions of the evolution of the recrystallized microstructure and fraction recrystallized from a CA simulation assuming site saturation conditions are shown in Fig. 10.

Refinements of CA static recrystallization models include the incorporation of the effect of inert second-phase particles on recrystallization as well as the extension to the modeling procedure to DDRX conditions. For example, in the former instance (Ref 28), a CA routine has been developed in which the effect of particles is purely geometric; that is, growth of a recrystallized grain was assumed to cease on impingement with a particle. The simulations lead to noticeable deviations from Avrami-like behavior (e.g., Eq 12) with increasing particle-aspect ratio and, to a lesser extent, with increasing particle fraction and size. In the CA simulation of dynamic recrystallization (Ref 29), rules for dislocation input and recovery, nucleation, and growth are specified. Specific sites at which recovery and nucleation occur may be chosen randomly or using deterministic methods based on local dislocation density. It has been shown that CA can replicate a number of the important features of DDRX such as the effect of initial grain size on the flow curve and necklace-recrystallization behavior (Fig. 11).

One of the major challenges of CA (and other kinds of discrete models) lies in the calibration of the simulations relative to real time and spatial scales. To address some of these difficulties, Davies (Ref 26) and Raabe (Ref 30) have used two different methods to fix the interface migration rate in CA simulations of static recrystallization. One is based on the theoretical relation between grain boundary velocity, grain boundary mobility (as a function of diffusion parameters), and the dislocation density difference (i.e., "pressure difference") across the moving interface. The other is semiempirical and is based on the Cahn-Hagel interface-averaged migration rate (Ref 31).

Monte-Carlo Techniques

Monte-Carlo (MC) techniques for simulating microstructure evolution were first developed in the 1980s to investigate problems such as grain growth, static recrystallization, and dynamic recrystallization. The foundation for MC methods lies in statistical mechanics (Ref 32) and the derivation of the Boltzmann factor. The Boltzmann factor describes the ratio of the probabilities, P, that a system can be found in two distinct states with energies of G_1 or G_2, namely:

$$P(G_2)/P(G_1) = [\exp(-G_2/k_bT)] / [\exp(-G_1/k_bT)] = \exp(-\Delta G/k_bT) \quad \text{(Eq 17)}$$

in which k_b is the Boltzmann constant.

Perhaps the simplest problem that can be addressed using the MC method is grain growth. In

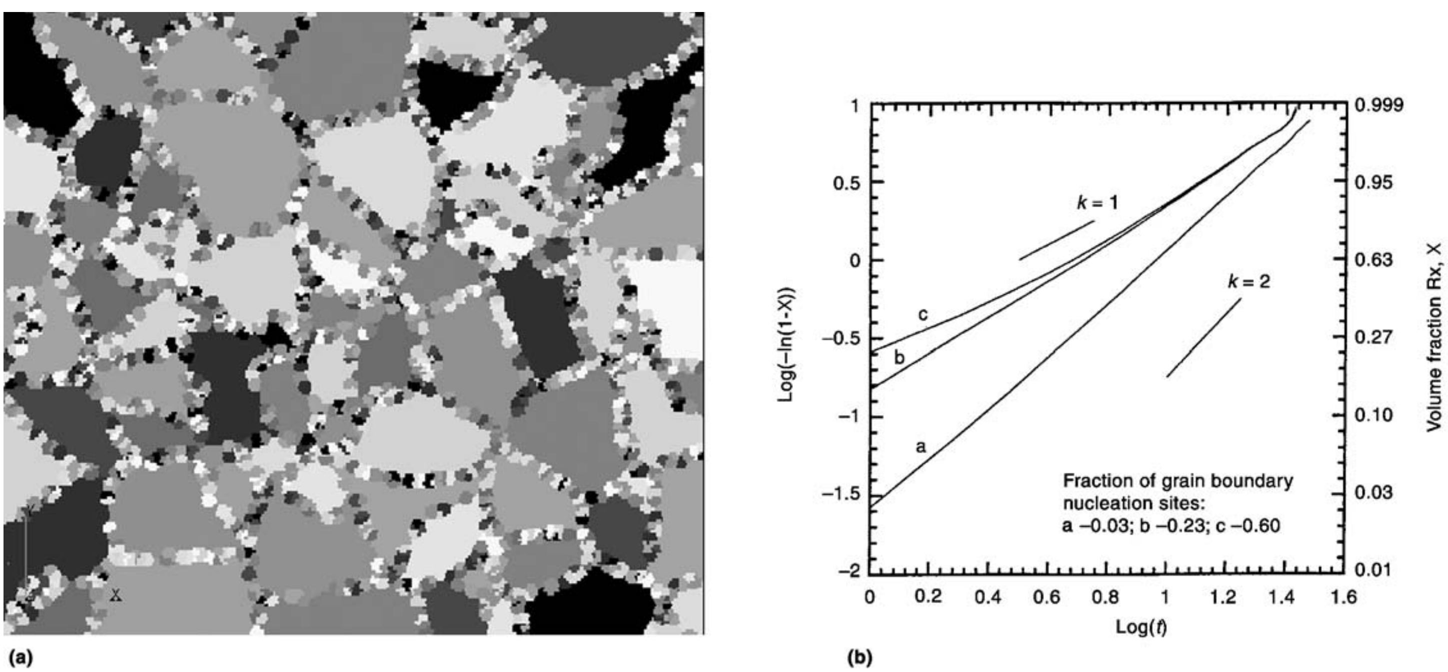


Fig. 10 Cellular-automata predictions for static recrystallization under grain-boundary nucleation, site-saturation conditions. (a) Microstructure evolution and (b) fraction recrystallized. Source: Ref 27

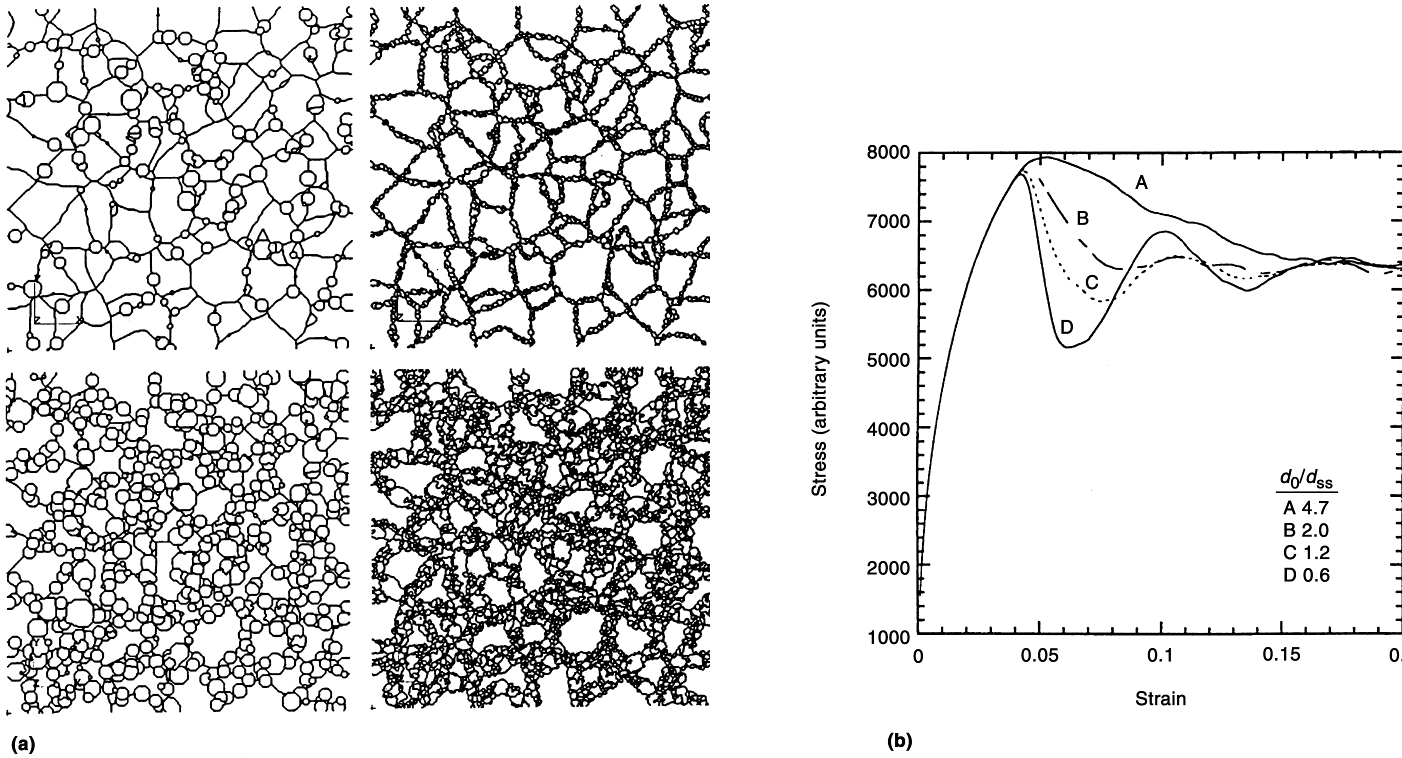


Fig. 11 Cellular-automata predictions for dynamic recrystallization. (a) Microstructure evolution and (b) stress-strain curves at various strain rates. Source: Ref 29

these simulations, the material domain, be it 2D or 3D, is discretized into material sites or points, each of which has one of a set of L possible crystallographic orientations. Grain boundaries divide groups of sites with identical orientation. To simulate the kinetics of boundary motion, a lattice site is selected at random, and a new trial orientation is also chosen at random from one of the orientations of the first or second nearest neighborhood sites. The energy of the system before and after the "flip" is calculated from the Hamiltonian H:

$$H = -J(\delta_{S_iS_j} - 1) \qquad \text{(Eq 18)}$$

in which S_i is one of the L orientations on site i ($1 \leq S_i \leq L$) and δ_{hl} is the Kronecker delta. Thus, nearest neighbor pairs contribute J to the system energy when they are of unlike orientation and zero otherwise. The change in free energy ΔG before and after the flip is calculated as the difference between the two Hamiltonians. The transition probability, W, is then given by:

$$W = \exp(-\Delta G/k_bT) \qquad \Delta G > 0 \qquad \text{(Eq 19a)}$$

$$W = 1 \qquad \Delta G \leq 0 \qquad \text{(Eq 19b)}$$

For $\Delta G > 0$, the transition probability is compared to a number between 0 and 1 obtained from a random-number generator. If W is greater than that number (as well as if $\Delta G \leq 0$), the flip is accepted. Successful transitions at a grain boundary to the orientation of a nearest neighbor grain correspond to boundary migration.

The time increment in the MC technique corresponds to the Monte Carlo Step (MCS), which is equivalent to N_s trial flips, where N_s corresponds to the number of sites in the simulation.

In initial MC investigations of grain growth (Ref 33), attention was focused on 2D simulations. Furthermore, the variation of grain-boundary energy and mobility with misorientation was not taken into account, and boundary migration was determined solely from the application of Eq 18 and 19. Nevertheless, normal growth was well simulated, yielding a growth exponent (Eq 15) of $n = 2.38$. Subsequent work (Ref 34) considered boundary-energy variations in calculating ΔG as well as the effect of relative boundary mobility and relative boundary energy on the flip probability. Such simulations predicted grain-growth exponents between 2.38 and 4. More recently, MC grain-growth approaches have been extended to 3D problems and have incorporated crystallographic texture effects. For example, using 3D MC, Ivasishin and his colleagues (Ref 35) have shown that texture influences can give rise to values of n that vary during the growth process (Fig. 12); such phenomena were correlated with the interaction of texture evolution and grain growth. In addition, in certain cases, n was shown to reach levels as high as 6.

The Monte-Carlo technique has also been applied to the analysis of both static and dynamic recrystallization. The approach for static recrystallization (Ref 36) is similar to that for MC simulation of grain growth in that material sites at a grain boundary are assigned an energy (per Eq 18). An additional energy term (which may vary from grain to grain, depending on crystal orientation) accounts for the difference in stored dislocation energy across boundaries. Furthermore, nucleation sites are selected at random, usually assuming site saturation or a constant nucleation rate. The probability of an (unrecrystallized to recrystallized) transition at a given site is then determined as before.

The MC approach for dynamic recrystallization (Ref 37) is similar to that for static recrystallization except that there is yet another term accounting for the *input* of dislocation energy at a specified rate. Refined MC dynamic recrystallization models (Ref 38) also incorporate the reduction of dislocation density/stored energy due to dynamic recovery in addition to the specified dislocation input. Hybrid MC-CA models that incorporate the influence of curvature driven and stored-energy-driven transformation are also under development (Ref 39).

Models for Grain Growth

A number of analytical models have been developed to predict grain growth. Most of these are based on the assumption of curvature-driven boundary migration. In such cases, the driving pressure is equal to γ_s/R, where R is the radius of curvature of the grains. The velocity v ($= dR/dt$) is then taken as proportional to the driving pressure with the proportionality constant being the mobility, M (Ref 1):

$$dR/dt = M(\gamma_s/R) \qquad \text{(Eq 20)}$$

If $2R$ represents the average grain size, Eq 20 is easily integrated to yield a relation of the form of Eq 15 with $n = 2$.

Hillert (Ref 40) derived an expression similar to Eq 20 that was capable of describing the growth of grains whose size is not uniform. The differential equation describing the growth of a grain of radius R is:

$$dR/dt = M\gamma_s[(1/R_{crit}) - (1/R)] \quad \text{(Eq 21)}$$

Grains with $R > R_{crit}$ were found to grow, and those with $R < R_{crit}$ shrink. The critical grain size R_{crit}, which separated the two regimes, was shown to be equal to the average grain size. Abbruzzese and Lucke (Ref 41), Eichelkraut et al. (Ref 42), and Humphreys (Ref 43), among others, extended the Hillert model to take into account the effect of texture/boundary misorientation on mobility and boundary energy in the analysis of grain and subgrain growth.

Relations similar to Eq 20 have also been developed to describe the kinetics of grain growth in the presence of pinning particles. In such cases, the curvature driving pressure is offset by a restraining, or drag, pressure. For small volume fractions of particles, the drag pressure is that associated with Zener pinning and is proportional to $\gamma_s f_v/d_p$. This assumption, in addition to the generalization of Eq 20 to allow for grain growth characterized by $n = p \neq 2$, yields a differential equation of the following form for the average grain diameter, D, as a function of time t (Ref 44, 45):

$$dD/dt = 2M(C^*\gamma_s)^{(p-1)}[(1/D) - (C'f_v/d_p)]^{(p-1)} \quad \text{(Eq 22)}$$

in which C^*, C', and p are constants. The relation shows that the growth rate may be high when D is small but decreases as D approaches the stable grain size ($= d_p/C'f_v$). The application of Eq 22 to model the growth of alpha grains in the presence of stable second-phase gamma particles in a gamma titanium aluminide alloy is illustrated in Fig. 13 (Ref 46).

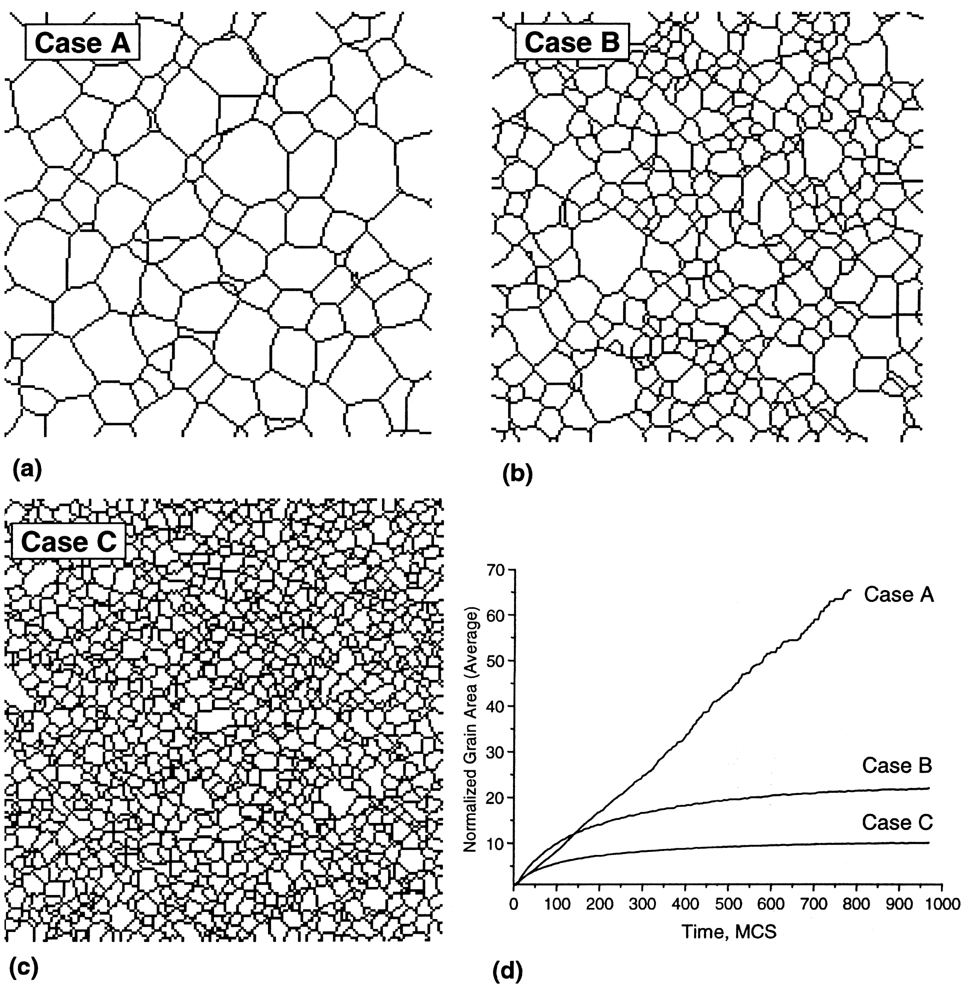


Fig. 12 Monte Carlo (3D) model predictions of (a, b, c) grain structure (2D sections after 1000 MC steps) and (d) grain-growth behavior for materials with various starting textures and assumed grain-boundary properties. (a) Case A, isotropic starting texture and isotropic boundary properties (normal grain-growth case); (b) Case B, initial, single component texture, weakly anisotropic grain-boundary properties; and (c) Case C, initial, single-component texture, strongly anisotropic grain-boundary properties. Source: Ref 35

Texture Evolution Models

Texture evolution models fall into two main categories, those principally for the prediction of either deformation textures or recrystallization textures. Deformation texture modeling is much more advanced compared to efforts for predicting recrystallization textures.

Deformation texture modeling treats the slip and twinning processes and the associated crystal rotations to predict anisotropic plastic flow and texture evolution. Models of this sort include lower- and upper-bound approaches. The lower-bound (or Sachs) method assumes that each grain of an aggregate is subjected to the same stress state as the imposed macroscopic stresses. The most highly stressed slip (or twinning) system in each grain is used to establish yielding; a simple average of these stresses for all grains determines aggregate behavior. By contrast, the upper-bound method, introduced by Taylor and Bishop and Hill (Ref 47, 48), assumes that each grain undergoes the same strain as that imposed on the aggregate and that the strain is uniform within the grains. The accommodation of an arbitrary imposed strain state (with five arbitrary strain components) requires the activation of at least five slip (or twinning) systems. In the Taylor approach, the selection of the specific set of systems is based on the principle of minimum *internal* (virtual) work. On the other hand, the Bishop-and-Hill approach makes the selection of the required systems on the basis of the maximum *external* work, a requirement that ensures that the strain-increment vector is normal to the yield locus so determined. It has been shown that the two different approaches give the same result. The Taylor/Bishop-Hill methods have been applied to a very large number of metals with different crystal structures. A number of software codes are available to conduct these analyses. One of the most popular codes is the polycrystal plasticity program developed at Los Alamos National Laboratory and known as LApp (Ref 49).

Although the upper-bound models give reasonable estimates of deformation texture evolu-

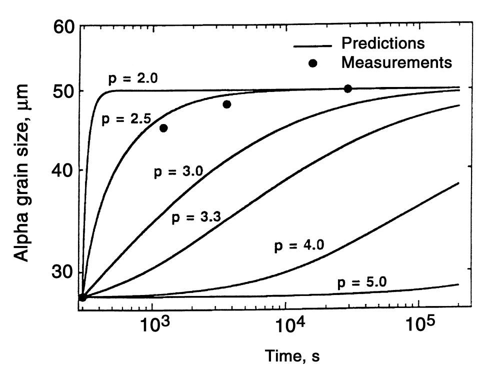


Fig. 13 Comparison of measured and predicted equiaxed alpha grain-growth kinetics for a near-gamma titanium aluminide alloy annealed in the alpha + gamma phase field. Source: Ref 46

tion in many cases, a number of refinements have been made to such techniques (Ref 50). These include the development of a relaxed-constraints, upper-bound approach in which the requirement of five independent deformation systems is relaxed for grain structures that are (or have become during deformation) elongated. A further refinement relates to the fact that classical upper-bound models enforce strain compatibility, but typically give rise to stress states that are not totally compatible with the macroscopic stress state. To remedy this deficiency, self-consistent approaches, which allow the strain to vary from grain to grain while enforcing the desired overall imposed strain, have been developed.

Another refinement to deformation texture modeling is that based on the finite element method (FEM) and thus known as crystal-plasticity FEM (Ref 51, 52). In this technique, each grain is divided into a number of elements. Each of the elements has its own constitutive equation depending on its prior history (e.g., level of hardening or softening) and instantaneous crystallographic orientation. With these local constitutive equations, the FEM analysis is then performed in a manner analogous to conventional, continuum simulations. Although considerably more computation intensive than other texture-calculation techniques, the crystal-plasticity FEM approach allows deformation nonuniformities *within* grains, as well as from grain to grain, to be established in addition to local textures. The approach is also useful for the determination of local conditions that may give rise to cavitation, provided that the physics associated with such a process can be quantified in terms of the field variables used in the code.

Recrystallization Texture Models. A relatively recent development in texture modeling is that associated with the recrystallization phenomena during (or following) hot deformation. During hot working, the textures that evolve are a result of both dislocation glide and dynamic recrystallization. Texture evolution can be quantified by mechanisms such as oriented nucleation and selective growth. In the former mechanism, recrystallization nuclei are formed in those grains that have suffered the least shear strain (i.e., dislocation glide). Such crystal orientations are usually established from Taylor/Bishop-Hill models. Selective (i.e., faster) growth is then assumed to occur for nuclei of particular misorientations with respect to the matrix. The details of this approach have been reviewed (Ref 53). Models incorporating oriented nucleation and selective growth have also been applied to static recrystallization (Ref 54). Such models are currently being integrated with cellular automata and crystal-plasticity FEM techniques to predict the spatial and temporal evolution of static recrystallization.

REFERENCES

1. F.J. Humphreys and M. Hatherly, *Recrystallization and Related Phenomena,* Elsevier, 1995
2. J.J. Jonas and H.J. McQueen, "Recovery and Recrystallization during High Temperature Deformation," *Treatise on Materials Science and Technology,* R.J. Arsenault, Ed., Academic Press, 1975, p 394
3. H.J. McQueen and J.E. Hockett, "Microstructures of Aluminum Compressed at Various Rates and Temperatures," *Metall. Trans.,* Vol 1, 1970, p 2997
4. V.M. Segal, "Equal Channel Angular Extrusion: From Macromechanics to Structure Formation," *Mater. Sci. Eng. A,* Vol 271, 1999, p 322
5. R.Z. Valiev, R.K. Islamgaliev, and I.V. Alexandrov, "Bulk Nanostructured Materials from Severe Plastic Deformation," *Prog. Mater. Sci.,* Vol 45, 2000, p 103
6. D.A. Hughes and N. Hansen, "Characterization of Sub-Micrometer Structures in Heavily Deformed Metals Over the Entire Misorientation Angle Range," *Ultrafine Grained Materials*, R.S. Mishra, S.L. Semiatin, C. Suryanarayanan, N.N. Thadhani, and T.C. Lowe, Ed., TMS, 2000, p 195
7. R.D. Doherty, D.A. Hughes, F.J. Humphreys, J.J. Jonas, D. Juul Jensen, M.E. Kassner, W.E. King, T.R. McNelley, H.J. McQueen, and A.D. Rollett, "Current Issues in Recrystallization: A Review," *Mater. Sci. Eng. A,* Vol 238, 1997, p 219
8. T. Sakai, M. Ohashi, K. Chiba, and J.J. Jonas, "Recovery and Recrystallization of Polycrystalline Nickel after Hot Working," *Acta Metall.,* Vol 36, 1988, p 1781
9. J.-P.A. Immarigeon and J.J. Jonas, "The Deformation of Armco Iron and Silicon Steel in the Vicinity of the Curie Temperature," *Acta Metall.,* Vol 22, 1974, p 1235
10. A. Yoshie, H. Mirikawa, and Y. Onoe, "Formulation of Static Recrystallization of Austenite in Hot Rolling Process of Steel Plate," *Trans. ISIJ,* Vol 27, 1987, p 425
11. A. Laasraoui and J.J. Jonas, "Prediction of Steel Flow Stresses at High Temperatures and Strain Rates," *Metall. Trans. A,* Vol 22, 1991, p 1545
12. U.F. Kocks, "Laws for Work Hardening and Low-Temperature Creep," *J. Eng. Mater. Techn., Trans. ASME,* Vol 98, 1976, p 76
13. W. Roberts, "Dynamic Changes That Occur During Hot Working and Their Significance Regarding Microstructural Development and Hot Workability," *Deformation, Processing, and Structure,* G. Krauss, Ed., ASM, 1984, p 109
14. H.D. Ryan and H.J. McQueen, "Comparative Hot Working Characteristics of 304, 316, and 317 Steels, Both Cast and Worked," *International Conf. on New Developments in Stainless Steel Technology,* R.A. Lula, Ed., ASM, 1985, p 293
15. C.M. Sellars, "Modeling Microstructure Evolution," *Mater. Sci. Technol.,* Vol 6, 1990, p 1072
16. C. Devadas, I.V. Samarasekara, and E.B. Hawbolt, "Thermal and Metallurgical State of Steel Strip during Hot Rolling," *Metall. Trans. A,* Vol 22, 1991, p 335
17. G. Shen, S.L. Semiatin, and R. Shivpuri, "Modeling Microstructural Development During the Forging of Waspaloy," *Metall. Mater. Trans. A,* Vol 26, 1995, p 1795
18. T. Sakai and J.J. Jonas, "Dynamic Recrystallization: Mechanical and Microstructural Considerations," *Acta Metall.,* Vol 32, 1984, p 189
19. B. Derby, "The Dependence of Grain Size on Stress During Dynamic Recrystallization," *Acta Metall. Mater.,* Vol 39, 1991, p 955
20. M. Avrami, "Kinetics of Phase Change: I. General Theory," *J. Chem. Phys.,* Vol 7, 1939, p 1103
21. M. Avrami, "Kinetics of Phase Change: II. Transformation-Time Relations for Random Distribution of Nuclei," *J. Phys. Chem.,* Vol 8, 1940, p 212
22. J.W. Cahn, "The Kinetics of Grain Boundary Nucleated Reactions," *Acta Metall.,* Vol 4, 1956, p 449
23. P. Hellman and M. Hillert, "On the Effect of Second-Phase Particles on Grain Growth," *Scand. J. Metall.,* Vol 4, 1975, p 211
24. P.M. Hazzledine and R.D.J. Oldershaw, "Computer Simulation of Zener Pinning," *Philos. Mag.,* Vol 61A, 1990, p 579
25. H.W. Hesselbarth and I.R. Gobel, "Simulation of Recrystallization by Cellular Automata," *Acta Metall. Mater.,* Vol 39, 1991, p 2135
26. C.H.J. Davies, "Growth of Nuclei in a Cellular Automaton Simulation of Recrystallization," *Scr. Mater.,* Vol 36, 1997, p 35
27. R.L. Goetz and V. Seetharaman, "Static Recrystallization Kinetics with Homogeneous and Heterogeneous Nucleation Using Cellular Automata," *Metall. Mater. Trans. A,* Vol 29, 1998, p 2307
28. C.F. Pezzee and D.C. Dunand, "The Impingement Effect of an Inert, Immobile Second Phase on the Recrystallization of a Matrix," *Acta Metall. Mater.,* Vol 42, 1994, p 1509
29. R.L. Goetz and V. Seetharaman, "Modeling Dynamic Recrystallization Using Cellular Automata," *Scr. Mater.,* Vol 38, 1998, p 405
30. D. Raabe, "Introduction of a Scalable Three-Dimensional Cellular Automaton with a Probabilistic Switching Rule for the Discrete Mesoscale Simulation of Recrystallization Phenomena," *Philos. Mag.,* Vol 79, 1999, p 2339
31. J.W. Cahn and W. Hagel, "Theory of the Pearlite Reaction," *Decomposition of Austenite by Diffusional Processes,* V.F. Zackay and H.I. Aaronson, Ed., Interscience Publishers, 1962, p 131
32. C. Kittel, *Thermal Physics,* Wiley, 1969
33. M.P. Anderson, D.J. Srolovitz, G.S. Grest, and P.S. Sahni, "Computer Simulation of Grain Growth–I. Kinetics," *Acta Metall.,* Vol 32, 1984, p 783
34. G.S. Grest, D.J. Srolovitz, and M.P.

Anderson, "Computer Simulation of Grain Growth–IV. Anisotropic Grain Boundary Energies," *Acta Metall.,* Vol 33, 1985, p 509
35. O.M. Ivasishin, S.V. Shevchenko, N.L. Vasiliev, and S.L. Semiatin, "3D Monte-Carlo Simulation of Texture Evolution and Grain Growth during Annealing," *Metal Physics and Advanced Technologies,* 2001, Vol 23, p 1569
36. D.J. Srolovitz, G.S. Grest, and M.P. Anderson, "Computer Simulation of Recrystallization, Part I: Homogeneous Nucleation and Growth," *Acta Metall.,* Vol 34, 1986, p 1833
37. A.D. Rollett, M.J. Luton, and D.J. Srolovitz, "Microstructural Simulation of Dynamic Recrystallization," *Acta Metall.,* Vol 40, 1992, p 43
38. P. Peczak and M.J. Luton, "A Monte Carlo Study of the Influence of Dynamic Recovery on Dynamic Recrystallization," *Acta Metall. Mater.,* Vol 41, 1993, p 59
39. A.D. Rollett and D. Raabe, "A Hybrid Model for Mesoscopic Simulation of Recrystallization," *Fourth International Conf. on Recrystallization and Related Phenomena,* T. Sakai and H.G. Suzuki, Ed., Japan Institute of Metals, 1999, p 623
40. M. Hillert, "On the Theory of Normal and Abnormal Grain Growth," *Acta Mater.,* Vol 13, 1965, p 227
41. G. Abbruzzese and K. Lucke, "A Theory of Texture Controlled Grain Growth–Derivation and General Discussion of the Model," *Acta Metall.,* Vol 34, 1986, p 905
42. H. Eichelkraut, G. Abbruzzese, and K. Lucke, "A Theory of Texture Controlled Grain Growth, Part II: Numerical and Analytical Treatment of Grain Growth in the Presence of Two Texture Components," *Acta Metall.,* Vol 36, 1986, p 55
43. F.J. Humphreys, "A Unified Theory of Recovery, Recrystallization, and Grain Growth Based on the Stability and Growth of Cellular Microstructures, Part I: The Basic Model," *Acta Mater.,* Vol 45, 1997, p 4231
44. I. Andersen and O. Grong, "Analytical Modeling of Grain Growth in Metals and Alloys in the Presence of Growing and Dissolving Precipitates, Part I: Normal Grain Growth," *Acta Metall. Mater.,* Vol 43, 1995, p 2673
45. G. Grewel and S. Ankem, "Modeling Matrix Grain Growth in the Presence of Growing Second Phase Particles in Two Phase Alloys," *Acta Metall. Mater.,* Vol 38, 1990, p 1607
46. V. Seetharaman and S.L. Semiatin, "Analysis of Grain Growth in a Two-Phase Gamma Titanium Aluminide Alloy," *Metall. Mater. Trans. A,* Vol 28, 1997, p 947
47. G.I. Taylor, "Analysis of Plastic Strain in a Cubic Crystal," *Stephen Timoshenko 60th Anniversary Volume,* Macmillan Company, 1938, p 218
48. J.F.W. Bishop and R. Hill, "A Theory of Plastic Distortion of a Polycrystalline Aggregate Under Combined Stresses," *Philos. Mag.,* Vol 42, 1951, p 414
49. J.S. Kallend, U.F. Kocks, A.D. Rollett, and H.-R. Wenk, "Operational Texture Analysis," *Mater. Sci. Eng. A,* Vol 132, 1991, p 1
50. U.F. Kocks, C.N. Tome, and H.-R. Wenk, *Texture and Anisotropy,* Cambridge University Press, 1998
51. G.B. Sarma and P.R. Dawson, "Effect of Interaction among Crystals on the Inhomogeneous Deformations of Polycrystals," *Acta Mater.,* Vol 44, 1996, p 1937
52. P. Bate, "Modeling Deformation Microstructure with the Crystal Plasticity Finite-Element Method," *Philos. Trans. R. Soc. (London) A,* Vol 357, 1999, p 1589
53. J.J. Jonas, "Modeling the Length Changes That Take Place During The Torsion Test," *Int. J. Mech. Sci.,* Vol 35, 1993, p 1065
54. G. Gottstein and R. Sebald, "Modeling of Recrystallization Textures," *Thermomechanical Processing of Steel,* S. Yue and E. Essadiqi, Ed., Canadian Institute of Mining, Metallurgy, and Petroleum, Montreal, 2000, p 21

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 47-56
DOI:10.1361/hwpd2003p047

Chapter 4

Bulk Workability Testing

George E. Dieter, University of Maryland

WORKABILITY is a complex property of a material, as indicated in Chapter 2, "Bulk Workability of Metals." It is difficult to isolate the intrinsic workability, because this property is strongly influenced by stress state, which is in turn affected by friction and the geometry of the tools and the workpiece. It has also been shown that the workability of a material is strongly influenced by metallurgical structure, which can be a complex function of temperature and strain rate. At the current state of development, the ability to model a forging process by calculating stress, strain rate, and temperature throughout a deforming workpiece with a computer-based finite-element technique exceeds the ability to predict the workability of the material.

A large number of tests are currently used to evaluate the workability of a material. The primary tests—tension, torsion, compression, and bend—are discussed in this chapter. These are tests for which the state of stress is well defined and controlled. Of these four tests, the compression test has been the most highly developed as a workability test.

Specialized workability tests that have been developed from the four primary tests are also covered. Each of these tests provides information that is not readily available from the primary tests.

A number of workability tests that are especially applicable to the forging process are discussed in Chapter 13, "Workability in Forging."

Primary Tests

The primary tests for workability are those for which the stress state is well known and controlled. Generally, these are small laboratory simulation tests.

Tension Test

The tension test is widely used to determine the mechanical properties of a material (Ref 1). Uniform elongation, total elongation, and reduction in area at fracture are frequently used as indexes of ductility. The extent of deformation possible in a tension test, however, is limited by the formation of a necked region in the tension specimen. This introduces a triaxial tensile stress state and leads to fracture.

For most metals, the uniform strain that precedes necking rarely exceeds a true strain of 0.5. For hot working temperatures, this uniform strain is frequently less than 0.1. Although tension tests are easily performed, necking makes control of strain rate difficult and leads to uncertainties about the value of strain at fracture because of the complex stresses that result from necking. Therefore, the utility of the tension test is limited in workability testing. This test is primarily used under special, high-strain-rate, hot tension test conditions to establish the range of hot working temperatures. A description of this test method can be found in Chapter 7, "Hot Tension Testing."

Torsion Test

In the torsion test, deformation is caused by pure shear, and large strains can be achieved without the limitations imposed by necking (Ref 2, 3). Because the strain rate is proportional to rotational speed, high strain rates are readily obtained (Table 1). Moreover, friction has no effect on the test, as it does in compression testing. The stress state in torsion may represent the typical stress in metalworking processes, but deformation in the torsion test is not an accurate simulation of metalworking processes, because of excessive material reorientation at large strains.

Because of the previously mentioned advantages, the torsion test is frequently used to measure the flow stress and the stress-strain curve (flow curve) under hot working conditions. Figure 1 shows typical flow curves as a function of temperature and strain rate. In the torsion test, measurements are made of the moment, *M,* to deform the specimen and the angle of twist, θ, or the number of turns ($\theta = 2\pi$ rad per turn). The shear stress, τ, on the outer surface of the specimen is given by the following:

$$\tau = \frac{M(3+m+n)}{2\pi r^3} \quad \text{(Eq 1)}$$

where r is the specimen radius, m is the strain-rate sensitivity found from plots of log M versus log $\dot{\theta}$ at fixed values of θ, and n is the strain-hardening exponent obtained from the instantaneous slope of log M versus log θ.

The engineering shear strain, Γ, and shear strain rate, $\dot{\Gamma}$, are given by:

$$\Gamma = \frac{r\,\theta}{L} \quad \text{(Eq 2)}$$

and

$$\dot{\Gamma} = \frac{r\dot{\theta}}{L} \quad \text{(Eq 3)}$$

where r is the radius of the specimen and L is the gage length. These values of shear stress and shear strain are typically converted to effective stress, $\bar{\sigma}$, and effective strain, $\bar{\varepsilon}$, by means of the von Mises yielding criterion (see Chapter 2, "Bulk Workability of Metals"):

$$\bar{\sigma} = \sqrt{3}\tau \quad \text{(Eq 4)}$$

and

$$\bar{\varepsilon} = \frac{\Gamma}{\sqrt{3}} \quad \text{(Eq 5)}$$

Figure 2 shows agreement in plots of $\bar{\sigma}$ versus $\bar{\varepsilon}$ for stress-strain data determined in torsion,

Table 1 Torsional rotation rates corresponding to various metalworking operations

Operation	von Mises effective strain rate ($\dot{\varepsilon}$)(a), s^{-1}	Corresponding surface shear strain rate in torsion ($\dot{\varepsilon}$), s^{-1}	Rotation rate(b), rpm
Isothermal forging	10^{-3}	1.73×10^{-3}	0.02
Hydraulic press forging	1	1.73	16.5
Extrusion	20	34.6	330.4
Mechanical press forging	50	86.6	827.0
Sheet rolling	200	346.4	3307.9
Wire drawing	500	866.0	8269.7

(a) $\dot{\varepsilon} = \dot{\Gamma}\sqrt{3}$. (b) Assuming specimen geometry with $r/L = 1.0$

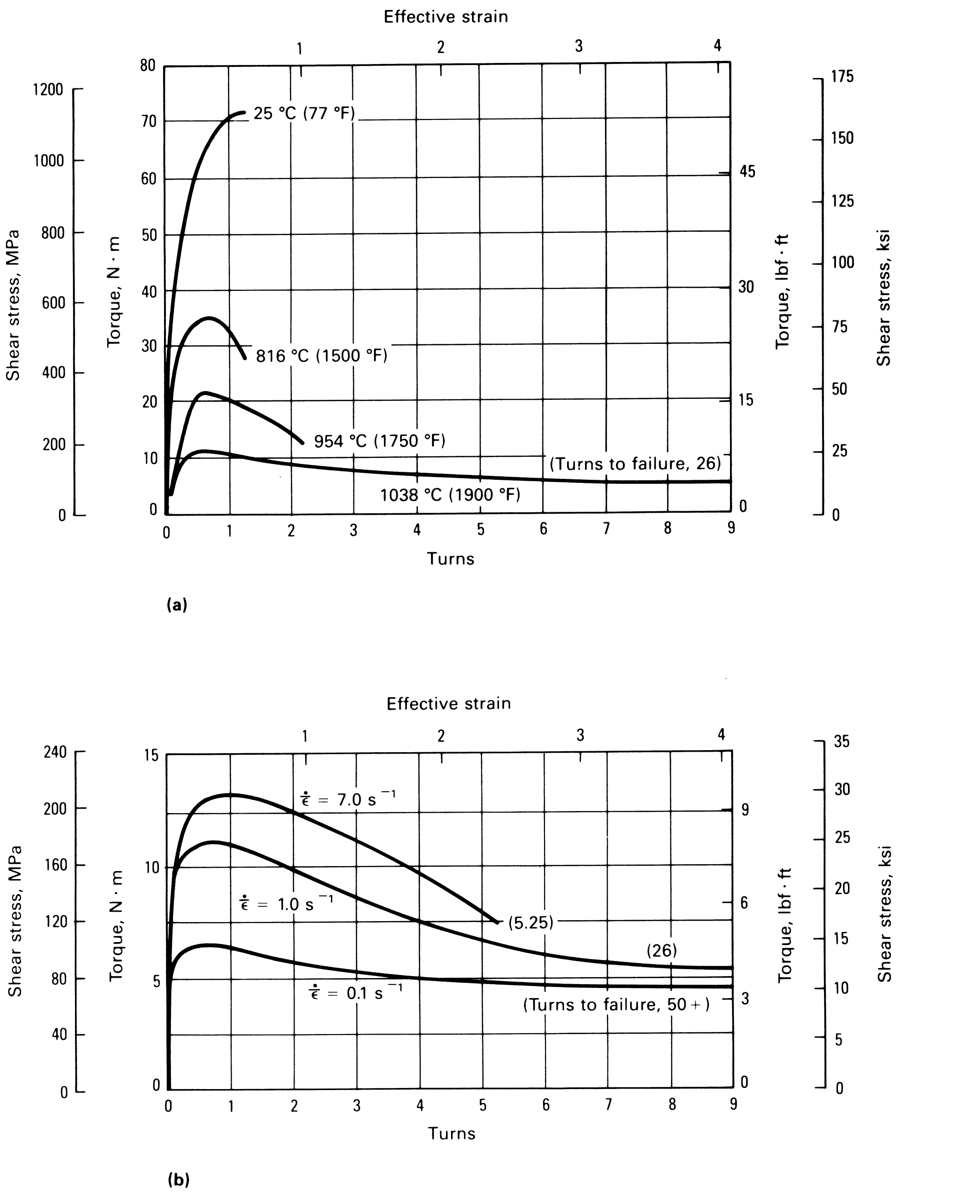


Fig. 1 Flow curves for Waspaloy. (a) Effect of temperature at a fixed effective strain rate of 1 s^{-1}. (b) Effect of strain rate at a fixed test temperature of 1038 °C (1900 °F). Flow softening at the higher temperature is a result of dynamic recrystallization. Source: Ref 4

tension, and compression. The agreement becomes much better at hot working temperatures.

Fracture data from torsion tests are usually reported in terms of the numbers of twists to failure or the surface fracture strain to failure. Figure 3 shows the relative hot workability of a number of steels and nickel-base superalloys, as indicated by the torsion test. The test identifies the optimal hot working temperature. For details on torsion testing methods, see Chapters 8 and 9, "Torsion Testing to Assess Bulk Workability" and "Hot Working Simulation by Hot Torsion Testing," respectively.

Compression Test

The compression test, in which a cylindrical specimen is upset into a flat pancake, is usually considered to be a standard bulk workability test. The average stress state during testing is similar to that in many bulk deformation processes, without introducing the problems of necking (in tension) or material reorientation (in torsion). Therefore, a large amount of deformation can be achieved before fracture occurs. The stress state can be varied over wide limits by controlling the barreling of the specimen through variations in geometry and by reducing friction between the specimen ends and the anvil with lubricants (see Chapter 5, "Cold Upset Testing").

Compression testing has developed into a highly sophisticated test for workability in cold upset forging, and it is a common quality-control test in hot forging operations. Compression forging is a useful method of assessing the frictional conditions in hot working. The principal disadvantage of the compression test is that tests at a constant, true strain rate require special equipment.

Compression Test Conditions. Unless the lubrication at the ends of the specimen is very good, frictional restraint retards the outward motion of the end face, and part of the end face is formed by a folding over of the sides of the original cylinder onto the end face in contact with the platens. The barreling that results introduces a complex stress state, which is beneficial in fracture testing but detrimental when the compression test is used to measure flow stress. The frictional restraint also causes internal inhomogeneity of plastic deformation. Slightly deforming zones develop adjacent to the platens, while severe deformation is concentrated in zones that occupy roughly diagonal positions between opposing edges of the specimen (see Fig. 17 of Chapter 2, "Bulk Workability of Metals").

Figure 4 shows the hot upsetting of a cylinder under conditions of poor lubrication in which the platens are cooler than the specimen. The cooling at the ends restricts the flow so that the deformation is concentrated in a central zone, with dead-metal zones forming adjacent to the platen surfaces (Fig. 4a).

As deformation proceeds, severe inhomogeneity develops, and the growth of the end faces is attributed entirely to the folding over of the sides (Fig. 4b). When the diameter-to-height ratio, D/h, exceeds approximately 3, expansion of the end faces occurs (Fig. 4c).

The conditions described previously are extreme and should not be allowed to occur in hot compression testing unless the objective is to simulate cracking under forging conditions. Adequate lubrication cannot improve the situation so that homogeneous deformation occurs; however, with glass lubricants and isothermal conditions, it is possible to conduct hot compression testing without appreciable barreling (Ref 6). Isothermal test conditions can be achieved by using a heated subassembly, such as that shown in Fig. 5, or heated dies that provide isothermal conditions (Ref 8).

The true strain rate, $\dot{\varepsilon}$, in a compression test is:

$$\dot{\varepsilon} = \frac{d\varepsilon}{dt} = \frac{-dh/h}{dt} = -\frac{1dh}{hdt} = -\frac{v}{h} \qquad \text{(Eq 6)}$$

where v is the velocity of the platen, and h is the height of the specimen at time, t. Because h decreases continuously with time, the velocity must decrease in proportion to $(-h)$ if $\dot{\varepsilon}$ is to be held constant. In a normal test, if v is held constant, the engineering strain rate, $\dot{e}$, remains constant:

$$\dot{e} = \frac{de}{dt} = \frac{-dh/h_0}{dt} = -\frac{1}{h_0}\frac{dh}{dt} = \frac{-v}{h_0} \qquad \text{(Eq 7)}$$

The true strain rate, however, is not constant. A machine called a cam plastometer can be used to cause the bottom platen to compress the specimen through cam action at a constant true strain rate to a strain limit of $\varepsilon = 0.7$. The use of cam

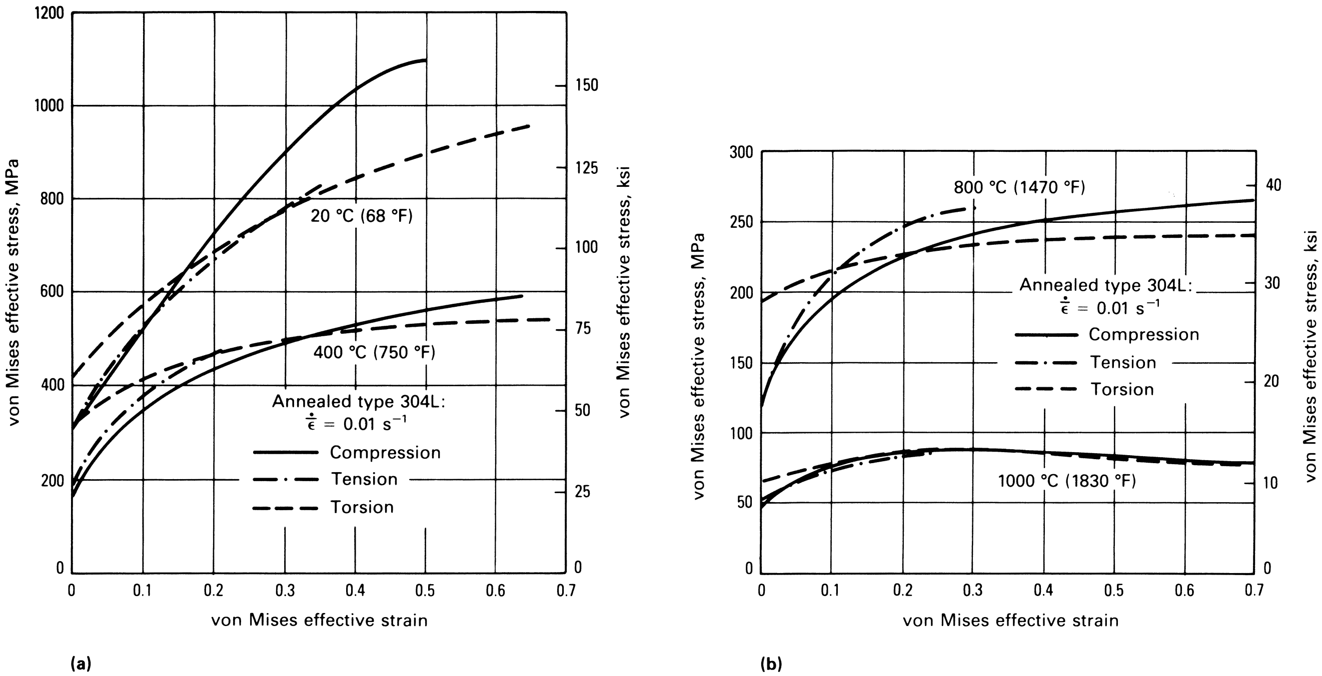


Fig. 2 Comparison of effective stress-strain curves determined for type 304L stainless steel in compression, tension, and torsion. (a) Cold working and warm working temperatures. (b) Hot working temperatures. Source: Ref 2

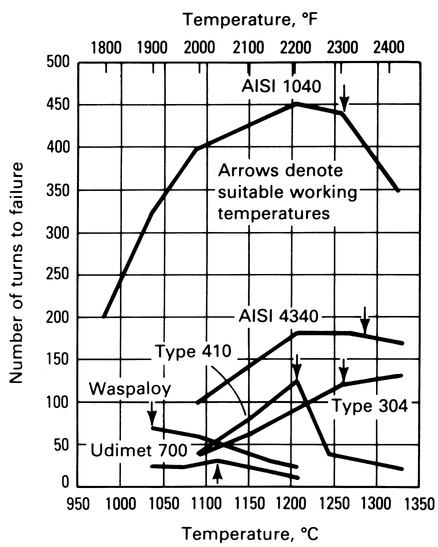


Fig. 3 Ductility determined in hot torsion tests. AISI, American Iron and Steel Institute. Source: Ref 2

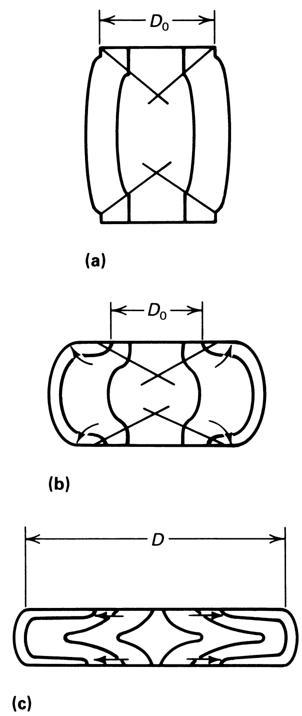


Fig. 4 Deformation patterns in nonlubricated, nonisothermal hot forging. (a) Initial barreling. (b) Barreling and folding over. (c) Beginning of end face expansion. Source: Ref 5

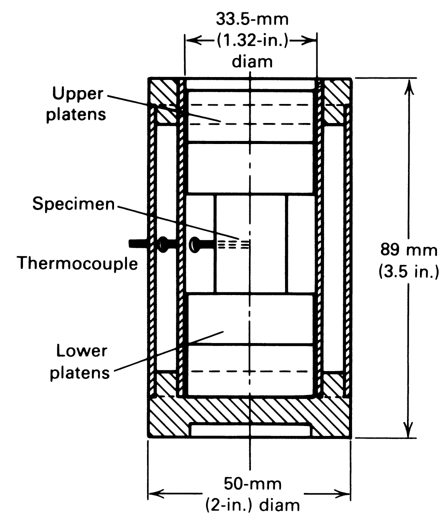


Fig. 5 Heated subassembly with specimen in position used to achieve isothermal test conditions. Thermocouple is removed prior to compression. Source: Ref 7

plastometers is limited; there probably are not more than ten in existence. However, an essentially constant true strain rate can be achieved on a standard closed-loop servocontrolled testing machine. Strain rates up to 100 s^{-1} have been achieved (Ref 6, 9).

When a constant true strain rate cannot be obtained, the mean strain rate may be adequate. The mean true strain rate $\langle\dot{\bar{\varepsilon}}\rangle$ for constant velocity, v_0, when the specimen is reduced in height from h_0 to h, is given by:

$$\left\langle\dot{\bar{\varepsilon}}\right\rangle = \frac{v_0}{2}\frac{\ln(h_0/h)}{(h_0 - h)} \qquad \text{(Eq 8)}$$

Flow Stress in Compression. Ideally, the determination of flow stress in compression is carried out under isothermal conditions (no die chilling) at a constant strain rate and with a minimum of friction in order to minimize barreling. These conditions can be met with conventional servohydraulic testing machines. For an essentially homogeneous upsetting test, a cylinder of diameter D_0 and initial height h_0 is compressed to height h and spread out to diameter D_1 according to the law of constancy of volume:

$$D_0^2 h_0 = D^2 h \qquad \text{(Eq 9)}$$

If friction can be neglected, the uniaxial compressive stress (flow stress), σ, corresponding to a deformation force P is:

$$\sigma_0 = \frac{P}{A} = \frac{4P}{\pi D^2} = \frac{4Ph}{\pi D_0^2 h_0} \qquad \text{(Eq 10)}$$

If substantial friction is present, the average pressure, $\bar{p}$, required to deform the cylinder is greater than the flow stress of the material, σ_0:

$$\frac{\bar{p}}{\sigma_0} = \left(\frac{h}{4\mu a}\right)^2 \left(e^{2\mu a/h} - \frac{2\mu a}{h} - 1\right) \quad \text{(Eq 11)}$$

where a is the radius of the cylinder, and μ is the Coulomb coefficient of friction. The true compressive strain, ε, is given by:

$$\varepsilon = \ln\left(\frac{h_0}{h}\right) \quad \text{(Eq 12)}$$

The effects of friction and die chilling can be minimized through the use of a long, thin specimen. Therefore, most of the specimen volume is unaffected by the dead-metal zones at the platens. However, this approach is limited, because buckling of the specimen occurs if h/D exceeds approximately 2.

An extrapolation method involves testing cylinders of equal diameters but varying heights, so that the D_0/h_0 ratio ranges from approximately 0.5 to 3.0 (Ref 10). A specific load is applied to the specimen, and the load is removed. The new height is determined in order to calculate a true strain. On relubrication, the specimen is subjected to an increased load, unloaded, and measured. The cycle is then repeated.

The same test procedure is followed with each specimen, so that the particular load levels are duplicated. The results are illustrated in Fig. 6. For the same load, the actual strain (due to height reduction) is plotted against the D_0/h_0 ratio for each test cylinder. A line drawn through the points is extrapolated to a value of $D_0/h_0 = 0$. This would be the anticipated ratio for a specimen of infinite initial height for which the end effects would be restricted to a small region of the full test height. The true stress corresponding to each of these true strains is given by Eq 10, with h values taken from Fig. 6 at $D_0/h_0 = 0$.

Advanced Analysis. The analysis of the compression of a cylinder so as to account for the influence of bulging on the axial flow stress has become a classical problem in mechanics. A recent analysis provides useful information on the problem (Ref 11). Starting with Hill's variational principle (Ref 12), the authors developed the following expression for the uniaxial compressive flow stress. This assumes a von Mises yield criterion and an interface friction factor $m = \sqrt{3}\tau_i/\sigma_0$ where τ_i is the interfacial shear (friction) component and where σ_0 is:

$$\sigma_0 = P\left(\bar{A} + \frac{2\pi R_1^3 m}{3\sqrt{3}h}\right)^{-1} \quad \text{(Eq 13)}$$

$\bar{A}$ is the mean cross-sectional area obtained by dividing the volume of the specimen by the instantaneous height of the compression specimen. R_1 is the radius of the specimen at the die-specimen interface, and P is the compression load to produce a specimen height h. Equation 13 assumes a rigid plastic material that follows a power-law relationship for strain hardening. It gives excellent results in room-temperature comparisons but makes no allowance for strain-rate-sensitive behavior such as would be found in metals at elevated temperature. This subject is pursued further in Chapter 6, "Hot Compression Testing."

Ductility Testing. The basic hot ductility test consists of compressing a series of cylindrical or square specimens to various thicknesses, or to the same thickness with varying specimen length-to-diameter (length-to-width) ratios. The limit for compression without failure by radial or peripheral cracking is considered to be a measure of workability. This type of test has been widely used in the forging industry. Longitudinal notches are sometimes machined into the specimens before compression, because the notches apparently cause more severe stress concentrations, thus providing a more reliable index of the workability to be expected in a complex forging operation.

Plastic Instability in Compression. Several types of plastic instabilities can be developed in the compression test. The first type is associated with a maximum in the true stress-strain curve. The second type concerns inhomogeneous deformation and shear band formation. Figure 7 shows the type of plastic instability that occurs in some materials in hot compression testing. At certain temperatures and strain rates, some of the typical strengthening mechanisms become unstable (see Chapter 3, "Evolution of Microstructure during Hot Working"). Because the rate of flow softening exceeds the rate of area increase as the specimen is compressed, a maximum results in the flow stress curve.

Analysis of the compression process indicates that the plastic deformation is stable (no maximum in the flow curve) as long as $(\gamma + m) \leq 1$, where γ is the dimensionless work-hardening coefficient, and m is strain-rate sensitivity. Both of these material parameters are defined subsequently (Ref 13,14). A material with a high strain-rate sensivity is more resistant to flow localization in the tension test (necking), but in compression testing, a higher rate sensitivity leads to earlier flow localization.

Flow softening or negative strain hardening can also produce flow localization effects in compression independent of the effects of die chilling or high friction. The constant strain-rate, isothermal hot compression test is useful for detecting and predicting flow localization. Nonuniform flow in compression is likely if a flow parameter, α_c, exceeds a certain value:

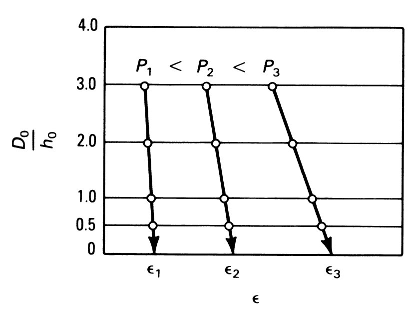


Fig. 6 Extrapolation method to correct for end effects in compressive loading. Source: Ref 10

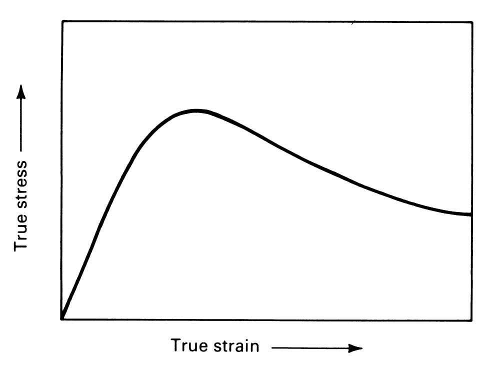


Fig. 7 Example of compressive flow stress curve showing strain softening

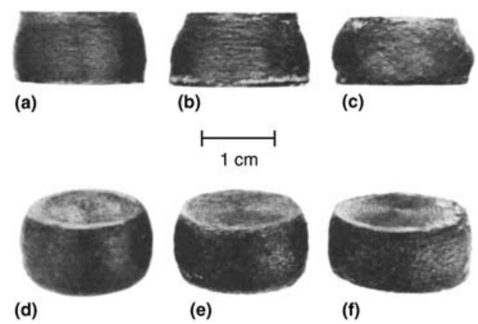


Fig. 8 Specimens of Ti-10V-2Fe-3Al from isothermal hot compression tests. (a) to (c) Tested at 704 °C (1300 °F). (d) to (f) Tested at 816 °C (1500 °F). Strain rates were 10^{-3} s^{-1} (a, d), 10^{-1} s^{-1} (b, e), and 10 s^{-1} (c, f). Before testing, the alloy had been β annealed to yield an equiaxed β starting microstructure. Source: Ref 16

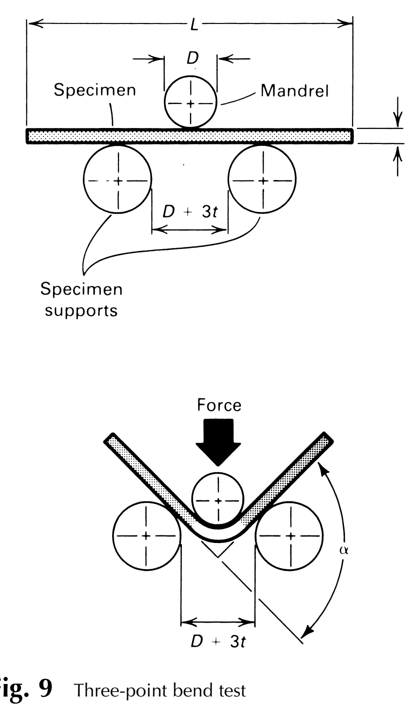


Fig. 9 Three-point bend test

$$\alpha_c = \frac{(\gamma - 1)}{m} \geq 5 \qquad \text{(Eq 14)}$$

where

$$\gamma = \left(\frac{1}{\sigma}\right)\left(\frac{d\sigma}{d\varepsilon}\right) \qquad \text{(Eq 15)}$$

and

$$m = \frac{\partial \ln \sigma}{\partial \ln \dot{\varepsilon}}\Big|_{T,\varepsilon} \approx \frac{\Delta \log \sigma}{\Delta \ln \dot{\varepsilon}}\Big|_{T,\varepsilon} \qquad \text{(Eq 16)}$$

Figure 8 illustrates the differences in deformation of titanium alloy samples. The specimens in Fig. 8(a) to (c) were deformed at a temperature, T, at which α_c was high. In Fig. 8(a), $\dot{\varepsilon} = 10^{-3}$ s^{-1}, and $\alpha_c = 2$. In Fig. 8(b), $\dot{\varepsilon} = 10^{-1}$ s^{-1}, and $\alpha_c = 5$. In Fig. 8(c), $\dot{\varepsilon} = 10$ s^{-1}, and $\alpha_c = 5$. However, the specimens in Fig. 8(d) to (f) were deformed at a temperature at which α_c was less than 0.

Bend Test

The bend test is useful for assessing the workability of thick sheet and plate. Generally, this test is most applicable to cold working operations. Figure 9 shows a plate deformed in three-point bending. The principal stress and strains developed during bending are defined in Fig. 10. The critical parameter is width-to-thickness ratio, w/t. if $w/t > 8$, bending occurs under plane-strain conditions ($\varepsilon_2 = 0$), and $\sigma_2/\sigma_1 = 0.5$. If $w/t > 8$, the bend ductility is independent of the exact w/t ratio. If $w/t < 8$, then stress state and bend ductility depend strongly on the width-to-thickness ratio.

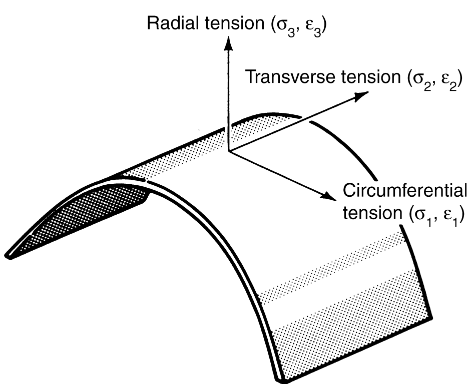


Fig. 10 Schematic of the bend region defining direction of principal stresses and strains

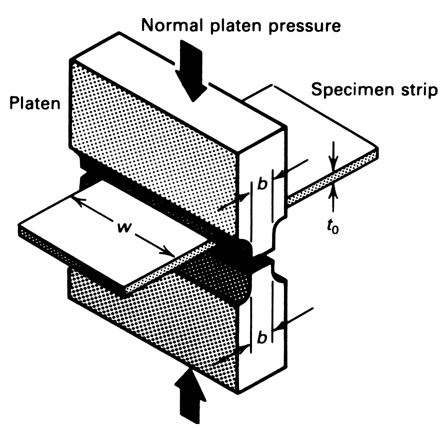


Fig. 11 Plane-strain compression test

For pure plastic bending, in which elastic deformation can be ignored, the maximum tensile fiber strain is:

$$\varepsilon_0 = \ln \sqrt{\frac{R_o}{R_i}} \qquad \text{(Eq 17)}$$

where R_o is the radius of curvature on the outer (tensile) surface, and R_i is the radius of curvature on the inner (compressive) surface (Ref 15). When this strain is entered into the stress-strain equation or curve for the material, it gives the flow stress for the material, $\bar{\sigma}$. Because of the plane-strain condition, the maximum fiber stress is $2\bar{\sigma}/\sqrt{3}$.

Specialized Tests

In the plane-strain compression test, the difficulties encountered with bulging and high friction at the platens in the compression of cylinders can be minimized (Ref 10). As shown in Fig. 11, the specimen is a thin plate or sheet that is compressed across the width, w, of the strip by narrow platens that are wider than the strip. The elastic constraints of the undeformed shoulders of material on each side of the platens prevent extension of the strip in the width dimension; hence, the term *plane strain.*

Deformation occurs in the direction of platen motion and in the direction normal to the length of the platen. To ensure that lateral spread is negligible, the width of the strip should be at least six to ten times the breadth of the platens. To ensure that deformation beneath the platens is essentially homogeneous, the ratio of platen breadth to strip thickness, b/t, should be between 2 and 4 at all times. It may be necessary to change the platens during testing to maintain this condition. True strains of 2 can be achieved by carrying out the test in increments in order to provide good lubrication and to maintain the proper b/t ratio. Although the plane-strain compression test is primarily used to measure flow properties at room temperature, it can also be used for elevated-temperature tests (see Chapter 6, "Hot Compression Testing").

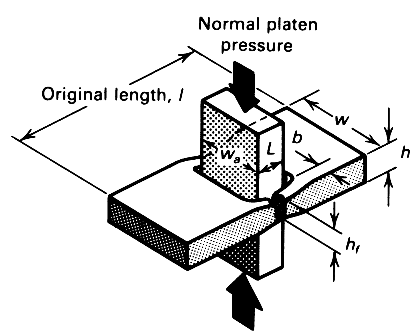


Fig. 12 Partial-width indentation test. $L \cong h$; $b = h/2$; $w_a = 2L$; $l = 4L$

Because of its plane-strain geometry, this test is more applicable to rolling operations than to forging.

The true stress and true strain determined from the test for plane-strain compression, ε_{pc}, can be expressed as:

$$p = \frac{P}{wb} \qquad \text{(Eq 18)}$$

$$\varepsilon_{pc} = \ln \frac{t_0}{t} \qquad \text{(Eq 19)}$$

Because of the stress state associated with plane-strain deformation, the mean pressure on the platens is 15.5% higher in the plane-strain compression test than in uniaxial compression testing. The true stress-strain curve in uniaxial compression (σ_0 versus ε) can be obtained from the corresponding plane-strain compression curve (p versus ε_{pc}) by:

$$\sigma_0 = \frac{\sqrt{3}}{2} p = \frac{p}{1.155} \qquad \text{(Eq 20)}$$

and

$$\varepsilon = \frac{2}{\sqrt{3}} \varepsilon_{pc} = 1.155 \varepsilon_{pc} \qquad \text{(Eq 21)}$$

The plane-strain compression test is generally accepted as one of the most reliable methods for the generation of flow stress data (σ versus ε) for thermomechanical processing. For reproducible measurements, corrections must be made for lateral spread of the specimen and deviation from ideal plane-strain behavior (Ref 17).

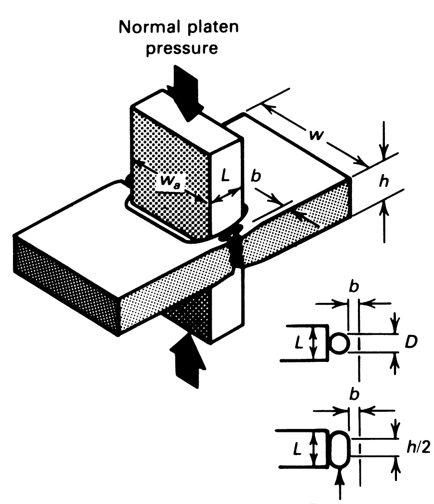


Fig. 13 Secondary-tension test showing the geometries of holes and slots. $L \cong h$; $w_a \geq 2 = h$; $b = h/4$; $D = h/2$

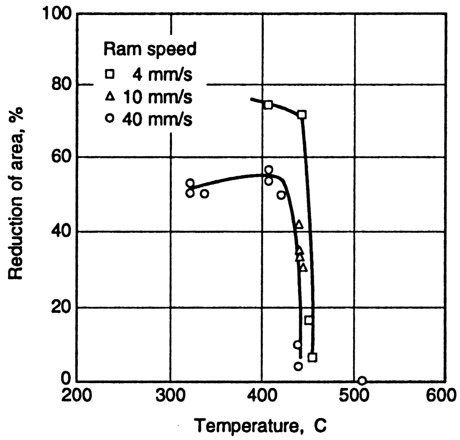


Fig. 14 Results of secondary-tension test on aluminum alloy 7075. Source: Ref 20

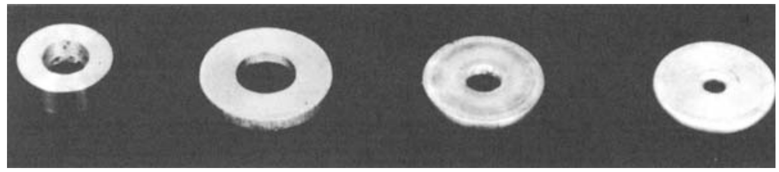


Fig. 15 Variation in shape of ring test specimens deformed the same amount under different frictional conditions. Left to right: undeformed specimen; deformed 50%, low friction; deformed 50%, medium friction; deformed 50%, high friction

The partial-width indentation test is a test that is similar to the plane-strain compression test for evaluating the workability of metals. However, the partial-width indentation test does not subject the test specimen to true plane-strain conditions (Ref 18). In this test, a simple slab-shaped specimen is deformed over part of its width by two opposing rectangular anvils having widths smaller than that of the specimen. On penetrating the workpiece, the anvils longitudinally displace metal from the center, creating overhangs (ribs) that are subjected to secondary, nearly uniaxial tensile straining. The material ductility under these conditions is indicated by the reduction in the rib height at fracture. The test geometry has been standardized (Fig. 12).

One advantage of this test is that it uses a specimen of simple shape. In addition, as-cast materials can be readily tested. One edge of the specimen can contain original surface defects. The test can be conducted hot or cold. The partial-width indentation test, therefore, is suitable not only for determining the intrinsic ductilities of materials but also for evaluating the inhomogeneous aspects of workability. This test has been used to establish the fracture-limit loci for ductile metals (Ref 19).

The secondary-tension test, a modification of the partial-width indentation test, imposes more severe strain in the rib for testing highly ductile materials. In this test, a hole or a slot is machined in the slab-type specimen adjacent to where the anvils indent the specimen. Preferred dimensions of the hole and slot are given in Fig. 13. With this design, the ribs are sufficiently stretched to ensure fracture in even the most ductile materials. The fracture strain is based on reduction in area where the rib is cut out, so that the fracture area can be photographed or traced on an optical comparator.

The secondary-tension test (STS) was used to assess the workability in hot rolling of two aluminum alloys, alloy 5182 and 7075 (Ref 20). Figure 14 shows STS results as a function of temperature for the 7075 alloy. These data indicate that a loss of workability occurs at approximately 440 °C. Rolling tests on the material showed that edge cracking was just beginning when rolled at a temperature of 450 °C. Good correlation between loss of ductility in the STS and incipient edge cracking was similarly obtained for the 5182 alloy.

Ring Compression Test. When a flat, ring-shaped specimen is upset in the axial direction, the resulting change in shape depends only on the amount of compression in the thickness direction and the frictional conditions at the die-ring interfaces. If the interfacial friction was zero, the ring would deform in the same manner as a solid disk, with each element flowing outward radially at a rate proportional to its distance from the center.

In the case of small, but finite, interfacial friction, the outside diameter is smaller than in the zero-friction case. If the friction exceeds a critical value, frictional resistance to outward flow becomes so high that some of the ring material flows inward to the center. Measurements of the inside diameters of compressed rings provide a particularly sensitive means of studying interfacial friction, because the inside diameter increases if the friction is low and decreases if the friction is higher (Fig. 15).

The ring test, then, is a compression test with a built-in frictional measurement. Therefore, it is possible to measure the ring dimensions and compute both the friction value and the basic flow stress of the ring material at the strain under the given deformation conditions.

Analysis of Ring Compression. The mechanics of the compression of flat, ring-shaped specimens between flat dies have been analyzed using an upper bound plasticity technique (Ref 21, 22). Values of p/σ_0 (where p is the average forging pressure on the ring, and σ_0 is the flow stress of the ring material) can be calculated in terms of ring geometry and the interfacial shear factor, m. In these calculations, neither σ_0 nor the interfacial shear stress, τ, appears in terms of independent absolute values but only as the ratio m.

The plasticity equations have been solved for several ring geometries over a complete range of m values from 0 to unity (Ref 23), as shown in Fig. 16. The friction factor can be determined by measuring the change in internal diameter of the ring.

The ring thickness is usually expressed in relation to the inside diameter (ID) and outside diameter (OD). The maximum thickness that can be used while still satisfying the mathematical assumption of thin-specimen conditions varies,

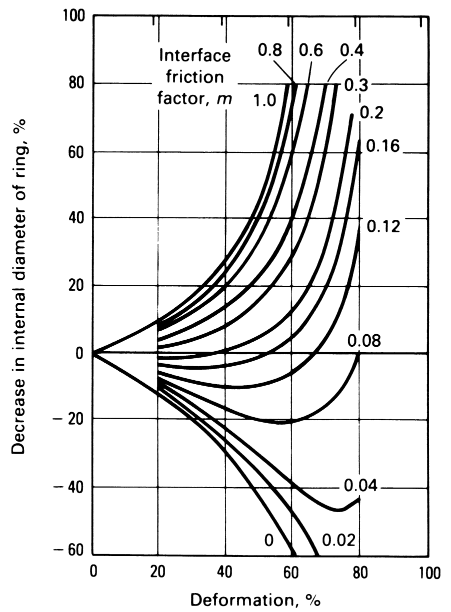
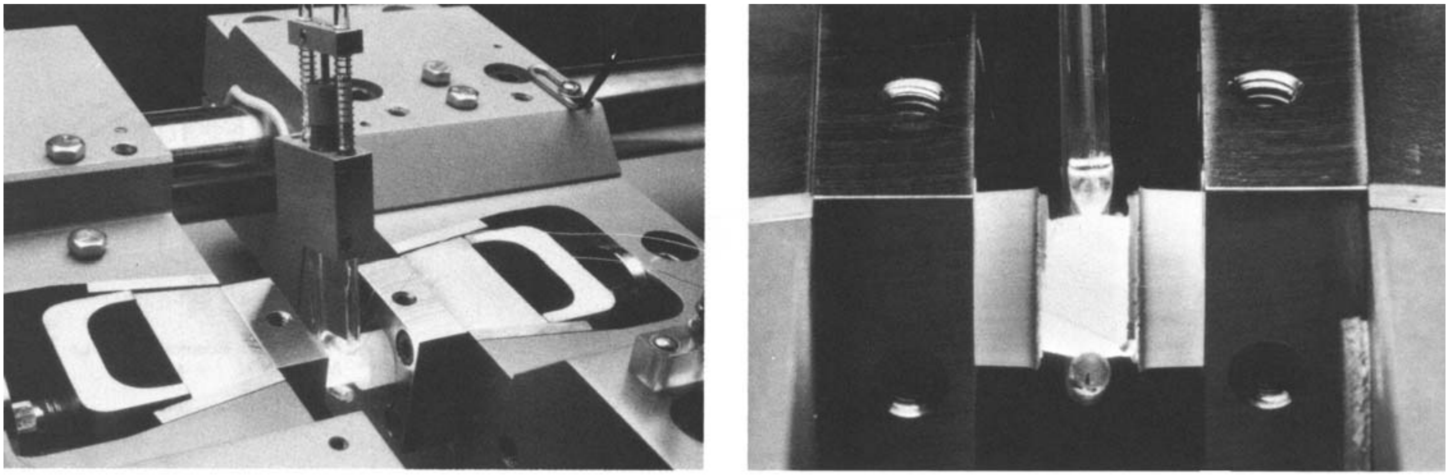


Fig. 16 Theoretical calibration curve for standard right with an outside-diameter-to-inside-diameter-to-thickness ratio of 6 to 3 to 2

depending on the actual friction conditions. Under conditions of maximum friction, the largest usable specimen height is obtainable with rings of dimensions in the OD-to-ID-to-thickness ratio of 6 to 3 to 1. Under conditions of low friction, thicker specimens can be used while still satisfying the previously mentioned assumption. For normal lubricated conditions, a geometry of 6 to 3 to 2 can be used to obtain results of sufficient accuracy for most applications.

For experimental conditions in which specimen thicknesses are greater than those permitted by a geometry of 6 to 3 to 1 and/or the interface friction is relatively high, the resulting side barreling or bulging must be considered. Analytical treatment of this more complex situation is available in Ref 24.

The ring compression test can be used to measure the flow stress under high-strain practical forming conditions. The only instrumentation required is that for measuring the force needed to produce the reduction in height. The change in diameter of the 6-to-3-to-1 ring is measured to obtain a value of the ratio p/σ_0 by solving the analytical expression for the deformation of the ring or by using computer solutions for the ring (Ref 25). Measurement of the



(a) (b)

Fig. 17 The Gleeble test unit used for hot tension and compression testing. (a) Specimen in grips showing attached thermocouple wires and linear variable differential transformer for measuring strain. (b) Closeup of a compression test specimen. Courtesy of Dynamics Systems, Inc.

area of the ring surface formerly in contact with the die and knowledge of the deformation load facilitate calculation of p and, therefore, the value of the material flow stress, σ_0, for a given amount of deformation. Repetition of this process with other ring specimens over a range of deformation allows the generation of a complete flow stress-strain curve for a given material under particular temperature and strain-rate deformation conditions. This method is particularly attractive for hot deformation processes where die chilling of the workpiece makes it difficult to simulate the test conditions to measure the flow stress.

Hot Tension Testing. Although necking is a fundamental limitation in tension testing, the tension test is nevertheless useful for establishing the temperature limits for hot working. The principal advantage of this test for industrial applications is that it clearly establishes maximum and minimum hot working temperatures (see Chapter 7, "Hot Tension Testing").

Most commercial hot tensile testing is done with a Gleeble unit, which is a high-strain-rate, high-temperature testing machine (Ref 26). A solid buttonhead specimen that has a reduced diameter of 6.4 mm (0.250 in.) and an overall length of 89 mm (3.5 in.) is held horizontally by water-cooled copper jaws (grips) through which electric power is introduced to resistance heat the test specimen (Fig. 17). Specimen temperature is monitored by a thermocouple welded to the specimen surface at its midlength. The thermocouple, with a function generator, controls the heat fed into the specimen according to a programmed cycle. Therefore, a specimen can be tested under time-temperature conditions that simulate hot working sequences.

The specimen is loaded by a pneumatic-hydraulic system. The load can be applied at any desired time in the thermal cycle. Temperature, load, and crosshead displacement are measured as a function of time. In the Gleeble test, the crosshead speed can be maintained as a constant throughout the test, but the true strain rate decreases until necking occurs, according to the relationship:

$$\dot{\varepsilon} = \frac{d\varepsilon}{dt} = \frac{1}{L}\frac{dL}{dt} \quad \text{(Eq 22)}$$

When the specimen necks, the strain rate increases suddenly in the deforming region, because deformation is concentrated in a narrow zone. Although this variable strain-rate history introduces some uncertainty into the determination of strength and ductility values, it does not negate the utility of the hot tension test. Moreover, a procedure has been developed that corrects for the change in strain rate with strain so that stress-strain curves can be constructed (Ref 27).

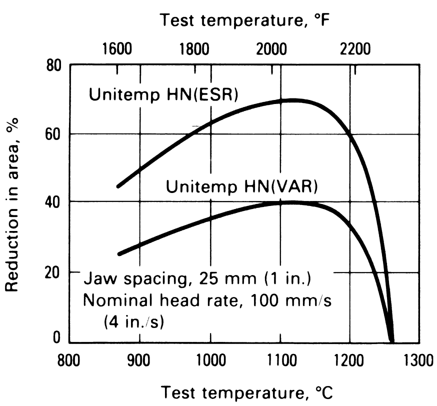
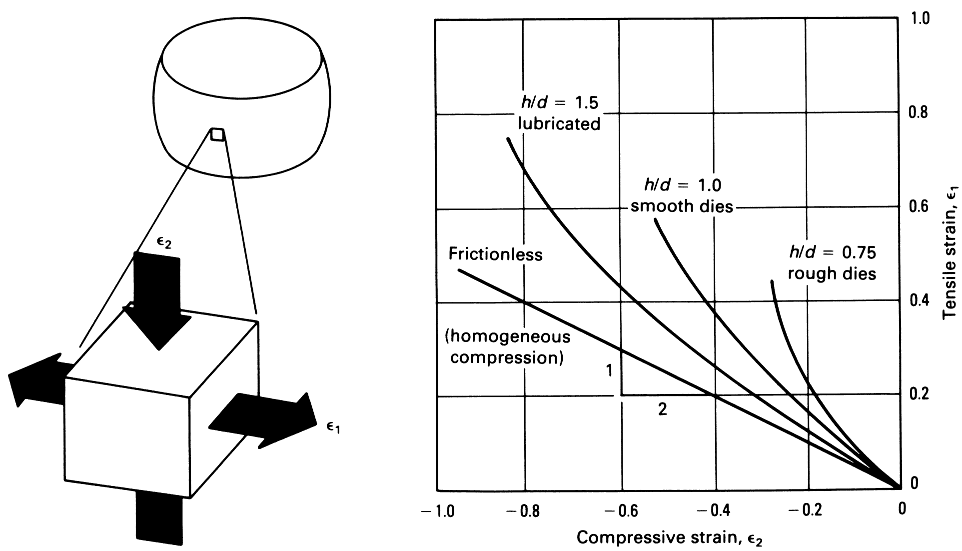


Fig. 18 Reduction in area versus test temperature obtained by hot tension testing on heating. Specimens were heated to the test temperature, held 5 min, and pulled to fracture.



(a) (b)

Fig. 19 (a) Localized strains on the bulging cylindrical surface of a compression test specimen. (b) Variation of strains with aspect ratio (h/D) of specimen and friction conditions. Source: Ref 28

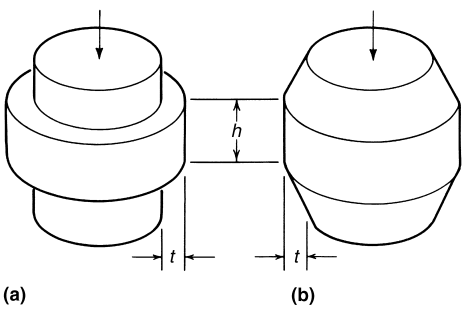


Fig. 20 (a) Flanged and (b) tapered prebulged compression test specimens. The lateral spread of the interior material under compression expands the rim circumferentially while little axial compression is applied.

The percent reduction in area is the primary result obtained from the hot tension test. This measure of ductility is used to assess the ability of the material to withstand crack propagation. Reduction in area adequately detects small ductility variations in materials caused by composition or processing when the material is of low-to-moderate ductility. It does not reveal small ductility variations in materials of very high ductility.

A general qualitative rating scale between reduction in area and workability is given in Table 2. This correlation was originally based on superalloys. In addition to ductility measurement, the ultimate tensile strength can be determined with the Gleeble test. This gives a measure of the force required to deform the material. Typical hot tension test curves are shown in Fig. 18.

Workability Analysis Using the Fracture Limit Line

Workability is determined by two main factors: the ability to deform without fracture, and the stress state and friction conditions present in the bulk deformation process. These two factors are brought together in an experimental workability analysis that has been further advanced (Ref 28). The method is most applicable when workability is limited by surface cracking, as in edge cracks in rolling or contact surface cracks in forging.

The first step is to determine the fracture limit line for the material in question at the temperature of interest. Details are given in Chapter 5, "Cold Upset Testing." While much work has been done for room-temperature deformation processes, it is possible to measure the curve at elevated temperature. Small compression specimens are used to establish the line. Variation of the *h*/*D* ratio and the end lubrication are used to change the ratio of the circumferential tensile strain, ε_1, to the axial compressive strain, ε_2, in the specimen. Figure 19 shows that for frictionless (ideal) compression, the strain ratio is $\varepsilon_1/\varepsilon_2 = -1/2$, and $\sigma_1 = 0$. The deformation in this case is referred to as homogeneous compression, because the only stress acting is an axial stress, and it is uniform throughout the specimen. This is the ideal condition to achieve when the compression test is

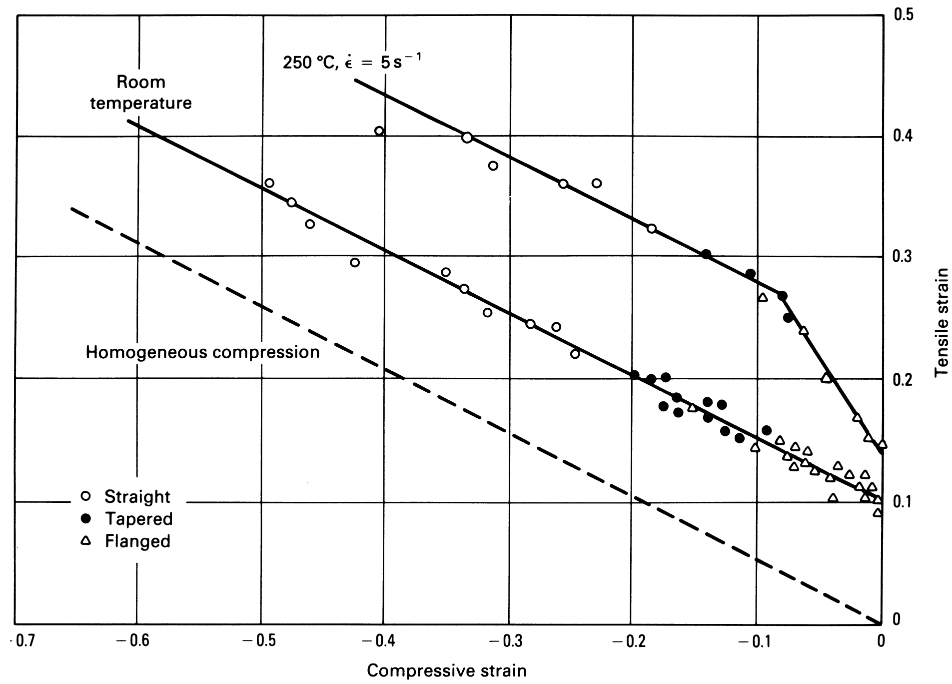


Fig. 21 Fracture limit lines for 2024 aluminum alloy in the T351 temper, measured by compression tests at room temperature and at 250 °C (480 °F). Source: Ref 28

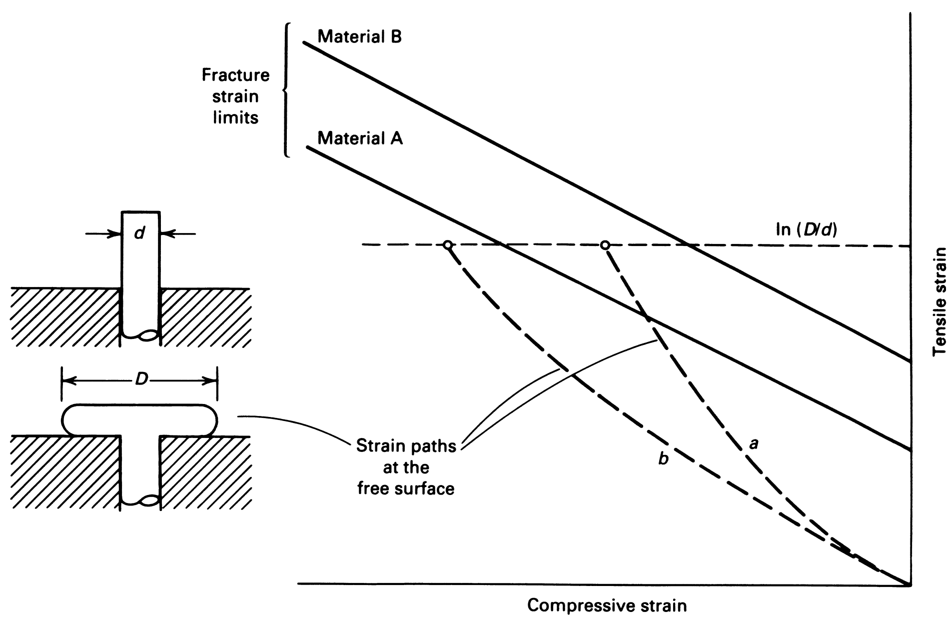


Fig. 22 Example of workability analysis. (a) Upsetting of a bar with diameter *d* to head with diameter *D* (b) Material fracture limit lines are superimposed on the strain paths by which the process achieves the final required strain. Strain path b (low friction) prevents fracture for both materials. Material B, with higher ductility, avoids fracture for either strain path. Source: Ref 28

used to measure the flow stress. When friction is present at the die contact surfaces, bulging occurs, and the strain path curves upward, as shown in Fig. 19.

The fracture limit line results from establishing the values of ε_1 and ε_2 at which a surface fracture can be just observed. Usually, this requires compressing the specimen incrementally with increasing strain or the testing of a series of identical specimens to increasing levels of ε_1. Changes in the geometry of the compression specimen are used to extend the range of surface strains toward the vertical (tensile) strain axis. This is done by prebulging the specimens by machining a flange or a taper on the cylinders (Fig. 20). Fracture limit curves take one of two forms: a straight line parallel to the line for homogeneous compression, but displaced along the tensile strain axis, or a curve with two straight segments. At low compressive strains,

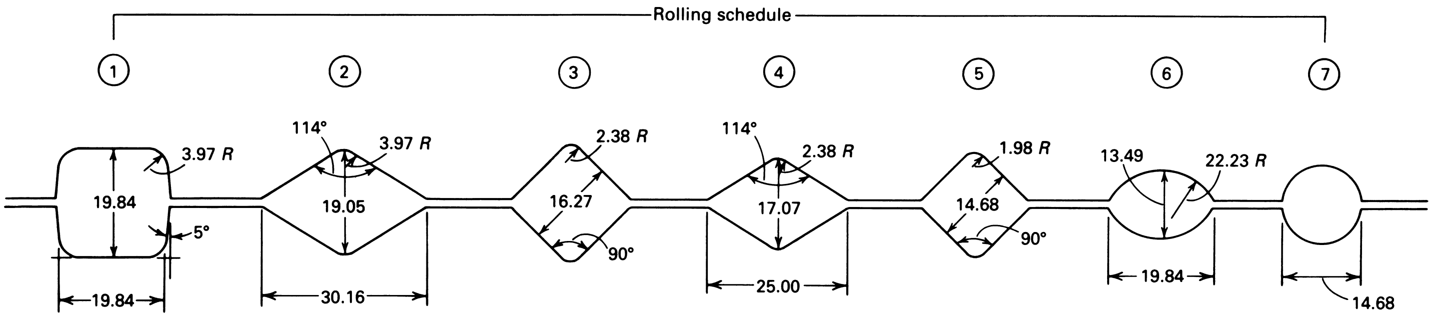


Fig. 23 Roll groove geometry for rolling square bars into round wire. Dimensions given in millimeters. Source: Ref 28

the line has a slope of −1, while at larger strains, the slope is $-^1/_2$. Figure 21 shows fracture limit lines for 2024 aluminum alloy that are determined at room temperature and 250 °C. Note how the tapered and flanged specimens are useful in the strain region close to the y-axis.

Workability Analysis. As a simplified example of the use of the fracture locus curve in workability analysis, consider a bolt-heading process (Fig. 22). A bolt head, D, must be formed from a rod with diameter, d. This requires achieving a tensile circumferential true strain of $\ln(D/d)$. However, the strain paths to achieve this strain depend on process parameters, such as h/d and friction, shown in Fig. 19. The fracture limit line for two materials, A and B, is shown. Now, if the strain path at the free surface is denoted by curve b, working with material A, the required strain is obtained without crossing the fracture limit line. Therefore, there is no fracture. However, if the lubrication in the cold-heading process is inadequate, the strain path may be curve a, and to reach $\ln(D/d)$ requires crossing the fracture limit line. Fracture ensues. There are two ways to solve this problem. Shifting to a material with a higher fracture limit, B, solves the problem. If this is impossible, then improving the lubrication to move the strain path back to curve b is called for. In other more complex situations, the process variables that can be changed include the die design, the workpiece (preform) design, and the distribution of lubricant.

A more realistic example concerns the production of aluminum 2024 round bar by hot rolling. Fracture occurs as a result of bulging of the free edges of the bars during rolling. Thus, it should be possible to predict fracture in bar rolling from compression tests used to develop the fracture limit curve. Figure 23 shows the sequence of passes required to roll a round bar. To determine the strains developed on the surface of the edge of the bar, the rolling process was modeled with lead bars rolled at room temperature. Surface strains were determined from measurement of a grid scribed on the edges with a sharp tool. In other situations, the strains could be calculated from analytical methods or by finite-element analysis (Ref 29). Figure 24 shows the measured localized strains for the various types of roll passes. The greatest tensile strain is produced in going from a square to a diamond cross section, and the least strain occurs on going from a square to an oval one. To the right in Fig. 24, the correlation between tensile strain and area reduction is given. The fracture limit lines for the material were given in Fig. 21. Superposition

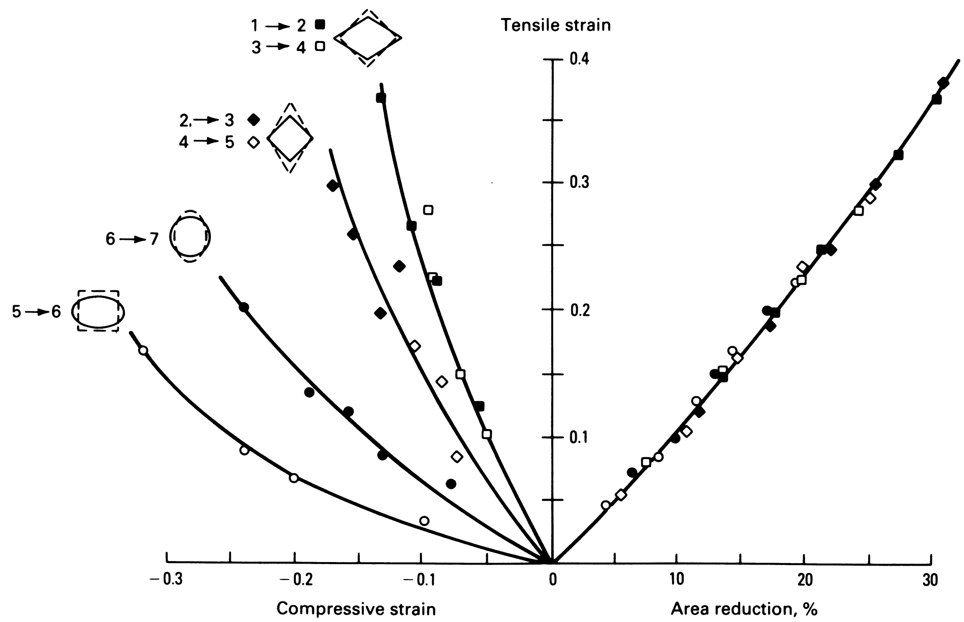


Fig. 24 Measured localized strains during the rolling of lead bars. Left side shows longitudinal strain versus vertical compressive strain. Right side shows longitudinal strain versus cross-section reduction of area at room temperature. Source: Ref 28

Table 2 Qualitative hot workability ratings for specialty steels and superalloys

Hot tensile reduction in area(a), %	Expected alloy behavior under normal hot reductions in open-die forging or rolling	Remarks regarding alloy hot working practice
<30	Poor hot workability, abundant cracks	Preferably not rolled or open-die forged; extrusion may be feasible; rolling or forging should be attempted only with light reductions, low strain rates, and an insulating coating.
30–40	Marginal hot workability, numerous cracks	This ductility range usually signals the minimum hot working temperature; rolled or press forged with light reductions and lower-than-usual strain rates
40–50	Acceptable hot workability, few cracks	Rolled or press forged with moderate reductions and strain rates
50–60	Good hot workability, very few cracks	Rolled or press forged with normal reductions and strain rates
60–70	Excellent hot workability, occasional cracks	Rolled or press forged with heavier reductions and higher strain rates than normal, if desired
>70	Superior hot workability, rare cracks. Ductile ruptures can occur if strength is too low.	Rolled or press forged with heavier reductions and higher strain rates than normal, if alloy strength is sufficiently high to prevent ductile ruptures

(a) Ratings apply for Gleeble tension testing of 6.4 mm (0.250 in.) diam specimens with 25 mm (1 in.) head separation. Source: Ref 26

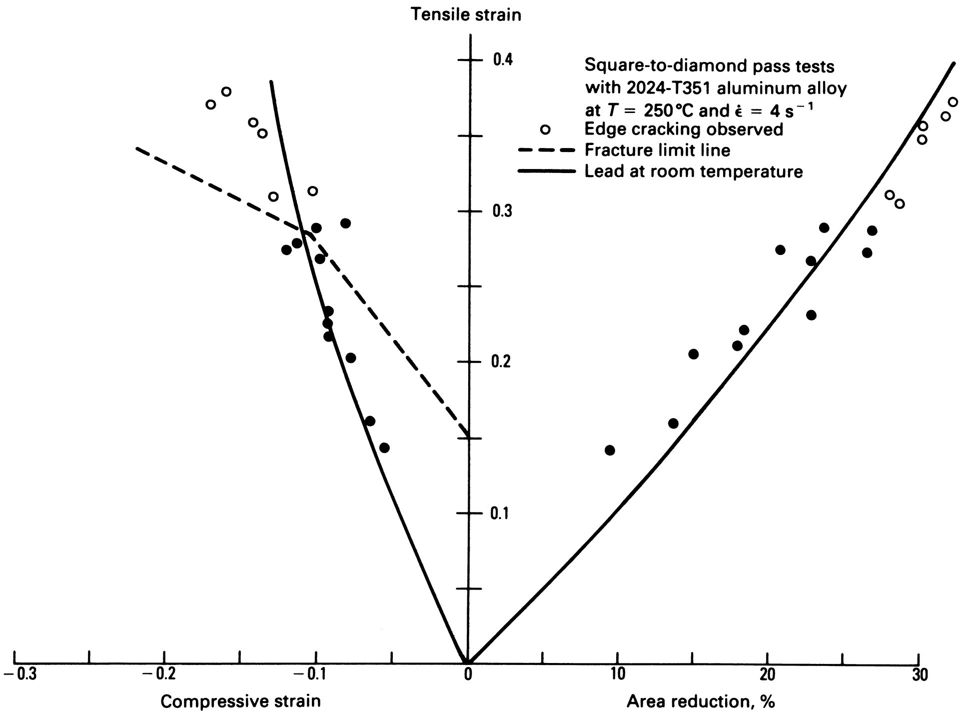


Fig. 25 Superposition of fracture limit line (dashed line) from Fig. 21 for 250 °C on the measured strains in rolling lead at room temperature (model material). Source: Ref 28

of the curve for 250 °C on the plot of strain path for the square-to-diamond pass yields Fig. 25. This shows that the limiting reduction for hot rolling of a square-to-diamond pass is approximately 27%.

Many other examples of the use of workability analysis in forging and extrusion, and examples of workability analysis applied to other forms of fracture, such as central burst and surface fir-tree defect, are given (Ref 28). Workability analysis can be used as a troubleshooting methodology when problems arise in production, or better yet, as a tool for the design of new processes.

REFERENCES

1. P. Han, Ed., *Tesile Testing,* ASM International, 1992
2. S.L. Semiatin, G.D. Lahoti, and J.J. Jonas, Application of the Torsion Test to Determine Workability, *Mechanical Testing,* Vol 8, *Metals Handbook,* 9th ed., American Society for Metals, 1985, p 154–184
3. S.L. Semiatin and J. Jonas, Torsion Testing to Assess Bulk Workability, *Workability and Process Design,* ASM International, 2003
4. S. Fulop, K.C. Cadien, M.J. Luton, and H.J. McQueen, *J. Test. Eval.,* Vol 5, 1977, p 419
5. J.A. Schey, T.R. Venner, and S.L. Takomana, Shape Changes in the Upsetting of Slender Cylinders, *J. Eng. Ind. (Trans. ASME),* Vol 104, 1982, p 79
6. G. Fitzsimmons, H.A. Kuhn, and R. Venkateshwar, Deformation and Fracture Testing for Hot Working Processes, *J. Met.*, May 1981, p 11–17
7. J.F. Alder and V.A. Phillips, The Effect of Strain Rate and Temperature on the Resistance of Aluminum, Copper and Steel to Compression, *J. Inst. Met.,* Vol 83, 1954–1955, p 80–86
8. F.J. Gurney and D.J. Abson, Heated Dies for Forging and Friction Studies on a Modified Hydraulic Forge Press, *Metall. Mater.,* Vol 7, 1973, p 535
9. J.G. Lenard, Development of an Experimental Facility for Single and Multistage Constant Strain Rate Compression, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 107, 1985, p 126–131
10. A.B. Watts and H. Ford, On the Basic Yield Stress Curve for a Metal, *Proc. Inst. Mech. Eng.,* Vol 169, 1955, p 1141–1149
11. F.-K. Chen and C.-J. Chen, On the Nonuniform Deformation of the Cylinder Compression Test, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 122, 2000, p 192–197
12. R. Hill, A General Method of Analysis for Metal-Working Processes, *J. Mech. Phys. Solids,* Vol 11, 1963, p 305–326
13. J.J. Jonas, R.A. Holt, and C.E. Coleman, Plastic Stability in Tension and Compression, *Acta Metall.,* Vol 24, 1976, p 911
14. S.L. Semiatin and J.J. Jonas, *Formability and Workability of Metals: Plastic Instability and Flow Localization,* American Society for Metals, 1984
15. P. Dadras, Stress-Strain Behavior in Bending, *Mechanical Testing and Evaluation,* Vol 8, *ASM Handbook,* ASM International, 2000, p 109–114
16. S.L. Semiatin, Workability in Forging, *Workability and Process Design,* ASM International, 2003
17. N.J. Silk and M.R. van der Winden, Interpretation of Hot Plane Strain Compression Testing of Aluminum Specimens, *Mater. Sci. Technol.,* Vol 15, 1999, p 295–300
18. S.M. Woodall and J.A. Schey, Development of New Workability Test Techniques, *J. Mech. Work. Technol.,* Vol 2, 1979, p 367–384
19. S.M. Woodall and J.A. Schey, Determination of Ductility for Bulk Deformation, *Formability Topics—Metallic Materials,* STP 647, ASTM, 1978, p 191–205
20. D. Duly, J.G. Lenard, and J.A. Schey, Applicability of Indentation Tests to Assess Ductility in Hot Rolling of Aluminum Alloys, *J. Mater. Process. Technol.,* Vol 75, 1998, p 143–151
21. B. Avitzur, *Metal Forming: Processes and Analysis,* McGraw-Hill, 1968
22. B. Avitzur and C.J. Van Tyne, Ring Forming: An Upper Bound Approach, *J. Eng. Ind. (Trans. ASME),* Vol 104, 1982, p 231–252
23. A.T. Male and V. DePierre, The Validity of Mathematical Solutions for Determining Friction from the Ring Compression Test, *J. Lubr. Technol. (Trans. ASME),* Vol 92, 1970, p 389–397
24. V. DePierre, F.J. Gurney, and A.T. Male, "Mathematical Calibration of the Ring Test with Bulge Formation," Technical Report AFML-TR-37, U.S. Air Force Materials Laboratory, March 1972
25. A.T. Male, V. dePierre, and G. Saul, *ASLE Trans.,* Vol 16, 1973, p 177–184
26. E.F. Nippes, W.F. Savage, B.J. Bastian, H.F. Mason, and R.M. Curran, An Investigation of the Hot Ductility of High Temperature Alloys, *Weld. J.,* Vol 34, April 1955, p 183–196s; see also http://www2.gleeble.com/gleeble/
27. R.L. Plaut and C.M. Sellars, Analysis of Hot Tension Test Data to Obtain Stress-Strain Curves to High Strains, *J. Test Eval.,* Vol 13, 1985, p 39–45
28. H.A. Kuhn, Workability Theory and Application in Bulk Forming Processes, *Forming and Forging,* Vol 14, *ASM Handbook,* ASM International, 1988, p 388–404
29. D.L. Dewhirst, Finite Element Analysis, *Materials Selection and Design,* Vol 20, *ASM Handbook,* ASM International, 1997, p 176–185

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 57-60
DOI:10.1361/hwpd2003p057

Chapter 5

Cold Upset Testing

Howard A. Kuhn, Consultant

DUCTILE FRACTURE is an extremely complex phenomenon that is dependent on the microstructure of the material and the complex stress and strain states developed during the deformation process (Ref 1, 2). The mechanism of ductile fracture is not completely understood but can be described by the process of void coalescence under shear deformation (Ref 3, 4). There is also some evidence that local plastic instability is associated with ductile fracture in metalworking (Ref 5, 6).

In evaluating the potential for fracture in a material during a particular process, it is generally necessary to combine a workability test with a fracture criterion. The workability test establishes the ductility of the material under standard conditions. The function of the fracture criterion is to extend this result to the stress and strain conditions existing in the deformation process of interest. Attention must be given, then, to the validity of the workability test in simulating the type of fracture that occurs in the actual process.

Most metalworking processes involve compressive deformation, and so the uniaxial compression test has been widely used for studying deformation behavior. However, the presence of friction between workpiece and tool in the compression test makes analysis complicated and restricts its application as a basic test of mechanical behavior. A cylindrical specimen compressed with friction at the die surfaces does not remain cylindrical in shape but becomes bulged or barreled. Severity of the bulged cylindrical surface increases with increasing friction and decreasing height-to-diameter ratio (Ref 7). Tensile stresses associated with the bulging surface make the upset test a candidate for workability testing (Fig. 1).

The effects of the bulged surface in cylindrical compression tests can be enhanced by the use of shaped compression specimens. These specimens contain flanged or tapered sections at midheight, which increase the circumferential tensile strain and decrease the axial compressive strain (Ref 8).

Collectively, the cylindrical, tapered, and flanged compression-test specimens provide a wide range of circumferential tension/axial compression strain states.

Strains at the equatorial surfaces of the compression-test specimens can be measured from grid marks. Measurement of strains at fracture for a wide range of test conditions leads to construction of a fracture-limit line for the material. Comparison of the strains in a material during an actual deformation process with the fracture-limit line for the material indicates the possibility of fracture (Ref 9).

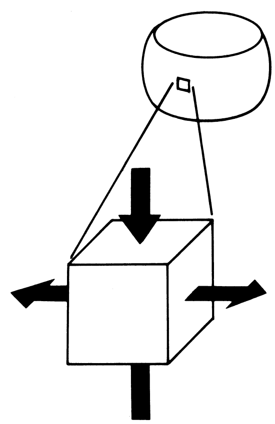


Fig. 1 Circumferential tension and axial compression stresses at the equator of an upset cylinder

Upset Test Technique

Test-Specimen Geometry and Friction Conditions. The aspect (height-to-diameter) ratio affects the strains occurring at the bulging free surface. The upper limit on this ratio is 2.0 because of the possibility of buckling. The lower limit is based on a convenient height for application of the grid marks. Normally, specimen aspect ratios range between 0.75 and 1.75.

Tapered and flanged test specimens are machined cylinders with a section of the original surface remaining at midheight, as shown in comparison with the cylindrical compression-test specimen in Fig. 2. The height of the cylindrical surface at midheight ranges from 0.2 to 0.75 times the specimen height. The reduced diameter of the flanged compression-test specimens is 0.8 times the original cylinder diameter, and the angle of the tapered compression specimens is 20°. In no case should the end-face diameter be less than one-half of the overall specimen height.

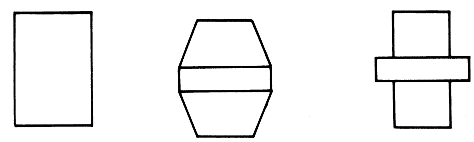


Fig. 2 Cylindrical (left), tapered (center), and flanged (right) compression-test specimens

Friction at the die contact surface also has a significant effect on the strains occurring at the free surface, particularly for cylindrical compression specimens. In the present work, the end faces of each specimen were polished to avoid lubricant entrapment. Die conditions include dies with knurled surfaces, dies with polished surfaces (2 μin. finish), and polished dies with MoS_2 lubricant. In the Male and Cockroft ring compression test (Ref 10), these friction conditions give friction factors, respectively, of 0.45, 0.21, and 0.085.

Each combination of friction condition and cylindrical aspect ratio gives a different combination of axial compressive strain and circumferential tensile strain at the free surface of a cylindrical specimen. Friction has the greatest effect on these strains. For the flanged and tapered compression-test specimens, the height of the midplane cylindrical section affected the strains, but friction had very little effect. Each test, with its own combination of surface strains, provides a point on the fracture-strain-limit line.

Strain Measurements. Use of the upset test for fracture studies, as first described by Kudo and Aoi (Ref 11), involves measurement of the axial and hoop strains at the equatorial surface. Preliminary tests showed that the hoop strain was uniform around the circumference of the barreled surface. The hoopstrains, therefore, can be determined from measurements of the equatorial diameter.

The axial strain requires measurement of the separation of two gage marks placed symmetrically above and below the midplane (Fig. 3). One of the following methods can be used for applying gage marks:

- Two strips of cellophane tape are wrapped around the cylindrical surface, such that there is a thin band of bare metal around the circumference at midheight. The specimens are then sprayed with blue steel dye, and the strips of tape are removed when the dye has dried. The gage band is approximately 3 mm (0.12 in.) wide and has sharply defined edges that deform with the specimen surface without cracking as the test proceeds.
- The gage band is produced by electrochemical etching using the electromark system or by the photo-resist method.
- The gage band is marked by scribing fine lines approximately 3 mm (0.12 in.) apart with a carbide-tipped height gage.
- Four gage points are indented at midheight of the cylindrical specimen with a Vickers or diamond brale hardness indenter.

Scribed lines or hardness indentations should be applied only to very ductile materials, because these gage marks could be potential stress concentrations, leading to premature fracture during upset testing.

Test specimens complete with gage bands are compressed incrementally between flat dies under one of the friction conditions until cracks at the bulged surface are observed visually. For each incremental deformation, displacement of the gage band is measured with a tool-maker's microscope to an accuracy of +0.013 mm (+0.0005 in.), and the diameter is measured with a micrometer to the same degree of accuracy. Average values are calculated from three measurements in an effort to obtain more accurate data.

The axial and circumferential true strains on the barreled surface are calculated by using the relations given in Eq 1 and 2:

Axial $$\varepsilon_z = \ln\left(\frac{h}{h_0}\right) \quad \text{(Eq 1)}$$

Circumferential $$\varepsilon_\theta = \ln\left(\frac{W}{W_0}\right) \text{ or } \ln\left(\frac{D}{D_0}\right) \quad \text{(Eq 2)}$$

where h_0 and h are initial and final gage heights, respectively; W_0 and W are initial and final gage widths, respectively; and D_0 and D are initial and final diameters, respectively (Fig. 3).

Crack Detection. Most materials exhibit some orange peel or surface roughening shortly after plastic flow begins. Fine networks of microcracks may also appear if the surface is observed at a magnification of approximately 30×. These cracks are stable and do not grow in size as deformation progresses. They should be ignored, because they do not represent a limitation to useful deformation of the material. After further deformation, one or more large cracks form and grow rapidly; these cracks can easily be seen with the unaided eye. The strains are measured, in accordance with Eq 1 and 2, when such large cracks are initiated.

There is always some uncertainty regarding the initiation of a large crack, and this uncertainty affects the accuracy of the results. Because large cracks grow rapidly, however, the strains measured at the first sight of a crack are very close to those that existed when the crack was actually initiated.

Material Considerations. Because cracks are initiated at the bulged surface, the ductility of the specimen during upsetting is significantly influenced by surface-condition factors such as surface defects, decarburization, and residual stress. It has been shown (Ref 12) that longitudinal surface defects on a spheroidized high-carbon steel sample drastically reduce its ductility during uniaxial compression testing due to circumferential tensile-stress concentrations at surface defects..

Material inhomogeneities of any type have an adverse effect on workability. Localized variations that can reduce the workability of a steel include ferrite grain size, nonmetallic inclusion content, banding, center segregation, and decarburization. In all cases, therefore, the test-specimen free surfaces should contain the same surface structural features as those of the actual material to be used in the deformation process of interest.

Another condition to be taken into account is the effect of mechanical anisotropy on ductile fracture: the direction of the inclusion alignment relative to the secondary tensile hoop stress generated by friction during the uniaxial compression test strongly influences the strain to fracture. It is necessary, therefore, that the test specimen be prepared so that the direction of tensile stresses, relative to the direction of inclusion alignment, is the same during testing as it is in the actual process.

Test Characteristics

Deformation. In a cylindrical compression test carried out under frictionless conditions, deformation of the test specimen is uniform, and no bulging is produced, as depicted in the sample on the right in Fig. 4. Frictionless conditions can be achieved through the use of a synthetic fluorine-containing resin film under laboratory conditions. In this way, the material flow-stress characteristics can be determined, unadulterated by the effects of friction.

Frictional constraint at the end faces of the specimen under conventional compression prevents uniform deformation and leads to a bulged surface at the midheight of the specimen, as shown in the center specimen in Fig. 4. Such bulging leads to circumferential tensile stresses (depicted in Fig. 1) and the possibility of fracture.

The inhomogeneous deformation occurring during compression with friction leads to three zones of deformation, as depicted in Fig. 5:

- Metal in contact with the top and bottom surfaces of the compression platens remains almost stationary. Such areas are also called dead-metal zones, or DMZs (region I in Fig. 5).
- Metal near the outer surface of the cylinder deforms as a result of compression, resulting in a bulged surface (region III in Fig. 5).
- The most severe deformation is concentrated in zones of shear just outside the DMZs near each contact surface (region II in Fig. 5).

In the case of tapered or flanged compression-test specimens, the interior deformation of the cylinder expands the central region, accentuating the circumferential tension. Because the free surface at midheight is not directly in contact with the die surfaces along a straight line, compression of this section is less than that in cylindrical compression.

Free-Surface Strains. At the free surfaces of the compressed cylinders, the strains consist of circumferential tension and axial compression. For frictionless (homogeneous) compression of cylindrical specimens, the tensile strain is equal to one-half the compressive strain. With increas-

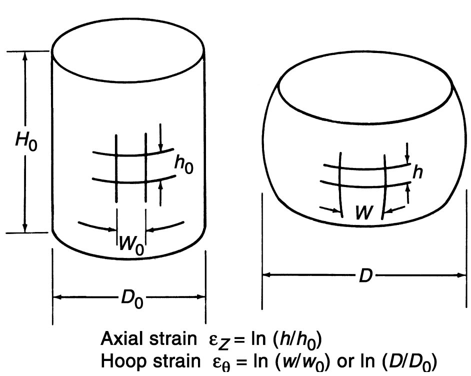


Fig. 3 Grids for strain measurements on upset cylinders

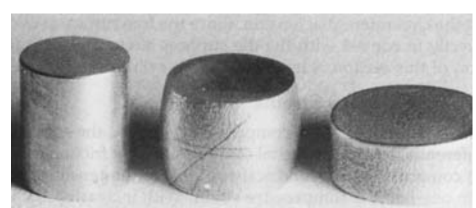


Fig. 4 Cylindrical compression-test specimens: (left) undeformed specimen; (center) specimen compressed with friction (note crack); (right) specimen compressed without friction

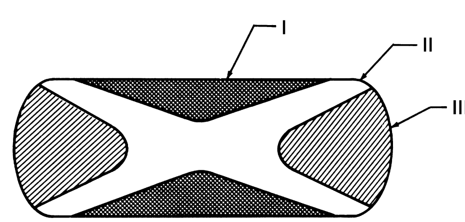


Fig. 5 Deformation zones in a longitudinal section of an upset cylinder (see text)

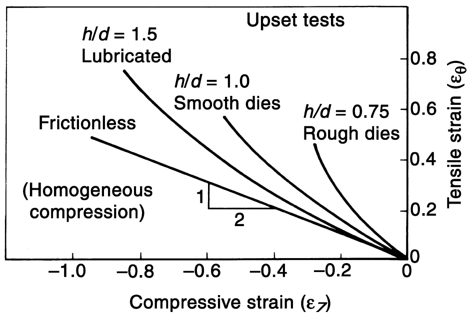


Fig. 6 Strain paths in upset-test specimens

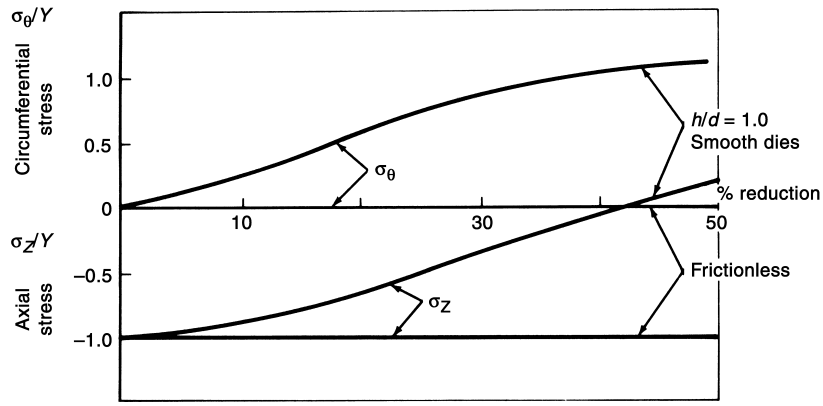


Fig. 7 Stresses at the equatorial surface of an upset-test specimen

ing frictional constraint, bulge severity increases, the tensile strain becomes larger, and the compressive strain decreases. Effects of the bulged profile, which develops naturally in cylindrical compression specimens, are imposed artificially through tapered and flanged compression tests.

Figure 6 summarizes the effects of friction, aspect ratio, and specimen profile on the measured free-surface strains at midheight. Measured strain paths are shown in terms of circumferential versus axial strain. Starting with the strain ratio of one-half for the case of homogeneous deformation, the strain-path slope increases with increasing friction. For a given value of friction, a decreasing aspect ratio slightly increases the strain-path slope. Tapered compression specimens further increase the strain-path slope, and flanged compression specimens result in strain paths that lie nearly along the circumferential tensile-strain axis.

Empirical relationships between test conditions and the strain paths shown in Fig. 6 have been derived in Ref 13. By approximating the strain paths as second order equations, the coefficients of the linear and quadratic terms are correlated statistically to the specimen aspect ratio (H/D) and friction. In addition, finite element analyses of the upsetting process (e.g., Ref 14) have been performed to determine the strain paths as a function of the process parameters.

Stress States. Stresses at the free surfaces of compressed specimens can be calculated from measured strains and plasticity equations. The Levy-Mises equations relating plastic strain increments to stress for an isotropic material, expressed in cylindrical coordinates, are:

$$d\varepsilon_r = d\lambda[\sigma_r - (\sigma_\theta + \sigma_z)/2] \qquad \text{(Eq 3)}$$

$$d\varepsilon_\theta = d\lambda[\sigma_\theta - (\sigma_r + \sigma_z)/2] \qquad \text{(Eq 4)}$$

$$d\varepsilon_z = d\lambda[\sigma_z - (\sigma_\theta + \sigma_r)/2] \qquad \text{(Eq 5)}$$

The equivalent strain increment, $d\bar{\varepsilon}$, and equivalent stress, $\bar{\sigma}$, are given as:

$$d\bar{\varepsilon} = \frac{\sqrt{2}}{3}\left[(d\varepsilon_r - d\varepsilon_z)^2 + (d\varepsilon_z - d\varepsilon_\theta) + (d\varepsilon_\theta - d\varepsilon_r)^2\right]^{1/2} \qquad \text{(Eq 6)}$$

$$\bar{\sigma} = \frac{1}{\sqrt{2}}\left[(\sigma_r - \sigma_z)^2 + (\sigma_z - \sigma_\theta)^2 + (\sigma_\theta - \sigma_r)^2\right]^{1/2} \qquad \text{(Eq 7)}$$

where $d\varepsilon_r$, $d\varepsilon_\theta$, and $d\varepsilon_z$ are incremental strains in the $\bar{r}$, $\bar{\theta}$, and z-directions; σ_r, σ_θ, and σ_z are stresses in the $\bar{r}$, $\bar{\theta}$, and z-directions; and $d\lambda$ is a proportionality constant that depends on material and strain level and that is given by $d\bar{\varepsilon}/\bar{\sigma}$. For the surface area under consideration (Fig. 1), the stress in the r-direction is zero ($\sigma_r = 0$), and the stress state becomes plane stress.

By manipulation of Eq 3–5, the stress components during deformation can be calculated according to the following equations:

$$\sigma_z = \frac{\bar{\sigma}}{\sqrt{3}}\left(\frac{2\alpha + 1}{\sqrt{\alpha^2 + \alpha + 1}}\right) \qquad \text{(Eq 8)}$$

$$\sigma_\theta = \frac{\bar{\sigma}}{\sqrt{3}}\left(\frac{\alpha + 2}{\sqrt{\alpha^2 + \alpha + 1}}\right) \qquad \text{(Eq 9)}$$

where $\alpha = d\varepsilon_z/d\varepsilon_\theta$, which is determined graphically from the strain paths (Fig. 6). The effective stress ($\bar{\sigma}$) is taken from the stress-strain curve of the material as obtained by a homogeneous compression test.

From Eq 8 and 9, the stress ratio, β, is defined as:

$$\beta = \frac{\sigma_\theta}{\sigma_z} = \frac{\alpha + 2}{2\alpha + 1}$$

For example, in homogeneous deformation, $\alpha = -2$, and β becomes zero. Consequently, the hoop stress becomes zero, and no fractures occur in this case. At the other extreme, when $\alpha > -½$, β becomes positive, and both σ_θ and σ_z are tensile. When $\alpha = 0$, $\sigma_z = ½\sigma_\theta$, which is the stress state in plane strain. Figure 7 illustrates typical changes in σ_θ and σ_z in a cylindrical compression test. Note that, at large deformations, as the bulge severity increases, σ_z becomes tensile.

Fracture Limits

As indicated in the previous sections, a wide range of stress and strain conditions can be generated at the free surfaces of cylindrical, tapered, or flanged test specimens. This range of conditions permits evaluation of the effects of variations in stress and strain states on the occurrence of fracture. The most convenient representation of fracture limits is a plot of circumferential and axial strains at fracture.

Previous studies have shown that the results of homogeneous cylindrical compression tests fit a straight line having a slope of one-half for all materials (Ref 7, 9). The projected intercept with the vertical axis depends on the material. Some examples are shown in Fig. 8. In addition, fracture lines have been generated for a wide variety of materials by various investigators. A summary of these results is given in Ref 15, indicating the value of the intercept for each material. As expected, materials having greater ductility have higher values of the fracture line intercept.

Use of flanged and tapered upset-test specimens expands the range of strains available for testing. Figures 9 and 10 show the combined results of cylindrical, tapered, and flanged upset tests on two alloys. The combination of results from all tests provides a continuous spectrum of strain states, with the tapered and flanged specimens adjoining, in sequence, the cylindrical upset-test results. For the aluminum alloy (Fig. 9), the data fit a straight line of constant slope, but the medium-carbon steel (Fig. 10) shows bilinear behavior, with a larger slope in the small strain region (Ref 8).

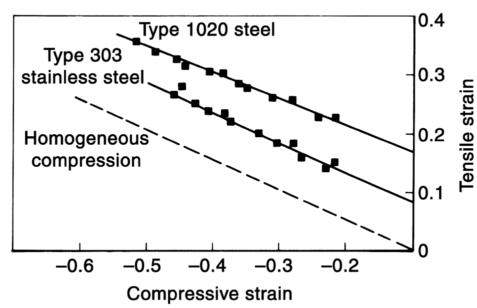


Fig. 8 Fracture loci in cylindrical upset-test specimens of two materials

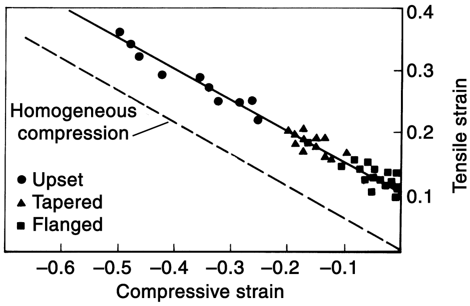


Fig. 9 Fracture loci in cylindrical, tapered, and flanged upset-test specimens of aluminum alloy 2024-T351

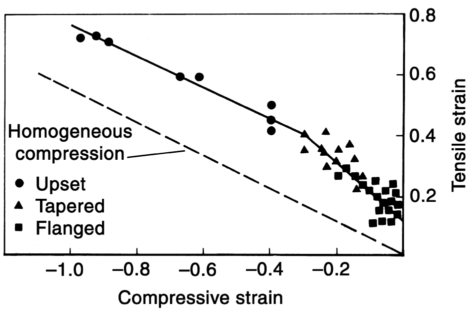


Fig. 10 Fracture loci in cylindrical, tapered, and flanged upset-test specimens of type 1045 cold-finished steel

Workability Diagram

The results of Fig. 8–10 constitute workability diagrams for free-surface fracture in bulk forming operations. They embody, graphically, both the material influence in the fracture-strain-limit line and the process influence through the variation of strain paths with die geometry, workpiece geometry, and friction. To illustrate this interrelationship, consider the bolt-heading process in Fig. 11. If material A is used for the product, strain path *a* crosses the fracture line on its way to its position in the final deformed geometry, and cracking is likely. Two options are open for avoiding defects: use material B, which has a higher forming-limit line; or alter the strain path to follow *b,* which, in this case, is accomplished through improved lubrication. Applications of this workability concept in deformation process design are given in Chapter 12 for a wide variety of materials and processes.

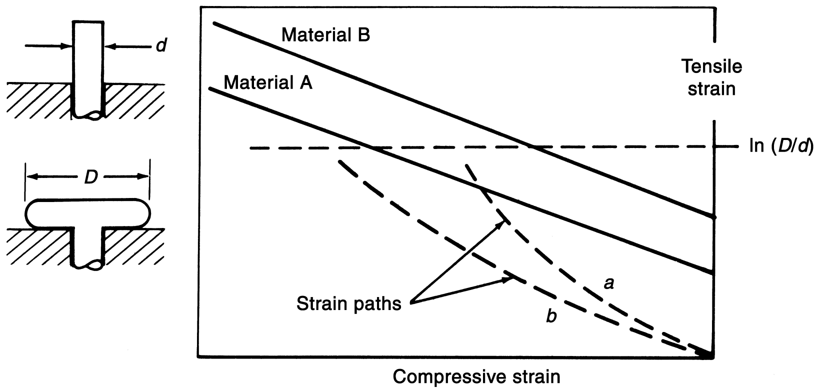


Fig. 11 Comparison of strain paths and fracture locus lines

Conclusions

A fracture criterion based on the linear relations between total surface strains at fracture has been presented. Plotting of the fracture-strain line on a graph of tensile surface strain (circumferential strain, ε_θ) versus compressive surface strain (axial strain, ε_z) yields a straight line with a slope of one-half, except in some exceptional cases where a bilinear relationship exists at the low-strain region. In any case, the fracture line represents the material limits to deformation. The strain paths (relation between tensile and compressive strains during deformation) are functions of the process parameters. Together, the fracture line and strain paths give a graphical representation of the material and process factors in workability. Comparison of the strain-path terminations required for complete formation of the part with the limiting fracture strains of the material indicates whether or not fracture is likely and suggests possible alterations of the material or process to avoid fracture.

ACKNOWLEDGMENT

This chapter, originally published in *Workability and Testing Techniques,* American Society for Metals, 1984, is revised from the 1984 text written by Peter W. Lee (deceased) and Howard Kuhn.

REFERENCES

1. G.E. Dieter, in *Ductility,* American Society for Metals, 1968, p 1
2. A.L. Hoffmanner, *Metal Forming: Interrelation Between Theory and Practice,* Plenum Press, 1971, p 349
3. H.C. Rogers, in *Ductility,* American Society for Metals, 1968, p 31
4. K.E. Puttick, *Philos. Mag.,* Vol 4 (Ser. 8), 1959, p 964
5. P.F. Thomason, *Int. J. Mech. Sci.,* Vol 11, 1969, p 189
6. H.A. Kuhn and P.W. Lee, *Metall. Trans.,* Vol 2, 1971, p 3197
7. P.W. Lee and H.A. Kuhn, *Metall. Trans.,* Vol 4, 1974, p 969
8. E. Erman and H.A. Kuhn, "Novel Test Specimens for Workability Measurement," Proc. ASTM Conf. on Compression Testing, ASTM, 3–5 March 1982
9. H.A. Kuhn, P.W. Lee, and T. Erturk, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 95H, 1973, p 213
10. A.T. Male and M.G. Cockroft, *J. Inst. Met.,* Vol 93, 1964, p 38
11. H. Kudo and K. Aoi, *J. Jpn. Soc. Techol. Plast.,* Vol 8, 1967, p 17
12. P.F. Thomason, *Int. J. Mech. Sci.,* Vol 11, 1969, p 187
13. J.J. Shah and H.A. Kuhn, Application of a Forming Limit Concept to Upsetting and Bolt Heading, *Proc. ASTM Conf. on Compression Testing.* ASTM, 3–5 March 1982
14. H.P. Ganser, A.G. Atkins, O. Kolednik, F.D. Fischer, and O. Richard, Upsetting of Cylinders: A Comparison of Two Different Damage Indicators, *Trans. ASME, J. Eng. Mater. Technol.,* Vol 123, 2001, p 94–99
15. A. Jenner and B. Dodd, Cold Upsetting and Free Surface Ductility, *J. Mech. Work. Technol.*, Vol 5, 1981, p 31–43

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 61-67
DOI:10.1361/hwpd2003p061

Chapter 6

Hot-Compression Testing

George E. Dieter, University of Maryland

THE UTILITY OF THE COMPRESSION TEST in workability studies has been shown in previous chapters. A cylindrical specimen is relatively easy to prepare, and loading in compression simplifies the design of grips. Moreover, the absence of necking, as occurs in a tension test, means that a compression test can be conducted to large strains that are comparable to those found in deformation processes such as forging or extrusion. However, as discussed in Chapter 4, "Bulk Workability Testing," barreling of the specimen due to friction at the platens complicates the determination of the true axial compression stress when the test is used to measure the stress-strain curve (flow properties) of the material. However, as Chapter 5, "Cold Upset Testing," describes, this feature of the compression test can be used as a workability test by measuring the propensity for surface cracking in deformation processes.

This chapter discusses the use of the compression test at elevated temperatures, a topic that is just touched on in previous chapters. Elevated temperature introduces the obvious experimental issues of providing a uniform and controlled temperature, as well as dealing with more difficult problems of lubrication of the platens. Incremental loading, which can be used for precise flow-stress measurements at room temperature, becomes problematic in elevated-temperature testing. In addition, the behavior of metals at elevated temperature becomes more varied. While most metals become softer as temperature increases, and often more ductile, it is possible to encounter temperature regions of embrittlement. Also, because of thermomechanical processes such as dynamic recovery and recrystallization that take place at hot working temperatures (see Chapter 3, "Evolution of Microstructure during Hot Working") plastic instabilities can develop in the compression test. Finally, metals become much more strain-rate sensitive at elevated temperature (Ref 1). This makes strain-rate control in testing very important.

Three forms of compression tests are discussed in this chapter:

- The cylindrical compression test, or hot upset test
- The hot compression of thin ring specimens
- The hot plane-strain compression test, in which sheet or thin plates are indented with long narrow platens

Cylindrical Compression Test

Testing Apparatus (Ref 2). Uniaxial compression testing for workability analysis is usually carried out in a servohydraulic machine such as that shown schematically in Fig. 1. Because strain-rate effects are large in high-temperature testing, it is important to use an apparatus with high machine stiffness. The furnace that surrounds the specimen is usually of resistance, radiant-heating, or induction type. It is important to protect the dynamic load cell and the hydraulic actuator with cooling blocks.

The top and bottom anvils are usually made of stainless steel for tests below 1000 °C (1832 °F), while nickel-base superalloys or Ti-Zr-Mo (TZM) alloy are used for higher temperatures. The diameter of the anvils should be at least three times that of the specimen. The compression platens are made from tool steel, tungsten carbide, TZM, or ceramic composite, depending on the temperature. Compression platens should be flat and parallel to within 0.0002 (ASTM E209).

Lubrication at the interface between the platen and the specimen is important in achieving uniform deformation. However, at elevated temperature, this can be difficult. Typical lubricants are graphite sheet, water-base graphite, boron nitride solutions, molybdenum disulfide, and low melting glasses. To retain the lubricant at the ends of the specimen, various types of grooved patterns are used. Application of the lubricant to the specimen must be done carefully. Usually this is done in two steps, allowing time for drying between applications.

Effect of Temperature. The effect of temperature on the compressive flow curve for metals is shown in Fig. 2. Curves A and B are for deformation at cold working temperatures, $T < 0.3T_m$, where T_m is melting temperature. Curve A, tested at a low strain rate, shows extensive strain hardening, while curve B, tested at a high strain rate, shows flow softening from a temperature rise due to deformation heating. Curves C and D are illustrative of deformation at hot working temperatures, $T > 0.6T_m$. Curve C is typical of metals that undergo dynamic recovery,

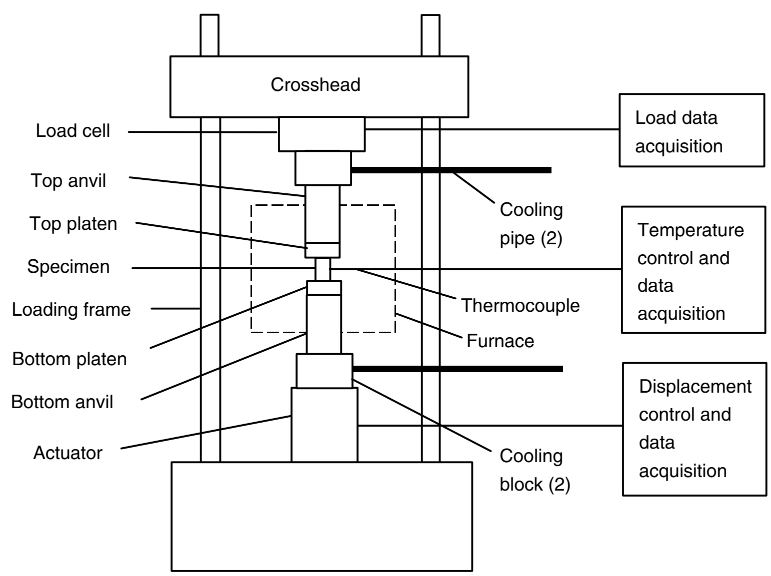


Fig. 1 Compression-testing system. Source: Ref 2

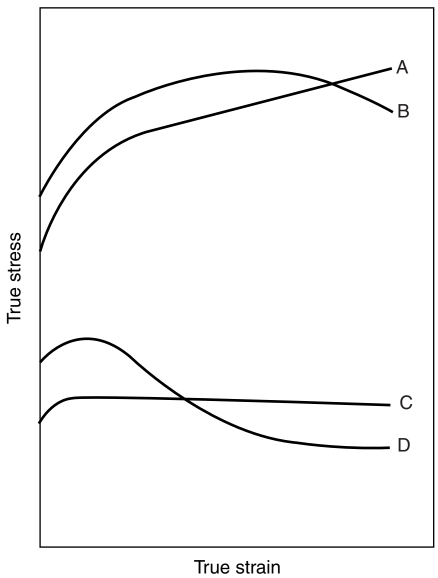


Fig. 2 Typical flow curves for metals deformed at cold working temperatures (A and B) and at hot working temperatures (C and D). Source: Ref 3

while Curve D exhibits flow softening due to deformation heating and/or microstructural instabilities such as the generation of a softer texture during deformation or dynamic spheroidization of second phases.

To accurately measure the temperature of the specimen, a thermocouple should be welded to the surface of the specimen or inserted in a small hole drilled into the specimen (Ref 3). In setting up the tests, it is important to establish the soak time needed for the specimen to come to equilibrium with the test temperature. The use of a heated guard ring box (Fig. 3) minimizes heat loss in inserting the compression specimen into the testing machine. Other ideas for conducting tests isothermally are given in Ref 4 and 5.

Effect of Strain Rate. The extent to which strain rate affects the flow stress of a material is given by its strain-rate sensitivity, m. It is the increase in stress (σ) needed to cause a certain increase in plastic-strain rate ($\dot{\varepsilon}$) at a given level of plastic strain (ε) and a given temperature (T).

$$m = \left(\frac{\Delta \log \sigma}{\Delta \log \dot{\varepsilon}}\right)_{\varepsilon,T} \quad \text{(Eq 1)}$$

Figure 4 shows how the dependence of flow stress on strain-rate sensitivity increases rapidly when the homologous temperature (T/T_m) exceeds a ratio of 0.6. Note that the symbol m is also used below, in accordance with standard convention, to denote the interface friction factor. Usually the context of the situation will help identify which concept is meant by the symbol.

The true strain in a compression test is a function of the height of the specimen:

$$\varepsilon = \ln\frac{h}{h_0} = \ln\left(1 - \frac{\Delta h}{h_0}\right) \quad \text{(Eq 2)}$$

In Eq 2, h and h_0 are the instantaneous and initial specimen height, respectively. $\Delta h = h_0 - h$ is the distance the crosshead of the testing machine has traveled since the compression test started. The negative sign obtained when evaluating Eq 2 denotes that this is a compressive strain.

The true strain rate in the compression test is given by:

$$\dot{\varepsilon} = \frac{d\varepsilon}{dt} = \frac{d(\ln h/h_0)}{dt} = \frac{1}{h}\frac{dh}{dt} = \frac{v}{h} \quad \text{(Eq 3)}$$

where v is the crosshead velocity. Equation 3 shows that for a constant crosshead velocity the true strain rate will increase as the specimen height decreases. Thus, to conduct a constant-strain-rate test the velocity of the moving platen should decrease continuously, while maintaining a constant (v/h) ratio. To see how this is done, Eq 3 is written as:

$$\varepsilon = \dot{\varepsilon}t \quad \text{(Eq 4)}$$

Combining this with Eq 2 gives:

$$\Delta h = h_0[1 - \exp(\dot{\varepsilon}t)] \quad \text{(Eq 5)}$$

Equation 5 is the basic equation for machine control for constant strain rate in compression testing since it expresses the position of the platen in terms of strain rate. The distance of crosshead travel is controlled by the voltage input as a function of time. Constant strain rate up to about 100 s^{-1} can be achieved (Ref 6, 7).

Effect of Deformation Heating. Only about 3 to 5% of the energy of plastic deformation stays in the material as stored energy. Thus, most of the energy that is required to cause plastic deformation appears as deformation heating that is available to raise the temperature of the test specimen. At slow rates of deformation, most of this heat escapes through the anvils and to the environment, but at high strain rates there is not time for much heat flow and the specimen temperature rises. Since one aims to measure flow curves at constant temperature and high strain rate, it is necessary to make a correction for deformation heating.

An expression for the temperature increase due to deformation heating, including an estimate of the heat generated by friction at the platen/specimen interface, was developed by Zhao (Ref 8):

$$\Delta T = \frac{h_a\sigma_a}{h_T}\left(\alpha\dot{\varepsilon} + \frac{mv_a}{h_a}\right)\left[1 - \exp\left(-\frac{h_T t}{c\rho h_a}\right)\right] \quad \text{(Eq 6)}$$

where σ_a is the measured average stress, h_a is the average height of the specimen, h_T is the heat-transfer coefficient between the platen and the specimen, m is the friction factor between the platen and the specimen, t is the time, and v_a is the average velocity at the platen/specimen interface. Also, $\alpha \approx 0.95$ is the fraction of deformation energy converted to heat, c is the material specific heat, and ρ is its density.

Knowing the temperature rise due to deformation heating, one can estimate the correction that must be added to σ_a. First make a plot of σ_a versus the corrected temperature ($T_{test} - \Delta T$) (Fig. 5). The slope of this plot gives $d\sigma/dT$ to be used in Eq 7 to calculate the stress increase:

$$\Delta\sigma = \Delta T \frac{d\sigma}{dT}\bigg|_{\varepsilon,\dot{\varepsilon}} \quad \text{(Eq 7)}$$

Finally, the isothermal stress-strain curve is obtained by adding the stress increase from Eq 7 to the average true stress σ_a. Figure 6 shows the

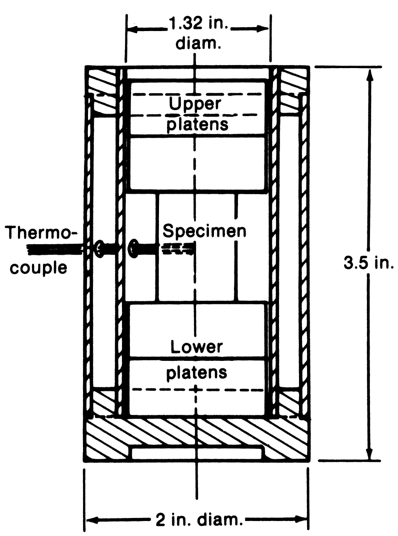


Fig. 3 Guard-ring box, with specimen in position. 1 in. = 25.4 mm. Source: Ref 4

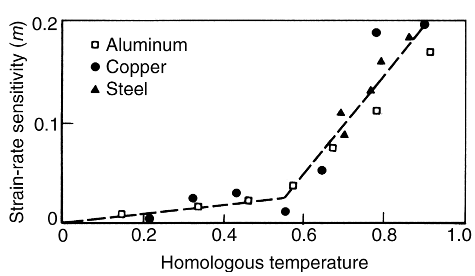


Fig. 4 Dependence of strain-rate effect on homologous temperature, T/T_m for specimens strained to 40% reduction. Source: Ref 4

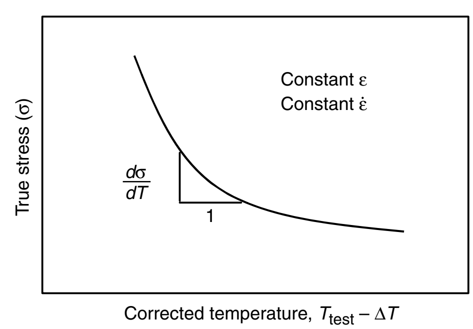


Fig. 5 The determination of derivative of stress in terms of temperature. Source: Ref 2

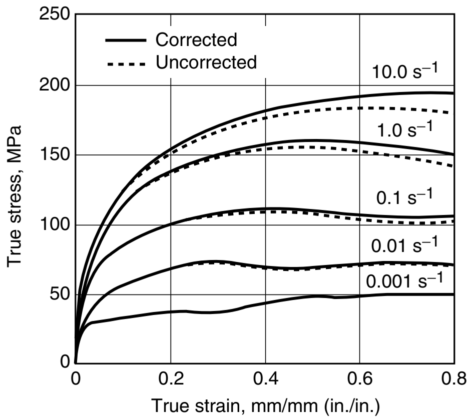


Fig. 6 True-stress/true-strain curves in compression, before and after temperature correction. Tests are for HY-100 steel at 1000 °C (1832 °F) and various strain rates. Note effect of strain rate on both level of flow curves and the amount of temperature correction. Source: Ref 9

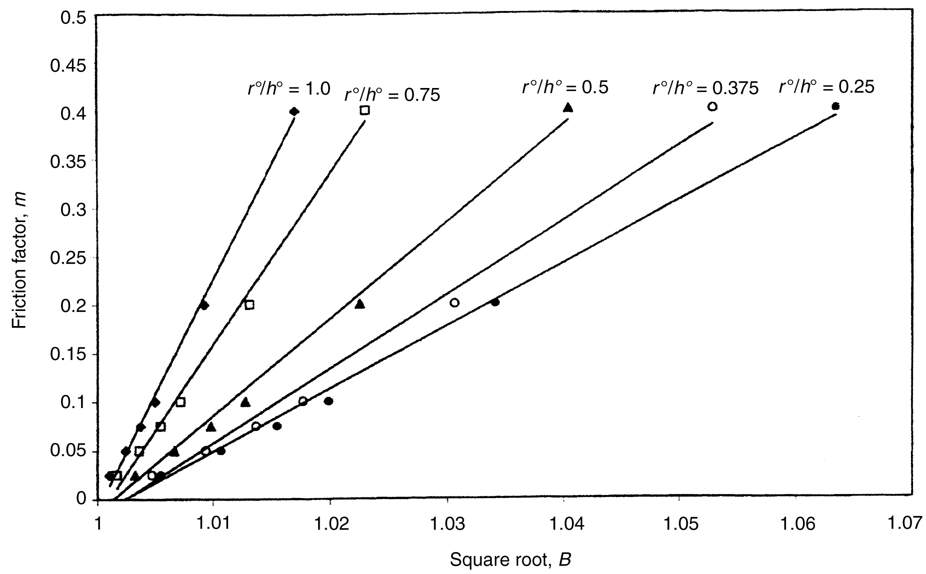


Fig. 7 Relationship between friction factor m and barreling coefficient B, for various values of specimen aspect ratio r/h. Data are for $T/T_m = 0.8$ and $\dot{\varepsilon} = 0.01$ s^{-1}. Source: Ref 11

compression stress-strain curves for a high-strength structural steel tested at 1000 °C (1832 °F). The corrected curves used the method described previously. Note how the influence of deformation heating on the flow curves is most marked at higher strain rates.

Determination of the Flow Curve. The basic data obtained from the compression test are the load and displacement (stroke). The true stress is the load P divided by the instantaneous cross-sectional area with diameter (D). If the compression deformation is homogeneous, then this is given by:

$$\sigma_a = \sigma_0 = \frac{4P}{\pi D^2} = \frac{4Ph}{\pi D_0^2 h_0} \quad \text{(Eq 8)}$$

where σ_a is the average true axial stress, which is equal to the flow stress σ_0 or effective stress $\bar{\sigma}$ for homogeneous deformation. When the friction at the specimen/platen interface is described by the interface friction factor $m = \sqrt{3}\,\tau_i/\sigma_0$, both the slab method of plasticity analysis and the upper-bound analysis (Ref 10) give the average axial flow stress as:

$$\sigma_a = \frac{4P}{\pi D^2} = \sigma_0\left(1 + \frac{mD}{3\sqrt{3}h}\right) \quad \text{(Eq 9)}$$

Equation 9 shows how the measured flow stress is in excess of the effective or flow stress when friction, and barreling, occur in the test.

Extensive finite-element modeling of the hot-compression test using realistic parameters for hot-worked alloys showed that an observed barreling coefficient can be used to determine the friction factor for use in Eq 9 (Ref 11). The barreling coefficient is defined as:

$$B = \frac{h(r_{max})^2}{h_0(r_0)^2} \quad \text{(Eq 10)}$$

where r_{max} is the maximum diameter of a specimen deformed to height h. The interface friction factor is linearly related to the square root of B (Fig. 7). This figure shows that barreling is strongly dependent on the slenderness ratio, r/h, even for low values of interface friction. The data in Fig. 7 can be used with Eq 9 to provide a correction to the measured axial stress.

Evans and Scharning (Ref 11) studied the systematic errors in flow-stress determination in the hot-compression test due to frictional forces at the interfaces and deformation heating during straining. They studied more than 3000 finite-element analyses using the six variables D/h, specimen volume, friction factor, homologous temperature, strain rate, and strain. As might be expected, the most important variables are friction and specimen geometry. Strain and strain rate are of intermediate importance, with specimen volume and temperature least important. A general interpolation function was determined so that the relative errors in stress can be calculated for any of the values in the experimental conditions. It is suggested that, with further work for validation, these equations could be used to correct the measured σ_a to the value of flow stress σ_0.

Discussion so far about determining the stress-strain curve has focused on the determination of stress. The true strain in the hot-compression test is found from Eq 2, making sure that correction is made for the elastic deflection of the testing machine (Ref 12). Since barreling leads to nonuniform deformation, this raises the question of whether Eq 2 is suitable to express the effective strain. Fortunately, it has been shown that the mean effective strain for a barreled specimen is the same, to a first approximation, as in the case of a compression specimen that did not barrel (Ref 13). This result provides the theoretical justification for using the axial compressive strain as the effective strain in constructing the stress-strain curve.

Plastic Instability in Compression. In tension testing the onset of necking indicates unstable flow, characterized by a rapid decrease in diameter localized to the neck region. It is this plastic instability that limits the use of the tension test for determining the flow stress at strains much greater than about $\varepsilon = 0.5$. In the tension test, necking occurs when the rate of strain hardening is no longer able to compensate for the decrease in cross-sectional area of the specimen.

In compression testing, a similar behavior occurs when strain softening is prevalent. Under normal conditions during compression, the cross-sectional area of the specimen increases, which increases the load-carrying capacity of the specimen. However, when strain softening occurs its load-carrying capability is decreased. When the rate of decrease in the strength of the material due to strain softening exceeds the rate of increase in the area of the specimen, an unstable mode of deformation occurs in which the specimen "kinks" as it is compressed (Fig. 8).

Instability in tension and compression can be described with the aid of the Considére construction (Fig. 9). Instability occurs when the slope of the load-elongation (P-e) curve becomes zero:

$$dP = d(\sigma A) = \sigma dA + A d\sigma = 0$$

or

$$d\sigma/\sigma = -dA/A = d\varepsilon = de(1 + e)$$

from the relationship between true strain, ε, and engineering strain, e, $\varepsilon = \ln(1 + e)$. Thus:

$$d\sigma/de = \sigma/(1 + e) \quad \text{(Eq 11)}$$

at the point of instability. Note that Fig. 9 has axes of true stress versus engineering strain. Instability

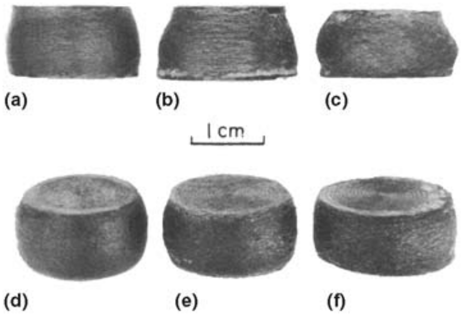


Fig. 8 Hot-compression test specimens of titanium alloy Ti-10V-2Fe-3Al. Specimens (a), (b), and (c) were tested at 704 °C and (d), (e), and (f) at 816 °C. Test strain rates were 10^{-3} s^{-1} (a) and (d), 10^{-1} s^{-1} (b) and (e), and 10 s^{-1} (c) and (f). All specimens had an equiaxed-β microstructure before testing. When deformed at 704 °C (top row) the values of α_c were 2.5, 5, 5 (left to right). At 816 °C, the value of α_c was less than zero at all strain rates. Note the correlation with nonuniform flow and flow localization and α_c. Source: Ref 14

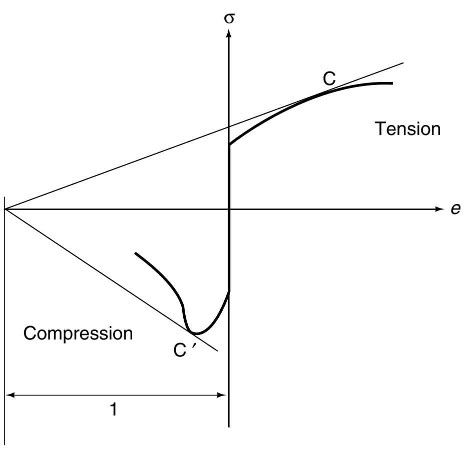


Fig. 9 Considére's construction showing point of instability in tension testing (due to decreasing strain-hardening rate) and in compression testing (due to strain softening)

occurs when the slope of the curve equals the ratio of the true stress to $(1 + e)$. For the tension curve, necking occurs when the stress-strain curve reaches point C. This defines the ultimate strength of the material in tension. For the compression curve, Fig. 9 shows that for a strain-softening material unstable flow occurs at point C′. The ultimate strength of the material for this case of work softening is defined by the stress at point C′.

Another manifestation of strain softening is the formation of internal shear bands. Figure 10 shows schematically how these form upon increasing compressive deformation. See Chapter 13, "Workability in Forging," Fig. 20 and 22 for actual microstructures. The two parameters that predict the propensity for shear banding are the normalized flow-softening rate, γ, and the strain-rate sensitivity index, m.

$$\gamma = \frac{1}{\sigma}\frac{d\sigma}{d\varepsilon} \qquad \text{(Eq 12)}$$

where $\sigma(\varepsilon)$ is the flow curve determined under constant $\dot{\varepsilon}$.

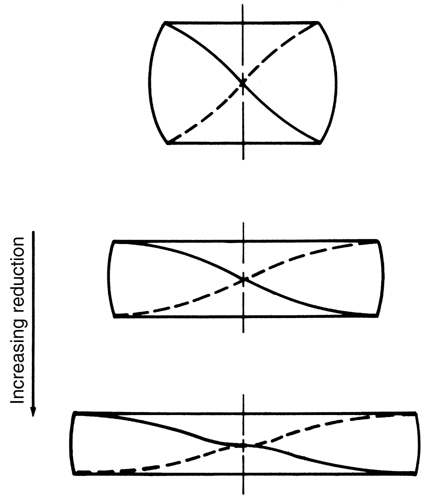


Fig. 10 Formation of shear bands in compression

$$m = \frac{d\ln\sigma}{d\ln\dot{\varepsilon}} = \frac{\dot{\varepsilon}}{\sigma}\frac{d\sigma}{d\dot{\varepsilon}} \qquad \text{(Eq 13)}$$

determined at fixed ε and T. Nonuniform plastic deformation in compression occurs at strain, strain rate, and temperature conditions when the parameter $\alpha_c \geq 5$ (Ref 15, 16).

$$\alpha_c = \frac{\gamma - 1}{m} \qquad \text{(Eq 14)}$$

The conditions that promote flow softening include dynamic recovery and dynamic recrystallization during hot working. Another source is the rapid spheroidization of pearlite and other lamellar microstructures and the coarsening of precipitates during dynamic hot deformation. Much more detail on flow-softening mechanisms can be found in Chapter 3, "Evolution of Microstructure during Hot Working."

Ring Compression Test

Background. The ring compression test was developed to provide a measurement of the interface friction between the specimen and the platens, but it also can be used to provide reasonable values for the flow stress in compression. When a flat, ring-shaped specimen is compressed in the axial direction, the change in dimensions depends on the amount of compression in the thickness direction and the frictional conditions at the platen/ring interfaces. If the interfacial friction were zero, the ring would deform in the same way as a solid disk, with each element flowing radially outward at a rate proportional to its distance from the center; that is, the internal diameter would increase. In the case of small but finite interfacial friction, the outside diameter of the ring is smaller than for the zero friction case. If friction exceeds a critical value, it is energetically favorable for only part of the ring to flow outward and for the remainder of the ring to flow to the center, decreasing the inside diameter. Thus, measurements of the inside diameter of compressed rings provide a particularly sensitive means of studying interfacial friction, because the inside diameter increases if the friction is low and decreases if the friction is larger, as shown in Fig. 11.

The ring compression test is particularly attractive for the measurement of friction in metalworking in that no direct measurement of force is required and no flow stress values for the deforming material are needed.

Analytical Basis. Male and Cockcroft (Ref 17, 18) showed the potential of the ring compression test, and Avitzur provided an analysis for the problem of the axial compression of flat, ring-shape specimens between flat dies (Ref 19, 20). The upper-bound solution developed by Avitzur is quite complex and is usually presented as curves of percent change in inside diameter versus percent reduction in the thickness of the ring (Fig. 12). Different calibration curves are found for ring specimens of different ratios of outside diameter to inside diameter to thickness. The most-used configurations are when these dimensions are in the ratio 6 to 3 to 2 or 6 to 3 to 1. An example of the use of the test is given in Ref 21.

From the Avitzur analysis it is possible to calculate values of p_a/σ_0 at the instant when deformation stops in terms of the ring geometry and the interfacial friction factor, m. In these equations neither the basic yield stress σ_0 nor the interfacial shear stress, τ_i, appears as independent value, but only as the ratio m:

$$m = \sqrt{3}\mu = \sqrt{3}\,\frac{\tau_i}{p_a} \approx \sqrt{3}\,\frac{\tau_i}{\sigma_0} \qquad \text{(Eq 15)}$$

where m is the friction factor, μ is the Coulomb coefficient of friction, τ_i is the interfacial shear stress, p_a is the average axial pressure, and σ_0 is the axial flow stress. The assumption in this analysis is that this ratio remains constant for the

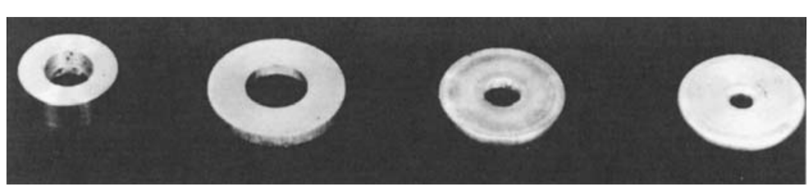


Fig. 11 Variation in shape of ring test specimens deformed the same amount under different frictional conditions. Left to right: undeformed specimen; deformed 50%, low friction; deformed 50%, medium friction; deformed 50%, high friction

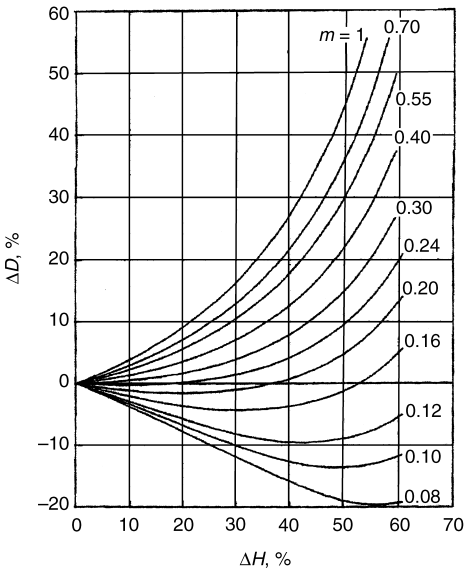


Fig. 12 A typical calibration curve for the ring compression test. Change in internal ring diameter versus change in specimen height, for a 6:3:2 ring. Source: Ref 21

given material and deformation conditions. If the analysis is carried out for a small increment of deformation, σ_0 and τ_i can be assumed to be approximately constant for this increment, and the solution is valid. On the assumption that the friction factor is constant for the entire deformation, it is justifiable to continue the analysis in a series of small deformation increments using the final ring geometry from one increment as the initial geometry for the subsequent increment. Moreover, using this procedure strain hardening can be accommodated, although a high strain-rate sensitivity may not be (Ref 22).

The analysis of Avitzur was modified to develop a reliable method for treating results from the ring compression test with bulge formation and varying friction factor (Ref 23). It was shown how the ring compression test can be used for valid measurement of flow stress. A ring specimen is deformed to a fixed h under the required conditions of temperature and strain rate, and the deformation load is recorded. The change in shape of the ring is measured at room temperature and the value of the ratio p_a/σ_0 is obtained from the computer solutions to Avitzur's equations. Reference 23 gives the equations, which are detailed and require a computer program for explicit solution of the ratio p_a/σ_0. Measurement of the area of the ring surface formerly in contact with the platens, together with knowledge of the deformation load, allows calculation of p_a and σ_0 is obtained. True strain is determined from Eq 2. Repeating this process with other ring specimens over a range of deformations allows the generation of a complete stress-strain curve for the material under the particular conditions of temperature and strain rate. Calibration curves similar to Fig.12 have also been established by finite-element modeling of hot ring compression tests (Ref 24).

Detailed comparisons of the results of many investigators who have used the hot ring compression test are given in Ref 25.

Plane-Strain Compression Test

Testing Conditions. The fundamental difficulties associated with high friction at the platen/specimen interface and barreling of the free surface can be minimized to a large degree in the plane-strain compression test (Ref 26). Moreover, as Fig. 13 shows, this test is ideally suited for testing sheet or thin plate. In the plane-strain compression test, a metal sheet is compressed across the width of the sheet by narrow platens that are wider than the strip. The elastic constraints of the undeformed shoulders of the material on each side of the platens prevent extension of the strip in its width dimension, hence the name plane strain. There is deformation in the direction of the platen movement and in the direction normal to the length of the platen.

To ensure that plane-strain deformation is achieved, the width of the strip should be at least 6 to 10 times the breadth of the platens. To ensure that deformation under the platens is essentially homogeneous, the ratio of platen width to strip thickness (w/h) at any instant in the test should be between 2 and 4. When the plane-strain compression test is done at room temperature, it is often carried out incrementally so as to provide replenishment of lubrication and to change the platens so as to maintain the proper w/h ratio. In this way it is possible to achieve true strains of around 2.

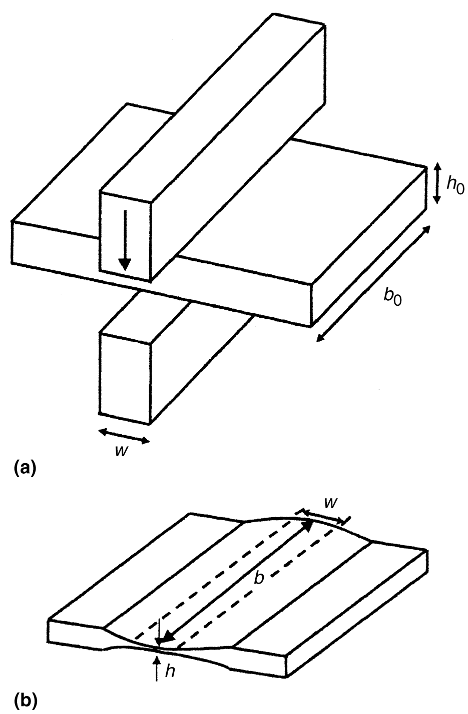


Fig. 13 Plane-strain compression (PSC) test. (a) Dimensions before deformation. (b) Dimensions after deformation. Note bulge that is prominent in the hot PSC test. Source: Ref 27

The true stress and true strain are determined from the plane-strain compression test by:

$$\sigma_1 = p_a = \frac{P}{wb}$$

and

$$\varepsilon_1 = \varepsilon_{pc} = \ln\frac{h_0}{h}$$

where the dimensions are indicated in Fig. 13 and the one-direction is taken parallel to the platen motion. However, because the stress and strain state in the test is that of plane strain, the mean pressure on the platens, p_a, is 15.5% higher than it would be in a uniaxial compression test. To use these values to construct a uniaxial compression flow curve ($\bar{\sigma}$ versus $\bar{\varepsilon}$), they must be converted to effective stress and strain (see Chapter 2, "Bulk Workability of Metals").

$$\bar{\sigma} = \sigma_0 = \frac{\sqrt{3}}{2} p_a = \frac{p_a}{1.155} \qquad \text{(Eq 16)}$$

and

$$\bar{\varepsilon} = \frac{2}{\sqrt{3}} \varepsilon_{pc} = 1.155\varepsilon_{pc} \qquad \text{(Eq 17)}$$

Note that Eq 16 assumes that friction between the platen and the strip is negligible. This can be achieved by incrementally applying lubrication during the test or by using a very low friction material such as polytetrafluoroethylene tape on the faces of the platens. Friction becomes a much more serious consideration in the hot plane-strain compression test, discussed next.

Hot Plane-Strain Compression Test. The plane-strain compression (PSC) test is finding growing use for making reliable and reproducible measurements of flow curves at elevated temperature. It is interesting to note that steps are underway to develop a good practice guide for the test through the offices of the National Physical Laboratory (United Kingdom). A nice feature of the test is that since a reasonably large specimen can be tested, it provides a good opportunity for studying microstructure development. Incremental tests are difficult to do in a hot PSC test, and since lubrication also is more difficult at elevated temperature, it is not surprising that load-displacement data for hot PSC tests require corrections for the friction between the platens and the specimen, and for lateral spread (Ref 27). Figure 13(b) defines the dimensions of interest. The dimension of concern is the breadth (width) b. Lateral spread is in the width direction and is evidenced by a bulge at the free surfaces. Thus, the stress state is not truly plane strain, and a correction needs to be made to the measured p_a.

The development of lateral spread means that the PSC test is not being carried out under truly plane-strain conditions. Therefore, corrections for spread need to be applied to p_a and ε_{pc} to give values close to the true axial values. Detailed finite-element analysis of the hot PSC test was carried out to find the effects of the test variables on spread (Ref 27). A thermomechanically coupled model that allowed for deformation heating and modeled the interfacial friction using Coulomb friction was used. The chief findings were:

- For $b_0/w \leq 5$ the spread increases almost exponentially with decreasing instantaneous thickness. However, raising initial width to large values incurs a strong penalty on the load required for deforming the specimen.
- In addition, lateral spread decreases with increasing friction coefficient and decreasing initial specimen thickness.
- There is a strong dependence of side curvature on friction coefficient.
- The spread is not sensitive to the strain rate or to the material.

An empirical relationship was established to evaluate instantaneous breadth b as a function of instantaneous thickness, h.

$$b = b_0\left[1 + C - C\left(\frac{h}{h_0}\right)^s\right] \quad \text{(Eq 18)}$$

where s is the spread exponent, with a typical value of 0.2, and C is defined as the spread coefficient determined from the final breadth b_f and thickness h_f of a specimen, according to:

$$C = \frac{b_f / b_0 - 1}{1 - (h_f / h_0)^s} \quad \text{(Eq 19)}$$

The flow stress in plane strain is by convention denoted as $2k$, where k is the flow stress in pure shear (torsion). The average pressure on the platens, p_a, will be greater than $2k$, depending on friction conditions. The instantaneous value of average pressure is given by:

$$p_a = P/wb \quad \text{(Eq 20)}$$

where P is the load on the platens and b is given by Eq 18. There are three types of friction conditions that need to be considered, sliding friction, sticking friction, and a combination of the two.

When the Coulomb friction is low, the average pressure relative to the plane-strain flow stress (Ref 28) is given by:

$$\frac{p_a}{2k} = \frac{1}{bw}\left[\frac{2h^2}{\mu^2} + \frac{(b-w)h}{\mu}\right]\left[\exp\left(\frac{\mu w}{h}\right) - 1\right] - \frac{2h}{\mu b} \quad \text{(Eq 21)}$$

where μ is the friction coefficient. For sticking friction, relative motion between the metal and the platens occurs by shearing of the metal at the interface, rather than sliding. Under these conditions, $p_a/2k$ is given by:

$$\frac{p_a}{2k} = 1 + \frac{w}{4h} - \frac{w^2}{12hb} \quad \text{(Eq 22)}$$

Equations 21 and 22 are only valid when sliding or sticking conditions apply over the entire face of the platens. However, it is very likely to have an intermediate situation where sticking exists in the central region of the platens and sliding friction conditions exist at the outer edges. The location of this transition, measured from the centerline of the platen, is given by z_o:

$$z_0 = \left(\frac{h}{2\mu}\right)\ln\left(\frac{1}{2\mu}\right) \quad \text{(Eq 23)}$$

The 12-term equation for this situation is not given here. It should be noted that Eq 21 and 22 contain extra terms over the usual equations for yielding in plane-strain compression because they have been corrected with the Orowan and Pascoe equations (Ref 29) that account for spread in rolling when $b < 6h$. This correction cannot be ignored for the hot PSC test because it is common to use specimens with $b = 5h$ (Ref 28).

The equations that relate frictional conditions to the flow stress, although complex, can be handled rather simply with modern computers. The value of z_0 is computed for each increment of deformation by using the instantaneous thickness h. This determines which of the friction equations prevail, and this allows a calculation of $p_a/2k$. With p_a known from Eq 20, the flow stress in plane strain $2k$ is readily determined. The final step is to convert the plane-strain flow stress into the equivalent axial flow stress:

$$\bar{\sigma} = \sigma_0 = 2k\frac{\sqrt{3}}{2}\frac{1}{f} \quad \text{(Eq 24)}$$

The term f is a correction to allow for the fact that because of bulging the stress state in the hot PSC test is not truly plane strain. This is discussed below.

The displacement measurements Δh from which strain is calculated must be corrected for errors in the zero position of displacement. These occur due to slight misalignment of the tool and/or specimen faces not being perfectly parallel (Ref 28). Also, zero displacement error can occur when a thick glass lubricant film is squeezed out from the platens during initial stages of deformation.

Finally, the axial strain measurement should be corrected with the f factor to determine the nominal equivalent strain. To convert the thickness strain by the factor $2/\sqrt{3}$, as in Eq 17, would be an overestimate because only part of the metal under the platen is in plane strain. This correction is based on the concept that at the ends of the platens at the free surface the material is almost free to move in both the length and width directions. Therefore, the deformation in these narrow regions is more like axisymmetric compression. Thus, a plane-strain condition exists only in the central (b-w) region of the specimen (Ref 30).

$$f = \frac{\frac{2}{\sqrt{3}}(b-w) + w}{b} \quad \text{(Eq 25)}$$

where b is the instantaneous breadth obtained from Eq 18. The width of the platen, w, does not change appreciably with deformation.

$$\bar{\varepsilon} = f\left(\frac{2}{\sqrt{3}}\ln h_0 / h\right) \quad \text{(Eq 26)}$$

Thus, the flow stress determined with the hot plane-strain compression test is obtained by plotting $\bar{\sigma}$ from Eq 24 versus $\bar{\varepsilon}$ obtained from Eq 26. When the corrections discussed previously are made to the data, tests between different laboratories agree to $\pm 5\%$ (Ref 30).

The plane-strain compression test is useful for measuring the workability of metals that are sensitive to shear. The deformation between the platens includes crossed shear bands from each corner of the platens. This intense shear can cause unstable deformation of hot deformed material in much the same way as it occurs in the cylindrical compression tests referred to previously.

Conclusions

The compression test is the most general-purpose test for workability. Carrying out the test at elevated temperature imposes issues with friction and temperature and strain-rate control that must be addressed with proper test equipment and experimental planning. In addition, the experimenter must be cognizant of the forms of plastic instability that can arise in hot-compression tests. The ring compression test provides for the best routine measurement of friction in hot deformation processing. However, it is an average value and cannot detect variations in friction over an area. Hot plane-strain compression offers distinct advantages when the objective is to provide flow curves for modeling plane-strain deformation processes such as rolling (Ref 31).

ACKNOWLEDGMENT

The material in this article is largely taken from:

- H.A. Kuhn, Uniaxial Compression Testing, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 143–151
- A.T. Male and G.E. Dieter, Hot Compression Testing, *Workability Test Techniques*, American Society for Metals, 1984, p 50–72
- D. Zhao, Testing for Deformation Modeling, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 798–810

REFERENCES

1. S.L. Semiatin and J. Jonas, *Formability and Workability of Metals*, American Society for Metals, 1984
2. D. Zhao, Testing for Deformation Modeling, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 798–810
3. D. Zhao and S. Lampman, Hot Tension and Compression Testing, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 152–163
4. J.F. Alder and V.A. Phillips, *J. Inst. Met.*, Vol 83, 1954–1955, p 80
5. F.J. Gurney and D.J. Abson, *Met. Mater.*, Vol 7, 1973, p 535
6. G. Fitzsimmons, H.A. Kuhn, and R. Venkateshwar, *J. Met.*, May 1981, p 11
7. High Strain Rate Tension and Compression Tests, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 429–446
8. D. Zhao, Temperature Correction in Compression Tests, *J. Mater. Process. Technol.*, Vol 36, 1993, p 467–471
9. M. Thirukkonda, D. Zhao, and A.T. Male, "Materials Modeling Effort for HY-100 Steel," Technical Report TR 96-027, NCEMT, National Center for Excellence in Metalworking Technologies, Johnstown, PA, March 1996
10. E.M. Mielnik, *Metalworking Science and Engineering*, McGraw-Hill, 1991, p 241–243
11. R.W. Evans and P.J. Scharning, Axisymmetric Compression Test and Hot Working Properties of Alloys, *Mater. Sci. Technol.*, Vol 17, 2001, p 995–1004
12. J.W. House and P.P. Gillis, Testing Machines and Strain Sensors, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 79–92
13. F.-K. Chen and C.-J. Chen, On the Nonuniform Deformation of the Cylinder Compression Test, *J. Eng. Mater. Technol. (Trans. ASME)*, Vol 122, 2000, p 192–197
14. S.L. Semiatin, Workability in Forging, Chapter 13, *Workability and Process Design*, ASM International, 2003
15. J.J. Jonas, R.A. Holt, and C.E. Coleman, Plastic Stability in Tension and Compression, *Acta Metall.*, Vol 24, 1976, p 911
16. S.L. Semiatin and G.D. Lahoti, The Occurrence of Shear Bands in Isothermal Hot Forging, *Metall. Trans. A*, Vol 13A, 1982, p 275
17. A.T. Male and M.G. Cockcroft, A Method for the Determination of the Coefficient of Friction of Metals under Conditions of Bulk Plastic Deformation, *J. Inst Met.*, Vol 93, 1964–1965, p 38–46
18. A.T. Male, Variations in Friction Coefficients of Metals During Compressive Deformation, *J. Inst. Met.*, Vol 94, 1966, p 121–125
19. B. Avitzur, *Metal Forming Processes and Analysis*, McGraw-Hill, 1968, p 81–93
20. B. Avitzur and C.J. Van Tyne, Ring Forming: An Upper Bound Approach, *J. Eng. Ind. (Trans. ASME)*, Vol 104, 1982, p 231–252
21. F. Wang and J.G. Lenard, An Experimental Study of Interfacial Friction—Hot Ring Compression, *J. Eng. Mater. Technol.*, Vol 114, 1992, p 13–18
22. G. Garmong, N.E. Paton, J.C. Chestnutt, and L.F. Nevarez, *Metall. Trans. A*, Vol 8A, 1977, p 2026
23. V. DePierre and F.J. Gurney, A Method for Determination of Constant and Varying Factors During Ring Compression Tests, *J. Lubr. Technol. (Trans. ASME)*, Vol 96, 1974, p 482–488
24. N.T. Rudkins, P. Hartley, I. Pillinger, and D. Petty, Friction Modelling and Experimental Observations of Hot Ring Compression Tests, *J. Mater. Process. Technol.*, Vol 60, 1995, p 349–353
25. K.P. Rao and K. Sivaram, A Review of Ring-Compression Testing and Applicability of the Calibration Curves, *J. Met. Process. Technol.*, Vol 37, 1993, p 295–318
26. A.B. Watts and H. Ford, An Experimental Investigation of the Yielding of Strip between Smooth Dies, *Proc. Inst. Mech. Eng.*, Vol B1, 1952, p 448–453
27. M.S. Mirza and C.M. Sellars, Modelling the Hot Plane Strain Compression Test-Effect of Friction and Specimen Geometry on Spread, *Mater. Sci. Technol.*, Vol 17, 2001, p 1142–1148
28. N.J. Silk and M.R. van der Winden, Interpretation of Hot Plane Strain Compression Testing of Aluminum Specimens, *Met. Sci. Technol.*, Vol 15, 1999, p 295–300
29. E. Orowan and K.J. Pascoe, Special Report No. 34, Iron and Steel Institute, 1946
30. H. Shi, A.J. McLaren, C.M. Sellars, R. Shahani, and R. Bolingbroke, Hot Plane Strain Compression Testing of Aluminum Alloys, *J. Test. Eval.*, Vol 25 (No. 1), 1997, p 61–73
31. S.B. Davenport, N.J. Silk, C.N. Sparks, and C.M. Sellars, Development of Constitutive Equations for Modelling of Hot Rolling, *Mater. Sci. Technol.*, Vol 16, 2000, p 539–546

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 68-85
DOI:10.1361/hwpd2003p068

Chapter 7

Hot-Tension Testing

P.D. Nicolaou, R.E. Bailey, and S.L. Semiatin

THE DEVELOPMENT of successful manufacturing techniques for metallic materials requires reliable information regarding hot-working characteristics. The proper hot-working temperature and deformation rate must be established to produce high-quality wrought products of complicated geometries. It is also important that product yield losses (from either grinding to remove surface cracks or excessive cropping to remove end splits) be held to a minimum, while avoiding the formation of internal cavities (pores). Severe cracking is ordinarily the result of high *surface* tensile stresses introduced when hot working is conducted either above or below the temperature range of satisfactory ductility. Similarly, cavitation is associated with *internal* tensile stresses, which, for a given material, depend on the temperature, the deformation rate, and workpiece/die geometry.

The first and most important step in specifying appropriate hot-working practice is to determine suitable hot-working conditions. In particular, the tensile ductility (e.g., fracture strain), the flow stress, and cavity formation conditions should be established as a function of temperature and strain rate. A curve of ductility versus temperature or strain rate shows what degree of deformation the material can tolerate without failure. On the other hand, a plot of the flow stress versus temperature, along with workpiece size and strain rate, indicates the force levels required of the hot-working equipment. Last, a curve of cavity volume fraction versus strain, strain rate, and temperature shows what processing parameters should be selected in order to produce high-quality products.

Although commercial metalworking operations cannot be analyzed in terms of a simple stress state, workpiece failures are caused by localized tensile stresses in most instances (Ref 1–4). In rolling of plate, for example, edge cracking is caused by tensile stresses that form at bulged (unrestrained) edges (Ref 1, 3, 4). The geometry of these unrestrained surfaces affects the magnitude of tensile stresses at these locations. Moreover, tensile stresses are also created on the unrestrained surface of a round billet being deformed with open dies. Therefore, to obtain a practical understanding of how well a material will hot work during primary processing, it is essential to know how it will respond to tensile loading at the strain rates to be imposed by the specific hot-working operation.

The ideal hot-workability test is one in which the metal is deformed uniformly, without instability, at constant true strain rate under well-controlled temperature conditions with continuous measurement of stress, strain, and temperature during deformation followed by instantaneous quenching to room temperature. Two types of hot-tension tests are discussed in this chapter: the Gleeble test and the conventional isothermal hot-tension test. The major advantage of the hot-tension test is that its stress/strain state simulates the conditions that promote cracking in most industrial metalworking operations. However, even though the tension test is simple in nature, it may provide misleading information if not properly designed. Specifically, parameters such as the specimen geometry, tension machine characteristics, and strain rate and temperature control all influence the results of the tension test. Therefore, the tension test should be designed and conducted carefully, and testing procedures should be well documented when data are reported.

Equipment and Testing Procedures

The apparatus used to conduct hot-tension tests comprises a mechanical loading system and equipment for sample heating. A variety of equipment types are used for applying forces (loads) to test specimens. These types range from very simple devices to complex systems that are controlled by a digital computer. The most common test configurations utilize universal testing machines, which have the capability to test material in tension, compression, or bending. The word *universal* refers to the variety of stress states that can be applied by the machine, in contrast to other conventional test machines that may be limited to either tensile loading or compressive loading, but not both. Universal test machines or tension-only test frames may apply loads by a gear (screw)-driven mechanism or hydraulic mechanisms, as discussed in more detail in the section "Frame-Furnace Tension-Testing Equipment" in this chapter.

The heating method used for hot-tension testing varies with the application. The most common heating techniques are direct-resistance heating (in the case of Gleeble systems) and indirect-resistance or induction heating with conventional load frames. In some cases, universal test machines may include a special chamber for testing in either vacuum or controlled atmosphere. Specimen (testpiece) temperatures typically are monitored and controlled by thermocouples, which may be attached on the specimen surface or located very close to the specimen. In some cases, temperature is measured by optical or infrared pyrometers. Accurate measurement and control is very critical for obtaining reliable data. To this end, the use of closed-loop temperature controllers is indispensable.

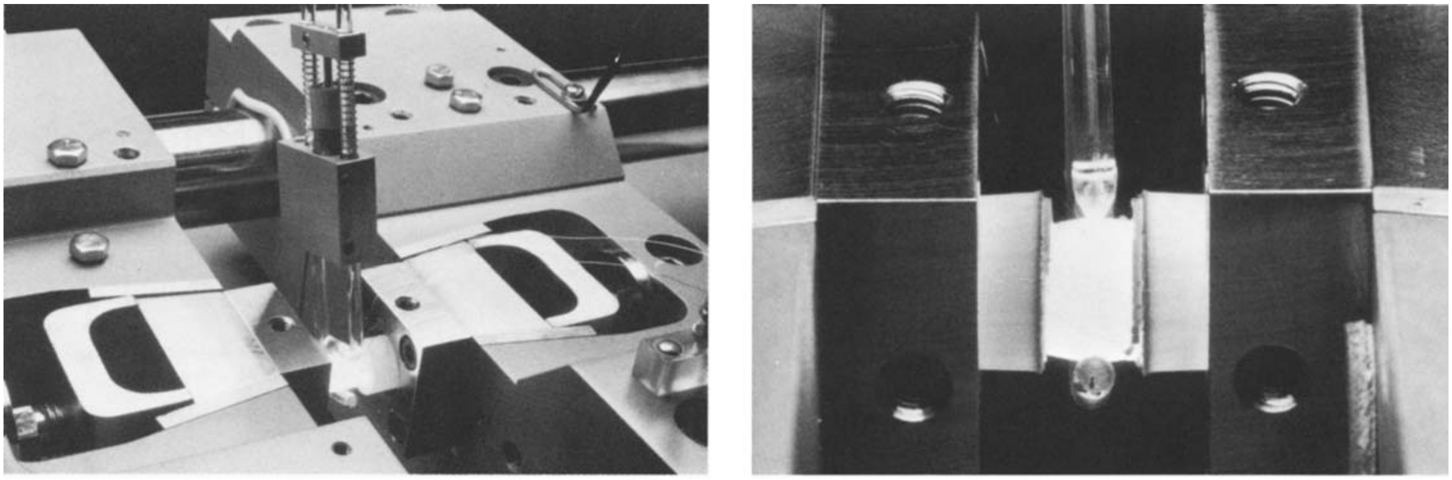
The occurrence of deformation heating may also be an important consideration, especially at high strain rates, because it can significantly raise the specimen temperature.

Gleeble Testing Equipment

The Gleeble system (Ref 5) has been used since the 1950s to investigate the hot-tension behavior of materials and thus to generate important information for the selection of hot-working parameters. A Gleeble unit is a high-strain-rate, high-temperature testing machine where a solid, buttonhead specimen is held horizontally by water-cooled grips, through which electric power is introduced to resistance heat the test specimen (Fig. 1). Specimen temperature is monitored by a thermocouple welded to the specimen surface at the middle of its length. The thermocouple, with a function generator, controls the heat fed into the specimen according to a programmed cycle. Therefore, a specimen can be tested under time and temperature conditions that simulate hot-working sequences.

Contemporary Gleeble systems (e.g., see www.gleeble.com) are fully integrated servohydraulic setups that are capable of applying as much as 90 kN (10 tons) of force in tension at displacement rates up to 2000 mm/s (80 in./s). Different load cells allow static-load measurement to be tailored to the specific application. Control modes that are available include displacement, force, true stress, true strain, engineering stress, and engineering strain.

The direct-resistance heating system of the Gleeble machine can heat specimens at rates of



(a) (b)

Fig. 1 Gleeble test unit used for hot-tension and hot-compression testing. (a) Specimen in grips showing attached thermocouple wires and linear variable differential transformer (LVDT) for measuring strain. (b) Close-up of a test specimen. Courtesy of Duffers Scientific, Inc.

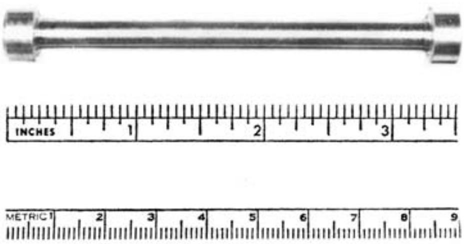


Fig. 2 Typical specimen used for Gleeble testing

up to 10,000 °C/s (18,000 °F/s). Grips with high thermal conductivity (e.g., copper) hold the specimen, thus making the system capable of high cooling rates as well. Thermocouples or pyrometers provide signals for accurate feedback control of specimen temperature. Because of the unique high-speed heating method, Gleeble systems typically can run hot-tension tests several times faster than conventional systems based on indirect-resistance (furnace) heating methods.

A digital-control system provides all the signals necessary to control thermal and mechanical test variables simultaneously through the digital closed-loop thermal and mechanical servo systems. The Gleeble machine can be operated totally by computer, by manual control, or by any combination of computer and manual control needed to provide maximum versatility in materials testing.

Sample Design. A typical specimen configuration used in Gleeble testing is shown in Fig. 2. This solid buttonhead specimen, with an overall length of 88.9 mm (3.5 in.), has an unreduced test-specimen diameter of 6.25 mm (0.25 in.). The length of the sample between the grips at the beginning of the test is also an important consideration. Generally, this length is 25.4 mm (1 in.). Shorter lengths produce a narrow hot zone and restrict hot deformation to a smaller, constrained region; consequently, the apparent reduction of area is diminished. On the other hand, a long sample length generally produces higher apparent ductility/elongation values. For example, Smith, et al. (Ref 6) have shown that a grip separation of 36.8 mm (1.45 in.) produces a hot zone about 12.7 mm (0.5 in.) long. When specimen diameter is increased, as is necessary in testing of extremely coarse-grain materials, the grip separation should also be increased proportionately to maintain a constant ratio of hot-zone length to specimen diameter.

Test Procedures. It is essential that hot-tension tests be conducted at accurately controlled temperatures because of the usually strong dependence of tensile ductility on this process variable. To this end, temperature is monitored by a thermocouple percussion welded to the specimen surface. Using a function generator, heat input to the specimen is controlled according to a predetermined programmed cycle chosen by the investigator. However, the temperature measured from this thermocouple junction does not coincide exactly with the specimen temperature because (a) heat is conducted away from the junction by the thermocouple wires, and (b) the junction resides above the specimen surface and radiates heat at a rate higher than that of the specimen itself. Consequently, the thermocouple junction is slightly colder than the test specimen. Furthermore, specimen temperature is highest midway between the grips and decreases toward the grips. In general, the specimen will fracture in the hottest plane perpendicular to the specimen axis. Therefore, it is important to place the thermocouple junction midway between the grips in order that the hottest zone of the specimen, which will be the zone of fracture, is monitored. The longitudinal thermal gradient does not present a serious problem because the specimen deforms in the localized region where the temperature is monitored. Consequently, the measured values of reduction of area and ultimate tensile strength represent the zone where the thermocouple is attached.

Strain rate is another important variable in the hot-tension test. However, strain rate varies during hot-tension testing under constant-crosshead-speed conditions and must be taken into account when interpreting test data. An analysis of the strain-rate variation during the hot-tension test and how it correlates to the strain rates in actual metalworking operations is presented later in this chapter.

The load may be applied at any desired time in the thermal cycle. Temperature, load, and crosshead displacement are measured versus time and captured by the data acquisition system. From these measurements, standard mechanical properties such as yield and ultimate tensile strength can be determined. The reduction of area at failure is also readily established from tested samples.

If hot-working practices are to be determined for an alloy for which little or no hot-working information is available, the preliminary test procedure usually comprises the measurement of data "on heating." In such tests, samples are heated directly to the test temperature, held for 1 to 10 min, and then pulled to fracture at a strain rate approximating the rate calculated for the metalworking operation of interest. The reduction of area for each specimen is plotted as a function of test temperature; the resulting "on-heating" curve will indicate the most suitable temperature range to be evaluated to determine the optimal preheat* temperature. This temperature, as indicated from the plot in Fig. 3, lies between the peak-ductility (PDT) and zero-ductility (ZDT) temperatures.

To confirm the appropriate selection of preheat temperature, specimen blanks should be heat treated at the proposed preheat temperature

*In the context of this chapter, preheat temperature is the temperature at which the test specimen or workpiece is held prior to deformation at *lower* temperatures. In actual metalworking operations, preheat temperature usually refers to the actual furnace temperature.

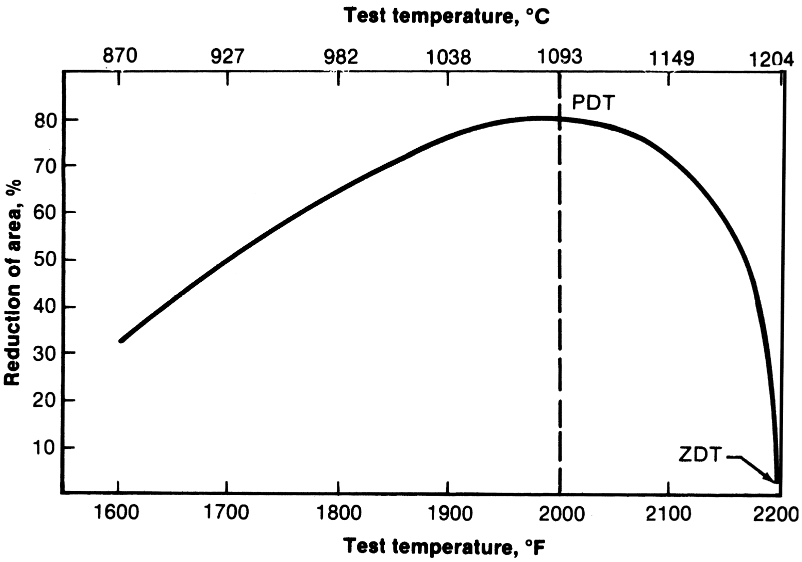


Fig. 3 Hypothetical "on-heating" Gleeble curve of specimen reduction of area as a function of test temperature

for a time period equal to that of a furnace soak commensurate with the intended workpiece size and hot-working operation. These specimens should be water quenched to eliminate any structural changes that could result from slow cooling. Subsequently, the specimens should be tested by heating to the proposed furnace temperature, holding at this preheat temperature for a moderate period of time (1 to 10 min) to redissolve any phases that may have precipitated, cooling to various temperatures at intervals of 25 or 50 °C (45 or 90 °F) below the preheat temperature, holding for a few seconds at the desired test temperature, and finally pulling in tension to fracture at the calculated strain rate. These "on-cooling" data demonstrate how the material will behave after being preheated at a higher temperature. Testing "on cooling" is necessary because the relatively short hold times during testing "on heating" may not develop a grain size representative of that hold temperature and may be insufficient to dissolve or precipitate a phase that may occur during an actual furnace soak prior to hot working. Also, most industrial hot metalworking operations are conducted as workpiece temperature is decreasing. The "on-cooling" data will indicate how closely the ZDT can be approached before hot ductility is seriously or permanently impaired. In addition, if deformation heating (Ref 7) during "onheating" tests has resulted in a marked underestimation of the maximum preheat temperature, this will be revealed and can be rectified by examination of "on-cooling" data.

Frame-Furnace Tension-Testing Equipment

Universal testing machines and tension-test frames can be used for hot-tension tests by attaching a heating system to the machine frame. The frame may impart loading by either a screw-driven mechanism or servohydraulic actuator. Screw-driven (or gear-driven) machines are typically electromechanical devices that use a large actuator screw threaded through a moving crosshead. The screws can turn in either direction, and their rotation moves a crosshead that applies a load to the specimen. A simple balance system is used to measure the magnitude of the force applied.

Loads may also be applied using the pressure of oil pumped into a hydraulic piston. In this case, the oil pressure provides a simple means of measuring the force applied. Closed-loop servohydraulic testing machines form the basis for the most advanced test systems in use today. Integrated electronic circuitry has increased the sophistication of these systems. Also, digital computer control and monitoring of such test systems have steadily developed since their introduction around 1965. Servohydraulic test machines offer a wider range of crosshead speeds of force ranges with the ability to provide economically forces of 4450 kN (10^6 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 become increasingly more difficult to produce as the force rating goes up.

For either a screw-driven or servohydraulic machine, the hot-tension test system is a load frame with a heating system attached. A typical servohydraulic universal testing machine with a high-temperature chamber is shown in Fig. 4. The system is the same as that used at room temperature, except for the high-temperature capabilities, including the furnace, cooling system, grips, and extensometer. In this system, the grips are inside the chamber but partly protected by refractory from heating elements. Heating elements are positioned around a tensile specimen. Thermocouple and extensometer edges touch the specimen. The grip design and the specimen geometry depend on the specific features of the frame and the heating unit as well as the testing conditions. Temperature is measured by thermocouples attached on or located very near to the specimen. In some cases, a pyrometer can also be used.

The most common methods of heating include induction heating and indirect-resistance heating in chamber. Typical examples are shown in Fig. 5. Induction heating (Fig. 5a) usually allows faster heating rates than indirect heating does, but accurate temperature control requires extra care. Induction-heating systems can reach testing temperatures within seconds. Induction heating heats up the outer layer of the specimen first. Furnaces with a lower frequency have better penetration capability. Coupling the heating coil and the specimen also plays an important role in heating efficiency. The interior of the specimen is heated through conduction. With the rapid heating rate, the temperature is often overshot and nonuniform heating often occurs.

Indirect-resistance heating may provide better temperature control/monitoring than induction heating can. Indirect-resistance heating can be combined readily with specially designed chambers for testing either in vacuum or in a controlled atmosphere (e.g., argon, nitrogen, etc.). Vacuum furnaces are expensive and have high maintenance costs. The furnace has to be mounted on the machine permanently, making it inconvenient if another type of heating device is to be used. The heating element is expensive and oxidizes easily. The furnace can only be opened at relatively lower temperatures to avoid oxidation. Quenching has to be performed with an inert gas, such as helium.

Environmental chambers (Fig. 5b), which are less expensive than vacuum chambers, have a circulation system to maintain uniform temperature inside the furnace. Inert gas can flow through the chamber to keep the specimen from oxidizing. Temperature inside the chamber can be kept to close tolerance (e.g., about ±1 °C, or ±2 °F). However, the maximum temperature of an environmental chamber is usually 550 °C (1000 °F), while that of a vacuum furnace can be as high as 2500 °C (4500 °F). The chamber can either be mounted on the machine or rolled in and out on a cart. Split-furnace designs (Fig. 5c) are also cost effective and easy to use. When not in use, it can be swung to the side. The split furnace shown in Fig. 5(c) has only one heating zone. More sophisticated split furnaces have three heating zones for better temperature control. Heating rate is also programmable. When furnace heating is used, it is a common practice to use a low heating rate. In addition, the specimen is typically "soaked" at the test temperature for about 10 to 30 min prior to the application of the load.

The mechanical and thermal control systems are similar to those described in the previous section on the Gleeble testing apparatus. The main advantage of hot-tension test machines is that the test specimen is heated uniformly along its entire gage length, and hence other useful materials properties such as total tensile elonga-

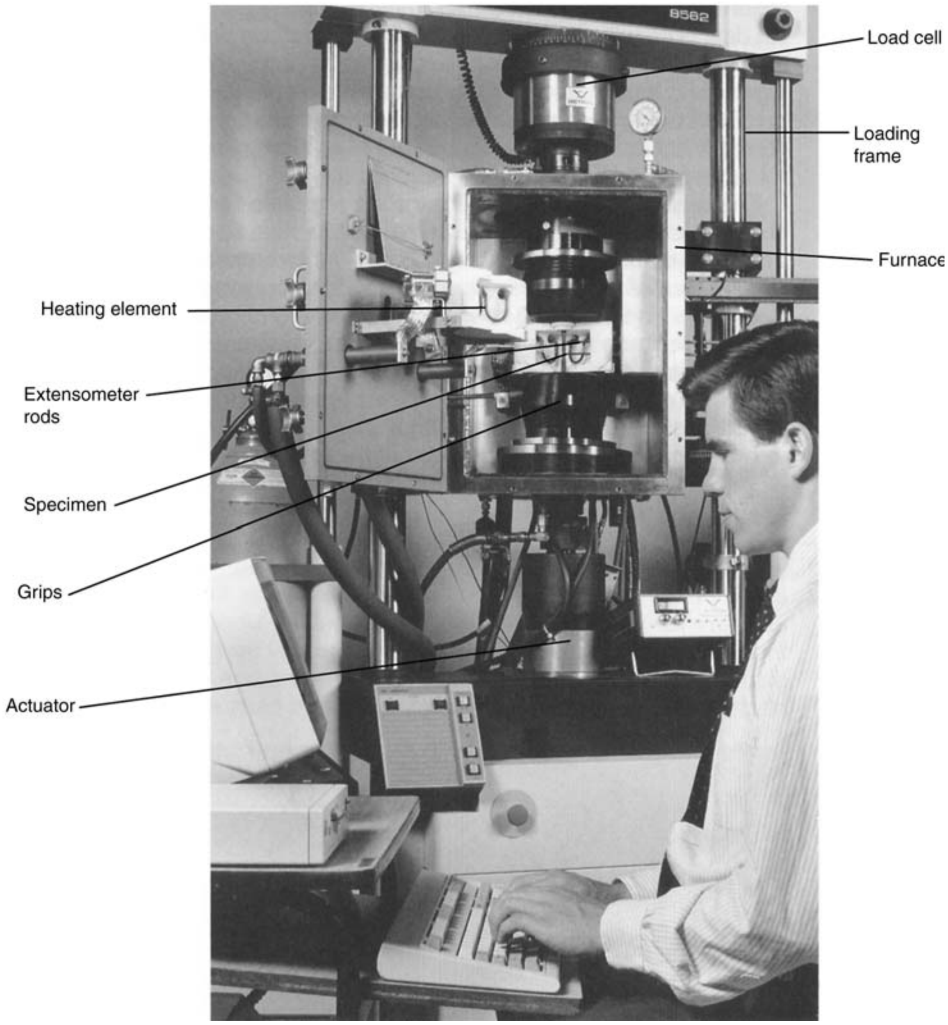


Fig. 4 Typical servohydraulic universal testing machine with a chamber and instrumentation for high-temperature testing

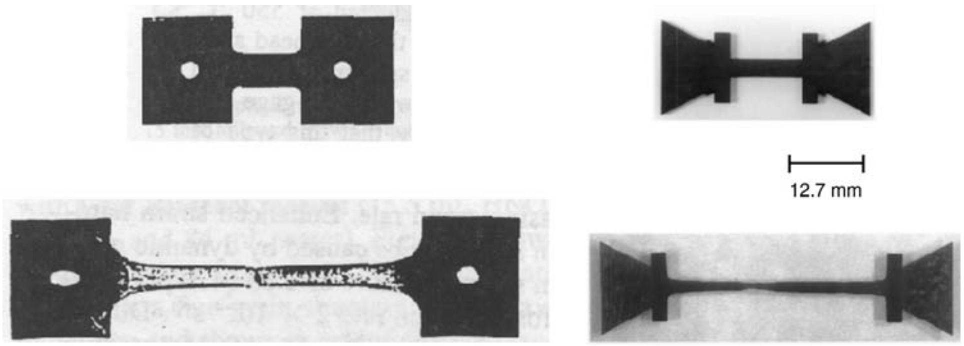
tion, plastic anisotropy parameter, cavity formation, and so forth can be determined in addition to yield/ultimate tensile strength and reduction of area. On the other hand, the overall time needed to conduct a single test may be longer than in the Gleeble test method.

Specimen Geometry. In tension testing, elongation values are influenced by gage length (see Chapter 2 "Bulk Workability of Metals"). It is thus necessary to state the gage length over which elongation values are measured. When the ductilities of different materials (or of a single material tested under different conditions) are compared in terms of total elongation, the specimen gage length also should be adjusted in proportion to the cross-sectional area. This is of great importance in the case of small elongations, because the neck strain contributes a significant portion to the total strain. On the other hand, the neck strain represents only a small portion of the elongation in the case of superplastic deformation. Unwin's equation (Ref 8) also shows the rationale for a fixed ratio of gage length with cross-sectional area, expressed as a fixed ratio of gage length to diameter (for round bars) or gage length to the square root of the cross-sectional area (sheet specimens). This reinforces the importance of stating the gage length used in measuring elongation values. Usually, the length-to-diameter ratio is between 4 and 6.

In addition to the gage length, the specimen shoulder geometry and hence the gripping system are also important design considerations. It is desirable that specimen deformation takes place only within its gage length; the shoulder should remain undeformed. This is not always the case, as can be seen from Fig. 6. The macrographs of Fig. 6(a) (Ref 9) show the initial and deformed condition of a specimen in which measurable deformation has occurred within the

(a) (b) (c)

Fig. 5 Typical examples of heating methods for load-frame tension testing. (a) Induction heating. (b) Environmental chamber. (c) Split-furnace setup



(a) (b)

Fig. 6 Initial specimen geometry and deformed specimen for cases in which (a) shoulder deformation occurred (Ref 9) or (b) the shoulder remained undeformed (Ref 10)

grip area. On the other hand, the specimen with a different shoulder geometry (Fig. 6b) deformed essentially only along its gage length (Ref 10). The shoulder-deformation problem is not insurmountable. In this regard, analyses and techniques, such as those developed by Friedman and Ghosh (Ref 9), should be applied in order to eliminate the effect of shoulder deformation from measured hot-tension data.

Hot Ductility and Strength Data from the Gleeble Test

The reduction of area (RA) and strength are the key parameters measured in hot-tension tests conducted with a Gleeble machine (Ref 1, 11, 12). Because RA is a very structure-sensitive property, it can be used to detect small ductility variations in materials of low to moderate ductility, such as specialty steels and superalloys. However, it should be recognized that RA will not effectively reveal small variations in materials of extremely high ductility (Ref 2). Yield and tensile strength can be used to select required load capacity of production processing equipment.

Ductility Ratings

Experience has indicated that the qualitative ratings given in Table 1 for hot ductility as a function of Gleeble reduction-of-area data can be used to predict hot workability, select hot-working temperature ranges, and establish hot-reduction parameters. "Normal reductions"* may be taken on superalloys when the reduction of area exceeds 50%, but lighter reductions are necessary when ductility falls below this level. Thus, in this rating system, the *minimum* hot-working temperature is designated by the temperature at which the reduction of area falls below approximately 30 to 40%. The *maximum* hot-working temperature is determined from "on-cooling" data. The objective is to determine which preheat temperature provides the highest ductility over the broadest temperature range without risking permanent structural damage by overheating.

An alloy with hot-tensile ductility rated as marginal or poor may be hot worked, but smaller reductions and fewer passes per heating are required, perhaps in combination with insulating coatings and/or coverings. In extreme instances, it may be necessary to minimize development of tensile strains by employing special dies for deforming under a strain state that more nearly approaches hydrostatic compression (e.g., extrusion).**

It should be emphasized that the hot-tension test reflects the inherent hot ductility of a material, that is, its natural ability to deform under deformation conditions. If a workpiece possesses defects or flaws, it may crack due to localized stress concentration in spite of good inherent hot ductility.

Figure 3 illustrates how hot-tension data are used to select a hot-working temperature. The safe, maximum hot-working temperature lies between the PDT and the ZDT. In this hypothetical curve of "on-heating" data, the PDT is 1095 °C (2000 °F) and the ZDT is 1200 °C (2200 °F). "On-cooling" data should be determined using preheat temperatures between the PDT and the ZDT. For example, 1095, 1150, 1175, and 1200 °C (2000, 2100, 2150, and 2200 °F) would be good preheat temperatures for "on-cooling" studies. Typical "on-cooling" results are depicted in Fig. 7. A 1200 °C (2200 °F) preheat temperature results in marginal or poor hot workability over the possible working range, whereas an 1175 °C (2150 °F) preheat temperature results in acceptable hot workability over a relatively narrow temperature range. Both 1150 and 1095 °C (2100 and 2000 °F) preheats result in good hot ductility over a relatively narrow temperature range. The 1150 °C preheat temperature is preferred over the 1095 °C preheat temperature because it provides good hot ductility over a broader temperature range.

Hot workability usually is enhanced by greater amounts of prior hot deformation. This occurs because second phases and segregation-prone elements are distributed more uniformly and the grain structure is refined. Deformation at high and intermediate temperatures during commercial hot-working operations often refines the grain structure by dynamic (or static) recrystallization, thereby augmenting subsequent hot ductility at lower temperatures. Because a specimen tested "on cooling" to the low-temperature end of the hot-working range has not been deformed at a temperature where grains dynamically recrystallize, the grain structure is unrefined. Thus, ductility values will tend to be somewhat lower than those experienced in an actual metalworking operation in which deformation at higher temperatures has refined the structure. The fact that the low-temperature end of the "on-cooling" ductility range is lower than the values that would result in plant metalwork-

*"Normal reductions" as used in this chapter depend on both the alloy system being hot worked and the equipment being used. For example, normal reductions for low-carbon steels would be much greater than those for superalloy systems.

**E. Siebel, *Steel,* Vol 93, 1933

Table 1 Qualitative hot-workability ratings for specialty steels and superalloys

Hot-tension reduction of area(a), %	Expected alloy behavior under normal hot reductions in open die	Remarks regarding alloy hot-working practice
<30	Poor hot workability. Abundant cracks	Preferably not rolled or open-die forged. Extrusion may be feasible. Rolling or forging should be attempted only with light reductions, low strain rates, and an insulating coating.
30–40	Marginal hot workability. Numerous cracks	This ductility range usually signals the minimum hot-working temperature. Rolled or press forged with light reductions and lower-than-usual strain rates.
40–50	Acceptable hot workability. Few cracks	Rolled or press forged with moderate reductions and strain rates
50–60	Good hot workability. Very few cracks	Rolled or press forged with normal reductions and strain rates
60–70	Excellent hot workability. Occasional cracks	Rolled or press forged with heavier reductions and strain rates.
>70	Superior hot workability. Rare cracks. Ductile ruptures can occur if strength is too low.	Rolled or press forged with heavier reductions and higher strain rates than normal provided that alloy strength is sufficiently high to prevent ductile ruptures.

(a) Ratings apply for Gleeble tension testing of 6.25 mm (0.250 in.) diam specimens with 25.4 mm (1 in.) head separation.

ing operations is not sufficient to alter the practical translation of the results. This feature serves as a safety factor for establishing the minimum hot-working temperature.

Although some alloys will recover hot ductility when cooled from temperatures in the vicinity of the ZDT, it is nonetheless wise to avoid hot-working preheat temperatures approaching the ZDT in plant practice because interior regions of the workpiece may not cool sufficiently to allow recovery of ductility, thereby causing center bursting. Because industrial furnaces do not control closer than approximately ±14 °C (±25 °F), the recommended furnace temperature ordinarily should be at least 14 °C (25 °F) lower than the maximum temperature indicated by testing "on cooling."

Strength Data

In the hot-working temperature range, strength generally decreases with increasing temperature. However, the strength data plotted in Fig. 8 demonstrate that deformation resistance does not vary with preheat temperature to the same degree as does ductility. Furthermore, strength measured "on heating" is usually greater than that measured "on cooling."

To calculate the force required to deform a metal in an industrial hot-working operation, accurate measurement of *flow stress* is desirable. Ultimate tensile stress measured in the hot-tensile test is only slightly greater than flow stress at the high-temperature end of the hot-working range because work hardening is negligible. However, the difference between ultimate tensile stress and flow stress increases as temperature decreases because restoration processes cease. Furthermore, the Gleeble tensile test does not accurately determine flow stress because the strain rate is not constant. Nonetheless, the test still provides useful, comparative information concerning how the strength of an alloy varies as a function of temperature within a given strain-rate range. For example, by analyzing strength values for common alloys in relation to the load-bearing capacity of a given mill, it may be possible to use test data for a new or unfamiliar alloy in judging whether the equipment is capable of forming the new or unfamiliar alloy.

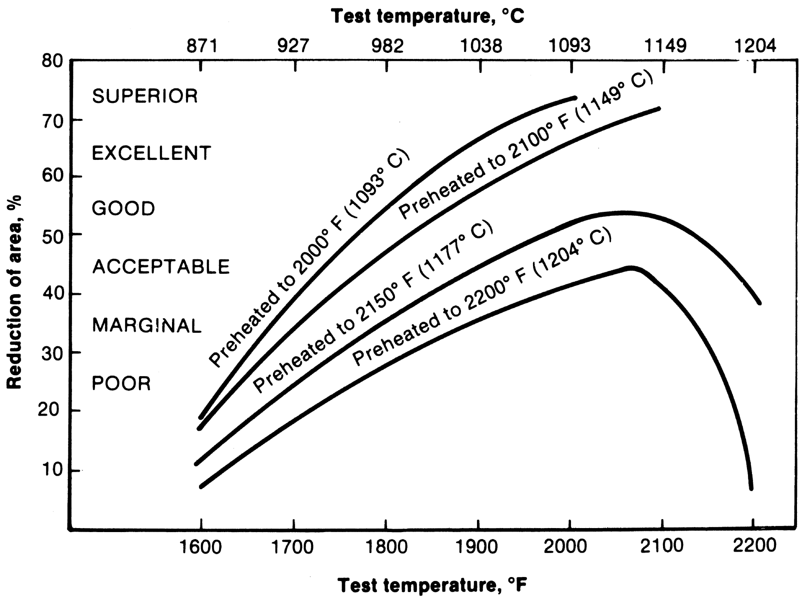


Fig. 7 Typical "on-cooling" Gleeble curves of specimen reduction of area as a function of test and preheat temperatures with typical hot-workability ratings indicated

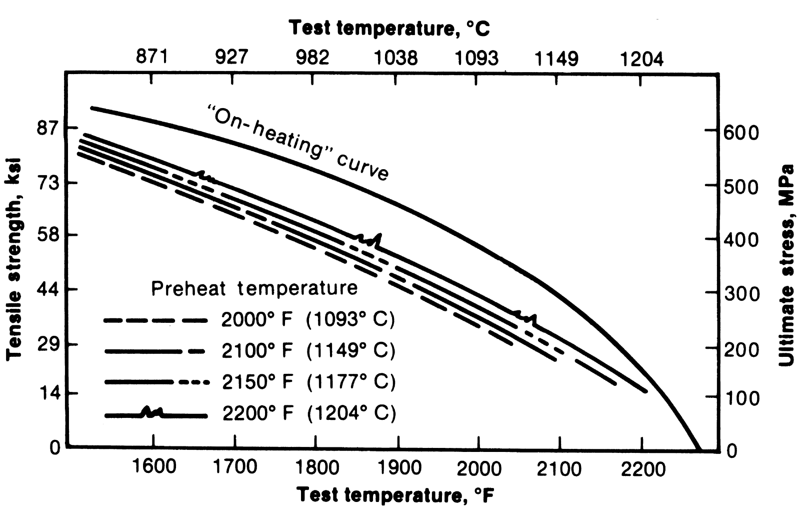


Fig. 8 Typical "on-heating" deformation-resistance data obtained in Gleeble testing

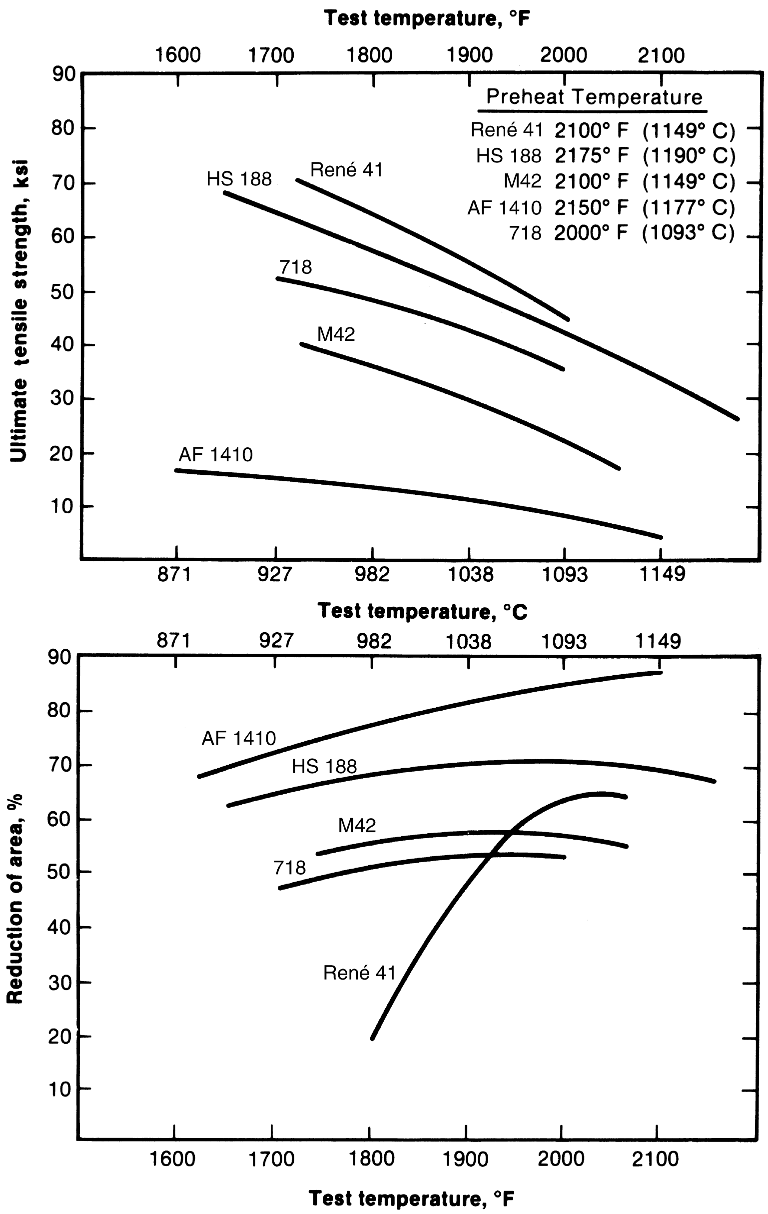
Hot-Tension Data for Commercial Alloys

For illustrative and comparative purposes, Gleeble hot ductility and strength curves for some commercial alloys are presented in Fig. 9. The nominal compositions of these materials are given in the table accompanying Fig. 9.

The hot-tensile strengths for the cobalt- and nickel-base superalloys over the hot-working temperature range are substantially higher than those for the high-speed tool steel and the high-strength alloy, which are iron-base materials. Furthermore, the ductility data reveal that René 41 has the narrowest hot-working temperature range (ΔT) of 140 °C (250 °F) of the three superalloys. Hot working of this alloy below 1010 °C (1850 °F) will lead to severe cracking. This characteristic, coupled with its high deformation resistance, makes this alloy relatively difficult to hot work. On the other hand, HS 188 has high deformation resistance, but it has high ductility over a broad hot-working temperature range from 1190 °C (2175 °F) to below 900 °C (1650 °F). Therefore, the permissible reduction per draft may be relatively small if the hot-working equipment is not capable of high loads, but HS 188 can be hot worked over a broader temperature range than René 41. However, if the equipment has high load capacity, then heavier reductions can be taken on HS 188 than on René 41. From the hot-working curves established for the iron-base, high-strength alloy AF 1410, the low deformation resistance coupled with high ductility over a broad temperature range indicate that this material has extremely good hot workability. Mill experience has verified this. The curves shown for M42, the high-speed tool steel, reveal that it is intermediate in hot workability between the superalloys and AF 1410; this conclusion has also been verified by mill experience.

Figures 10 and 11 illustrate the variation of hot-tensile ductility values at various temperatures. These results were correlated to the fracture surfaces and structures of the test specimens. For the high-strength iron-base alloy AF 1410, the "on-heating" curve in Fig. 10 shows that the ZDT is never reached at practical upper-limit hot-working temperatures. At the highest temperature tested (1230 °C, or 2250 °F) and at the PDT (1120 °C, or 2050 °F), where hot-tensile ductility is extremely high, the fracture appearance is ductile and dynamic recrystallization occurs, leading to an equiaxed grain structure. At the higher temperature, a coarser grain structure results from grain growth, which accounts for the drop in ductility. At the opposite end of the hot-working temperature range (842 °C, or 1548 °F), the elongated grain structure reveals that dynamic recrystallization does not occur, and the fracture surfaces indicate a - less-ductile fracture mode. The correlation among fracture appearance, microstructure, and hot-tensile ductility was even more evident for a developmental solid-solution-strengthened cobalt-base superalloy (Fig. 11). At the PDT (1150 °C, or 2100 °F), dynamic recrystallization occurs, fracture was primarily transgranular, and the fracture appearance was ductile. At the ZDT (1200 °C, or 2200 °F), both static recrystallization and grain growth were obvious, but incipient melting was not evident in the microstructure. Microstructural evidence of incipient melting at the ZDT is observed for some alloys, but not for others (Ref 13).

An example of the sensitivity of the hot-tensile Gleeble test is shown in Fig. 12 for iron/nickel-base superalloy Alloy 901 (Ref 14). A small amount of lanthanum added to one heat (top curve) was sufficient to reduce the analyzed sulfur content to the 1 to 5 ppm range. This resulted in a small improvement in the hot-tensile ductility according to Gleeble hot-tensile data.



Material	Alloy type	Nominal composition, wt%
René 41	Nickel-base superalloy	0.09 C, 19 Cr, 10 Mo, 11 Co, 3 Ti, 1.5 Al, 1.35 Fe, bal Ni
Alloy 718	Nickel-base superalloy	0.05 C, 18 Cr, 3 Mo, 5 Nb, 1 Ti, 0.5 Al, 19 Fe, bal Ni
HS 188	Cobalt-base superalloy	0.08 C, 22 Cr, 22 Ni, 14 W, 1.5 Fe, 0.05 La, bal Co
M42	High-speed tool steel	1.10 C, 3.8 Cr, 9.5 Mo, 8 Co, 1.5 W, 1.2 V, bal Fe
AF 1410	High-strength steel	0.16 C, 2 Cr, 1 Mo, 14 Co, 10 Ni, bal Fe

Fig. 9 Typical "on-cooling" Gleeble curves of strength and ductility as functions of test temperature for several commercial alloys.

Isothermal Hot-Tension Test Data

From the *isothermal* hot-tension test, information can be obtained about a number of material parameters that are important with regard to metalworking process design. These include plastic-flow (stress-strain) behavior, plastic anisotropy, tensile ductility, and their variation with the test temperature and the strain rate.

Stress–Strain Curves

Engineering stress-strain curves from isothermal hot-tension tests are constructed from load-elongation measurements. The engineering, or nominal, stress is equal to the *average* axial stress and is obtained by dividing the instantaneous load by the *original* cross-sectional area of the specimen. Similarly, the engineering, or nominal, strain represents the *average* axial strain and is obtained by dividing the elongation of the gage length of the specimen by its *original* length. Hence, the form of the engineering stress-strain curve is exactly the same as that of the load-elongation curve. Examples of engineering stress-strain curves obtained from hot-tension testing of an orthorhombic titanium aluminide alloy (Ref 15) at 980 °C (1800 °F) and a range of nominal (initial) strain rates are shown in Fig. 13. The curves exhibit a stress maximum at strains less than 10%, a regime of quasi-stable flow during which a diffuse neck develops and the load drops gradually, and, lastly, a period of rapid load drop during which the flow is highly localized (usually in the center of the specimen gage length) and failure occurs. The engineering stress-strain curve does not give a true indication of the deformation characteristics of a metal because it is based entirely on the *original* dimensions of the specimen. These dimensions change continuously during the test. Such changes are very significant when testing is performed at elevated temperatures.

The true stress and true strain are based on *actual* (instantaneous) cross-sectional area and length measurements at any instant of deformation. The true-stress/true-strain curve is also known as the *flow curve* since it represents the basic plastic-flow characteristic of the material under the particular (temperature-strain rate) testing conditions. Any point on the flow curve can be considered the yield stress for a metal strained in tension by the amount shown on the curve. An example of the variation of the true stress versus true strain for Al-8090 alloy deformed under superplastic conditions (T = 520 °C, $\dot{\varepsilon} = 7.8 \times 10^{-4}\ s^{-1}$) is shown in Fig. 14 (Ref 16). Under these conditions, it is apparent that the flow stress is almost independent of strain. For ideal superplastic materials, the flow stress is independent of strain. A nearly constant, or steady-state, flow stress is also frequently observed at hot-working temperatures in materials that undergo dynamic recovery. In these cases, steady-state flow is achieved at strains of the order of 0.2, at which the rate of strain hardening due to dislocation multiplication is exactly balanced by the rate of dislocation annihilation by dynamic recovery.

The variation of true stress with true strain can also give insight into microstructural changes that occur during hot deformation. For example, for superplastic materials, an increase in the flow stress with strain is normally indicative of strain-enhanced grain growth. A decrease in flow stress, particularly at high strains, can often imply the development of cavitation damage (see the section "Cavitation During Hot-Tension Testing" in this chapter) or the occurrence of dynamic recrystallization. As an example, true-stress/true-strain curves for a γ-TiAl submicrocrystalline alloy deformed at temperatures between 600 and 900 °C (1110 and 1650 °F) and a nominal (initial) strain rate of $8.3 \times 10^{-4}\ s^{-1}$ are shown in Fig. 15 (Ref 17). These curves reveal that deformation at low temperatures, at which nonsuperplastic conditions prevail, is characterized by an increase of flow stress with strain due to the strain hardening. At higher temperatures, the effect of strain on the flow stress decreases until it becomes negligible at the highest test temperature, thus indicating the occurrence of superplastic flow.

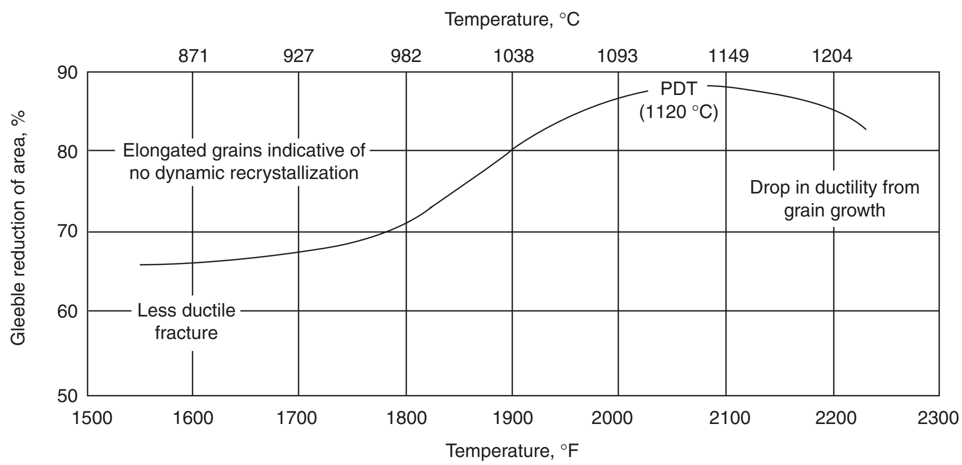


Fig. 10 Typical Gleeble curve of reduction of area versus test temperature for an aircraft structural steel (AF 1410). At the PDT, dynamic recrystallization occurs leading to an equiaxed grain structure. Fracture appearance is ductile.

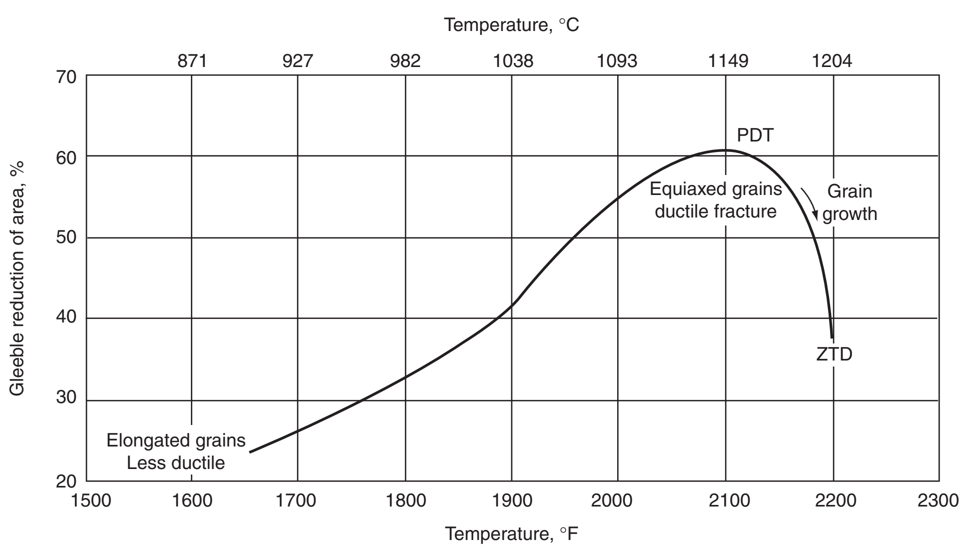


Fig. 11 Typical Gleeble curve of reduction area versus test temperature for a cobalt-base superalloy.

Material Coefficients from Isothermal Hot-Tension Tests

A number of material coefficients can be obtained from isothermal hot-tension tests. These include measures of strain and strain-rate hardening and plastic anisotropy. The strain-hardening exponent (usually denoted by the symbol n) describes the change of flow stress (with an effective stress, $\bar{\sigma}$) with respect to the effective strain, $\bar{\varepsilon}$, such that:

$$n = \frac{\partial \ln \bar{\sigma}}{\partial \ln \bar{\varepsilon}} \qquad \text{(Eq 1a)}$$

For a uniaxial tension test, and prior to the development of a neck, the distinction of effective stress and strain is not necessary because they are equal to the axial stress σ and strain ε, so that the expression is simply:

$$n = \frac{\partial \ln \sigma}{\partial \ln \varepsilon} \qquad \text{(Eq 1b)}$$

The strain-hardening exponent may have values from $n = 0$ for a perfectly plastic solid to $n = 1$ for an elastic solid; negative values of n may also be found for materials that undergo flow softening due to changes in microstructure or crystallographic texture during deformation. According to Eq 1, if the constitutive equation for stress-strain behavior is of the form $\sigma = K\varepsilon^n$, then a logarithmic plot of true stress versus true strain results in a straight line with a slope equal to n. However, this is not always found to be the case and reflects the fact that this relationship is only an empirical approximation. Thus, when the plot of ln(σ) versus ln(ε) [or the plot of log(σ) versus log(ε)] results in a nonlinear value of n, then the strain-hardening exponent is often defined at a particular strain value. In general, n increases with decreasing strength level and decreasing ease of dislocation cross slip in a polycrystalline material.

The strain-rate sensitivity exponent (usually denoted by the symbol m) describes the variation of the flow stress with the strain rate. In terms of effective stress ($\bar{\sigma}$) and effective strain rate ($\dot{\bar{\varepsilon}}$), it is determined from the following relationship:

$$m = \frac{\partial \ln \bar{\sigma}}{\partial \ln \dot{\bar{\varepsilon}}} \qquad \text{(Eq 2a)}$$

which is simplified for the condition of pure uniaxial tension as:

$$m = \frac{\partial \ln \sigma}{\partial \ln \dot{\varepsilon}} \qquad \text{(Eq 2b)}$$

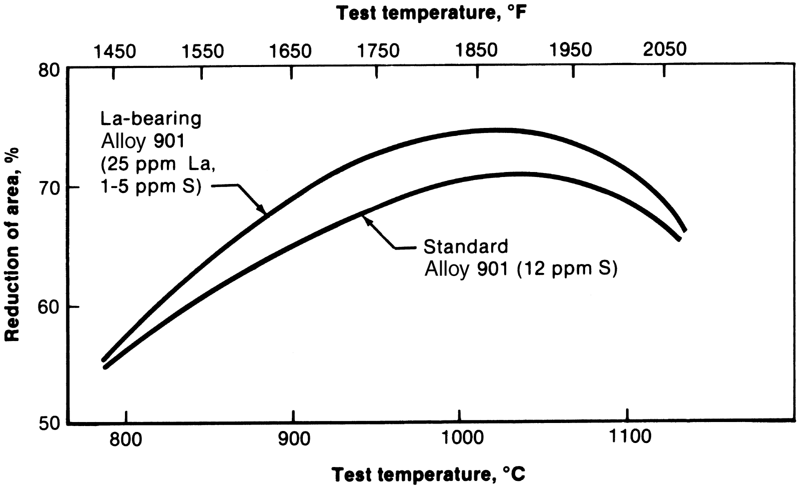


Fig. 12 Gleeble ductility curves for lanthanum-bearing and standard Alloy 901 tested on cooling from 1120 °C. Note that the lanthanum-bearing heat displays slightly higher ductility. Specimens represent transverse orientation on a nominal 25 cm square billet. Specimen blanks were heat treated at 1095 °C for 2 h and then water quenched prior to machining. Specimens were heated to 1120 °C, held for 5 min, cooled to test temperature and held for 10 s before being tested at a nominal strain rate of 20 s^{-1} (crosshead speed 5 cm/s; jaw spacing, 2.5 cm). After Ref 14

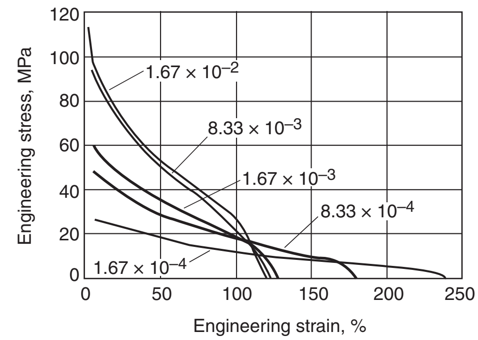


Fig. 13 Engineering stress-strain curves for an orthorhombic titanium alloy (Ti-21Al-22Nb) tested at 980 °C and a range of initial strain rates (s^{-1}). Source: Ref 15

Deformation tends to be stabilized in a material with a high *m* value. In particular, the presence of a neck in a material subject to tensile straining leads to a locally higher strain rate and thus to an increase in the flow stress in the necked region due to *strain-rate* hardening. Such strain-rate hardening inhibits further development of the strain concentration in the neck. Thus, a high strain-rate sensitivity imparts a high resistance to necking and leads to high tensile elongation or superplasticity. Materials with values of *m* equal to or greater than approximately 0.3 exhibit superplasticity, assuming cavitation and fracture do not intercede. An empirical relation between tensile elongation and the *m* value is revealed in the data collected by Woodford (Ref 18) shown in Fig. 16. In addition, a number of theoretical analyses have been conducted to relate *m* and tensile failure strain, ε_f (Ref 19–21). For example, Ghosh (Ref 19) derived:

$$\varepsilon_f = -m \ln(1 - f^{1/m}) \qquad \text{(Eq 3)}$$

in which *f* denotes the size of the initial geometric (area) defect at which flow localization occurs.

The plastic anisotropy parameter *R* characterizes the resistance to thinning of a *sheet* material during tension testing and is defined as the slope of a plot of width strain, ε_w, versus thickness strain, ε_t, (Ref 22), that is:

$$R = \frac{d\varepsilon_w}{d\varepsilon_t} \qquad \text{(Eq 4)}$$

A material that possesses a high *R* value has a high resistance to thinning and hence good formability, especially during deep-drawing operations. Materials with values of *R* greater than unity have higher strength in the thickness direction than in the plane of the sheet.

The plastic anisotropy parameter can be readily measured using specimens deformed in uniaxial tension. However, caution should be exercised when making such measurements to ensure that the stress state along the gage length is uniaxial. Therefore, measurements in regions near the sample shoulder and the failure site (where a stress state of hydrostatic tension may develop during necking) should be avoided. Figure 17 shows an example of such data from a Ti-21Al-22Nb sample pulled to failure in uniaxial tension at a nominal strain rate of 1.67 × 10^{-4} s^{-1} and at a temperature of 980 °C (1800 °F). Within experimental scatter, the *R* value is constant for the majority of deformation. Apparently, lower values of *R* at low strains (near the specimen shoulder) or very high strains (at the fracture tip) are invalid due to constraint or flow-localization effects, respectively, and hence conditions that are not uniaxial.

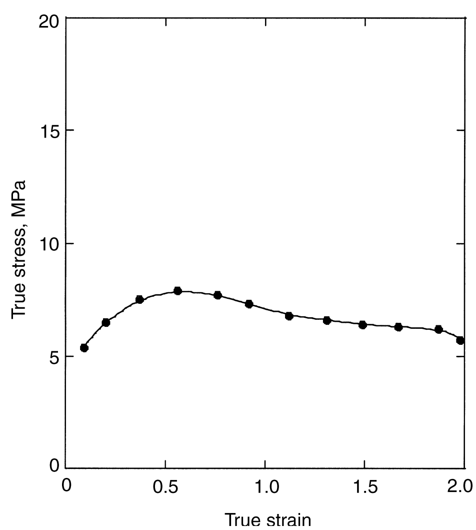


Fig. 14 True-stress/true-strain data for an Al-8090 alloy deformed in tension at 520 °C and a true strain rate of 7.8 × 0^{-4} s^{-1}. After Ref 16

The *R* value of a sheet material may be sensitive to the testing conditions and in particular to the strain rate and temperature. This is a result of variation of the mechanism that controls deformation (e.g., slip, grain-boundary sliding, etc.) with test conditions. For the orthorhombic titanium aluminide material discussed previously, the normal plastic anisotropy parameter shows a very weak dependence on strain, but a noticeable variation with strain rate (Fig. 18). This trend can be attributed to the presence of mechanical and crystallographic texture and the effect of strain rate on the operative deformation mechanism.

Effect of Test Conditions on Flow Behavior

When considering the effect of test conditions on flow behavior, it must be understood that testing for the modeling of deformation processes is very different from testing for static mechanical

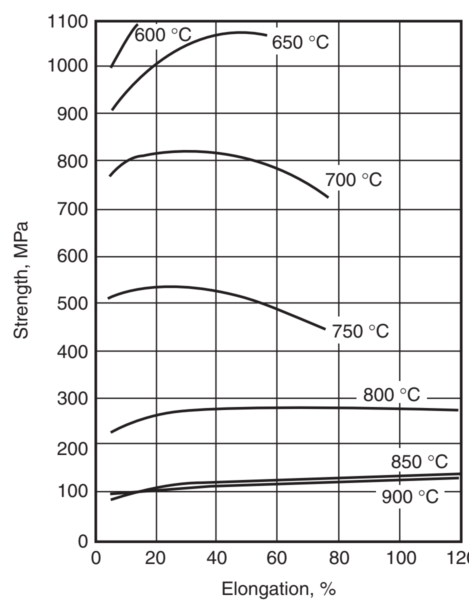


Fig. 15 True-stress/true-strain curves obtained from tension testing of submicrocrystalline TiAl samples. After Ref 17

properties at very low (quasi-static) loading rates. Testing conditions for deformation processes must cover a range of strain rates and may require high strain rates of 1000 s^{-1} or more. For tension testing, conventional test frames are applicable for strain of rates less than 0.1 s^{-1}, while special servohydraulic frames have a range from 0.1 to 100 s^{-1} (see "Introduction to High Strain Rate Testing" in *Mechanical Testing and Evaluation,* Volume 8 of the *ASM Handbook*, 2000, p 427). For strain rates from 100 to 1000 s^{-1}, the Hopkinson (Kolsky)-bar method is used. This chapter and the following discussions only consider isothermal conditions and strain rates below 0.1 s^{-1}, where inertial effects can be neglected.

Effect of Strain Rate and Temperature on Flow Stress. At hot-working temperatures, most metals exhibit a noticeable dependence of flow stress on strain rate and temperature. For instance, the variation of flow stress with strain rate for Ti-6Al-4V (with a fine equiaxed microstructure) deformed at 927 °C is shown in Fig. 19 (Ref 23). For the strain-rate range shown in Fig. 19, a sigmoidal variation of the flow stress with strain rate is observed. From these data, the strain-rate sensitivity (*m* value) can be readily calculated. The result of these calculations (Fig. 20) shows that *m* is low at low strain rates and then increases and passes through a maximum after which it decreases again. This behavior is typical of many metals with fine-grain microstructures and reveals that superplasticity is not manifested in either the low-stress, low-strain-rate region I or the high-stress, high-strain-rate region III. Rather, superplasticity is found only in region II in which the stress increases rapidly with increasing strain rate. The superplastic region II is displaced to higher strain rates as temperature is increased and/or grain size is decreased. Moreover, the maximum observed values of *m* increase with similar changes in these parameters.

The stress-strain curve and the flow and fracture properties derived from the hot-tension test are also strongly dependent on the temperature at which the test is conducted. In both single-crystal and polycrystalline materials, the strength decreases with temperature because the critical resolved shear stress decreases sharply with an increase in temperature. On the other hand, the tensile ductility *increases* with temperature because of the increasing ease of recovery and recrystallization during deformation. However, the increase in temperature may also cause microstructural changes such as precipitation, strain aging, or grain growth that may affect this general behavior.

The flow stress dependence on temperature and strain rate is generally given by a functional form that incorporates the Zener-Hollomon parameter, $Z = \dot{\varepsilon} \exp(Q/RT)$ (Ref 24) in which Q is the apparent activation energy for plastic flow, R the universal gas constant, and T is the absolute temperature.

Effect of Crosshead Speed Control on Hot Tension Data. The selection of constant-strain-rate versus constant-crosshead-speed control in conducting isothermal, hot-tension tests is an important consideration, especially for materials that are superplastic. When experiments are conducted under constant-crosshead-speed conditions, the specimen experiences a decreasing strain rate during the test, thus making the interpretation of results difficult, especially in the superplastic regime. A method to correct for the strain-rate variation involves continuously changing the crosshead speed during the tension test to achieve nearly constant strain rate. This approach assumes uniform deformation along the gage length and no end effects and leads to the following relation between crosshead speed $\dot{\delta}$, desired strain rate $\dot{\varepsilon}$, the initial gage length l_o, and time t:

$$\dot{\delta} = \dot{\varepsilon}\, l_o \exp(-\dot{\varepsilon} t) \qquad \text{(Eq 5)}$$

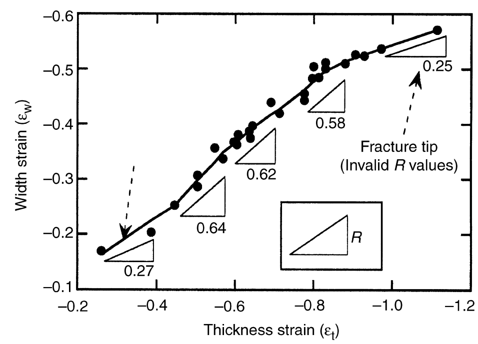


Fig. 17 Width versus thickness strain (ε_w versus ε_t) for an orthorhombic titanium aluminide specimen deformed at 980 °C and a nominal strain rate of 1.67×10^{-4} s^{-1}. Source: Ref 10

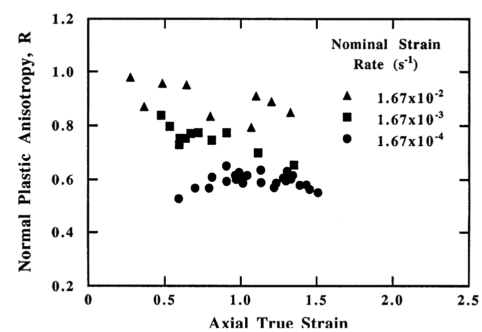


Fig. 18 Anisotropy parameter *R* versus the local axial true strain for various nominal strain rates. Data correspond to a Ti-21Al-22Nb alloy. Source: Ref 10

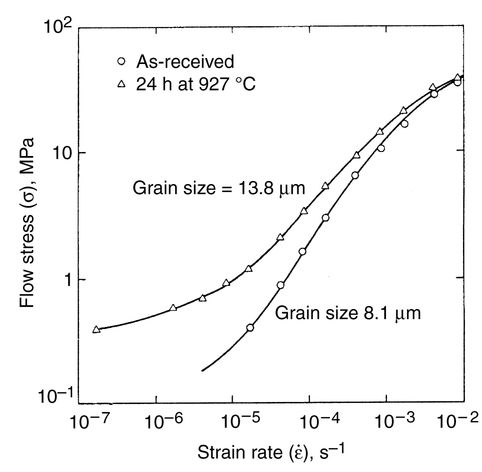


Fig. 19 Flow stress as a function of strain rate and grain size for a Ti-6Al-4V alloy deformed at 927 °C. The strain level was about 0.24. After Ref 23

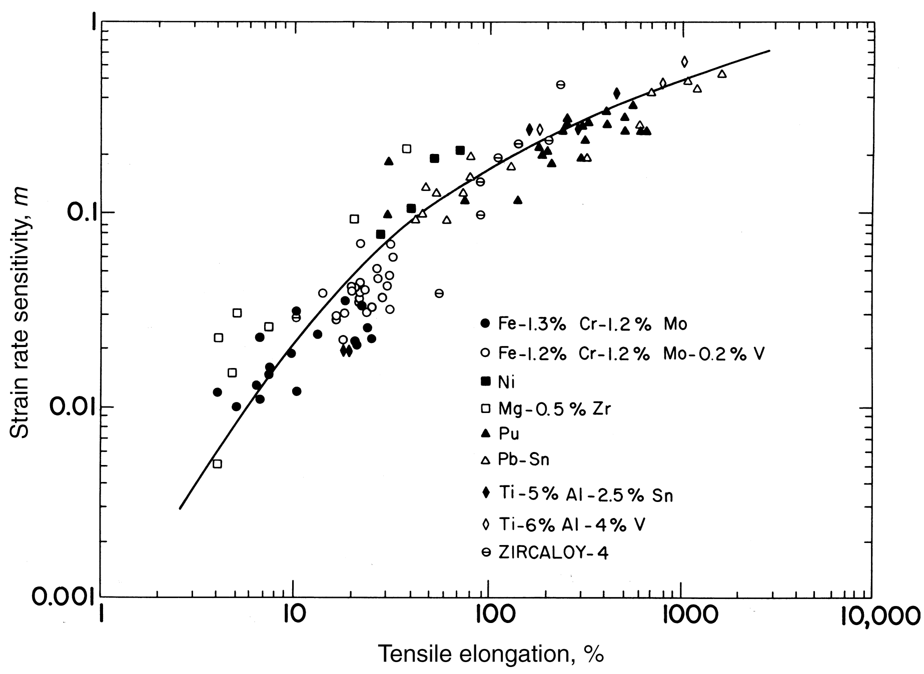


Fig. 16 Tensile elongation as a function of the strain-rate sensitivity. Source: Ref 18

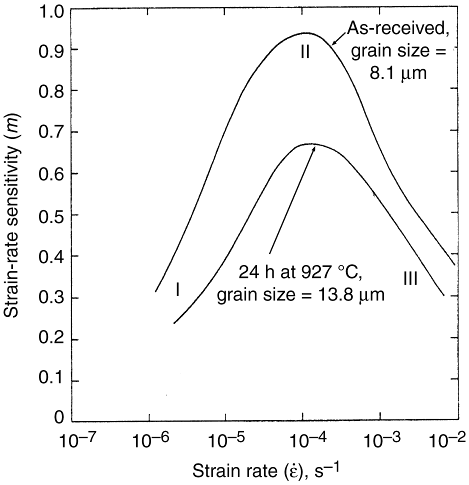


Fig. 20 Strain-rate sensitivity (m) versus strain rate ($\dot{\varepsilon}$) for the data corresponding to Fig. 19. After Ref 23

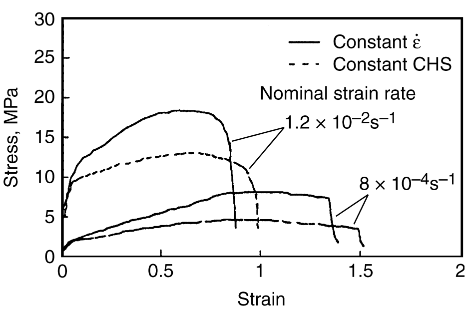


Fig. 21 Comparison of stress versus strain for constant nominal strain rate (constant crosshead speed, CHS) and constant true strain rate ($\dot{\varepsilon}$) for 5083 Al at 550 °C. Source: Ref 26

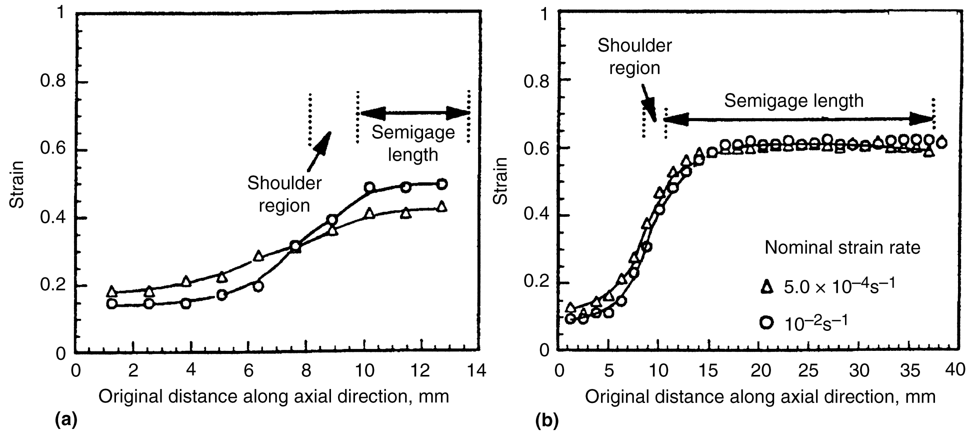


Fig. 22 Strain distribution for 12.7 mm (a) and 63.5 mm (b) gage length specimens for two different strain rates. Length strains are plotted versus original axial position along the gage length. Source: Ref 9

The crosshead-speed schedule embodied in Eq 5 has been used successfully for a test of Ti-6Al-4V (Ref 25). Verma et al. (Ref 26) have also shown the efficacy of this approach by conducting tension tests at constant crosshead speed as well as constant strain rate on superplastic 5083-Al specimens. Figure 21 compares stress-strain characteristics determined under constant-crosshead-speed conditions with those from constant-strain-rate tests for two different initial strain rates. Constant-crosshead-speed tests showed consistently lower strain hardening (lower flow stresses) and larger strain to failure (higher tensile elongations) than the corresponding constant-strain-rate tests did. The above finding highlights the importance of the test control mode; in addition, this mode should be clearly stated when elongation and/or flow stress data are reported.

Effect of Gage Length on Strain Distribution. Under superplastic deformation conditions, specimen geometry (especially shoulder design) plays an important role in the determination of hot-tension characteristics. In the section "Frame-Furnace Tension-Testing Equipment," two different specimen designs are discussed (Fig. 6). For one of these designs, deformation was limited essentially to the gage section, while the other had experienced deformation in the shoulder section. For the specimen geometry in Fig. 6(a), tension tests indicated that significant straining can occur in the grip regions and that large strain gradients exist within the gage section of the specimen. The strain gradient (variation) along the gage length and the deformation of the grip section depend on the gage length and tensile strain rate. As can be seen in Fig. 22, the strain gradient of the smaller gage length (12.7 mm, or 0.5 in.) specimen geometry is much steeper than that of the larger one (63.5 mm, or 2.5 in.). With regard to the smaller gage length specimen, it is observed that the strain gradient becomes steeper as the strain rate increases. Furthermore, a reduction of deformation in the shoulder can be achieved by decreasing the *width* of the gage section because of the decrease in deformation load and hence stress level generated in the shoulder. However, there are constraints in gage-width reductions arising from the microstructural characteristics of a particular material; in some cases, there may be an insufficient number of grains across the specimen section.

Modeling of the Isothermal Hot-Tension Test

The detailed interpretation of data from the isothermal hot-tension test frequently requires some form of mathematical analysis. This analysis is based on a description of the local stress state during tension testing and some form of numerical calculation. The approach is described briefly in this section.

Stress State at the Neck

Prior to necking, the stress state in the tension test is uniaxial. However, the onset of necking is accompanied by the development of a triaxial (hydrostatic*) state of stress in the neck. Because the flow stress of a material is strongly dependent on the state of stress, a correction must be introduced to convert the measured average axial stress into the effective uniaxial flow stress; that is:

$$\sigma_l^{av} = \bar{\sigma}/F_T \qquad \text{(Eq 6)}$$

in which σ_l^{av} denotes the average axial stress required to sustain further deformation, $\bar{\sigma}$ is the effective (flow) stress, and F_T is the stress triaxiality factor. The magnitude of F_T (which essentially determines the magnitude of the *average* hydrostatic stress within the neck) depends on the specimen shape (round bar or sheet) and the geometry of the neck. Bridgman (Ref 27) conducted a rigorous, theoretical analysis with regard to the stress state at the neck for both round-bar and for sheet specimen geometries. For a plastically isotropic material, the following equations were derived for the stress triaxiality factor of round-bar (F_T^r) and sheet (F_T^s) specimens in the *symmetry plane* of the neck:

$$F_T^r = \left\{\left[1+\left(2\frac{R}{a}\right)\right]\ln\left[1+\left(\frac{a}{2R}\right)\right]\right\}^{-1} \qquad \text{(Eq 7)}$$

$$F_T^s = \left\{\left(1+2\frac{R}{a}\right)^{1/2} \ln\left[1+\frac{a}{R}+\left(\frac{2a}{R}\right)^{1/2}\left(1+\frac{1}{2}\frac{a}{R}\right)^{1/2}\right]-1\right\}^{-1} \qquad \text{(Eq 8)}$$

*The term *hydrostatic stress* is defined as the mean value of the normal stresses. The term *triaxial stress* is often used to imply the presence of a hydrostatic stress. However, the term *triaxial stress* is not equivalent to *hydrostatic stress,* because the presence of a triaxial stress state could be a combination of shear stresses and/or normal stresses or only normal stresses. The term hydrostatic stress is thus preferred and more precise in describing solely normal stresses in three orthogonal directions.

in which a represents the specimen half radius or width, and R is the radius of curvature of the neck.*

The variation of the stress triaxiality factor for round-bar (F_T^r) and sheet (F_T^s) specimens as a function of the a/R ratio is shown in Fig. 23. For a positive a/R value (concave neck profiles), F_T is less than unity, thus promoting flow stabilization. On the other hand, for negative a/R (convex neck profiles), $F_T > 1$; thus, flow tends to be destabilized.

In a rigorous sense, the closed-form equations for F_T (Eq 7 and 8) are applicable only for the plane of symmetry at the neck. At other locations, the solution for the exact form of F_T is not available. However, as has been shown from finite-element method (FEM) analyses (Ref 28, 29), Eq 7 and 8 provide a good estimate for the stress triaxiality factor in regions away from the symmetry plane provided that the local values of a and R are inserted into the relations.

Numerical Modeling of the Hot-Tension Test

Two types of methods have been employed to model the tension behavior of materials: the FEM and the somewhat simpler finite-difference ("direct-equilibrium") method originally presented by G'sell et al. (Ref 30), Ghosh (Ref 31), and Semiatin et al. (Ref 32). Both approaches involve solutions that satisfy the axial force equilibrium equation and the appropriate boundary conditions. These models enable the prediction of important parameters such as neck profile, failure mode, axial-strain distribution, and ductility. A comparison of simulation results (e.g., nominal stress-strain curves, axial-strain distribution, and total elongation) obtained from FEM analyses to those of the direct-equilibrium method has shown that the latter approach gives realistic predictions (Ref 29). To this end, a brief description of this simpler method is given in the following paragraphs.

Model Formulation. The formulation of the direct-equilibrium method is based on discretization of the sample geometry, description of the material flow behavior, and development of the appropriate load-equilibrium equation. The specimen geometry (dimensions, cross-section shape, geometrical defects, etc.) is first specified. The specimen is divided along the axial direction into horizontal slices/elements (Fig. 24). For the material flow behavior, the simple engineering power-law formulation has been used in most modeling efforts, that is:

$$\bar\sigma = K\,\bar\varepsilon_s^n\dot{\bar\varepsilon}_s^m \qquad \text{(Eq 9)}$$

in which $\bar\sigma$, $\bar\varepsilon_s$, and $\dot{\bar\varepsilon}_s$ denote the effective stress, effective strain, and effective strain rate, respectively, of the material. K, n, and m represent the strength coefficient, strain-hardening exponent, and the strain-rate-sensitivity index, respectively.

At any instant of deformation, the axial load P should be the same in each element in order to maintain force equilibrium. The load borne by each slice is equal to the product of its load-bearing cross-sectional area and axial stress; the axial stress is equal to the flow stress corrected for stress triaxiality due to necking and evaluated at a strain rate corresponding to that which the material elements experience.* The load-equilibrium condition is thus described by:

$$\bar\sigma_i A_{lb}^i / F_T^i = \bar\sigma_j A_{lb}^j / F_T^j \qquad \text{(Eq 10)}$$

or, using Eq 9:

$$\dot{\bar\varepsilon}_{s_i}^m \bar\varepsilon_{s_i}^n A_{lb}^i / F_T^i = \dot{\bar\varepsilon}_{s_j}^m \bar\varepsilon_{s_j}^n A_{lb}^j / F_T^j \qquad \text{(Eq 11)}$$

in which the subscripts and/or superscripts i and j denote the corresponding parameters for elements i and j, respectively, F_T represents the stress triaxiality factor, and A_{lb} is the load-bearing area.

For the case in which the material cavitates during tension testing (see the section "Cavitation During Hot-Tension Testing" in this chapter), Eq 10 and 11 must be modified. In particular, the presence of cavities affects the external dimensions of the specimen (because they lead to a volume increase) and hence the load-bearing area, the stress triaxiality factor, and the strain rate at which the material deforms.

As discussed by Nicolaou et al. (Ref 33), spherical and uniformly distributed cavities increase each of the three dimensions (length, width, and thickness**) of the tension specimen by the same amount. The relationship between the macroscopic area (i.e., the external area of the specimen) A_m, the load-bearing area A_{lb}, and the initial (uncavitated) area A_o^{sp} is then simply:

$$A_m = A_{lb}/(1 - C_v)^{2/3} \qquad \text{(Eq 12)}$$

where C_v is the cavity volume fraction and A_{lb} is given by:

$$A_{lb} = A_o^{sp}\exp\left(-\bar\varepsilon_s\right) \qquad \text{(Eq 13)}$$

In addition, the length l, width w, and thickness t, for a sheet specimen, or radius r of a round-bar specimen increase according to:

$$d' = \frac{d}{(1-C_v)^{1/3}} \qquad \text{(Eq 14)}$$

in which d' denotes any of the dimensions (l', w', t', or r') for the case when cavities are present in the material, and d represents the respective dimension changes with strain alone.

The matrix strain rate $\dot{\bar\varepsilon}_s$, can also be related to the macroscopic sample strain rate $\dot{\bar\varepsilon}$. Using power-dissipation arguments, the relation between the two strain rates is found to be (Ref 34):

$$\dot{\bar\varepsilon}_s = (1/\phi\rho)\,\dot{\bar\varepsilon} \qquad \text{(Eq 15)}$$

in which ρ is the relative density of the specimen ($\rho = 1 - C_v$) and ϕ is the stress-intensification

*For $a/R < 0$ (in particular, for $-2 < a/R < 0$), the stress triaxility factor for sheet tensile specimesns with a *convex* curvature is given by $F_T\left\{\left[2\sqrt{Q}\arctan\left(1/\sqrt{Q}\right)\right]-1\right\}^{-1}$, where $Q = -(1 + 2R/a)$.

*In case of a cavitating material, the strain rate is that of the matrix-material element, *not* the matrix-cavity continuum.
**Diameter for a round-bar specimen geometry

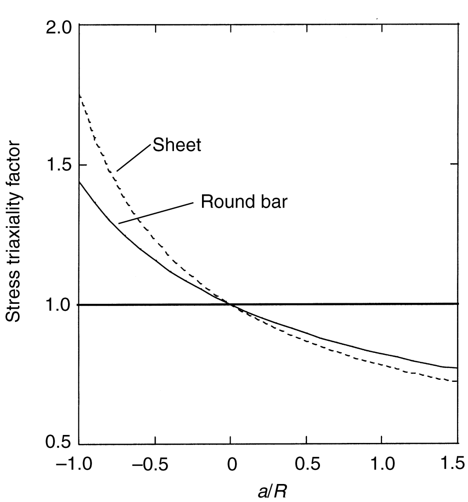


Fig. 23 Stress triaxiality factor for sheet and round-bar specimens

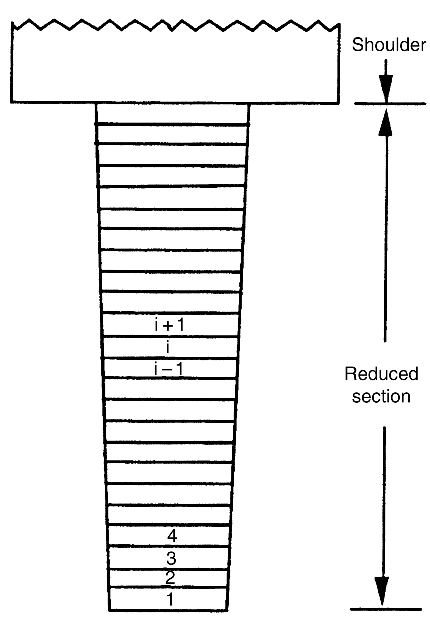


Fig. 24 Discretization of the sheet specimen for the simulations of the isothermal hot-tension test (Ref 33). The specimen geometry corresponds to the specimen shown in Fig. 6(b) (Ref 10, 15)

factor, which for spherical and uniformly distributed cavities is (Ref 33):

$$\phi = 1/\rho^{2/3} \quad \text{(Eq 16)}$$

Algorithm. After having specified the specimen geometry and the material-flow relation, the equations for model formulation can be inserted into an algorithm to simulate the tension test. At any instant of deformation, the axial variation in strain rate is calculated based on the load-equilibrium equation, which provides the ratios of the strain rates in the elements, and the boundary condition (e.g., constant crosshead speed), which provides the specific magnitudes of the strain rates. The strain rates are then used to update the macroscopic (and microscopic) strain and cavity volume fraction (for a cavitating material) in each element. The simulation steps are:

1. An increment of deformation is imposed, and a/R and F_T are calculated for each slice.
2. From the true strain and the cavity-growth rate (see the section "Cavitation During Hot Tension Testing"), the cavity volume fraction is determined.
3. The true-strain-rate distribution is calculated for each element, using the equilibrium equation and the boundary condition.
4. From the true-strain, cavity volume fraction, and strain-rate distributions, the engineering stress and strain are calculated.
5. Steps 1 to 4 are repeated until a sharp neck is formed (localization-controlled failure) or the cavity volume fraction at the central element reaches a value of 0.3 (fracture/cavitation-controlled failure).

Example Applications. Several results illustrate the types of behavior that can be quantified using the direct-equilibrium modeling approach. The first deals with the effect of specimen taper on tensile elongation. Tension-test specimens usually have a small (≤2%) reduction in the cross-sectional area from the end to the center of the reduced section in order to control the location of failure. The predicted effect of reduced-section taper on the engineering stress-strain curves for non-strain-hardening materials is shown in Fig. 25 (Ref 29). The effect of the absence of a taper on increased elongation is quite dramatic, especially as the strain-rate sensitivity increases from $m = 0.02$ to $m = 0.15$. For materials deformed at cold-working temperatures ($m < 0.02$), tensile flow will still localize in the absence of a taper because the reduced section itself acts as the defect relative to the greater cross-sectional area of the shoulder (Ref 31, 32). In contrast to the results for samples with and without a 2% taper, the predictions for samples with a 1% versus a 2% taper show much less difference.

With appropriate modification, the direct-equilibrium modeling approach may also be used to analyze the uniaxial hot-tension testing of sheet materials that exhibit normal plastic anisotropy (Ref 35). Selected results are shown in Fig. 26. The engineering stress-strain curves exhibit a load maximum, a regime of quasi-stable flow during which the diffuse neck develops and the stress decreases gradually, and finally a period of rapid load drop during which flow is highly localized in the center of the gage length. When the m value is low, an increase in R increases the amount of quasi-stable flow; that is, it stabilizes the deformation in a manner similar to the effect of strain-rate sensitivity. In addition, the simulation results reveal that the flow-stabilizing effect of R decreases as m increases and in fact becomes negligible for conditions that approach superplastic flow (i.e., $m > 0.3$).

Cavitation During Hot-Tension Testing

A large number of metallic materials form microscopic voids (or cavities) when subjected to large strains under tensile modes of loading. This formation of microscopic cavities, which primarily occurs in the grain boundaries during high-temperature deformation, is referred to as cavitation. In some cases, cavitation may lead to premature failure at levels of deformation far less than those at which flow-localization-controlled failure would occur. For a given material, the extent of cavitation depends on the specific deformation conditions (i.e., strain rate and temperature). A wide range of materials exhibit cavitation; these materials include aluminum alloys (Fig. 27a), conventional titanium alloys (Fig. 27b), titanium aluminides, copper alloys, lead alloys, and iron alloys (Ref 36–38).

An important requirement for cavitation during flow under either hot-working or superplastic conditions is the presence of a tensile stress. On the other hand, under conditions of homogeneous compression, cavitation is not observed; in fact, cavities that may be produced under tensile flow can be removed during subsequent compressive flow. In addition, it has also been demonstrated that the superposition of a hydrostatic pressure can reduce or eliminate cavitation (Ref 39). Hot isostatic pressing can also heal the deformation damage of nucleated cavities.

Cavitation is a very important phenomenon in hot working of materials because not only may it lead to premature failure during forming, but it also may yield inferior properties in the final part. Therefore, it has been studied extensively, primarily via the tension test.

Cavitation Mechanisms/Phenomenology

Cavitation occurs via three often-overlapping stages during tensile deformation: cavity nucleation, growth of individual cavities, and cavity coalescence. Each stage is briefly described in the following sections, while a more detailed

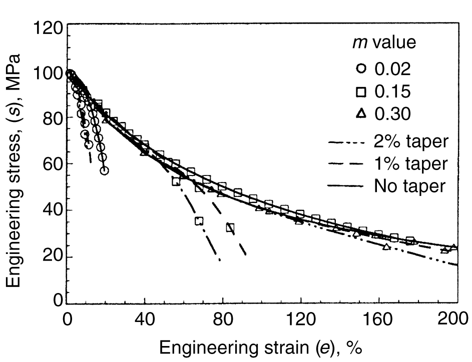


Fig. 25 Comparison of the engineering stress-strain curves for non-strain-hardening samples without or with a 1 or 2% taper predicted using the direct-equilibrium approach. Source: Ref 29

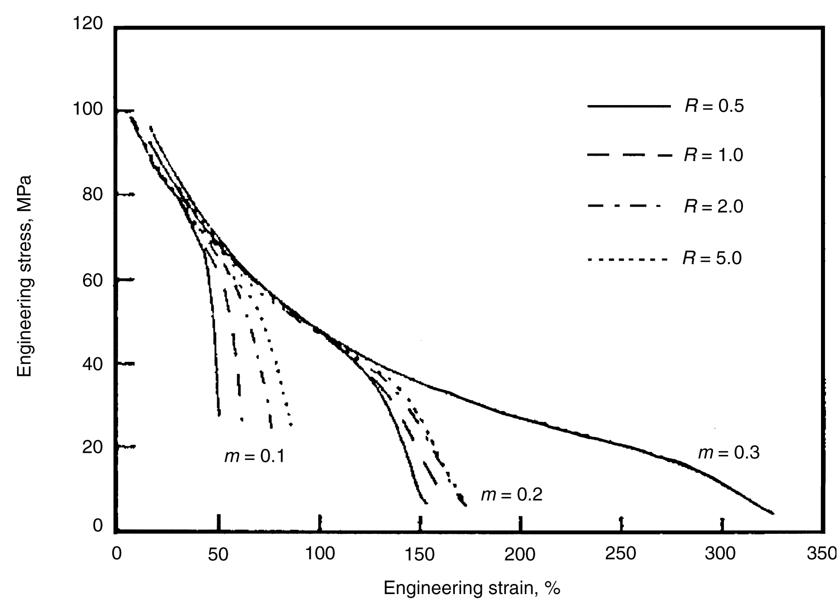


Fig. 26 Direct-equilibrium simulation predictions of engineering stress-strain curves at hot-working temperatures for various values of the strain-rate sensitivity and the normal plastic anisotropy parameter. Source: Ref 35

review of ductile fracture mechanisms is in the article "Mechanisms and Appearances of Ductile and Brittle Fracture in Metals" *Failure Analysis and Prevention*, Volume 11 of the *ASM Handbook* (2002, p 587–626).

Cavity Nucleation. Several possible cavity-nucleation mechanisms have been established including (a) intragranular slip intersections with nondeformable second-phase particles and grain boundaries, (b) sliding of grains along grain boundaries that is not fully accommodated by diffusional transport into those regions, and (c) vacancy condensation on grain boundaries (Ref 40). A frequently used cavity-nucleation criterion based on stress equilibrium at the cavity interface is:

$$r_c = 2(\gamma + \gamma_p - \gamma_i)/\sigma \qquad \text{(Eq 17)}$$

in which r_c is the critical cavity radius above which a cavity is stable; γ, γ_p, and γ_i denote the interfacial energies of the void, the particle, and the particle-matrix interface, respectively; and σ is the applied stress. This criterion requires flow hardening, which is minimal in superplastic materials except in cases of significant grain growth, in order to nucleate cavities at less favorable sites, such as smaller particles. In addition, such surface-energy considerations require stresses for initiation and early growth that are unrealistically high. Therefore, the development of other (constrained-plasticity) approaches based on nucleation and growth from inhomogeneities/regions of high local stress triaxiality has been undertaken (Ref 41).

The cavity-nucleation rate N is defined as the number of cavities nucleated per unit area and unit strain. N may either be constant or decrease or increase with strain. However, such strain dependencies are usually not strong. Measurements have shown than N can be bracketed between 10^4 and 10^6 cavities per mm^3 per unit increment of strain (Ref 16, 41, 42).

The constrained plasticity analysis suggests that a size distribution of second-phase constituents/imperfections may lead to a variety of cavity-growth rates at the nano/submicron cavity-size level. From an operational standpoint, this effect may thus lead one to conclude that cavities nucleate continuously rather than merely become microscopically observable continuously. Irrespective of the exact mechanism, it is thought that the assumption of continuous nucleation of cavities of a certain size (e.g., 1 μm) still produces the same "mechanical" effect on failure via cavitation or flow localization as the postulated actual physical phenomenon.

Cavity-growth mechanisms can be classified into two broad categories: diffusional growth and plasticity-controlled growth (Ref 43). Diffusional growth dominates when the cavity size is very small. As cavity size increases, diffusional growth decreases very quickly, and plastic flow of the surrounding matrix becomes the dominant cavity-growth mechanism. An illustration of a cavity-growth-mechanism map (Ref 44) is shown in Fig. 28. From an engineering viewpoint, plasticity controlled growth is of greatest interest. In such cases, the growth of an isolated, noninteracting cavity is described for the case of uniaxial tension deformation by:

$$V = V_o \exp(\eta\varepsilon)$$

or

$$r = r_o \exp\left(\frac{\eta}{3}\varepsilon\right) \qquad \text{(Eq 18)}$$

in which V and r are the cavity volume and radius, respectively, V_o and r_o are the volume and radius of the cavity when it becomes stable, ε denotes axial true strain, and η is the individual cavity-growth rate.

Several analyses have been conducted to correlate the cavity-growth rate η with material parameters and the deformation conditions. For example, Cocks and Ashby (Ref 45) derived the following relation between η and m for a planar array of spherical, noninteracting, grain-boundary cavities under tensile straining conditions:

$$\eta = 1.5\left(\frac{m+1}{m}\right)\sinh\left[\frac{2}{3}\frac{(2-m)}{(2+m)}\right] \qquad \text{(Eq 19)}$$

It should be noted that this theoretical relationship between m and the cavity-growth parameter η for an individual cavity follows the same general trend as the experimentally determined correlation between the strain-rate sensitivity and the *apparent* cavity-growth rate η_{APP}. The parameter η_{APP}, which is readily derived from experimental data (Ref 46), is defined by:

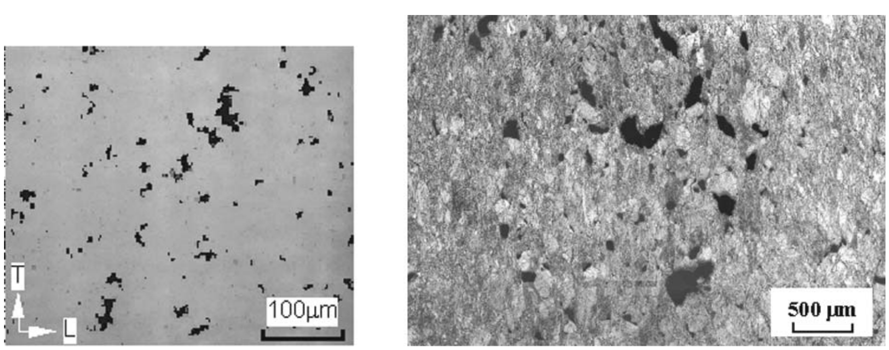
$$C_v = C_{v_o} \exp(\eta_{APP}\,\varepsilon) \qquad \text{(Eq 20)}$$

in which C_v is the cavity volume fraction at a true strain ε, η_{APP} is the apparent cavity-growth rate, and C_{v_o} is the so-called initial cavity volume fraction.

Cavity coalescence is the interlinkage of neighboring cavities due to a microscopic flow-localization process within the material ligament between them. Coalescence occurs when the width of the material ligament reaches a critical value that depends on initial cavity spacing and the strain-rate sensitivity. Coalescence can occur along both the longitudinal and transverse directions with the latter being more important because it eventually leads to failure. According to Pilling (Ref 47), cavity coalescence may be regarded as a process that in effect increases the mean cavity-growth rate. In particular, the effect of pairwise coalescence on the average cavity-growth rate $d\bar{r}/d\varepsilon$ can be estimated from:

$$\frac{d\bar{r}}{d\varepsilon} = \frac{8C_v\Phi\eta\left(0.13r - 0.37(dr/d\varepsilon)_i\,\delta\varepsilon\right) + (dr/d\varepsilon)_i}{1 - 4C_v\Phi\eta\delta\varepsilon} \qquad \text{(Eq 21)}$$

in which C_v is the instantaneous volume fraction of cavities, η is defined from Eq 18, $\delta\varepsilon$ ($= d\varepsilon$) is



(a) (b)

Fig. 27 Examples of cavitation. (a) In aluminum (Al-7475) alloy. Courtesy of A.K. Ghosh. (b) In titanium (Ti-6Al-4V) alloy. Source: Ref 37

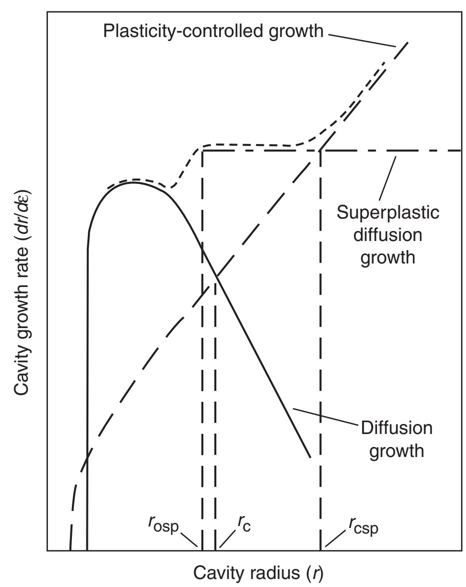


Fig. 28 Variation of the cavity-growth rate for different mechanisms. r_c, critical cavity radius; r_{osp}, cavity radius for onset of superplastic deformation; r_{csp}, critical cavity radius for superplastic deformation. Source: Ref 44

a small increment of strain, $(dr/d\varepsilon)_i$ is the rate of growth per unit strain of an isolated cavity ($=\eta r/3$ from Eq 18), and Φ is given by:

$$\Phi = (1 + \eta\delta\varepsilon/3 + (\eta\delta\varepsilon)^2/27) \quad \text{(Eq 22)}$$

The phenomenon of cavity coalescence was further investigated by Nicolaou and Semiatin (Ref 48, 49), who conducted a numerical analysis of the uniaxial tension test considering: the temporal and spatial location of the cavities inside the specimen and the temporal cavity radius. Two cases were considered: a stationary cavity array (similar to the analysis of Pilling) and continuous cavity nucleation. The analysis of the stationary cavity array led to a much simpler expression than Eq 21, that is:

$$\frac{d\bar{r}}{d\varepsilon} = \eta\bar{r}\left(\tfrac{1}{3} + C_v\right) \quad \text{(Eq 23)}$$

This simple equation gives predictions very similar to the more complex relation of Eq 21.

With regard to the continuous cavity nucleation case, it was found that the average cavity radius was described by:

$$\frac{d\bar{r}}{d\varepsilon} = \eta\bar{r}\left(0.2 + C_v\right) \quad \text{(Eq 24)}$$

A comparison with experimental cavity size measurements (e.g., Fig. 29) revealed that actual results are bounded by cavity-growth-and-coalescence models that assume either a constant, continuous nucleation rate (lower limit) or a preexisting cavity array with no nucleation of new cavities (upper limit), that is, Eq 24 and 23, respectively.

Stress–Strain Curves

The work of Nicolaou et al. (Ref 33) also shed light on the effect of cavitation on stress-strain behavior. Engineering stress-strain (S-e) curves for a range of strain-rate sensitivities (m values) and cavity-growth rates η_{APP} were predicted using the direct-equilibrium modeling approach (Fig. 30); for all of the cases, the strain-hardening exponent n was 0. The cavity volume fraction (C_V) in the central element at failure is also indicated in the plots.

Examination of the engineering stress-strain curves reveals that cavitation causes a noticeable reduction in total elongation; this reduction is quantified and discussed in more detail in the next section. Figure 30 also shows that the stress-strain curves for cavitating and noncavitating samples with the same value of m are very close to each other, except at elongations close to failure.

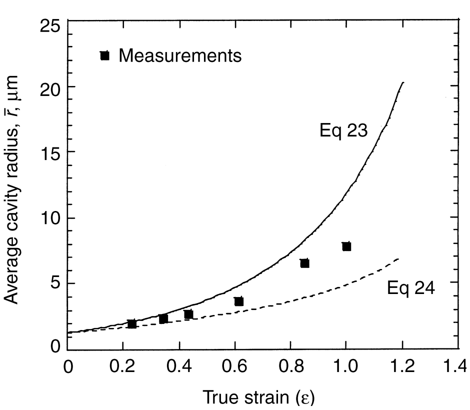


Fig. 29 Comparison of measurements and predictions of the evolution of average cavity radius with strain for an Al-7475 alloy assuming continuous nucleation (Eq 24) or a preexisting cavity array (Eq 23). After Ref 49

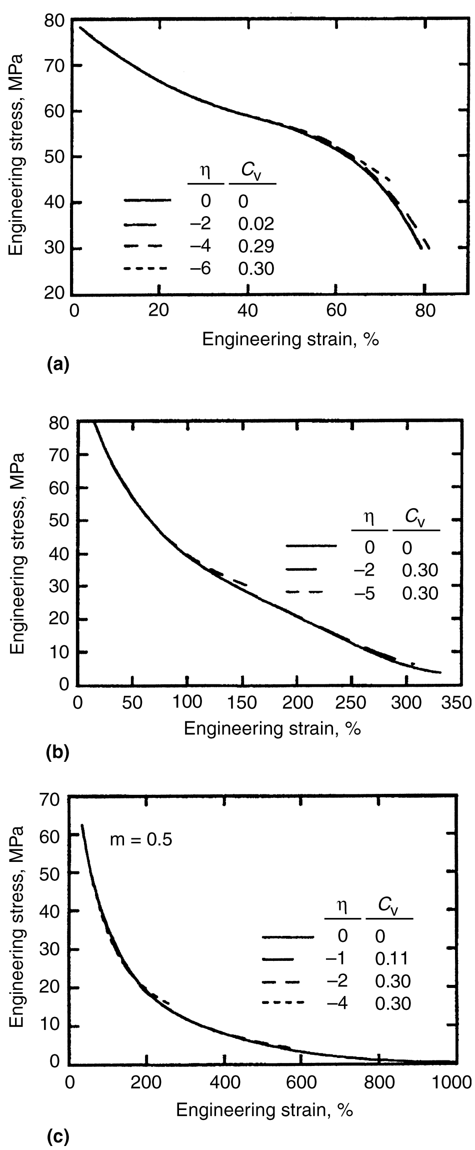


Fig. 30 Predicted engineering stress-strain curves for tension testing of sheet samples with a 2% taper, assuming strain-hardening exponent $n = 0$, initial cavity volume fraction $C_{v_0} = 10^{-3}$, various cavity-growth rates η, and a strain-rate sensitivity exponent m equal to (a) 0.1, (b) 0.3, or (c) 0.5. After Ref 33

Surprisingly, the engineering stress at a given elongation for a cavitating material is *higher* than the corresponding stress (at the same elongation) for a noncavitating material. This intuitively unexpected result was interpreted by the examination of the effective (load-bearing) area at the same elongation of a cavitating and a noncavitating specimen. In particular, the analysis of Nicolaou et al. (Ref 33) revealed that for a given elongation the effective area of the cavitating specimen is larger than the area of a noncavitating one. Therefore, the load and hence the engineering stress required to sustain deformation is higher in the case of a cavitating material.

As shown in Fig. 30, the difference between the engineering stress-strain curves of cavitating and noncavitating materials is not very large. Therefore, it can be deduced that Considere's criterion, if implemented in the usual fashion using data from a tension test (i.e., a plot of load versus the elongation of the gage section), can be used to test whether fracture of a tensile specimen occurs due to instability, regardless of the presence of extensive internal cavities in the material and whether the volume of the material is conserved (Ref 50).

Failure Modes During Hot-Tension Testing

Cavitating hot-tension specimens may fail by either localized necking ("flow localization") or fracture/cavitation. The second mode of failure occurs without flow localization in the neck and resembles a brittle type of fracture because the fracture tip has a considerable area. Micrographs of these modes of failure are presented in Fig. 31. The localization type of failure shown in Fig. 31(a) is for an orthorhombic titanium aluminide (Ti-21Al-22Nb) deformed at 980 °C (1795 °F) and a nominal strain rate of 1.6×10^{-3} s^{-1}. On the other hand, Fig. 31(b) displays the fracture-controlled failure of a γ titanium aluminide alloy (Ref 51) deformed in tension at 1200 °C (2190 °F) and a nominal strain rate of 10^{-3} s^{-1}.

The particular mode of failure of a material tested under tension conditions can be predicted by the magnitude of the strain-rate sensitivity m and the apparent cavity-growth rate η_{APP}. The corresponding failure-mechanism map for non-strain-hardening materials is plotted in Fig. 32. For deformation under superplastic conditions ($m > 0.3$) and $\eta_{APP} > 2$, the map shows that failure is fracture/cavitation-controlled. On the other hand, flow-localization-controlled failure is seen to predominate only for small values of the cavity-growth rate. In Fig. 32, experimental observations of the failure mode of γ and orthorhombic titanium aluminides are also plotted. The solid data points correspond to fracture-controlled failures, while the open data points correspond to localization-controlled failures. Given the assumptions made in deriving such maps, it can be concluded that the prediction of failure mode from the magnitudes of m and η_{APP} provides good agreement with actual behavior.

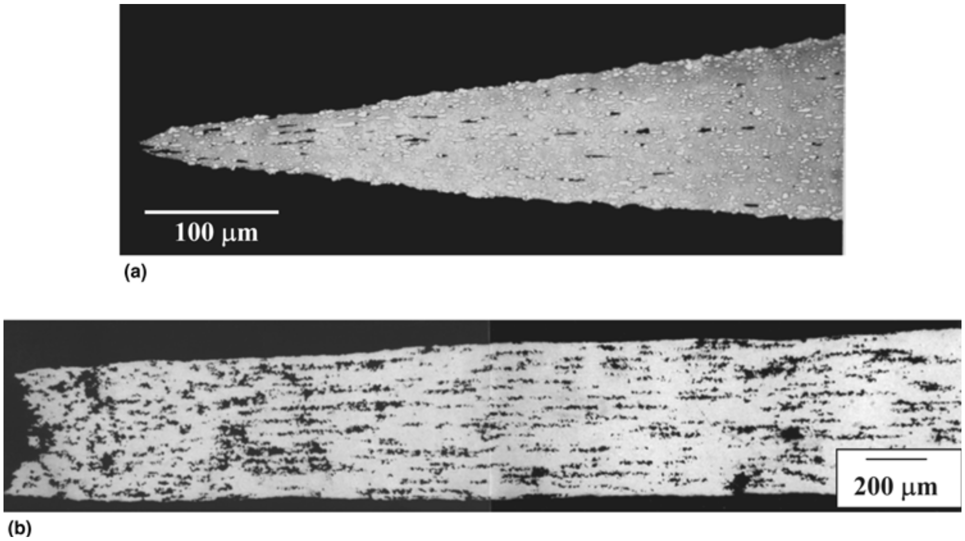


Fig. 31 Micrographs. Orthorhombic titanium aluminide that failed in tension by flow localization. Source: Ref 10. (b) Near-γ titanium aluminide that failed in tension by fracture (cavitation). Source: Ref 51

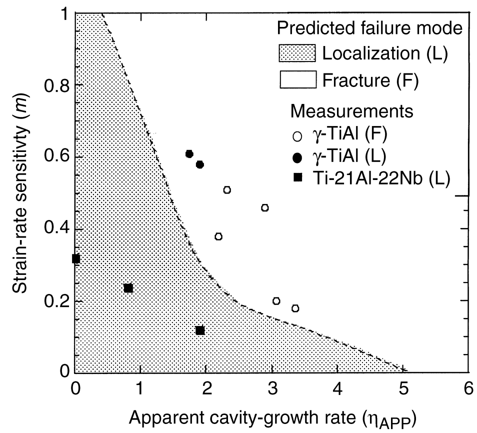


Fig. 32 Failure-mode map developed from simulations of the sheet tension test. Experimental data points are also shown on the map.

Total Tensile Elongations

As shown in Fig. 33, cavitation may lead to premature failure and thus to a significant reduction in the tensile elongation compared to that measured by Woodford (Ref 18) for noncavitating metals. For a fixed value of m, the reduction in elongation for fracture-controlled failure depends on several factors, such as the cavity-nucleation rate, cavity-growth rate, cavity shape and distribution, and the cavity architecture.

Several analyses have been conducted to quantify the effect of cavitation on the tensile ductility. These include the two-slice approach* by Lian and Suery (Ref 52), micromechanical approaches by Zaki (Ref 53), and Nicolaou and Semiatin (Ref 54), as well as approaches based on the direct-equilibrium approach described in the section "Numerical Modeling of the Hot-Tension Test" in this chapter. Several results from the direct-equilibrium model serve to illustrate the efficacy of such techniques.

The results shown in Fig. 33 correspond to a non-strain-hardening material ($n = 0$) with a 2% taper and $C_{vo} = 10^{-3}$. The topmost curve in this plot depicts the total elongation as a function of m for a noncavitating sample, that is, Woodford's trend line. For such a material, the elongation is controlled, of course, by the onset of localized necking. The remaining curves in Fig. 33 for *cavitating* samples indicate the decrement in elongation due to the occurrence of fracture prior to localized necking. For low values of η_{APP} (<2) and m (<0.3), the decrement is equal to zero because failure is still necking controlled. On the other hand, the decrement is largest for large values of η_{APP} and m, for which the critical volume fraction of cavities for fracture (assumed to be 0.30) is reached much before the elongation at which necking occurs. In fact, for $\eta_{APP} > 5$, the total elongation is almost independent of m for $m > 0.3$ because fracture in these cases intercedes during largely uniform, quasi-stable flow.

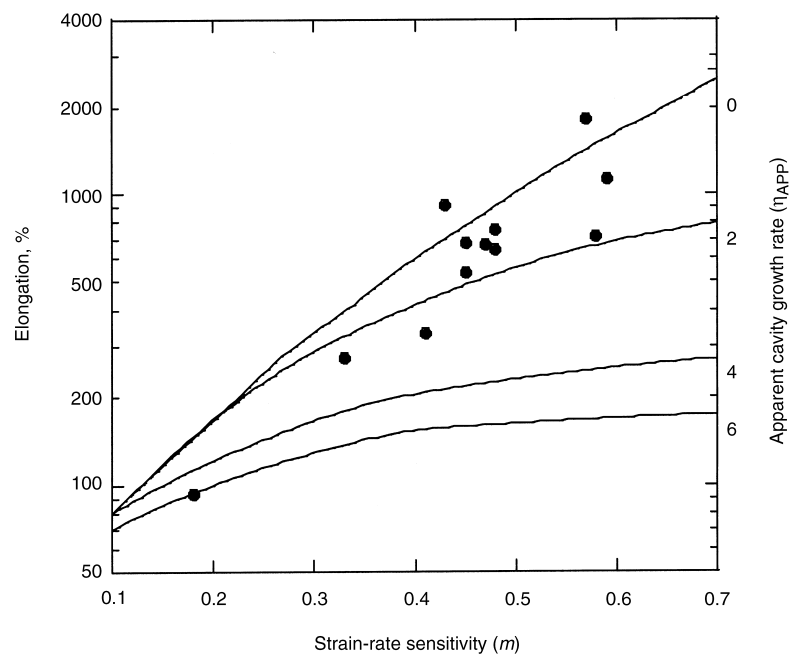


Fig. 33 Elongation as function of the strain-rate sensitivity and (apparent) cavity-growth rate predicted from direct-equilibrium simulations. The individual data points represent experimental data. Source: Ref 33

In most cases, rigorous comparisons of predicted tensile elongations and experimental data (Fig. 33) cannot be made because the cavity-growth rate and the cavity-size population and shape were not measured, while other important parameters such as the specimen geometry were not reported. Therefore, a general comparison based only on the value of m can be made. From the results of Fig. 33, it is seen that most of the data points overlie the predicted curves.

Comparisons of reported tensile elongations data (Table 2) to predictions of a microscopic model (Ref 54) in which the cavity architecture has been taken into account (through the value of G*) are shown in Fig. 34. With the exception of one data point (No. 9), the major deviations of the

*The two-slice approach assumes that the specimen comprises two regions, one of them consisting of the central plane of the specimen that contains an initial geometric or strength defect. The deformation of each region obeys the flow rule while at any instant of deformation the *load* is the same in both regions (slices).

*The parameter G is a factor that describes the geometry of the ligament between two cavities as a function of the cavity architecture within the specimen.

Table 2 Experimental data from the literature for the deformation and failure of cavitating materials

Data point	Material	m	η	Tensile elongation, %
1	γ-TiAl (as received)	0.38	2.2	219
2	γ-TiAl (as received)	0.51	2.3	350
3	γ-TiAl (as received)	0.62	1.8	446, 532
4	γ-TiAl (heat treated)	0.18	3.4	93, 104
5	γ-TiAl (heat treated)	0.15	8.0	51
6	5083 Al	0.50	5.2	172
7	Zn-22Al	0.45	1.5	400
8	α/β brass	0.60	2.3	425
9	Coronze 638	0.33	4.5	275

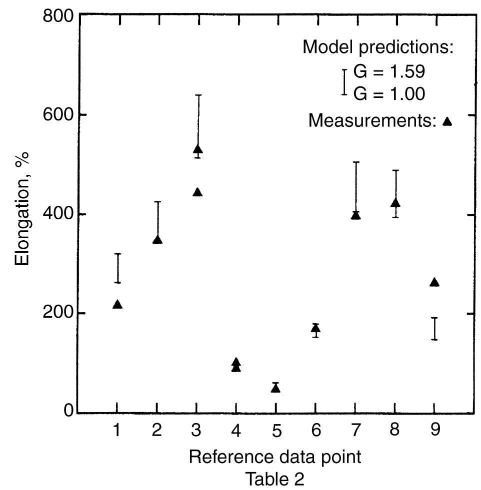


Fig. 34 Comparison of experimentally determined total elongations with (microscopic) model predictions that incorporate the cavity architecture. Source: Ref 54

predictions tend to be on the high side. These deviations could be a result of the neglect in the microscopic model of the macroscopic strain gradient in the diffuse neck of real tension samples. In addition, as mentioned previously, specimen geometry and deformation in the shoulder region of actual tension specimens have an effect on measured ductilities, which is difficult to quantify. Nevertheless, agreement between the measured and predicted ductilities is reasonably good.

REFERENCES

1. *Met. Ind.,* Vol 11, 1963, p 247–249
2. M.G. Cockroft and D.J. Latham, Ductility and the Workability of Metals, *J. Inst. Met.,* Vol 96, 1968, p 33–39
3. *J. Inst. Met.,* Vol 92, 1964, p 254–256
4. G. Cusminsky and F. Ellis, An Investigation into the Influence of Edge Shape on Cracking During Rolling, *J. Inst. Met.,* Vol 95 (No. 2), Feb 1967, p 33–37
5. E.F. Nippes, W.F. Savage, B.J. Bastian, and R.M. Curran, An Investigation of the Hot Ductility of High-Temperature Alloys, *Weld. J.,* Vol 34, April 1955, p 183s–196s
6. D.F. Smith, R.L. Bieber, and B.L. Lake, "A New Technique for Determining the Hot Workability of Ni-Base Superalloys," presented at IMD-TMS-AIME meeting, 22 Oct 1974
7. S.I. Oh, S.L. Semiatin, and J.J. Jonas, An Analysis of the Isothermal Hot Compression Test, *Metall. Trans. A*, Vol 23A (No. 3), March 1992, p 963–975
8. W.C. Unwin, *Proc. Inst. Civil Eng*., Vol 155, 1903, p 170–185
9. P.A. Friedman and A.K. Ghosh, Microstructural Evolution and Superplastic Deformation Behavior of Fine Grain 5083Al, *Metall. Mater. Trans. A,* Vol 27A (No. 12), Dec 1996, p 3827–3839
10. P.D. Nicolaou and S.L. Semiatin, High-Temperature Deformation and Failure of an Orthorhombic Titanium Aluminide Sheet Material, *Metall. Mater. Trans. A.*, Vol 27A (No. 11), Nov 1996, p 3675–3681
11. R. Pilkington, C.W. Willoughby, and J. Barford, The High-Temperature Ductility of Some Low-Alloy Ferritic Steels, *Metal Sci. J.*, Vol 5, Jan 1971, p 1
12. D.J. Latham, *J. Iron Steel Inst*., Vol 92, 1963–1964, p 255
13. R.E. Bailey, *Met. Eng. Quart.*, 1975, p 43–50
14. R.E. Bailey, R.R. Shiring, and R.J. Anderson, *Superalloys Metallurgy and Manufacture*, Proc. of the Third International Conference, Claitor's Publishing, Sept 1976, p 109–118
15. P.D. Nicolaou and S.L. Semiatin, An Investigation of the Effect of Texture on the High-Temperature Flow Behavior of an Orthorhombic Titanium Aluminide Alloy, *Metall. Mater. Trans. A.*, Vol 28A (No. 3A), March 1997, p 885–893
16. J. Pilling and N. Ridley, *Superplasticity in Crystalline Solids*, The Institute of Metals, London, UK, 1989
17. R.M. Imayev and V.M. Imayev, Mechanical Behaviour of TiAl Submicrocrystalline Intermetallic Compound at Elevated Temperatures, *Scr. Met. Mater.*, Vol 25 (No. 9), Sept 1991, p 2041– 2046
18. D.A. Woodford, Strain-Rate Sensitivity as a Measure of Ductility, *Trans. ASM*, vol 62 (No. 1), March 1969, p 291–293
19. A.K. Ghosh and R.A. Ayres, On Reported Anomalies in Relating Strain-Rate Sensitivity (m) to Ductility, *Metall. Trans. A*, Vol 7A, 1976, p 1589–1591
20. J.W. Hutchinson and K.W. Neale, Influence of Strain-Rate Sensitivity on Necking Under Uniaxial Tension, *Acta Metall.*, Vol 25 (No. 8), Aug 1977, p 839–846
21. F.A. Nichols, Plastic Instabilities and Uniaxial Tensile Ductilities, *Acta Mater.*, Vol 28 (No. 6), June 1980, p 663–673
22. M.A. Meyers and K.K. Chawla, *Mechanical Metallurgy Principles and Applications*, Prentice Hall, 1984.
23. N.E. Paton and C.H. Hamilton, Microstructural Influences on Superplasticity in Ti-6Al-4V, *Metall. Trans. A*, Vol 10A, 1979, p 241–250
24. G.E. Dieter, *Mechanical Metallurgy*, 3rd ed., McGraw-Hill, 1986
25. A.K. Ghosh and C.H. Hamilton, Mechanical Behavior and Hardening Characteristics of a Superplastic Ti-6Al-4V Alloy, *Metall. Trans. A*, Vol 10A, 1979, p 699–706
26. R. Verma, P.A. Friedman, A.K. Ghosh, S. Kim, and C. Kim, Characterization of Superplastic Deformation Behavior of a Fine Grain 5083 Al Alloy Sheet, *Metall. Mater. Trans. A*, Vol 27A (No. 7), July 1996, p 1889–1898
27. P.W. Bridgman, *Studies in Large Plastic Flow and Fracture*, McGraw-Hill, 1952
28. A.S. Argon, J. Im, and A. Needleman, Distribution of Plastic Strain and Negative Pressure in Necked Steel and Copper Bars, *Metall. Trans. A*, Vol 6A, 1975, p 815–824
29. C.M. Lombard, R.L. Goetz, and S.L. Semiatin, Numerical Analysis of the Hot Tension Test, *Metall. Trans. A*, Vol 24A (No. 9), Sept 1993, p 2039–2047
30. C. G'sell, N.A. Aly-Helal, and J.J. Jonas, *J. Mater. Sci.*, Vol 18, 1983, p 1731–1742
31. A.K. Ghosh, A Numerical Analysis of the Tensile Test for Sheet Metals, *Metall. Trans. A*, Vol 8A, 1977, p 1221–1232
32. S.L. Semiatin, A.K. Ghosh, and J.J. Jonas, A "Hydrogen Partitioning" Model for Hydrogen Assisted Crack Growth, *Metall. Trans. A*, Vol 16A, 1985, p 2039–47
33. P.D. Nicolaou, S.L. Semiatin, and C.M. Lombard, Simulation of the Hot-Tension Test under Cavitating Conditions, *Metall. Mater. Trans. A.*, Vol 27A (No. 10), Oct 1996, p 3112–3119
34. S.L. Semiatin, R.E. Dutton, and S. Shamasundar, Materials Modeling for the Hot Consolidation of Metal Powders and Metal-Matrix Composites, *Processing and Fabrication of Advanced Materials IV*, T.S. Srivatsan and J.J. Moore, Ed., The Minerals, Metals, and Materials Society, 1996, p 39–52
35. P.D. Nicolaou and S.L. Semiatin, *Scr. Mater.*,Vol 36, 1997, p 83–88
36. M.M.I. Ahmed and T.G. Langdon, Exceptional Ductility in the Superplastic Pb-62 Pct Sn Eutectic, *Metall. Trans. A*, Vol 8A, 1977, p 1832–1833
37. S.L. Semiatin, V. Seetharaman, A.K. Ghosh, E.B. Shell, M.P. Simon, and P.N. Fagin,

Cavitation During Hot Tension Testing of Ti-6Al-4V, *Mater. Sci. Eng. A*, Vol A256, 1998, p 92–110
38. C.C. Bampton and J.W. Edington, The Effect of Superplastic Deformation on Subsequent Service Properties of Fine-Grained 7475 Aluminum, *J. Eng. Mater. Tech*., Vol 105 (No. 1), Jan 1983, p 55–60
39. J. Pilling and N. Ridley, Cavitation in Aluminium Alloys During Superplastic Flow, *Superplasticity in Aerospace*, H.C. Heikkenen and T.R. McNelley, The Metallurgical Society, 1988, p 183–197
40. G.H. Edward and M.F. Ashby, Intergranular Fracture During Powder-Law Creep, *Acta Metall*., Vol 27, 1979, p 1505–1518
41. A.K. Ghosh, D.-H. Bae, and S.L. Semiatin, Initiation and Early Stages of Cavity Growth During Superplastic and Hot Deformation, *Mater. Sci. Forum*, Vol 304–306, 1999, p 609–616
42. S. Sagat and D.M.R. Taplin, Fracture of a Superplastic Ternary Brass, *Acta Metall*.,Vol 24 (No. 4), April 1976, p 307–315
43. A.H. Chokshi, The Development of Cavity Growth Maps for Superplastic Materials, *J. Mater. Sci.*, Vol 21, 1986, p 2073–2082
44. B.P. Kashyap and A.K. Mukherjee, Cavitation Behavior During High Temperature Deformation of Micrograined Superplastic Materials—A Review, *Res Mech.*, Vol 17, 1986, p 293–355
45. A.C.F. Cocks and M.F. Ashby, Creep Fracture by Coupled Power-Law Creep and Diffusion Under Multiaxial Stress, *Met. Sci*., Vol 16, 1982, p 465–478
46. P.D. Nicolaou, S.L. Semiatin, and A.K. Ghosh, An Analysis of the Effect of Cavity Nucleation Rate and Cavity Coalescence on the Tensile Behavior of Superplastic Materials, *Metall. Mater. Trans. A*, Vol 31A, 2000, p 1425–1434
47. J. Pilling, Effect of Coalescence on Cavity Growth During Superplastic Deformation, *Mater. Sci. Technol*., Vol 1 (No. 6), June 1985, p 461–466
48. P.D. Nicolaou and S.L. Semiatin, Modeling of Cavity Coalescence During Tensile Deformation, *Acta Mater*, Vol 47, 1999, p 3679–3686
49. P.D. Nicolaou and S.L. Semiatin, The Influence of Plastic Hardening on Surface Deformation Modes Around Vickers and Spherical Indents, *Acta Mater*, Vol 48, 2000, p 3441–3450
50. L. Weber, M. Kouzeli, C. San Marchi, and A. Mortensen, On the Use of Considere's Criterion in Tensile Testing of Materials Which Accumulate Internal Damage, *Scr. Mater.* Vol 41, 1999, p 549–551
51. C.M. Lombard, "Superplasticity in Near-Gamma Titanium Aluminides," Ph.D. Thesis, University of Michigan, Ann Arbor, MI, 2001
52. J. Lian and M. Suery, Effect of Strain Rate Sensitivity and Cavity Growth Rate on Failure of Superplastic Material, *Mater. Sci. Technol*., Vol 2, 1986, p 1093–1098
53. M. Zaki, Micronecking and Fracture in Cavitated and Superplastic Materials, *Metall. Mater. Trans. A*, Vol 27A, 1996, p 1043–1046
54. P.D. Nicolaou and S.L. Semiatin, A Theoretical Investigation of the Effect of Material Properties and Cavity Architecture/Shape on Ductile Failure During the Hot Tension Test, *Metall. Mater. Trans. A.*, Vol 29A, 1998, p 2621–2630

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 86-121
DOI:10.1361/hwpd2003p086

Chapter 8

Torsion Testing to Assess Bulk Workability

S.L. Semiatin, Air Force Research Laboratory, Materials and Manufacturing Directorate
J.J. Jonas, McGill University

WORKABILITY is generally defined as the ability to impart a particular shape to a piece of metal under the load capacity of the available tooling and equipment, without the introduction of fracture or the development of undesirable microstructures. A complete description of the workability of a material is therefore specified by its flow-stress dependence on processing variables (e.g., strain, strain rate, preheat temperature, and die temperature), its failure behavior, and the phase transformations that characterize the alloy system to which it belongs.

Very few mechanical tests are capable of providing information about all of these aspects, which is primarily a result of the large deformations that are common in massive forming processes such as forging, extrusion, and rolling. For most common mechanical tests—simple tension and compression tests—the maximum uniform strains achieved are rather low because of necking and barreling, respectively. By contrast, strains in excess of 0.3 to 0.7, the levels typical of these uniaxial tests, are readily achieved in the torsion test.

The torsion test has been used for some 50 years as a means of hot workability assessment in metals and alloys (Ref 1–5). The principal virtue of this straining mode is the fact that the test-pieces do not undergo significant shape change as they are deformed as long as the gage section is restrained to a fixed length. Because of the large strains that can be achieved under relatively uniform deformation conditions, the torsion test is the preferred test for obtaining flow-stress data and unambiguous indications of failure and microstructural response from deformation processing (Ref 6, 7). Provided failure does not intercede, it is not unusual to be able to obtain flow-stress results in torsion to deformation levels equivalent to a true axial strain of 5 or more in tension, or a reduction of 90 to 95% in compression. Another attractive feature of the torsion test is that a constant true strain rate can be imposed on any given annular region of the specimen by simply twisting one end relative to the other at a constant angular velocity. By contrast, more complex velocity control is required in axial loading tests. The interpretation of torsion-test data, however, is more complex than that used for axial-testing procedures because the strain and strain rate vary along the specimen radius.

When failure modes are important, torsion is also a valuable diagnostic test. These failures are usually divided into two broad categories: fracture-controlled failures, in which deformation is relatively uniformly distributed prior to failure, and flow-localization-controlled failures, in which plastic deformation has been concentrated in a particular area of the metal specimen or workpiece prior to the actual fracturing process.

Fracture of metals occurs by a variety of mechanisms that depend largely on deformation rate and temperature. At cold working temperatures, or temperatures corresponding approximately to one-fourth or less of the absolute melting or solidus temperature ($T/T_M \leq 0.25$), ductile fracture, which is characterized by void initiation at second-phase particles and inclusions, void growth, and final coalescence, is most prevalent. At warm ($0.25 \leq T/T_M \leq 0.6$) and hot ($T/T_M \geq 0.6$) working temperatures, processes such as wedge cracking (the opening of voids and their propagation at triple junctions, or the intersection of three grain boundaries) and cavitation (the formation of voids or cavities around second-phase particles, particularly at grain boundaries) are the most important fracture mechanisms. Frequently, the deformation rate determines which of these fracture mechanisms predominates. Typically, the former mechanism occurs at high strain rates, and the latter occurs at low strain rates.

The different fracture regimes are best illustrated by deformation-processing maps (Fig. 1), which are frequently determined by tension or torsion tests. Such maps may also depict regions in which large strains may be achieved without fracture, or so-called "safe" processing regimes, as well as those in which certain microstructural changes, such as dynamic recovery or recrystallization, can be expected during processing.

The torsion test can also be used to determine conditions under which flow-localization-controlled failures occur. These failures are manifested by shear bands or regions of localized shear deformation because torsion imposes a mode of simple shear (Fig. 2). During torsion, shear bands may originate at a point of material

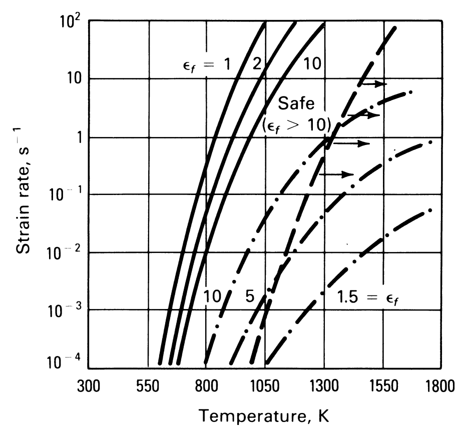


Fig. 1 Typical deformation-processing map for austenitic stainless steel, showings regions of ductile fracture, wedge cracking, dynamic recrystallization, and "safe" forming. Note that the boundaries for safe forming (i.e., the loci of processing conditions between ductile fracture and wedge cracking) vary depending on the required ductility (or fracture strain ε_f). Dash-dot line, wedge cracking; solid line, ductile fracture; dashed line, dynamic recrystallization. Source: Ref 8

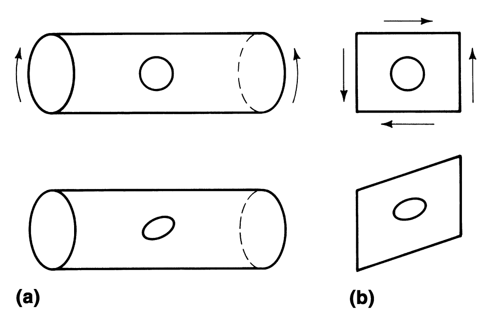


Fig. 2 Analogy between (a) torsion of a thin-wall tube and (b) simple shear. Note the equivalence between the deformations of initially circular grid elements in the two modes.

inhomogeneity or temperature nonuniformity. Under the former conditions, localization is driven by flow softening (a decrease of flow stress with increasing strain) due to:

- Dynamic recovery or dynamic recrystallization
- Microstructural instabilities, such as the breakup and spheroidization of lamellar microstructures
- Texture softening
- Deformation heating

When deformation heating occurs, very high strain rates prevent the dissipation of heat, which in conjunction with the large negative temperature dependence of flow stress found in many metals leads to particularly strong flow localizations sometimes referred to as adiabatic shear bands. At moderate strain rates, deformation heat generation and heat conduction can introduce substantial temperature gradients during torsion testing, as in conventional metalworking processes, and can cause noticeable flow localization in materials with flow stresses that are strongly temperature dependent.

Because failure mode and microstructural development depend on process variables, torsion-testing equipment and procedures to assess workability must be carefully designed. The ability to control and measure strain and strain rate (through monitoring of twist and twist rate) is important, as is test temperature. Furthermore, the ability to impose arbitrary deformation or loading histories should be considered when designing torsion apparatus to be used to simulate processes such as multistage rolling or wire drawing.

This chapter discusses equipment design, procedures, and interpretation of torsion tests to establish workability. Experimental considerations are discussed, along with the application of torsion testing to obtain flow-stress data and to gage fracture-controlled workability. Use of the torsion test to gage flow-localization-controlled failure and the effect of processing on microstructure using torsion is detailed as well.

Material Considerations

The microstructure of the material to be tested in torsion greatly affects specimen design and the interpretation of workability data. The most significant material characteristics are grain size, crystallographic texture, and mechanical texture. Because metals are composed of individual grains with specific crystallographic orientations and directional plastic deformation properties, the torsion-specimen gage-section diameter should include at least 20 to 30 grain diameters. Torsion measurement thus reflects an average of the flow and failure response of the material. This requirement on gage-section diameter may therefore result in specimens of varying sizes, depending on whether they are cut from fine-grained wrought products or coarse-grained cast products.

The presence of crystallographic texture must also be taken into consideration. In specimens taken from wrought bar stock, for example, wire textures (a specific crystallographic direction parallel to the bar axis) are common. The strong textures in plate materials should also be noted when torsion specimens are fabricated from these materials. Samples taken from forgings or extrusions often exhibit preferred crystallographic orientations, which also may vary from one location to another.

Mechanical texture, or the presence of inclusions, grain boundaries, or other microstructural features with a preferred direction, should also be taken into account when designing a torsion specimen. Such texture, or fibering as it is sometimes called, can greatly affect the torsional ductility, depending on the orientation relationship between the torsion axis and the mechanical "fiber." The relationship between ductility and mechanical texture is analogous to the relationship between these two properties in tensile tests performed on specimens cut from the rolling direction (higher ductility) and transverse direction (lower ductility) of rolled plate.

The mechanical behavior in torsion testing is somewhat more complex, however, because of the tendency of the fiber to rotate as twisting proceeds. Consider, for example, the torsion of a solid round bar in which inclusions are initially in the form of stringers parallel to the torsion axis (Fig. 3). After twisting, the inclusions at the center of the bar have not rotated, but those at the surface have rotated to form an angle, the magnitude of which depends on the gage length and diameter and the amount of twist. It is apparent that the generation of such a mechanical texture can substantially reorient the planes of weakness into the form of a "wolf's ear," as evidenced by tensile tests on bars prestrained in torsion (Fig. 4). Rotation of the mechanical texture and the general level of inclusions may have a significant effect on the torsional fracture strain. Because of this, torsion testing is frequently used to uncover lot-to-lot variations in workability for a given material.

Specimen Design

Although the torsion test has been used for many years, there are still no standards for specimen design. Typical specimen geometries are shown in Fig. 5. A certain sense of standardization does exist inasmuch as many researchers have used specimens with a gage-length-to-radius ratio of approximately 8 to 1 (Ref 10–31), although data have been reported for specimens with ratios from 0.67 to 1 (Ref 32) to 17 to 1 (Ref 33). Many have chosen to use solid-bar speci-

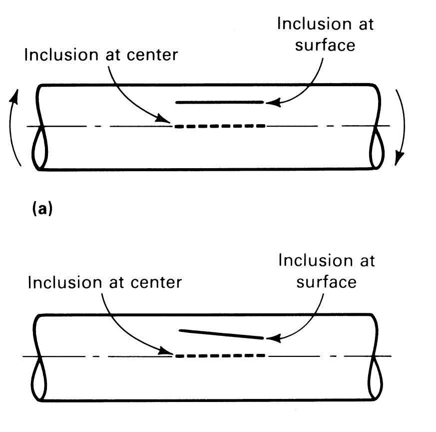


Fig. 3 Torsion specimen with two typical inclusions at the center and the surface. (a) Prior to twisting. (b) After twisting

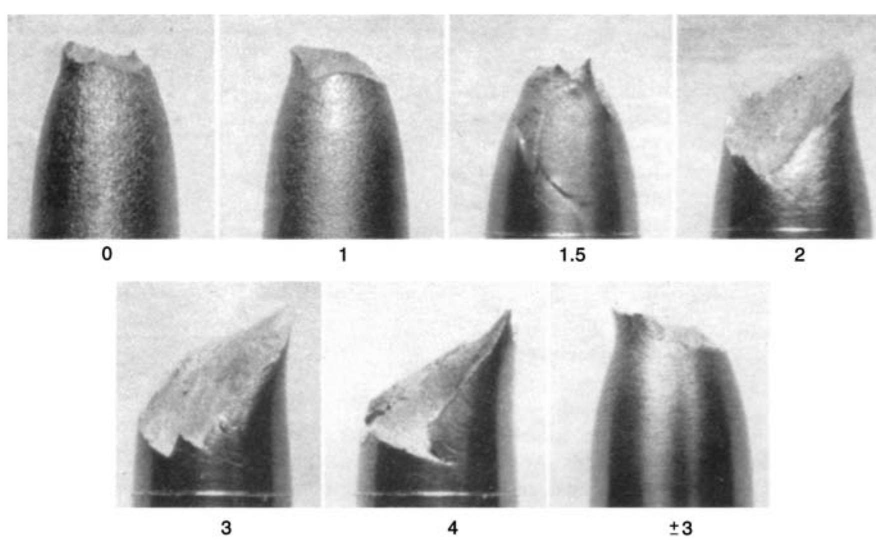


Fig. 4 Tensile fracture in torsionally prestrained copper specimens. The shear strain in the surface of each specimen is indicated below the photograph. Specimen marked ±3 shows the tensile fracture after twisting to a shear strain of 3 and then completely untwisting. Source: Ref 9

mens, some with a reduced gage section and transition fillets, whereas others have preferred tubular specimens. Solid specimens are favored for high-strain studies because thin-wall tubes have a tendency to buckle in this regime. Tubular specimens, on the other hand, have a distinct advantage from the point of view of the analysis for stress and strain and are particularly valuable for studies of work hardening at low strains.

The most important components of design are gage section, fillet radius, shoulder length, and grip design. Gage-section geometry determines the deformation level and deformation rate for a given amount of twist and a given twist rate, based on:

$$\gamma = \frac{r\theta}{L} \quad \text{(Eq 1)}$$

and

$$\dot{\gamma} = \frac{r\dot{\theta}}{L} \quad \text{(Eq 2)}$$

where γ is the engineering shear strain (equal to twice the tensorial shear strain or $\sqrt{3}$ times the von Mises effective strain), θ is twist angle (in radians), r is the radial position, L is the gage length, $\dot{\gamma}$ is the engineering-shear-strain rate, and $\dot{\theta}$ is the twist rate (in radians per unit time).

For a given amount of twist, large values of r and small values of L promote high values of γ. Similarly, large r and small L values result in large values of $\dot{\gamma}$ for a given $\dot{\theta}$. Values for γ and $\dot{\gamma}$ vary with radius in a solid bar or thick-wall tube; they are greatest at the surface and 0 at the center. Strain and strain rate are typically reported as the values that pertain to the surface of the specimen. When torsion data are compared to measurements obtained from other deformation modes, it is this strain and this strain rate that are usually important.

For wrought materials, typical geometry consists of a gage diameter of 6.5 mm (0.25 in.) and a length ranging from 6.5 to 50 mm (0.25 to 2.0 in.) for solid-bar samples. The shorter lengths are used to obtain higher strain rates. The gage length selected also influences the fillet radius that should be used in the torsion specimen because some deformation usually occurs in this portion of the specimen as well. When the gage length is short (e.g., 6.5 mm, or 0.25 in.), a sharp fillet radius (e.g., 0.25 mm, or 0.010 in.), which is blended to the shoulder area by a sharp conical taper of about 120° included angle, is often used to prevent excessive deformation outside the gage section. For longer gage-length specimens, a substantially larger fillet radii (e.g., 4 mm, or 0.15 in.) can be used.

Thin-wall tubular specimens provide an advantage over solid-bar specimens in that shear-strain and shear-strain-rate gradients are virtually eliminated in the design. This is desirable in workability testing because ambiguities associated with definition of the failure strain and strain rate are eliminated. Moreover, the use of thin-wall tubes eliminates the variation of temperature across the section thickness that occurs in solid bars twisted at high rates. These temperature variations are a consequence of the variations in shear strain and thus deformation work converted into heat. Thin-wall specimens must have a rather small length-to-diameter ratio to avoid buckling. Even small amounts of buckling can noticeably affect the magnitude of the plastic properties deduced from the hot torsion test (Ref 34). Typical dimensions for thin-wall specimens are 12.5 mm (0.50 in.) outer diameter, 11 mm (0.44 in.) inner diameter, and a gage length of 2.5 mm (0.10 in.).

Barraclough et al. (Ref 35) have examined the effects of variations in specimen geometry on the stress-strain curves obtained, as well as on microstructure and ductility data. They demonstrated that, for aluminum and silicon steel, which only exhibit dynamic recovery during straining, the equivalent stress-strain curves are reproducible to within normal experimental scatter when the length-to-radius ratio is greater than 2 to 1. It was also shown that the strain to maximum stress in types 304 and 321 stainless steels at temperatures at which they dynamically recrystallize decreases as the length-to-radius ratio increases up to 2 to 1 (Fig. 6). They attributed this latter effect to an increase in the homogeneity of recrystallization in the longer specimens. It was also noted that significant deformation occurs in the fillets between the gage section and the shoulders, which induces an error in strain and strain-rate measurement. If the fillet radius is the same for specimens of different length-to-radius ratios, the error in strain increases with a decrease in the ratio.

The effect of bore diameter in tubular specimens was examined by Barraclough et al. (Ref 35) and was found not to be of critical importance in determining the strain to maximum stress for type 304 stainless steel (Fig. 7). The choice of bore diameter, however, is influential in determining the strain to failure. This behavior is discussed in more detail later in this chapter.

In both solid bars and tubular specimens, the gage-length-to-diameter ratio may also have a marked effect on the actual specimen tempera-

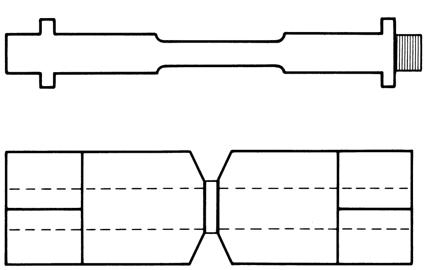


Fig. 5 Typical torsion specimen geometries used for workability testing. See text for discussion of dimensions.

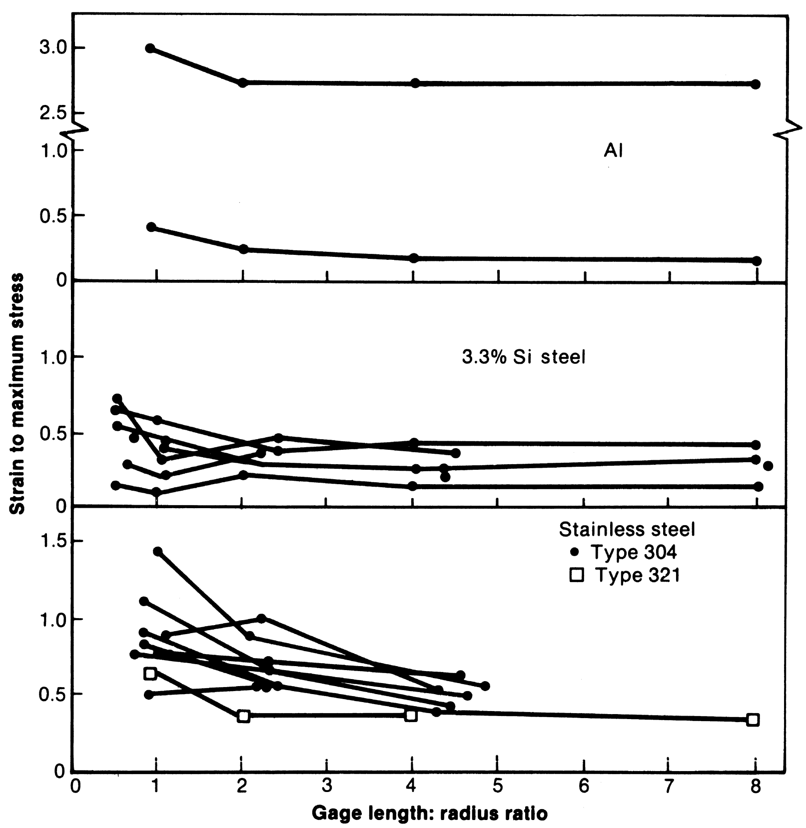


Fig. 6 Effect of gage-length-to-radius ratio on strain (ε_m) to maximum flow stress or to the onset of steady-state flow. The solid lines join data points obtained at similar strain rates and temperatures. Source: Ref 35

ture during moderate-speed ($\dot{\gamma} = 10^{-2}$ to 10 s^{-1}) torsion tests due to the effects of heat conduction. Because of this, flow curves derived from data obtained at these rates tend to show a dependence on the length-to-diameter ratio (*L*/*d*). Flow curves for large *L*/*d* specimens tend to fall below those for small *L*/*d* ratios, in which most of the deformation heat is dissipated into the shoulders (Fig. 8). Interpretation of fracture strain data from such tests should take into account not only the nominal (initial) test temperature, but also the temperature history during the test.

When workability is limited by flow localization, gage length can noticeably affect the failure process as well, particularly at moderate deformation rates. At these rates, heat conduction along the length of the gage section and into the shoulder can set up a substantial temperature gradient, the magnitude of which depends on the gage length, shoulder size, and deformation rate. In turn, the temperature gradient influences the deformation resistance along the length of the bar and hence the strain profile.

The design of the shoulder and grip ends of torsion specimens is determined largely by the method of heating and the type of torsion machine to be used. In all cases, the shoulder diameter should be at least one-and-one-half times the gage-section diameter and preferably two to three times as large to prevent plastic deformation as well as to minimize elastic distortion. Shoulder length can be very short for torsion tests conducted at room temperature. If testing is to be carried out at elevated temperature, the shoulder should be longer to ensure temperature uniformity, typically about 25 mm (1.0 in.) for specimens heated by induction and 25 to 50 mm (1.0 to 2.0 in.) for specimens heated in furnaces.

An interesting aspect of shoulder design for torsion testing is that there are no "end effects" in contrast to the considerations that apply to tension testing. This is often referred to as *St. Venant's principle* for tension, which signifies that stress concentrations arising from changes in cross section persist for a distance of about one diameter along the axis. In torsion, the principle does not apply to the *shear* stresses that are developed. Hence, very short, thin-wall tubes can be used without introducing errors in the stress measurement.

The grip ends of torsion specimens for workability studies are generally of two types: threaded and geometric cross sections. For threaded grip ends, a surface against which the torque can be reacted is provided by making the major diameter of the threads less than the shoulder diameter, or by including flanges between the threaded end and the shoulders in the specimen design. Triangular, hexagonal, or other simple geometric cross sections can also be used for grip ends. In these instances, the torque is reacted against the flat faces of the ends in slotted holders. Selection of threaded ends versus triangular or hexagonal cross-sectional ends is based largely on whether torsion testing is conducted under load control or stroke control. In load control, axial loads are applied to prevent buckling or to enable combined tension-torsion deformation; threaded ends are required for this type of test.

Triangular or hexagonal ends are most common when tests are to be run in stroke control to prevent changes in gage-section geometry that may result from the axial extension that characterizes the torsion testing of many metals. Frequently, threaded ends can also be used in such situations; however, triangular or hexagonal ends have the added advantage of allowing rapid specimen removal for purposes of quenching, which is beneficial in subsequent metallurgical analysis. The possible development of axial stresses and their effect on flow and failure response in fixed-end tests should be carefully considered, because these stresses can be a significant fraction of the torsional shear stress (Fig. 9). The subject of axial stresses is discussed in greater detail later in this chapter.

The final step in the preparation of torsion specimens for determination of workability is the application of a fine axial line on the gage-section surface. This may be done using a felt-tip pen for room-temperature tests or a fine-metal scribing instrument for elevated-temperature tests. After twisting to failure, the line may still be straight (i.e., it will form a helix at a fixed angle to the torsion axis), indicating fracture-controlled failure, or it will exhibit a "kink" at a larger angle to the torsion axis than the remainder of the line, indicating flow-localization-controlled failure. In the former case, the surface shear strain at fracture (γ_{sf}) is given by γ_{sf} = tan φ_f, where φf is the angle between the scribe line and torsion axis at failure. This value of γ_{sf} should agree with that obtained from the equation given above for shear strain as a function of *r*, θ, and L. When flow localization has occurred prior to fracture, the tangent of the angle between the line away from the localization region and the torsion axis provides an estimate of the workability of the material. However, the tangent of the angle between the line and the torsion axis in the localization region also yields an estimate of the fracture strain, provided the test has been carried out to this point.

Torsion Equipment

The torsional mode of mechanical testing, and hot torsion in particular, is not as widely applied as tension, so that there are few commercially available machines. Instead, the apparatus is commonly designed for the specific needs of the individual researcher. The testing machines described by early authors of hot-workability studies were designed to provide a range of constant twist rate only (Ref 1–5, 11–13, 16, 17, 33, 36–41). Later test machines were constructed that can deliver multistage deformation schedules with intermediate unloading of the specimen (Ref 29, 31, 42–51). These latter machines can be effective in simulating multistage working operations.

All torsion machines have certain features in common: a test frame, a drive system, twist and

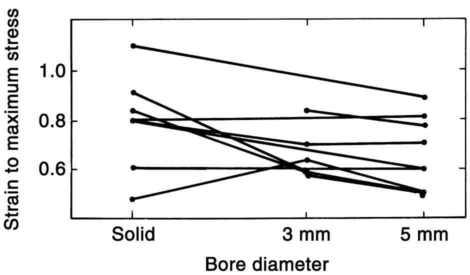


Fig. 7 Effect of bore diameter on strain (ε_m) to maximum stress of type 304 stainless steel specimens with a 7 mm gage length and a gage-length-to-radius ratio of 2 to 1. The solid lines join data points obtained at similar strain rates and temperatures. Source: Ref 35

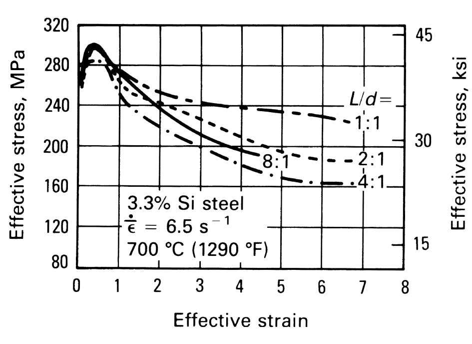


Fig. 8 Stress-strain curves for solid torsion specimens of 3.3% Si steel showing effect of gage-length-to-diameter ratio (*L*/*d*) on flow stress at high strain rates when adiabatic heating occurs. The flow curves are in terms of von Mises effective stress-strain ($\bar{\sigma} - \bar{\varepsilon}$), defined by $\bar{\sigma} = \sqrt{3}\tau$ and $\bar{\varepsilon} = \gamma/\sqrt{3}$, where τ-γ is the shear-stress/shear-strain curve obtained in torsion testing. Source: Ref 35

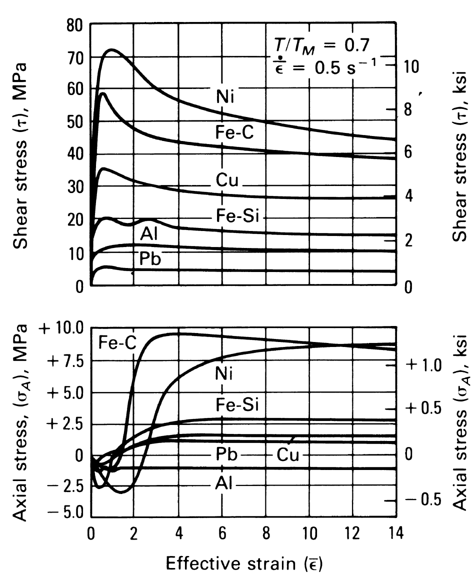


Fig. 9 Dependence of shear stress and mean axial stress on effective strain in fixed-end torsion tests at high temperatures. Source: Ref 12

torque monitoring devices, and a furnace. This section briefly reviews the choices made by some designers as a guide to the construction of a torsion facility.

Design of the Test Frame

An essential requirement of a hot torsion test is that the specimen be twisted about its axis with no induced bending. Therefore, the test frame should be of high rigidity and should have a facility for the accurate alignment of the rotating components. The two major types of setups comprise horizontal load trains mounted on lathe beds and vertical load trains mounted in modified test machines. Lathe beds (Fig. 10) provide adequate stiffness and precision-machined guideways that enable accurate alignment of the tooling components. Typically, for test specimens 2 to 6 cm (0.8 to 2.5 in.) in length, a bed of 2 to 3 m (6.5 to 10 ft) will accommodate all the necessary devices of the load train. However, this type of arrangement does not allow for control of axial loads. Tests can be conducted with the tailstock (saddle) either fixed or floating.

Control of axial loads can be achieved by mounting the components in a modified test machine, such as the electrohydraulic setup shown in Fig. 11. A hydraulic motor is in series with an independently controlled, linear, hydraulic actuator that may be run either in load or stroke control. The torque is reacted against the posts of the test frame using a reaction plate to which a set of cam followers is attached, which allow the ram-hydraulic motor assembly to translate up and down. A series of adapters and the torque and axial load cells are located at the top of the load train. The use of many adapters as well as individual load cells for axial load and torque is undesirable, because they increase the overall length of the load train. As the length increases, the stiffness of the torsion machine decreases, and the possibility of problems associated with wave effects and "ringing" in the load cells during high-strain-rate experiments increases. The top of the load train is attached to the upper crosshead by a set of bolts and spherical washers that provide the primary means of aligning the system. If the grips and other components are accurately machined and the system is accurately aligned, the total indicator runout on the specimen rotation can be kept to several thousandths of an inch with a modified testing machine system. Such tolerances are comparable to the tolerances and deflections maintained in a linear actuator and are usually acceptable for most torsion work. These tolerances can only be improved by placing special reinforcing structures or die sets between the crossheads of the machine, thereby substantially increasing the cost of the equipment.

Motors

Three types of motors are generally used for torsion testing to assess workability: electric, hydraulic, and hydraulic rotary actuator.

Electric Motors. Variable-speed electric motors are the simplest and were the first to be used for torsion testing. These motors can have speeds ranging from fractions of a revolution per minute (rpm) to several thousand rpm. However, the available speed range for a given motor may be limited. In addition, they are not readily adapted to provide a variation in strain rate during each deformation cycle, as is necessary for accurate simulation of plate- or strip-mill operations. Other limitations are the difficulty of maintaining a constant rotation rate if the torque requirement changes during a test and the generally low torques that such motors generate. Electric motors may be fitted with gearboxes and energy-storage devices to prevent these problems.

Perhaps the most publicized electric-drive machine is that designed by Rossard (Ref 39, 50) at IRSID in France. This machine has essentially all the features of devices of similar vintage (Ref 11, 13, 17, 40, 41). It consists of a 4 kW, 1500 rpm electric motor, a 30-to-1 continuously variable hydraulic speed reducer, a 30-to-1 gear box and an electropneumatic clutch-and-brake assembly. This arrangement provides a strain-rate range of ~0.5 to 20 s^{-1} for specimens measuring 24 by 6 mm (0.9 by 0.2 in.). Lower strain rates can be obtained by installation of other fixed-ratio gear boxes between the hydraulic speed reducer and the clutch assembly. The maximum torque available at the

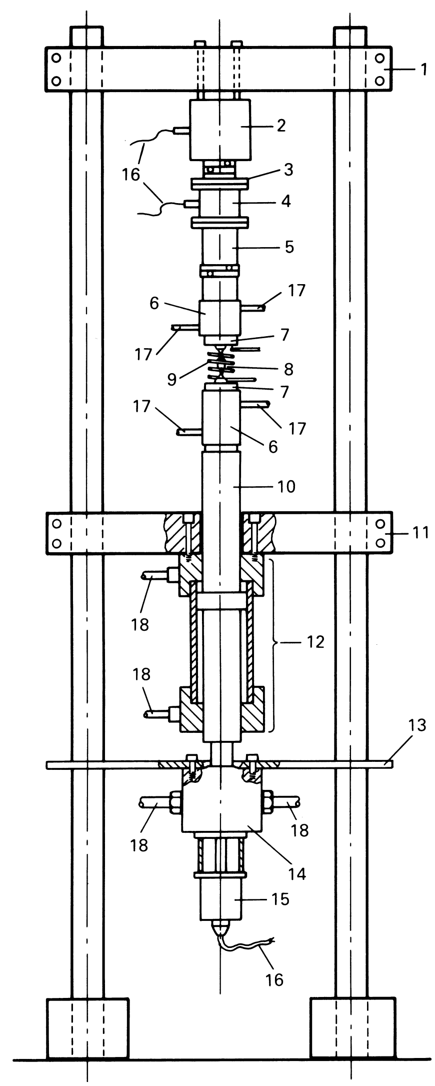


Fig. 11 Electrohydraulic testing machine modified for combined torsional and axial loading: 1, upper crosshead; 2, tension load cell; 3, torque cell adapter No. 1; 4, torque cell; 5, torque cell adapter No. 2; 6, water-cooled grip; 7, specimen holder; 8, specimen; 9, induction coil; 10, ram; 11, lower crosshead; 12, linear (tension-compression) actuator; 13, torque reaction plate; 14, hydraulic motor; 15, incremental optical encoder; 16, electrical leads; 17, water lines; 18, hydraulic lines. Source: Ref 52

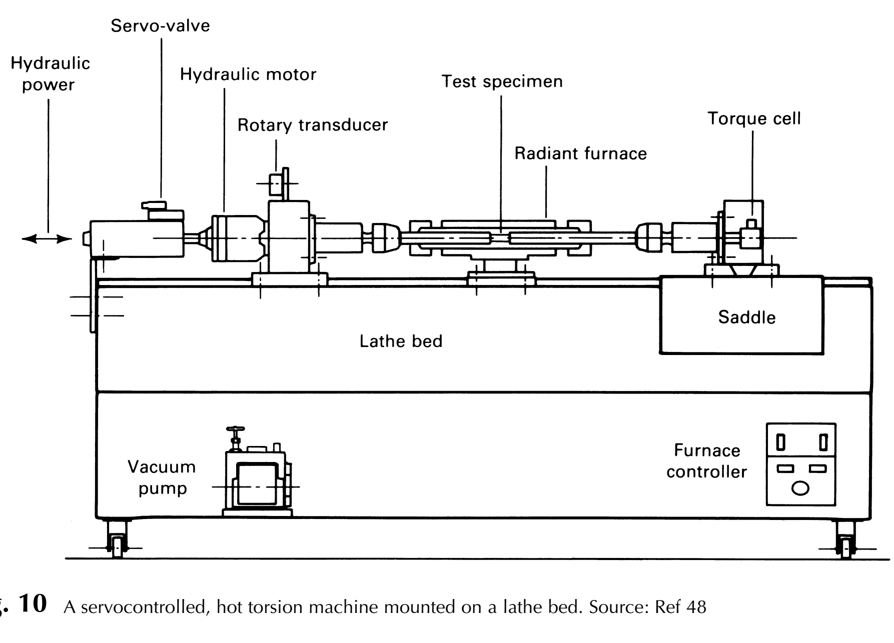


Fig. 10 A servocontrolled, hot torsion machine mounted on a lathe bed. Source: Ref 48

highest strain rate is limited by the power rating of the motor, which for the configuration described previously is about 10 N · m (7.5 lbf · ft), which may be insufficient for simulation of rolling of steel at lower temperatures. Although this difficulty can be solved by increasing the power of the motor, this increases the inertia of the drive system, which places more severe demands on the braking system. This is particularly troublesome at lower strain rates, where severe strain overshoot can occur.

Weiss et al. (Ref 47) overcame some of these difficulties by the incorporation of three independent drive motors rated at 11, 1.5, and 0.2 kW (Fig. 12). This system also includes a flywheel that ensures that only a 1.5% reduction of speed occurs when the clutch is engaged. The machine has a speed of 0.01 to 3000 rpm in 14 steps, which for the specimen dimension used corresponds to a range of shear-strain rates up to 300 s^{-1}, covering seven orders of magnitude. At the highest speed the maximum torque is limited by the 11 kW motor to 25 N · m (18 lbf · ft).

A critical component of all electric-drive systems is the clutch/brake assembly. Many of the early systems were not equipped with brakes, and so there was a tendency for the end of the specimen to continue to rotate after the clutch had been disengaged. Furthermore, the mechanical "dog" systems used on some machines had the tendency to cause "ringing" of the specimen after engagement so that a true constant rate of deformation was not achieved (Ref 11). The use of axial electromagnetic clutches eliminated this problem, but provided a less-positive linkage with a consequent loss of acceleration. Those machines that have been designed to carry out working-schedule simulation use conical electropneumatic clutch-and-brake assemblies (Ref 41, 43–45, 47). Weiss et al. (Ref 47) have described their design in detail (Fig. 13). The clutch develops a normal load of 25 kN (2.8 tonf) at the frictional interface, thereby enabling a torque of 1000 N · m (740 lbf · ft) to be transmitted. The system, however, has a very short response time, which allows the driven parts to be accelerated from 0 to 3000 rpm in 2 to 3 ms, corresponding to approximately 0.3 radians or 0.05 revolutions. The brake, which is of similar design, has comparable responsiveness.

Hydraulic Motors and Hydraulic Rotary Actuators. To overcome the drawbacks of electric motors, hydraulic motors and hydraulic rotary actuators are used in sophisticated torsion systems. Hydraulic motors are typically of the fixed-displacement type, in which the rotation and displacement of the shaft is proportional to the volume of oil flowing through the motor. The torque is produced by the oil as it pushes against a set of spring-loaded vanes that rotate around the motor housing or spherical pistons that ride up and down a contoured cam. In such motors, multiple turns are achieved by supplying a continuous flow of oil. Typical speeds range from 1 to about 1000 rpm. Slower speeds are obtained by using gear reducers in series with the motor.

Fulop et al. (Ref 48) describe a machine (Fig. 10) that uses a piston-type hydraulic drive controlled by a servovalve. The hydraulic motor can develop a maximum torque of 120 N · m (88 lbf · ft) at a maximum speed of 140 radians/s (~1325 rpm) with a fluid-delivery rate of 70 l/min.

One advantage of the servohydraulic drive system is that it eliminates the need for clutches, brakes, and gear boxes. The latter may be included, however, to extend the strain-rate range downward. The use of a closed-loop control system provides high effective stiffness, so that the specimen can be stopped and the specimen unloaded in a similar period. The ability to actually unload the specimen by reverse rotation is very important in mill-simulation tests and provides a high degree of precision to the imposed strains.

In hydraulic rotary actuators, the torque is also supplied by oil pressure against a set of vanes. However, the rotation is limited to only a fraction of a revolution. The specimen gage length therefore must be short or the diameter must be large to achieve large strains. Such actuators are suitable when loading histories involving rotation direction reversals are required or when the system is to be used for torsional fatigue testing as well.

In addition, rotary actuators can produce a greater range of rotation rates than hydraulic motors, particularly at the low end of the scale where fractions of an rpm are possible. The rotation rates required to simulate various metalworking operations are given in Table 1.

The flow of oil through either hydraulic motors or rotary actuators may be at a given rate that is determined by the hydraulic pump supplying the oil (an open-loop arrangement) or may be accurately controlled by a servovalve in a closed-loop circuit. In an open-loop system, the rotation rate of the motor can vary as the torque requirement changes. This difficulty is overcome in a closed-loop, servocontrolled system. The servovalve is an electrically actuated valve that forms part of a feedback circuit. A rotary transducer measures the rotation of the

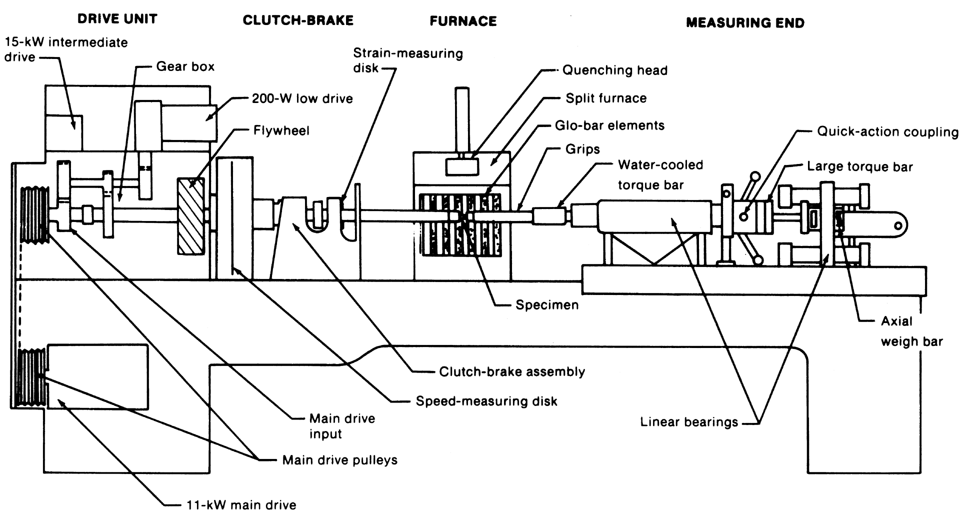


Fig. 12 Testing machine designed by Weiss et al. Source: Ref 47

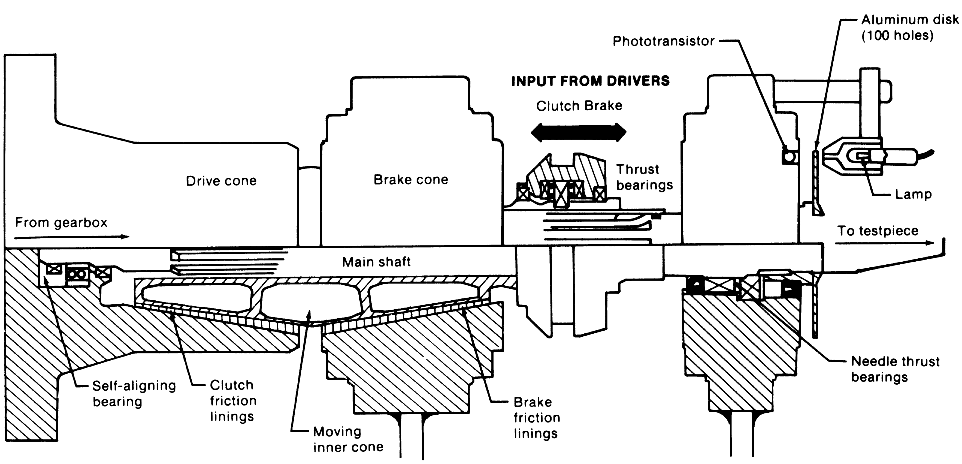


Fig. 13 Conical electropneumatic clutch-and-brake system after Weiss et al. (Ref 47)

Table 1 Torsional rotation rates corresponding to various metalworking operations

Operation	von Mises effective strain rate ($\dot{\bar{\varepsilon}}$)(a), s^{-1}	Corresponding surface shear strain rate in torsion ($\dot{\gamma}$), s^{-1}	Rotation rate(b), rpm
Isothermal forging	10^{-3}	1.73×10^{-3}	0.02
Hydraulic press forging	1	1.73	16.5
Extrusion	20	34.6	330.4
Mechanical press forging	50	86.6	827.0
Sheet rolling	200	346.4	3307.9
Wire drawing	500	866.0	8269.7

(a) $\dot{\bar{\varepsilon}} = \dot{\gamma}/\sqrt{3}$. (b) Assuming specimen geometry with $r/L = 1.0$

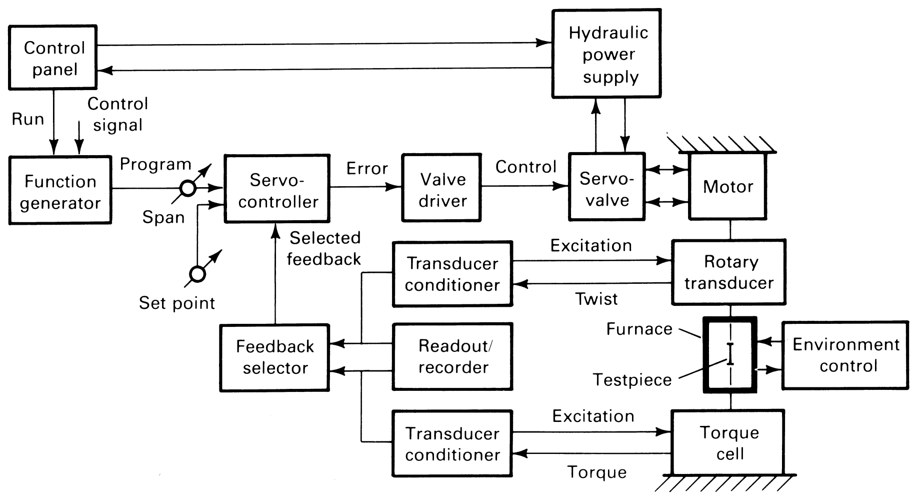


Fig. 14 Typical control and data acquisition system for torsion testing. Source: Ref 48

hydraulic motor or rotary actuator, and a function generator provides a predetermined twist rate (or twist history) in terms of an electrical voltage that varies with time. A sample circuit of this type is shown in Fig. 14.

During operation, oil is fed through the servovalve into the motor or rotary actuator. The rotation is sensed by a transducer that sends a voltage to the servocontroller. The voltage is then compared to the predetermined voltage from the function generator for that given point in time. Depending on the difference between the two voltages, the servovalve is either opened or closed to increase or decrease the flow of oil to the hydraulic motor or rotary actuator. This sequence of operations is performed continuously during the torsion test to ensure that the proper twist history is being applied.

Servocontrolled hydraulic systems can also be programmed for arbitrary strain-rate profiles. In this way, it is possible to impose deformation steps that accurately follow the strain-rate variations that occur, for instance, in the entry of a rolling mill. Indeed, it is possible to examine the effects of mill geometry and layout on the mill loads as well as the working characteristics of the material.

Twist Sensors

Specially designed transducers are necessary to measure the twist of the motor shaft and specimen. As mentioned previously, a rotation transducer is a component of the feedback circuitry required to control the hydraulic motor. These transducers are of three types: optical devices, variable resistors (rheostats), and rotary variable-differential transformers.

With the exception of the device described by Fulop et al. (Ref 48), measurement of angular displacement has typically been carried out optically. A typical example of such devices is that described by Weiss et al. (Ref 47), which uses a metal disk with holes that interrupts the path between a lamp and a phototransistor. A Schmidt trigger circuit, in turn, operates a digital frequency counter and provides a pulsed signal to a recorder. Two disks are described: one with 100 holes, which monitors specimen rotation, and one with 1000 holes, which measures that of the gearbox output. This device is capable of a 200 μm response at the specimen, which means that a shear strain of 0.8% can be resolved at 40 s^{-1}.

Improved strain resolution can be obtained by the use of optical encoders, which use a transparent disk on which a coded mask is etched. A series of phototransistors transmit the binary or BCD (binary coded decimal) position code directly to the display or recording system. A ten-bit shaft encoder can resolve position to about 20 minutes of arc or 6 milliradians, while a twelve-bit device can achieve a resolution of 1.5 milliradians, that is, a shear strain of 0.02% on a specimen with a length-to-diameter ratio of 4 to 1. An added advantage of an optical encoder is that the angular displacement of the specimen is determined absolutely because the rotational sense is also coded. When linked to a fast digital-to-analog (D/A) converter, it can produce an analog output that is proportional to displacement, which makes it ideal for position control in servocontrolled hydraulic drive systems.

Rheostat shafts may also be connected to one end of the hydraulic motor. This device consists of a dual-gang potentiometer, which is connected to voltage conditioning and switching circuits and an eight-bit D/A converter. For each half revolution, one of the two potentiometers is active. At the completion of each half turn, the signal input to the conditioner is switched from one potentiometer to the other, and, simultaneously, the D/A converter is incremented and its output summed with the scaled output of the active potentiometer. In this way a continuous output ramp is obtained that has a slope of 0.2 V/rev. The resolution of this device is ±1 mV or ±0.03 rad. The minimum rotation speed is determined by the smallest step of a digitally generated ramp and the resolution of the control electronics and is about 0.03 rad/s. When typical specimens are used, shear-strain rates from 4×10^{-3} to 100 s^{-1} can be obtained.

The rotary variable-differential transformer is perhaps the most accurate of the rotation-measuring devices and is used primarily in sophisticated rotary actuator systems. Like a linear variable-differential transformer, a rotary variable-differential transformer generates a variable voltage, as an iron core moves through a current-carrying coil.

When the rotation transducer is located in the load train close to the torsion specimen, the voltage produced by it is often used to measure the actual twist applied to the specimen. However, if the load train is long and the transducer is not close to the specimen, elastic deflections of the system can introduce systematic errors in the specimen-twist measurement. To overcome this difficulty, an auxiliary twist transducer is sometimes placed near the specimen. For example, a rheostat system to which a drive pulley is attached may be coupled to the specimen grip using a tight-fitting rubber O-ring or gear mechanism. This arrangement is satisfactory for rotation rates of about several hundred rpm.

Torque Transducers

Torque sensors for torsion testing are usually of the reaction-torque type and are usually placed on the side of the specimen opposite the hydraulic motor or rotary actuator. Load transducers that are capable of monitoring both axial and torsional loads are also available. In either case, the torque is measured through a structural member in the load cell, to which foil-resistance strain gages have been applied. As torque is applied, the structural member deflects elastically. The response is measured with the strain gages. In practice, the torque can be calibrated indirectly using a calibration resistance in a

Wheatstone-bridge circuit or directly by imposing fixed amounts of torque by hanging weights from a lever arm attached to the torque cell. To ensure that the measured torque is free of any frictional components, it is desirable that the torque cell be mounted between the specimen and the fixed-end bearing. Because this places the cell close to the furnace and the specimen extension bars, it is commonly necessary to water cool the load cell in order to stabilize the gage factor.

The capacity of the torque cell should accommodate the specimen geometry, test material, and test temperature. For typical specimen geometries, torque cells with ratings of 0 to 565 N · m (415 lbf · ft) are sufficient. When testing is carried out on very soft metals or on most metals at hot working temperatures (at which the flow stress is very low), a torque cell of one-tenth this capacity (55 N · m, or 40 lbf · ft) is required to maintain adequate signal-to-noise ratios for data acquisition. With modern electronics, most torque cells can be calibrated and used at various percentages (usually 10, 20, 50, or 100%) of their maximum rating, which allows a particular transducer to be used over a range of temperatures and strain rates. With commercially available torque cells, torque can be measured with a precision of 0.1% of the range and with a linearity of better than 0.05%.

In view of the amount of data processing required to extract data on shear stress and strain from the torque/twist curves, it is advantageous to directly digitize and store the data with the aid of a desktop computer. Computer-automated data acquisition greatly enhances the versatility of the testing system, particularly when complex working schedules are simulated.

Heating Systems and Water-Cooled Grips

Torsion tests to establish workability are frequently conducted at elevated temperatures. For this purpose, specially designed heating arrangements and specimen grips are necessary. Selection of a suitable specimen-heating system depends on the particular type of test that is envisioned. In general, the furnace type and arrangement should be such as to provide rapid changes in test temperature as well as a means of fast quenching. The two most common methods of heating are furnace and induction. For typical heating times, furnaces or induction generators should have power ratings of 2 to 10 kW.

Furnaces are usually electric-resistance or quartz-tube radiant heating types. The hot zone should be long enough to ensure uniform heating of the gage section. Initial temperature nonuniformities may give an incorrect picture of failure if the test material is sensitive to flow localization. This is particularly true of titanium- and nickel-base alloys, for which the flow stress is very sensitive to temperature in the hot-working regime, and the thermal conductivity is low.

Conventional resistance-wound furnaces have been used when long-term temperature stability is of primary importance, such as with relatively low-strain-rate hot strength or ductility tests (Ref 10, 25, 29, 30, 46, 53). With such units, rapid quenching is difficult because the specimen must first be removed from the furnace. Glover and Sellars (Ref 46, 53) obtained rapid quenching by withdrawing the specimen axially with the aid of a pneumatic actuator that positioned the specimen beneath a quench spray. This device could only be used either with free-end specimens or after specimen failure. On the other hand, Weiss et al. (Ref 47) used a complex apparatus that first removed a split radiant bar furnace from around the specimen and covered the elements before in situ quenching with water spray. Both schemes can quench the specimen within 0.1 to 1 s.

In order to facilitate specimen quenching as well as to provide arbitrary temperature/time profiles, many researchers have opted for induction or radiant heaters (Ref 12, 21, 36, 40, 45, 48, 54, 55). Both types have very low thermal capacity and thereby promote rapid heating and cooling. From the point of view of ease of quenching, induction heaters are preferred.

In induction systems, the coil for torsion specimen heating should overlap the shoulders of the test specimen to ensure temperature uniformity. In general, the most efficient application of induction heating is for ferrous alloys (carbon and ferritic alloy steels, austenitic stainless steels, etc.) and nickel- and titanium-base alloys. Generator frequencies for heating torsion specimens of these alloys are generally in the radio-frequency range (~100 to 400 kHz). Aluminum and copper alloys do not induction heat readily and thus should be heated by this method only with the use of susceptors.

To prevent contamination of torsion specimens during elevated-temperature torsion studies, specimens may be tested in specially designed chambers or within quartz tubes in which a controlled atmosphere, such as argon, is maintained. A simpler, less expensive technique is to apply protective coatings to the specimens. These coatings are usually vitreous and are supplied as frits or as commercial metalworking glass lubricants. Frits are applied at room temperature by grinding them into a fine powder, preparing a slurry with an alcohol carrier, and dipping the specimens into the coating preparation. Commercial lubricants used for specimen protection can be applied by a variety of methods such as dipping, brushing, and spraying. For both types of glass, the coating melts and forms a viscous, protective coating as the specimen is heated. Before using such coatings, however, initial trials should be conducted to verify that undesirable reactions with the workpiece do not occur at elevated temperatures.

Another problem associated with elevated-temperature testing is the necessity of preventing heat from being conducted into parts of the load train located away from the specimen grips (load cells, motors, etc.). This is accomplished by providing water cooling to the grips. This provision is relatively straightforward for the stationary grip; it is accomplished by brazing copper tubing that contains a water supply onto the grip.

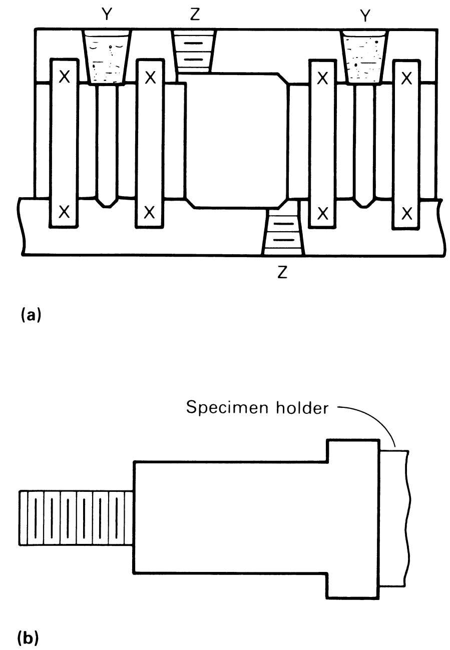


Fig. 15 Water-cooled rotating grip for high-temperature torsion testing. (a) Outer housing. (b) Inner core. Note that the outer housing, to which water lines are attached, is held stationary during testing by reaction rods against which the torque is reacted. X, O-ring grooves; Y, water inlet/outlet; Z, tapped hole for reaction rod

Water cooling is more difficult to accomplish in grips that rotate. In this case, the grip, or an adjacent fixture comprising a core that can rotate and an outer casing that is held stationary must be constructed. A typical example of this type of fixture is shown in Fig. 15. Water leakage is prevented by a set of rubber O-rings. Depending on the actual test temperature, the water-cooled fixtures should be constructed of tool steel or stainless steel. Grips or parts of tooling that are subjected to temperatures higher than approximately 540 °C (1000 °F) should be fabricated from high-temperature superalloys, such as Waspaloy, Inconel 718, or IN-100.

Flow-Stress Data

When failure is controlled by fracture without prior flow localization, flow curves and fracture strains can be readily derived from measurements of torque versus twist during testing. Frequently, these data indicate increasing or nondecreasing torque with increasing rotation. However, at hot working temperatures (and sometimes even at cold working temperatures), the material may show softening due to dynamic recovery, dynamic recrystallization, microstructural softening, or deformation heating. In these cases, the only reliable method of confirming that failure is fracture controlled is by examination of the surface scribe line or the specimen microstructure. If the deformation has been fracture controlled, the inclination of the scribe line to the torsion axis will remain uniform, and the microstructure in the gage section will be deformed uniformly. The tangent of the angle between the scribe line and the torsion axis should be equal to the shear strain corresponding to that value of twist at which the torque-twist record shows a sudden drop, indicating fracture.

Reduction of Torque-Twist Data to Shear Stress versus Shear Strain

Methods for reducing torque-twist (M-θ) data to the shear stress/shear strain (τ-γ) form vary with specimen geometry. Reduction of torque-twist data is easiest with thin-wall tubes. In this case, the shear strain is found from $\gamma = r^*\theta/L$, where r^* is the mean tube radius and L is the gage length. A relation between the torque and the (assumed) uniform shear stress is found from the torque-equilibrium relation $M = 2\pi r^{*2}t\tau$, which results in $\tau = M/2\pi r^{*2}t$. Here, t is the tube-wall thickness, assumed to be small in comparison to r^*.

For round solid bars, the derivation of τ-γ from M-θ is more complex. Several methods are available; however, all overlook the occurrence of deformation heating effects at high strain rates. Nadai (Ref 56) was apparently the first to formally extend the elastic analysis of a torsional bar from Ludwik (Ref 57) to include the case of fully plastic flow. He argued that the distortions produced in an isotropic medium by small angles of relative twist can be considered as simple shear strains that are proportional to the distance, r, of an element from the bar axis, that is, Eq 1. This assumes that all radii remain straight during deformation, which is borne out by observation and symmetry considerations even at high strains as well as finite-element modeling (Ref 58). The total torque developed in the testpiece can be obtained by noting that a solid bar can be considered as an assembly of concentric, thin-wall tubes.

When the specimen is twisted, the component of torque, dM, due to one such tube of radius r and thickness dr is given by:

$$dM = 2\pi r^2 \tau dr \qquad \text{(Eq 3)}$$

where τ is the current shear stress in the elemental tube. The total torque induced in the testpiece is then:

$$M = 2\pi \int_0^{a_2} \tau r^2 dr \qquad \text{(Eq 4)}$$

where a_2 is the outer radius of the bar. If the specimen is in the form of a tube of inner radius a_1 and outer radius a_2, Eq 4 becomes:

$$M = 2\pi \int_{a_1}^{a_2} \tau r^2 dr \qquad \text{(Eq 5)}$$

Now, as the shear strain γ is given by Eq 1 and $dr = Ld\gamma/\theta$, on changing variables, Equations 4 and 5 become:

$$M = \frac{2\pi}{\theta^3} \int_0^{\gamma_2} \tau\gamma^2 d\gamma \qquad \text{(Eq 6)}$$

or:

$$M = \frac{2\pi}{\theta^3} \int_{\gamma_1}^{\gamma_2} \tau\gamma^2 d\gamma \qquad \text{(Eq 7)}$$

where γ_2 and γ_1 are the shear strains corresponding to the radii a_2 and a_1, respectively.

Since the measured parameter in the torsion test is the torque, M, analytical schemes are required to extract the shear stress from the integrals in Eq 4 to 7. Nadai's approach (Ref 56) is to assume that the shear stress in the body depends only on the local value of the shear strain, namely:

$$\tau = \tau(\gamma) \qquad \text{(Eq 8)}$$

Then, since the value of the integral in Eq 6 depends on the upper limit, differentiation with respect to θ yields:

$$\tau_a = \frac{1}{2\pi a^3}\left(3M + \theta\frac{dM}{d\theta}\right) \qquad \text{(Eq 9)}$$

or

$$\tau_a = \frac{M}{2\pi a^3}\left(3 + d\ln M/d\ln\theta\right) \qquad \text{(Eq 10)}$$

This expression serves as the basis of the well-known graphical technique for obtaining shear stresses from torque-twist records for solid bars (Fig. 16). Clearly, $dM/d(\theta/L) = BC/DC$, and, since $DC = \theta/L$, then $BC = \theta\, dM/d\theta$, so that:

$$\tau_a = \frac{1}{2\pi a^3}(3BA + BC) \qquad \text{(Eq 11)}$$

In spite of its simplicity, Nadai's method suffers from two serious limitations. First, the identity in Eq 8 takes no account of the strain-rate sensitivity of the shear stress, so that Eq 9 cannot be applied at high temperatures at which the rate sensitivity is significant. Second, the accuracy with which the surface shear stress is determined depends critically on the precision with which the slope of the torque-twist curve is known.

The analysis for torsion of a solid bar of rate-sensitive material is somewhat different. The simplest method was developed by Fields and Backofen (Ref 36). In this technique, the shear strain in the τ-γ relation is taken to be that for the outer-diameter surface of the solid bar ($\gamma = r_s\theta/L$). The corresponding shear stress is obtained from:

$$\tau = (3 + n^* + m^*)M/2\pi r_s^3 \qquad \text{(Eq 12)}$$

The quantity n^* is the instantaneous slope of log M versus log θ. The value of m^* is found from plots of log M versus log $\dot{\theta}$ at fixed values of θ. Thus, tests at a variety of strain rates are required for proper evaluation. If the material exhibits power-law strain-rate hardening ($\tau \propto \dot{\Gamma}^m$), m^* is equal approximately to the strain-rate sensitivity exponent, m. At cold working temperatures, m^* is usually small and can often be neglected with respect to the value of $3 + n^*$. At hot working temperatures, m^* can take values comparable to the values of m at these temperatures, which typically range from 0.1 to 0.3, or values that are not negligible in comparison to $3 + n^*$.

If the specimen design is a thick-wall tube (i.e., the wall thickness is comparable to the radius), a variation of the Fields and Backofen formula may be used to derive the τ-γ relation. γ is taken to be the shear strain at the outer diameter, $\gamma = r_s\,\theta/L$. However, the shear stress in this case is:

$$\tau = \frac{(3 + n^* + m^*)M}{2\pi r_s^3}\left[1 - \left(\frac{r_i}{r_s}\right)^{3+n^*+m^*}\right] \qquad (13)$$

where n^* and m^* are the slopes of the two logarithmic plots, and r_i denotes the internal specimen radius.

The Fields and Backofen analysis has been widely used as a conversion technique, but in general the procedure has been abbreviated in that the torque-twist curves are simply scaled uniformly with constant values of m^* and n^*. For hot-torsion tests, the rate sensitivity m^* is usually that obtained from plots of log M (steady state) versus log $\dot{\theta}$, and n^* is taken as 0. This form of the procedure is convenient and provides reasonable values of steady-state shear stress, but it has the effect of greatly distorting the initial work-hardening portion of the flow curve and therefore renders invalid the measurements of work-hardening rates and yield stress.

Even when the analytical procedure of Eq 12 is fully applied, a significant source of error arises from the formalism itself. It may be shown that the relative error in shear stress is given by:

$$\frac{\Delta\tau}{\tau} = \frac{\Delta M}{M} + 2\theta\left(\frac{\Delta M}{\delta M \cdot M} + \frac{\Delta\theta}{M \cdot \delta\theta}\right)\frac{\delta M}{\delta\theta}\frac{1}{(3 + n^* + m^*)} \qquad \text{(Eq 14)}$$

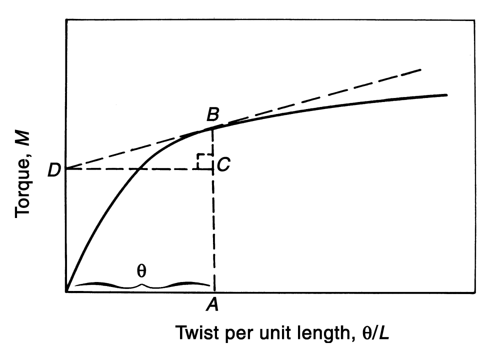


Fig. 16 Graphical determination of shear stress from torque-twist records for rate-insensitive materials. Source: Ref 56

where Δ indicates the absolute error of the measurement, and δ the differences as used in slope determinations. It is important to note that the imprecision in τ increases linearly with θ. This effect is demonstrated in Fig. 17, where Eq 12 is applied directly to torque-twist data for electrolytic tough-pitch copper. If the torque curve had been smoothed prior to the conversion, a smooth stress-strain curve would have been obtained, but the error in stress would still have been $\pm 10\%$ (Ref 59).

Another source of error that must be kept in mind when Eq 12 is used arises from the assumption that the shear stress is uniquely defined when the shear strain and strain rate are known. In their original paper (Ref 36), Fields and Backofen pointed to the importance of strain-rate history in determining stress, but failed to mention the inappropriateness of strain as a "state" parameter. Canova (Ref 60) have considered the error to be attributable to the choice of strain rate and strain as the appropriate "state" variables, rather than the more realistic selection of strain rate and "hardness state" (Ref 61). They state that the relative error in stress is given by:

$$\frac{\Delta\tau}{\tau} = \frac{(m^* - \bar{m})}{(3 + n^* + m^*)} \qquad \text{(Eq 15)}$$

where m^* and n^* have the same meaning as in Eq 12 and where $\bar{m} = (\partial \ln M / \partial \ln \dot{\theta})_{\text{state}}$. This latter quantity is the rate sensitivity measured by the instantaneous difference in torque induced by a sudden change in $\dot{\theta}$. Although $\bar{m}$ may differ from m^* by a factor as large as 1 to 2, $\Delta\tau/\tau$ is nevertheless small compared with other intrinsic errors arising from this method and, for all practical purposes, can be neglected.

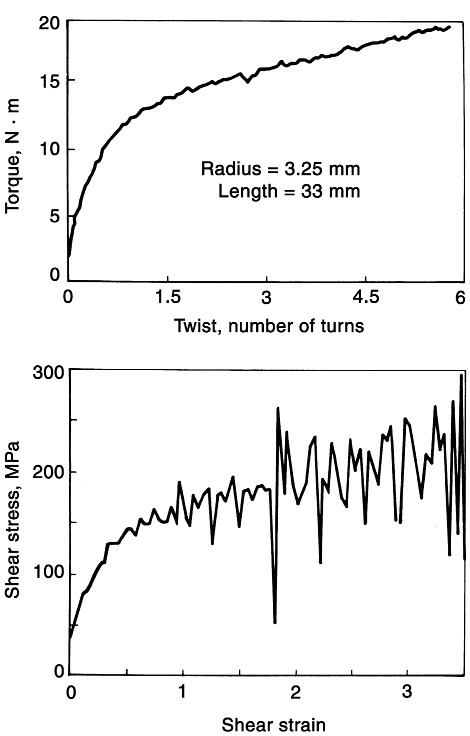


Fig. 17 Shear-stress/shear-strain curve calculated directly from experimental torque-twist curve for electrolytic tough pitch copper at room temperature and 0.01 turns/s using the method of Fields and Backofen. Source: Ref 36, 59

Another important consideration in the reduction of torque-twist data is the occurrence of deformation in the fillet radius. A simple method to correct for such deformation in a solid-bar sample utilizes the definition of an *effective* gage length equal to the length of the reduced section plus an increment in length whose magnitude is a function of the gage diameter, fillet radius, m^*, and n^*. Pohlandt (Ref 62) derived a semiempirical expression for the effective gage length. Subsequently, Khoddam and his coworkers (Ref 63, 64) applied the finite-element method (FEM) to validate the Pohlandt equation for materials whose constitutive behavior is of the power-law form. By contrast, FEM simulations revealed that the effective gage length varies with deformation when the constitutive behavior is of the hyperbolic sine form.

Shear-Stress Derivations for Arbitrary Flow Laws

The methods described previously for converting the measured torque to shear stress at the outer radius of solid-bar samples and thick-wall tubes are based on the analysis of Ref 36 and therefore apply only to materials obeying simple constitutive relations such as the parabolic (power) law. When accurate determinations of the large-strain effective stress versus effective strain relations are required for materials that exhibit more complex behaviors, FEM or alternate experimental methods must be used. The former approach (Ref 65–67) involves inverse numerical calculations and the application of standard optimization techniques. The alternate experimental methods comprise:

- Thin-wall tube testing
- Differential testing
- Multiple testpiece method

These techniques are more accurate than the methods discussed previously in this chapter for metals subject to flow softening, dynamic recrystallization, dynamic recovery, and so forth. Generally, they are appropriate when the flow stress at a given strain depends on the temperature and strain-rate history of the test, because these techniques are based on deducing the properties of a thin incremental layer of the sample without requiring any prior assumptions about the material behavior. The thin-wall tube method was described previously in this chapter and is not discussed here.

Differential testing is a modification of the thin-wall technique that requires the use of two solid samples of slightly different radius, R_1 and R_2. In contrast to the tube method, samples tested according to this technique are not subject to buckling because of axial compression at ambient temperatures or to rupture because of axial tension at elevated temperatures. Consequently, large maximum strains can be attained. In a sense, the R_1 bar serves as a rotating mandrel and supports the attached $(R_2 - R_1)$ layer.

The difference between the two measured torques M_1 and M_2 determined at the same twist rate is used to deduce the properties of the incremental layer. The average shear stress in the layer between R_1 and R_2 is then given by:

$$\bar{\tau} = \frac{3}{2\pi}\frac{M_2 - M_1}{R_2^3 - R_1^3} \qquad \text{(Eq 16)}$$

This shear stress is considered to be developed at the mean radius of the layer $(R_1 + R_2)/2$, so that the associated mean shear strain is given by $\gamma_{\text{ave}} = [(R_1 + R_2)/2](\theta/L)$, where θ/L is the twist per unit length of the specimen. However, because of the nature of the torque integral and the flow-hardening characteristics of most materials, τ generally corresponds to the stress acting at a radius greater than $(R_1 + R_2)/2$ and therefore to a shear strain greater than γ_{ave}. The magnitude of this error, which also affects the results of the thin-wall tube tests, depends on the details of the work-hardening relation as well as on $(R_2 - R_1)$ and varies during a given test.

The differential method suffers from errors that are inherent to the definitions of the stress and the strain introduced above. This method is therefore best suited to samples with a small radius difference $(R_2 - R_1)$. However, under these conditions, a large scatter in the values of the torque difference $M_2 - M_1$ exists due to the usual sources of experimental error.

Multiple Testpiece Method. The shortcomings of the differential method can be overcome by using multiple testpieces of increasing radius (Ref 59), at the cost, however, of increasing the complexity of the test. The multiple testpiece method is based on the following derivative of the torque integral with respect to the outer radius R at a given twist (θ), twist rate ($\dot{\theta}$), and outer radius (R_o):

$$\left.\frac{\partial M}{\partial R}\right|_{\theta,\dot{\theta}} = 2\pi\, \tau_{R_o} R_o^2 \qquad \text{(Eq 17)}$$

Thus, the precise value of the current shear stress at R_o can be determined from a knowledge of the slope of the torque/radius curve at this radius:

$$\tau_{R_o} = \frac{\left.\frac{\partial M}{\partial R}\right|_{\theta,\dot{\theta}}}{2\pi R_o^2} \qquad \text{(Eq 18)}$$

The magnitudes of the corresponding strain and strain rate are well defined at this position and are given by $\gamma = R_o\,\theta/L$ and $\dot{\gamma} = R_o\dot{\theta}/L$.

The results of a series of experiments analyzed in this way are given in Fig. 18. Ten samples of increasing diameter (5 to 10 mm, or 0.2

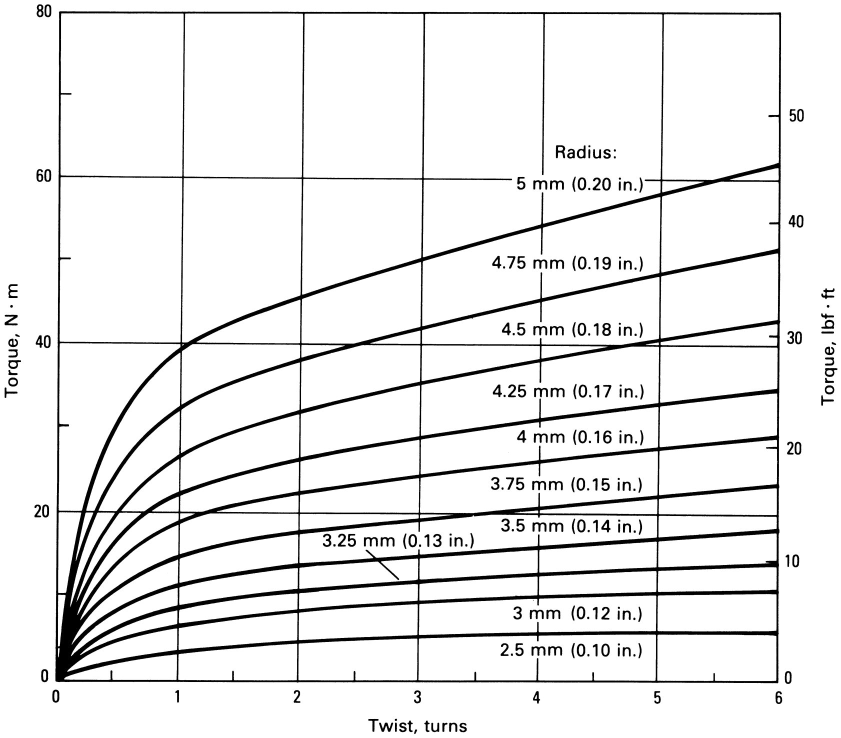


Fig. 18 Experimental torque/twist curves determined in torsion on annealed electrolytic tough pitch copper. Twisted at room temperature at 0.01 turns/s. Source: Ref 59

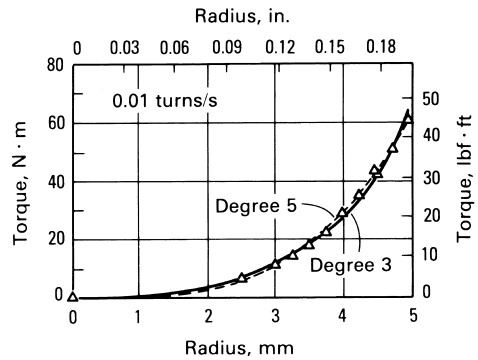


Fig. 19 Torque/radius data taken from the results of Fig. 18 at $\theta = 35$ radians. The influence of the degree of the log M versus R polynomial on the smoothing of the curve is apparent. Source: Ref 59

to 0.39 in.) were tested. These curves were used to construct a set of torque/radius curves at selected intervals of twist; the relation for $\theta = 35$ radians (5.6 turns) is illustrated in Fig. 19. Each curve was fitted by means of a log M versus R polynomial, where a polynomial of degree 5 was found to produce satisfactory results. The derivative of this polynomial gives the slope $(\partial M/\partial R)|_{\theta,\theta'}$, corresponding to a selected R_o (e.g., 3.5 mm, or 0.14 in., in Fig. 19), from which the value of τ_{R_o}, can be found from Eq 18. The complete τ versus γ curve is obtained from a series of such determinations at increasing angles of twist θ.

Torque and Angle of Elastic Unloading. When the torsion testing of solid bars is carried out at ambient temperatures (i.e., on rate-insensitive materials and in the absence of anelastic effects), a torque cell is not absolutely indispensible for the measurement of the moment or couple. The latter can be deduced from the elastic unloading angle per unit length θ^*/L, as determined from (Ref 68):

$$M = \left(\frac{\pi}{2}\right) G R_o^4 \left(\frac{\theta^*}{L}\right) \quad \text{(Eq 19)}$$

where G is the elastic shear modulus of the material and R_o is the outer radius. By progressively applying measured amounts of twist $\Delta\theta$ and determining the unloading angle θ^* at each step as the experiment proceeds, the full torsion stress-strain curve can be deduced. Here, the accumulated strain is given by $\gamma = \theta R_o/L$, where $\theta = \sum\Delta\theta$, and τ must be derived from the moment M by one of the methods described previously.

Deformation Heating

When the surface-shear-strain rate in torsion is less than approximately 0.01 s^{-1}, the τ-γ curves established from the measurement of torque and twist may be assumed to be representative of isothermal (constant temperature) behavior. Above this strain rate, thermal conduction is not sufficiently rapid to dissipate the deformation heat into the specimen shoulders (for typical metals, alloys, and specimen geometries) (Ref 69). Because of this, the τ-γ results are representative of flow behavior at temperatures that increase with the amount of deformation. Furthermore, because the temperature increase ΔT is a function of γ, a radial, as well as an axial, temperature gradient can be developed in solid-bar specimens. For this reason, tubular specimens, in which radial effects are negligible, are preferable for high-rate torsion tests.

The temperature increase that occurs during a torsion test is a function of the amount of heat generated and the amount of heat conducted into the shoulders. The amount of heat conducted into the shoulders depends on specimen geometry and requires numerical analysis.

During an extremely rapid test ($\dot{\gamma}$ typically greater than or equal to about 20 s^{-1}), heating is essentially adiabatic. Under these conditions, all of the deformation work converted to heat is assumed to be retained in the gage section, giving rise to a temperature increase of $\Delta T = (\eta \int \tau d\gamma)/\rho c$. The parameter η is the fraction of the deformation work per unit volume ($\int \tau d\gamma$) converted into heat, ρ is the material density, and c is the specific heat. The value of η usually ranges from 0.90 to 1.00; 0.95 is the value most frequently assumed. The plastic work per unit volume ($\int \tau d\Gamma$) is simply the area under the τ-γ curve. Once the temperature increase (ΔT) has been estimated, the shear stress associated with the test strain rate and particular values of γ and T (equal to nominal test temperature plus ΔT) can be determined.

By conducting torsion tests at several temperatures at a given "adiabatic" strain rate, a high-rate isothermal torsion flow curve can be determined by estimating the ΔT values and plotting τ versus the actual temperature at various levels of shear strain. This procedure is shown in Fig. 20. For each level of γ, a smooth curve is drawn through the data points. Equivalent isothermal

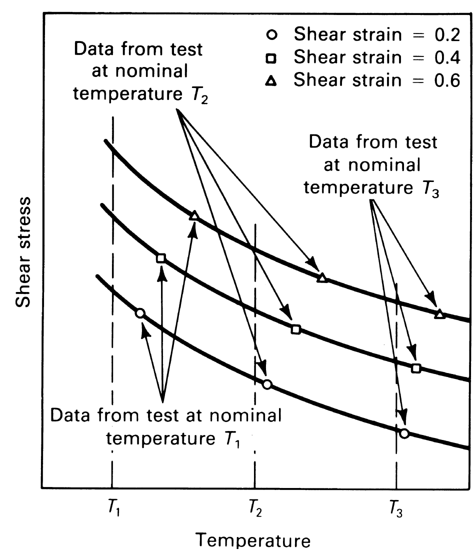


Fig. 20 Stress-temperature plots used to estimate equivalent isothermal high-strain-rate flow curves. The intersections of the dashed lines with the flow-stress/temperature curves yield estimates of the isothermal flow behavior at temperatures T_1, T_2, and T_3.

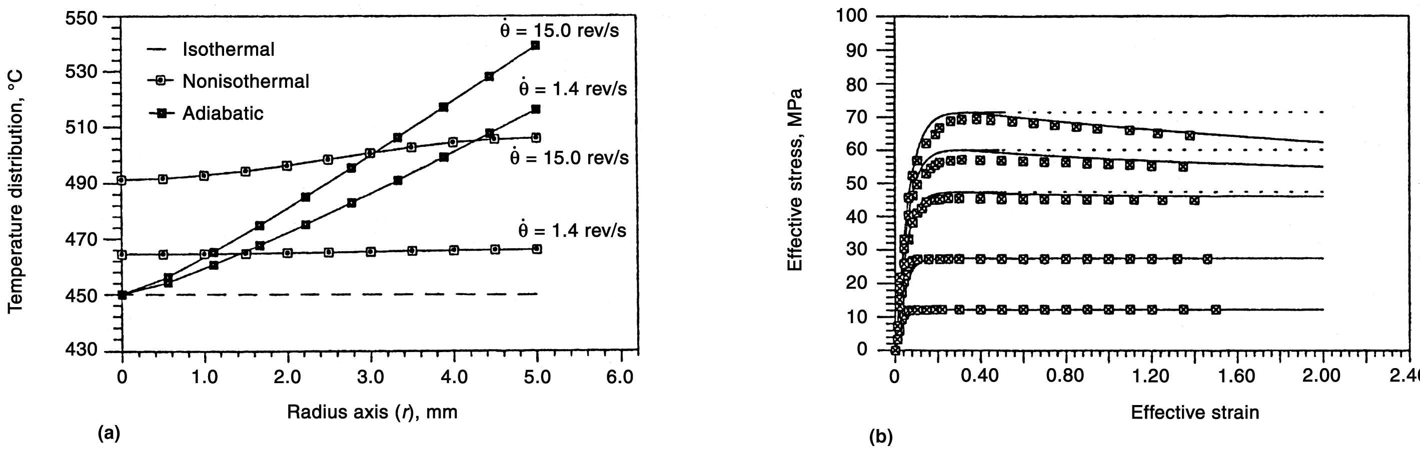


Fig. 21 Finite-element-method predictions of (a) temperature rise at the sample midlength and (b) effective stress-strain curves from simulations of the torsion of aluminum alloy 5252 under various heat-transfer conditions. In (b), the simulation results are compared to experimental data. (a) T_o = 450 °C (840 °F), γ_P = 4.8. (b) T_o = 500 °C (930 °F), D_g = 10 mm (0.4 in.), D_h = 20 mm (0.8 in.), L_g = 10 mm (0.4 in.). Source: Ref 72

flow curves are then obtained at various temperatures by "reading" points off the various curves at these fixed temperatures. This procedure assumes that shear strain is a state variable, that is, that the shear stress is uniquely determined by the value of γ, regardless of prior shear-strain history. Over small intervals of γ and ΔT, such an assumption is probably sound from an engineering viewpoint.

Finite-element analysis may also be used to deduce the temperature nonuniformity in torsion tests that are neither purely isothermal nor adiabatic. For example, Zhou and Clode (Ref 70–72) have applied such a technique for the analysis of the torsion of aluminum alloy 5252 to predict the temperature uniformity as well as constitutive behavior (Fig. 21), among other quantities. Heat *losses,* such as those that may occur during torsion tests on high-temperature alloys using an induction method for sample heating, have also been treated using the finite-element method (Ref 73).

Texture Development and Axial Stresses During Torsion Testing

The axial stresses developed during fixed-end torsion testing are related directly to the deformation textures produced by straining (Ref 74–85) as are the length changes observed during free-end testing. The sign and magnitude of these effects depend on the orientation of the mean value of all the Burgers (i.e., slip) vectors involved in a given increment of twist (Fig. 22). The applied shear stress causes slip (shear) to occur on a variety of crystallographic slip (shear) planes lying either in the transverse plane of the specimen or on planes that are inclined with respect to the transverse plane. Such shear displacements are produced as multiples of the elemental unit of shear, the magnitude and direction of which are given by the Burgers vector. In most metals, the latter corresponds to a single atom diameter.

When the average Burgers vector lies in the transverse (shear) plane of the specimen, there is no length change or axial stress induced. By contrast, when it is inclined either away from or toward the fixed end of the specimen, lengthening or shortening of the sample, respectively, occurs (free-end conditions). Alternatively, under fixed-end conditions, compressive or tensile stresses are developed, respectively.

The inclination of the mean Burgers vector depends in turn on the texture that is developed by twisting and, in particular, on the rotation of the ideal orientations away from strict coincidence with the axial (shear plane normal) and tangential (shear direction) directions of the sample. For example, when copper is deformed at 100 to 300 °C (212 to 570 °F), the $\{1\bar{1}1\}\langle 110\rangle/\{\bar{1}1\bar{1}\}\langle\bar{1}\bar{1}0\rangle$ component predominates at strains of 1 to 2. This set of ideal orientations indicates that two particular (111) glide planes lie in the transverse plane of the specimen (in different grains) and that these are oriented so that certain $\langle 110\rangle$ glide directions are aligned along the shear direction. Actually, however, this set of orientations is rotated slightly about the radial direction of the sample in the sense opposite to that of the shear, so that the (111) planes do not lie *exactly* in the transverse plane, and the $\langle 110\rangle$ directions are not aligned *exactly* with the shear direction. This inclination is responsible for the compressive stresses (fixed ends) and lengthening (free ends) observed in this strain range (Ref 74, 75).

At deformations of 2 to 4, the above component is replaced by the $\{001\}\langle 110\rangle$ orientation, which is rotated in the same sense as the shear. Such an inclination of this component also produces a compressive force (or lengthening). Finally, at strains greater than about 5, a steady state of flow is attained, in which the $\{1\bar{1}1\}\langle 110\rangle/\{\bar{1}1\bar{1}\}\langle\bar{1}\bar{1}0\rangle$ set again predominates, inclined in this case in the same sense as the shear. This inclination is in the opposite sense to the one observed at low strains and is responsible for the tensile stress that develops at large strains (Fig. 23). In a similar manner, specimen shortening is induced when testing is carried out under free-end conditions.

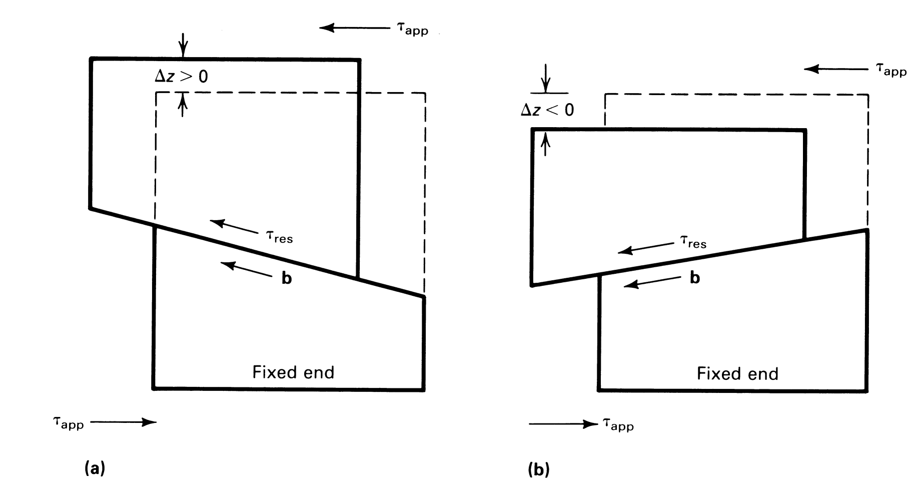


Fig. 22 Effect of the inclination of the mean Burgers vector **b** on specimen length change during simple shear deformation. (a) When the mean Burgers vector **b** is inclined away from the fixed end, specimen lengthening occurs. τ_{res} is resolved shear stress; τ_{app} is applied shear stress; Δz is change in specimen length. (b) When the mean Burgers vector **b** is inclined toward the fixed end, specimen shortening is observed.

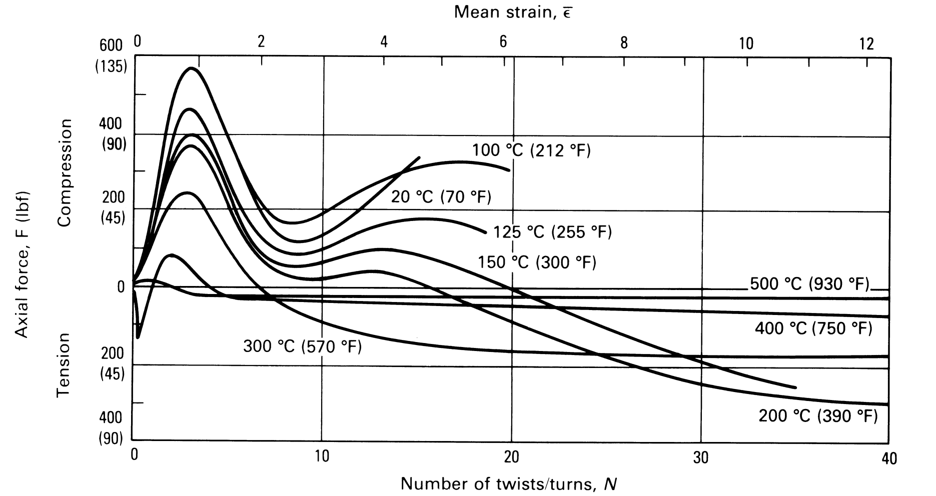


Fig. 23 Dependence of the axial force on temperature in copper. $\dot{\bar{\varepsilon}} = 5 \times 10^{-3}$ s^{-1} from 20 to 200 °C (70 to 390 °F); $\dot{\bar{\varepsilon}} = 5 \times 10^{-2}$ s^{-1} from 300 to 500 °C (570 to 930 °F). Source: Ref 74

Table 2 Common texture components and axial forces observed during the torsion testing of fcc metals

		Axial force(c)		
Type(a)	Miller indices(b)	Rotated in the sense opposite to the shear	No rotation	Rotated in the same sense as the shear
A	$\{1\bar{1}1\}\langle110\rangle$	C	0	T
$\bar{A}$	$\{\bar{1}1\bar{1}\}\langle\bar{1}\bar{1}0\rangle$	C	0	T
A_1^*	$\{\bar{1}\bar{1}1\}\langle112\rangle$	T	T	T
A_2^*	$\{11\bar{1}\}\langle112\rangle$	C	C	C
B	$\{1\bar{1}2\}\langle110\rangle$	0	0	0
$\bar{B}$	$\{\bar{1}1\bar{2}\}\langle\bar{1}\bar{1}0\rangle$	0	0	0
C	$\{001\}\langle110\rangle$	T	0	C

(a) The components $A/\bar{A}$, A_1^*/A_2^*, and $B/\bar{B}$ are observed in pairs, as required by the symmetry of the torsion test. The C component is self-symmetric. (b) The plane $\{hkl\}$ is the crystallographic plane parallel to the macroscopic shear (transverse) plane; the direction $\langle uvw\rangle$ is the crystallographic direction parallel to the macroscopic shear (circumferential) direction. (c) C, compression or lengthening; T, tension or shortening. Source: Ref 75

The principal textures developed in face-centered cubic (fcc) metals are listed in Table 2, with their associated axial forces or length changes. Some texture components, for example, the $\{1\bar{1}2\}\langle110\rangle/\{\bar{1}1\bar{2}\}\langle\bar{1}\bar{1}0\rangle$ set, do not lead to any axial force or length change, even when the pole figure is rotated about the radial direction away from full symmetry with respect to the shear plane normal (axial) and shear (tangential) directions. By contrast, other texture components, for example, the $\{\bar{1}\bar{1}1\}\langle112\rangle/\{11\bar{1}\}\langle112\rangle$ set, produce axial effects even when they are in the maximum symmetry position. Some examples of the textures produced in copper that is twisted at 20 to 500 °C (68 to 930 °F) are illustrated in Fig. 24.

The textures developed in twisted body-centered cubic (bcc) metals differ somewhat from those of fcc metals and are listed in Table 3. A given bcc texture component of the form (hkl) $\langle uvw\rangle$ produces an axial effect that is qualitatively similar to that of an fcc component of the form (uvw) $\langle hkl\rangle$. Rigorously, the stress deviator tensors for the bcc and fcc cases are related through a simple symmetry operation. The twisting of bcc iron, which induces a compressive force (lengthening) at small strains, as in the case of the fcc metals, is due to the inclination of the $\{0\bar{1}1\}\langle111\rangle/\{01\bar{1}\}\langle\bar{1}\bar{1}\bar{1}\rangle$ set in the sense opposed to the shear. In a similar manner, the tensile stresses (shortening) developed at large strains ($\varepsilon > 5$) are associated with the $\{\bar{1}\bar{1}2\}\langle111\rangle$ component, which in this case is rotated in the same sense as the shear.

Crystal plasticity models have been quite successful in accounting for the development of deformation textures and for the axial effects

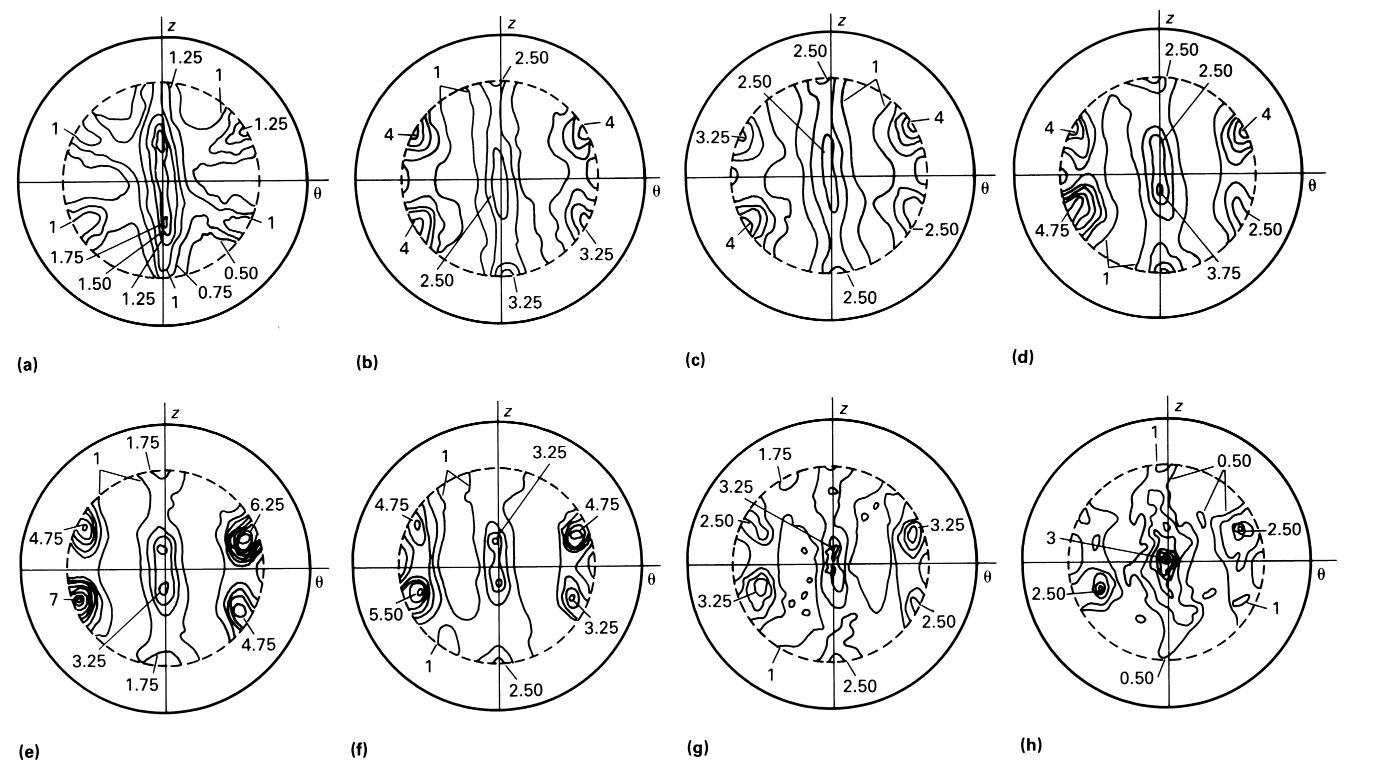


Fig. 24 Dependence of copper textures on temperature. (a) At 20 °C (68 °F), $\bar{\varepsilon} = 4.7$. (b) At 100 °C (212 °F), $\bar{\varepsilon} = 4.65$. (c) At 125 °C (225 °F), $\bar{\varepsilon} = 5.89$. (d) At 150 °C (300 °F), $\bar{\varepsilon} = 10.85$. (e) At 200 °C (390 °F), $\bar{\varepsilon} = 31$. (f) At 300 °C (570 °F), $\bar{\varepsilon} = 31$. (g) At 400 °C (750 °F), $\bar{\varepsilon} = 31$. (h) At 500 °C (930 °F), $\bar{\varepsilon} = 31$. Same strain rates as in Fig. 23. Source: Ref 74

observed at room temperature as well as at elevated temperatures prior to the initiation of dynamic recrystallization (Ref 76–79, 82, 83). This comment applies to both fcc and bcc metals in which lengthening (free-end testing) or the development of compressive stresses (fixed-end testing) is generally observed. Nevertheless, it should be kept in mind that the specific textures developed during torsion are usually not relatable to those that may occur under strain paths that characterize forging, extrusion, and so forth.

When fcc metals (except aluminum) are deformed at *elevated* temperatures, dynamic recrystallization is initiated after a critical strain. In these instances, the texture changes have been modeled for the case of free-end torsion (Ref 80–82). The textures that evolve are a result of both dislocation glide and dynamic recrystallization. The effect of recrystallization on texture evolution can be quantified by mechanisms such as oriented nucleation and selective growth. In the former mechanism, recrystallization nuclei are formed in those grains that have suffered the least shear strain (i.e., dislocation glide). Selective growth is assumed to occur for nuclei of particular misorientations with respect to the matrix. With this approach, Toth and Jonas (Ref 80–82) demonstrated that torsion textures can be predicted quantitatively for copper (Fig. 25). Specifically, it was shown that the $\{100\}\langle 011\rangle$ texture component is gradually replaced by the $\{1\bar{1}1\}\langle 011\rangle$ and $\{1\bar{1}\bar{1}\}\langle 211\rangle$ components. It is this substitution that is responsible for converting the behavior from lengthening to shortening (free-end) or from compression to tension (fixed-end) (Ref 84). Simulations of a comparable nature have also been carried out for bcc materials (Ref 85).

Table 3 Common texture components and axial forces observed during the torsion testing of bcc metals

		Axial force(c)		
Type(a)	Miller indices(b)	Rotated in the sense opposite to the shear	No rotation	Rotated in the same sense as the shear
D_1	$\{11\bar{2}\}\langle 111\rangle$	C	C	C
D_2	$\{\bar{1}\bar{1}2\}\langle 111\rangle$	T	T	T
E	$\{0\bar{1}1\}\langle 111\rangle$	C	0	T
$\bar{E}$	$\{01\bar{1}\}\langle\bar{1}\bar{1}\bar{1}\rangle$	C	0	T
F^*	$\{110\}\langle 001\rangle$	0	0	0

(a) The $E/\bar{E}$ orientations occur as a twin symmetric set, as required by the symmetry of the torsion test. The D_1, D_2, and F^* components are self-symmetric. (b) The plane $\{hkl\}$ is the crystallographic plane parallel to the macroscopic shear (transverse) plane; the direction $\langle uvw\rangle$ is the crystallographic direction parallel to the macroscopic shear (circumferential) direction. (c) C, compression or lengthening; T, tension or shortening. Source: Ref 75

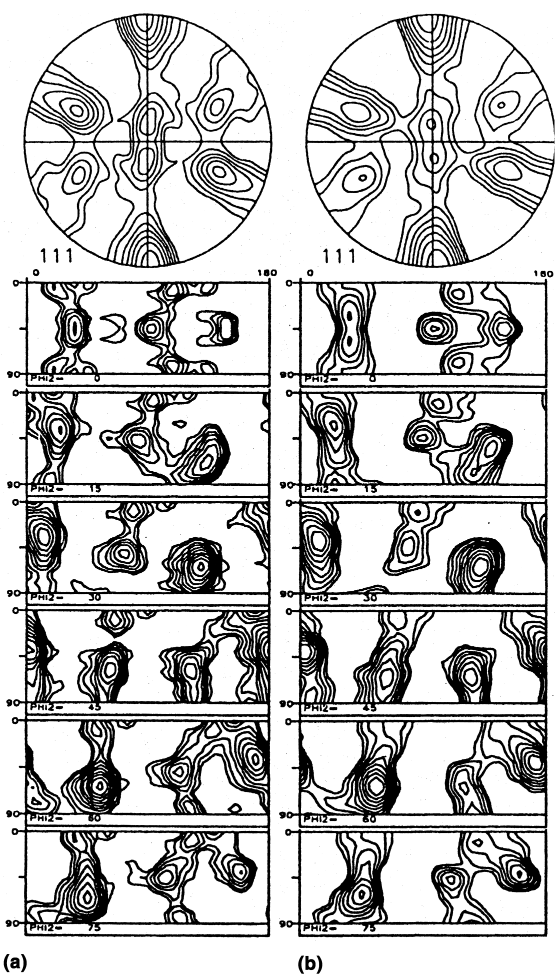


Fig. 25 Textures in terms of (111) pole figures and orientation distribution functions (ODFs) for copper deformed in torsion at 300 °C to a shear strain of 11. (a) Measurements and (b) predictions based on the initial texture and simulations of crystallographic slip and dynamic recrystallization. Isovalues on all plots are 0.8, 1.0, 1.3, 1.6, 2.0, 2.5, 3.2, 4.0, 5.0, and 6.4. Source: Ref 80, 82

Reduction of τ-γ Data to $\bar{\sigma} - \bar{\varepsilon}$

To make use of shear-stress/shear-strain (τ-γ) data for the prediction of load and the analysis of metal flow in actual metalworking operations, these quantities can be converted to effective stress ($\bar{\sigma}$) and effective strain ($\bar{\varepsilon}$). These stresses and strains are the equivalent quantities that would result in an identical amount of deformation work in a state of uniaxial compression or tension. For a given deformation level, then:

$$\int \tau d\gamma = \int \sigma d\varepsilon = \int \bar{\sigma} d\bar{\varepsilon} \qquad \text{(Eq 20)}$$

where σ-ε is the uniaxial flow curve. In uniaxial tension of a material with a random crystallographic texture, $\bar{\sigma} = \sigma$, and $\bar{\varepsilon} = \varepsilon$. For uniaxial compression, $\bar{\sigma} = -\sigma$, and $\bar{\varepsilon} = -\varepsilon$.

In torsion, the relation between τ-γ and $\bar{\sigma}$-$\bar{\varepsilon}$ depends on the plasticity theory formulation that is used. The most common concept, first proposed by von Mises (Ref 86) and extended by Shrivastava et al. (Ref 87), leads to the relations $\bar{\varepsilon} = \gamma/\sqrt{3}$ and $\bar{\sigma} = \sqrt{3}\tau$. Note that the concepts of effective stress and effective strain are also approximations, because the $\bar{\sigma}$-$\bar{\varepsilon}$ relations obtained from different mechanical tests are often different, depending on the particular material and test conditions. Nevertheless, for load estimation, these constructs are sufficiently accurate for engineering purposes. Unless otherwise specified, the effective stress, effective strain, and effective strain-rate definitions used in this chapter are based on the von Mises formulation.

Effective Stress for Free-End versus Fixed-End Testing. The simple expressions for effective stress and strain given in the previous section apply solely to the case in which sample geometry does not change and no axial stresses are developed during torsion testing. In practice, a choice must be made between the fixed- and free-end conditions of straining. Generally, fixed-end conditions are preferred. If the free-end method is chosen, the length can increase (at ambient temperatures) or decrease (at elevated temperatures) by about 10 to 40% (Ref 12). The derived shear stress τ, or $\sigma_{\theta z}$, is proportional to M/R^3 (see above), where M is the measured torque and R is the current sample radius. In most tests, the radius change is not monitored, and the initial value is used instead. Thus, a length change of 10%, which involves a radius change of about 5%, leads to an error in the derived shear stress $\sigma_{\theta z}$ of about 15%. This error increases with the 3/2 power of the length change and can become quite significant.

By contrast, if the length is held constant, axial stresses, σ_{zz}, develop, which range from 2 to 20% of the developed shear stress $\sigma_{\theta z}$. The axial stress $\sigma_{zz}(R)$ at the surface of a solid torsion bar sample can be estimated from the mean stress $\sigma_m = F/\pi R^2$ by using the formula derived in Ref 74:

$$\sigma_{zz}(R) = \frac{F}{\pi R^2} \times \left(1 + \frac{1}{2}\frac{\partial \ln F}{\partial \ln N} + \frac{1}{2}\frac{\partial \ln F}{\partial \ln \dot{N}}\right) \qquad \text{(Eq 21)}$$

where F is the axial force, and N and $\dot{N}$ are the number of revolutions and twist rate, respectively. This relation only applies rigorously to materials displaying simple hardening laws, as discussed previously in the section on the Fields and Backofen relation for deriving the shear stress $\sigma_{\theta z}$ (referred to as τ), from the moment M.

When σ_{zz} and the circumferential stress $\sigma_{\theta\theta}$ are nonzero (the radial stress σ_{rr} can be taken to be zero at the surface of a torsion specimen), the following relation applies for effective stress (Ref 74):

$$\bar{\sigma} = \left[3\sigma_{\theta z}^2 + (\sigma_{\theta\theta}^2 + \sigma_{zz}^2 - \sigma_{\theta\theta}\sigma_{zz})\right]^{1/2} \qquad \text{(Eq 22)}$$

Detailed calculations (Ref 74, 75) show that the correction associated with the inner bracketed term on the right side of Eq 22 is only about 1% for fcc metals and attains a maximum of about 8% for bcc metals, which tend to develop higher ratios of axial to transverse shear stress. Because the magnitude of the correction depends on the square of the $\sigma_{zz}/\sigma_{\theta z}$ ratio, where the latter is always $\ll 1$, the error introduced by neglecting σ_{zz} in the calculation of the effective stress is generally small.

Typical Flow Curves

Flow-stress results for type 304L austenitic stainless steel at two strain rates and a variety of cold, warm, and hot working temperatures are shown in Fig. 26. Note that the shear-strain levels in these curves are about 5, corresponding to von Mises effective strains in excess of 2.5 or to values much greater than those obtained in tension or compression tests. All tests shown in Fig. 26 were taken to failure. However, the flow stresses for the hot working tests are only shown for $\gamma \le 5$.

The flow-stress behavior observed in torsion is best interpreted in terms of test temperature and strain rate. For the type 304L alloy (Fig. 26), the flow stresses generally increase with strain (or remain constant with strain) at cold and warm working temperatures. Only at 20 °C (68 °F) and at 200 °C (390 °F) and a von Mises surface effective strain rate of $\dot{\bar{\varepsilon}}_s = \dot{\gamma}_s/\sqrt{3} = 10$ s^{-1} do the flow curves exhibit softening. Under these conditions, scribe-line measurements indicate that flow localization had occurred prior to fracture. All other scribe-line observations indicate fracture-controlled failure.

In contrast to these flow curves are those obtained at hot working temperatures, all of which reveal maxima followed by decreasing flow stresses and eventually steady-state flow. This trend is indicative of dynamic microstructural changes characterized by recrystallization phenomena. However, scribe-line measurements confirmed that, despite the flow-stress decrease, failure was fracture controlled in all cases, due to the stabilizing influence of the increased strain-rate sensitivity.

The flow curves for type 304L stainless steel shown in Fig. 26 also indicate the influence of strain rate on flow behavior. At cold and warm working temperatures, strain rate has only a slight effect on flow response. In fact, the 10 s^{-1} curve at a given temperature eventually drops below the curve measured at 0.01 s^{-1}. The lower rate can be considered isothermal and the higher rate adiabatic. Thus, the crossover of flow curves at the two rates is a result of deformation heating and a relatively small strain-rate sensitivity (as shown by the initial portions of the flow curves, in which thermal effects are unimportant). Isothermal flow curves for 10 s^{-1} can be deduced by estimating the associated ΔT values and by constructing σ-T plots. This leads to isothermal high-strain-rate flow curves that are consistently above the lower-strain-rate flow curves.

In contrast to the trends at cold and warm working temperatures, the type 304L flow response in the hot working regime reveals a noticeable strain-rate effect. Under these conditions, the high-strain-rate curves are considerably above their low-strain-rate counterparts at a given test temperature. Such a response is the result of the high strain-rate sensitivity of most metals at hot working temperatures. The strain-rate sensitivity effect offsets any possible crossover due to deformation heating at the higher strain rates. Flow stresses in this temperature regime are much lower than those at cold and warm working temperatures. Because of this, ΔT values associated with the higher strain rate, which vary with the magnitude of τ and γ, tend to be smaller at the higher temperatures.

Flow-stress data from the torsion test for other alloys (such as Waspaloy, Fig. 27) exhibit similar variations with temperature and strain rate.

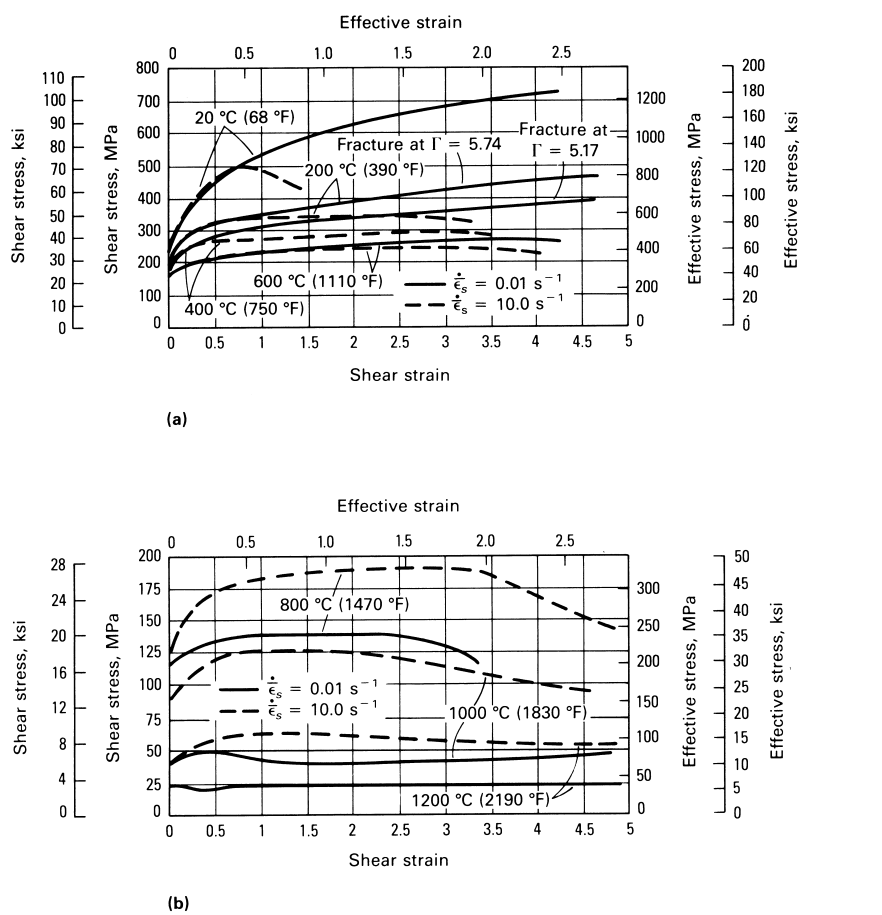


Fig. 26 Flow curves from type 304L stainless steel torsion tests. (a) Cold and warm working temperatures. (b) Hot working temperatures. Source: Ref 88

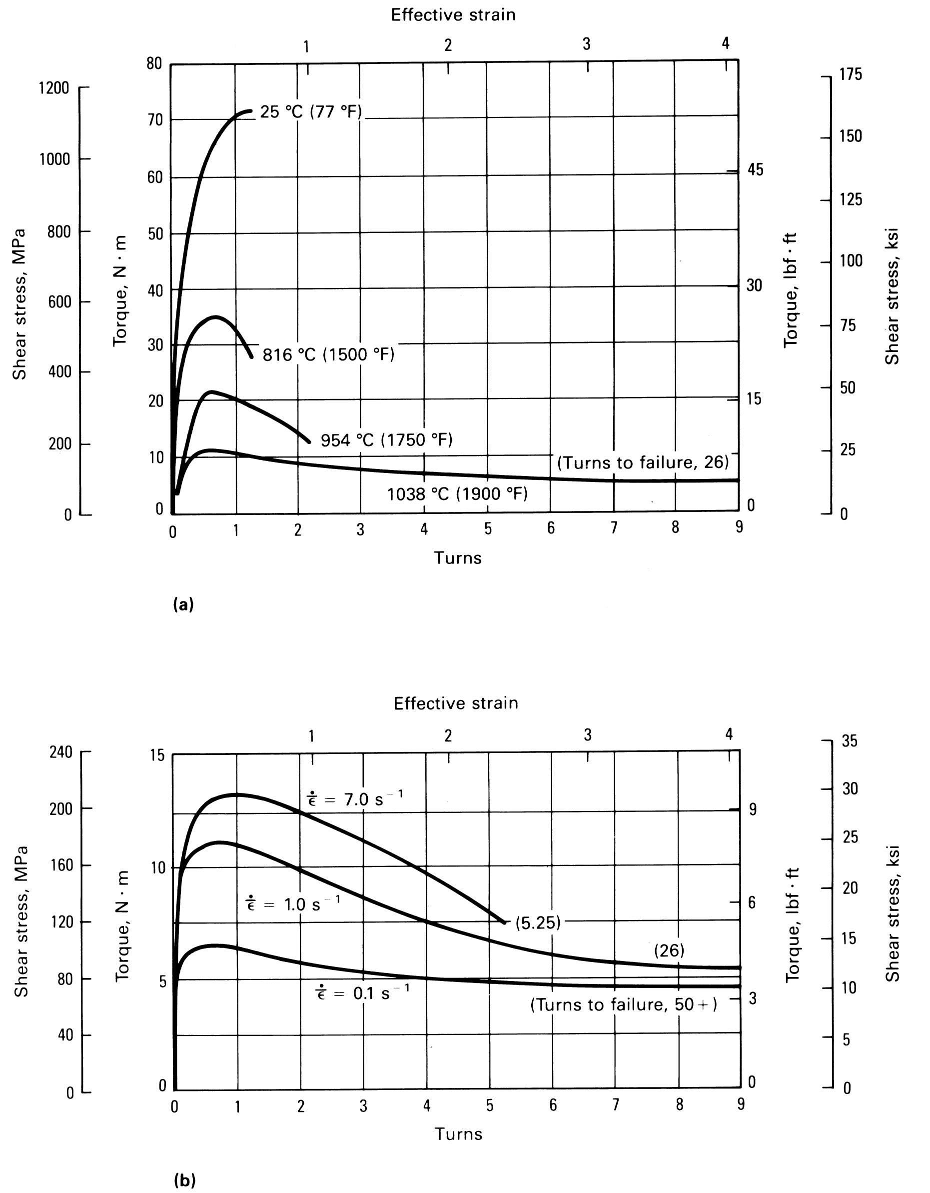


Fig. 27 Flow curves for Waspaloy. (a) Effect of temperature at a fixed effective strain rate of 1 s^{-1}. (b) Effect of strain rate at a fixed test temperature of 1038 °C (1900 °F). Flow softening at the higher temperature is a result of dynamic recrystallization. Source: Ref 48

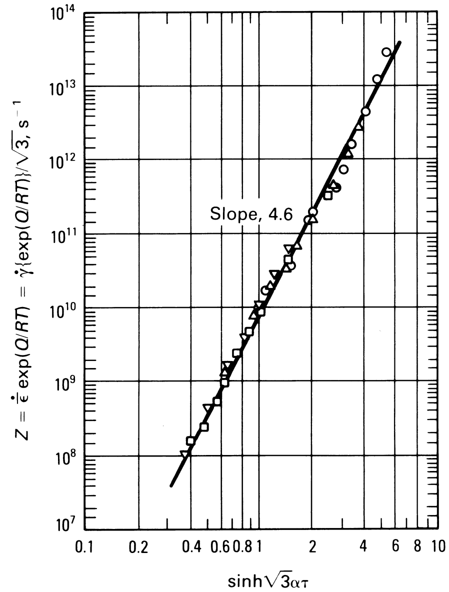


Fig. 28 Correlation of torsional flow stress data for a 0.25% C steel using a temperature-compensated strain-rate parameter (the Zener-Hollomon parameter, Z). Source: Ref 89

At cold and warm working temperatures, strain hardening often persists to large strains, except for high-rate tests, which are characterized by deformation heating and a general decrease in flow stress with temperature. At hot working temperatures, flow curves frequently show flow-stress maxima followed by flow softening and a steady-state flow stress. This behavior is associated with a variety of microstructural changes. Furthermore, the strain-rate sensitivity tends to be small ($m \leq 0.02$), except under hot working conditions, at which the value of m is typically between 0.1 and 0.3 for many metals and alloys.

The steady-state flow stress under hot working conditions measured in torsion tests (and a variety of other mechanical tests) frequently is a function of the Zener-Hollomon parameter, $Z = \dot{\bar{\varepsilon}} \exp(Q/RT)$, where Q is the apparent activation energy for the flow process involved, R is the gas constant, and $\dot{\bar{\varepsilon}} = \dot{\gamma}/\sqrt{3}$. For many single-phase materials, the relationship between Z and the steady-state values of τ is of the form:

$$(A \sinh \sqrt{3}\alpha'\tau)^{m'} = Z \qquad \text{(Eq 23)}$$

in which A, α', and m' are material constants. An example of the fit for a 0.25% C steel is shown in Fig. 28.

Comparison of Torsion Flow-Stress Data to Data from Other Workability Tests

Shear-stress/shear-strain data from torsion tests are often converted to "effective" flow-stress data using the von Mises relations. In this form, they are useful for load-prediction and metalworking analysis. However, even when converted to effective terms, flow-stress data obtained from torsion tests do not always show perfect agreement with data from other mechanical tests.

The differences are usually greatest at cold-working temperatures. They are a consequence of one or a combination of several factors, such as the effect of deformation path (tension, compression, etc.) on the development of crystallographic texture (Ref 90–92) and on the nature of the microscopic slip or twinning processes that control the observed macroscopic strain-hardening rate. The variability of such factors is less likely at hot working temperatures because of the dynamic restorative processes that may prevent sharp textures from being formed. Also, a significant amount of deformation occurs at grain boundaries as well as in the matrix of metals. The stress obtained in torsion often lies *below* that obtained by axisymmetric methods. Such differences have been noted at low (Ref 92–94) as well as hot working temperatures (Ref 32, 59, 96, 97). Nadai and Davis (Ref 95) used an expression for equivalent strain similar to that of Eichinger (Ref 98) that reduced the discrepancy.

Sakai and Jonas (Ref 99) have noted that the strain to the onset of dynamic recrystallization is delayed in torsion when compared with compression test results at the same equivalent strain rate. The maximum stress, however, does not differ markedly between the two testing methods. Inasmuch as a critical level of work hardening is thought to be required to initiate dynamic recrystallization at a given strain rate and temperature, an increase in the critical strain (by a factor of 2 to 3 times) can be understood if the

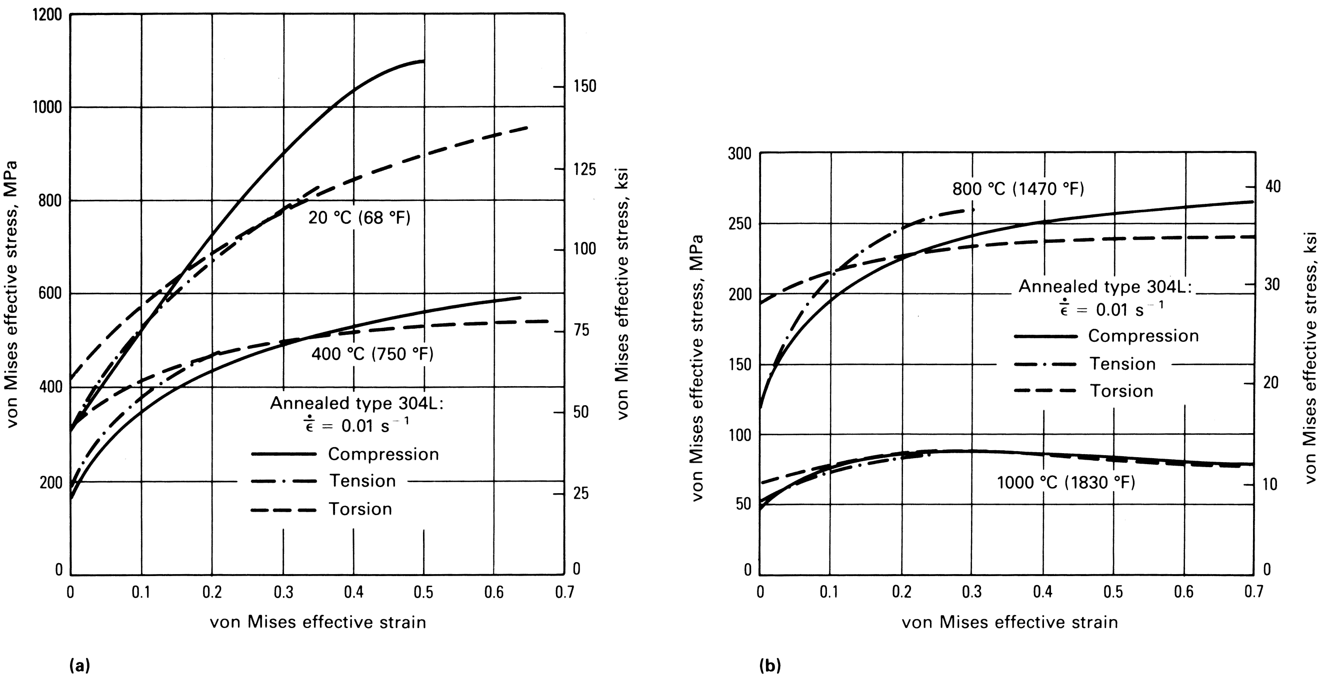


Fig. 29 Comparison of effective stress-strain curves determined for type 304L stainless steel in compression, tension, and torsion. (a) Cold and warm working temperatures. (b) Hot working temperatures. Source: Ref 100

work-hardening rates in torsion differ from those in tension and compression. Canova et al. (Ref 59) have shown that the material flow in the surface of a torsion bar follows the pattern of simple shear. These observations are consistent with the conclusion that the state of strain in torsion is one of simple shear (Ref 87). It can then be argued (Ref 59) that at large strains dislocation motion would tend to be restricted to transverse planes in the torsion bar and that the active Burgers vectors would be close to the circumferential direction. Accordingly, strain is essentially accumulated by macroscopic slip on a single set of parallel planes, and the work-hardening rate should be lower than in tension or compression, where two or more macroscopic slip planes are active. This view adequately accounts for the discrepancies between the flow behavior of materials in torsion and in tension and compression, and it preserves the validity of the expressions for equivalent stress and strain. It raises serious questions, however, as to the true equivalency of alternative testing procedures when the microstructural state is of prime importance, as it is in hot working. It also has impact on the question of hot-ductility determination.

In Fig. 29, torsion flow-stress data for type 304L are compared to compression and tension data in terms of von Mises effective stress and strain. At cold and warm working temperatures, as well as at low hot working temperatures (800 °C, or 1470 °F), the flow curves from the various tests do not coincide. Generally, there is a lower level of strain hardening in torsion. Thus, although the overall stress levels are similar, the actual shapes of the curves are quite different. Even if other definitions of effective stress and strain are employed, the differences between the curves cannot be eliminated. However, an estimate of the working loads can still be derived from torsion data plotted in von Mises terms. In contrast to the 20, 400, and 800 °C (68, 750, and 1470 °F) behaviors, comparison of type 304L torsion data to tension and compression data is quite good at the hot working temperature of 1000 °C (1830 °F) (Fig. 29b). This is most likely a result of the absence of marked textural and strain-hardening effects.

Other metals show the same divergence between torsion flow curves and those obtained by other test techniques. Data for copper and aluminum (tested at room temperature) are given in Fig. 30 and 31. Comparisons with tension and plane-strain compression results are given in Fig. 30 and 31, respectively. The effective stress-strain curves from torsion show consistently lower levels of strain hardening as well as lower flow stresses.

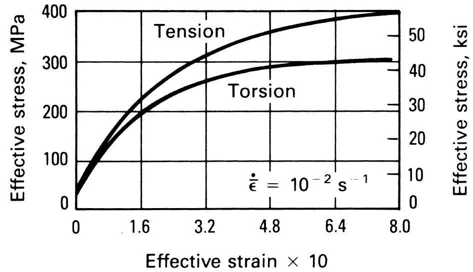


Fig. 30 Flow curves determined at room temperature in tension and torsion on oxygen-free high-conductivity copper. Source: Ref 87

Figure 32 illustrates results for a material with a crystal structure different from the fcc materials in Fig. 30 and 31. These effective stress-strain curves are for bcc carbon steels that were tested in torsion and tension. To achieve the high deformation levels in tension, samples were prestrained by wire drawing. The von Mises flow curves for torsion lie below those for tension for

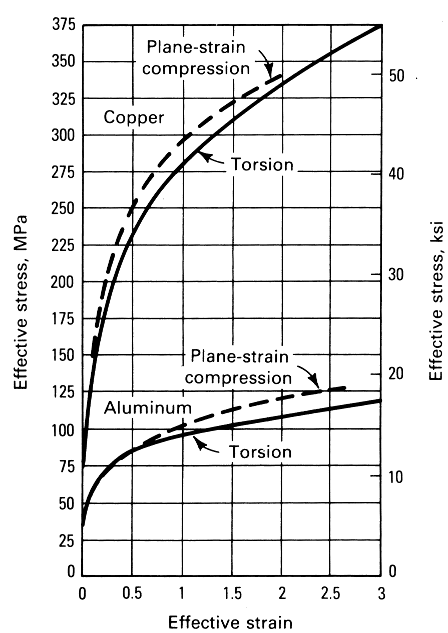


Fig. 31 Comparison of room-temperature flow curves from torsion and plane-strain compression tests on copper and aluminum. Source: Ref 32

both the low-carbon steel and the pearlitic, near-eutectoid high-carbon steel.

Also shown in Fig. 32 are effective stress-strain curves from torsion tests that were calculated on the basis of the Tresca criterion, in which the effective stress is equal to 2τ and the effective strain to $\gamma/2$. A divergence between torsion and tension is still present.

The differences between torsion flow stresses and results from other tests described previously have much less of an effect on load predictions than on the estimation of formability such as in sheet stretching. In these operations, workability is controlled by the onset of instability and flow localization, which are influenced strongly by the rate of strain hardening, as expressed by the strain-hardening exponent, n. Because the n values obtained from torsion tests are typically below values obtained from the other tests, they should be used with caution in formability modeling.

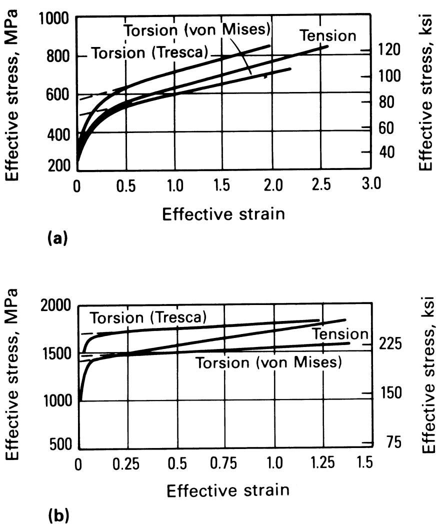


Fig. 32 Flow curves determined via torsion testing and tension testing (following wire drawing). (a) 0.06% C steel. (b) 0.85% C steel (in pearlitic condition). Note that the torsion data are expressed in terms of both the von Mises and the Tresca effective stress-strain definitions. Source: Ref 101

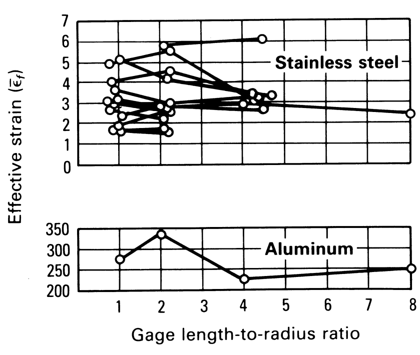


Fig. 33 Effect of the gage length-to-radius ratio on the effective strain to failure ($\bar{\varepsilon}_f$) in torsion tests. Lines join results at similar strain rate and temperature. Source: Ref 35

Interpretation of Torsion Fracture Data

The fracture strains measured in torsion are highly reproducible and provide a good quantitative measure of the ductility of materials undergoing high-strain deformation. Frequently, torsion tests are conducted to determine the effects of process conditions (e.g., strain rate and temperature) and material composition on workability. A large number of hot ductility tests have been performed successfully on a variety of materials over a wide range of temperatures and strain rates (Ref 3, 11–14, 24, 27, 29–32, 35, 37, 40, 45, 55, 91, 102–108). It is not the purpose of this section to review the test results obtained in these studies, but rather to discuss the test procedures, some typical results, and their application. For extensive reviews of these and other data, the reader is referred to a number of excellent reviews on the topic (Ref 22, 89, 109, 110).

Effect of Specimen Size

For round-bar samples, the surface strain at fracture is usually relatively independent of specimen design (e.g., length-to-radius ratio), as shown by the results from torsion tests at hot working temperatures for type 304 stainless steel and aluminum given in Fig. 33.

In another study of the hot ductility of ferrous alloys, Reynolds and Tegart (Ref 14) reported that the number of twists to failure increased with radius when specimens of constant length-to-radius ratio were used. It has been shown in a number of studies (Ref 27, 45, 55) that ductility in torsion is determined by the growth rate of surface cracks that propagate inward as straining proceeds. Because surface oxidation promotes initiation and propagation of cracks at elevated temperatures, the change in ductility with specimen radius could be rationalized in terms of the associated change in surface-to-volume ratio (Ref 21).

In tubular specimens, it has been found that the bore diameter and wall thickness do significantly influence the measured ductility (Fig. 34). This observation points to an effect that is possibly the source of the enhanced ductility sometimes observed in torsion: the constraint of the solid core. In thin-wall tubes, cracks generated at the outer surface, at the highest stress, only have to grow a distance equal to the wall thickness to produce complete failure. This occurs rapidly because the stresses are generally high throughout the wall. By contrast, surface-initiated cracks in solid-bar specimens must grow along the entire radius, under the influence of stresses that decrease markedly with distance. The central core of a solid specimen, therefore, acts to inhibit complete failure and thereby enhances ductility.

Variations with specimen geometry may also occur when experiments are run at moderate strain rates or at rates at which specimen geometry affects the temperature profile due to heat conduction. In these cases, the fracture strain, initial test temperature, and estimated temperature at fracture should be reported together.

Effect of Temperature and Alloying on Torsional Ductility

In torsion, the ductility of most metals is moderate at cold working temperatures ($\gamma_f \leq 1$ to 5), least at warm working temperatures, and greatest at hot working temperatures, at which γ at fracture often exceeds 10 or more. At cold working temperatures, fracture occurs by ductile fracture initiated at second-phase particles and inclusions. The decrease in ductility at warm working temperatures occurs because of the

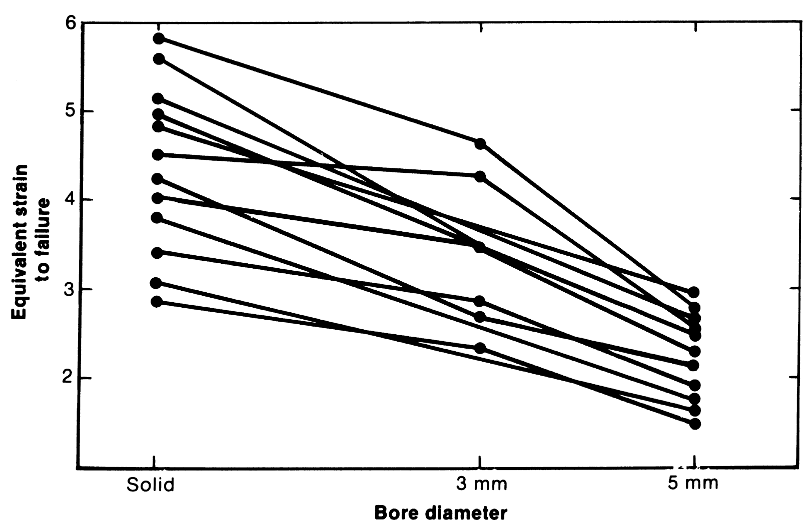


Fig. 34 Effect of bore diameter on strain to failure (ε_f) of type 304 stainless steel specimens with a 7 mm (0.3 in.) gage diameter and a gage-length-to-radius ratio of 2:1 (lines join data points obtained at similar strain rates and temperatures). Source: Ref 35

thermal activation of grain-boundary sliding, which culminates in brittle intergranular failures. At hot working temperatures, dynamic restoration processes such as recovery and recrystallization act to heal incipient voids (due to cavitation) and microcracks, thereby increasing ductility substantially.

These processes are most rapid in single-phase pure metals, which tend to have the largest ductilities at hot working temperatures. Alloys with solute elements tend to have lower ductilities because of the increased difficulty of dynamic restoration. Moreover, in such materials, the onset of true hot working conditions occurs at higher temperatures than in pure metals. This is also true for alloys with two or more phases in which deformation tends to be inhomogeneous because of the variation of properties between the two phases, which results in poor ductility.

In the hot working regime, ductility does not increase monotonically with temperature up to the melting or solidus temperature. Frequently, ductility passes through a maximum, which usually correlates well with the optimal temperature for forging or extrusion. Above this temperature, ductility may decrease because of effects such as grain growth, deformation heating, or incipient melting (hot shortness). Increases in grain size promoted by very high temperatures increase the tendency for intergranular fracture, while retained deformation heat at moderate to high strain rates may be sufficient to increase the instantaneous local temperature above the incipient melting or solidus temperature, thus resulting in poor ductility.

Figure 35 summarizes the qualitative dependence of workability on temperature for a wide range of alloys. These trends are classified by the number of phases, the types of compounds that are present at various temperatures, and so on and are a valuable reference tool when new alloys for which quantitative data are unavailable are being formed.

Quantitative fracture data from torsion tests are usually reported in terms of twists to failure or surface fracture strain. Although using twists to failure as a measure of ductility is somewhat qualitative unless the specimen geometry is specified, it is often used to document the effect of temperature and to establish optimal working conditions. Examples of this type of torsion data obtained from hot torsion experiments are shown in Fig. 36 for several steels and nickel alloys.

As shown in Fig. 36, carbon and alloy steels such as 1040 and 4340 are very workable. These steels fall into group VII of Fig. 35. As the temperature is decreased, a two-phase ferrite plus austenite structure ($\alpha + \gamma$) is formed from the single-phase austenite (γ), causing a sharp decrease in ductility. This is also shown for a high-oxygen iron alloy in Fig. 37.

Types 304 and 410 stainless steels are generally single-phase alloys at hot working temperatures and exhibit increasing ductility with temperature. However, at very high temperatures, δ ferrite is formed in type 410, causing a drop in workability. Therefore, type 304 exhibits a behavior like a group I material, and type 410 behaves like a group V alloy in Fig. 35.

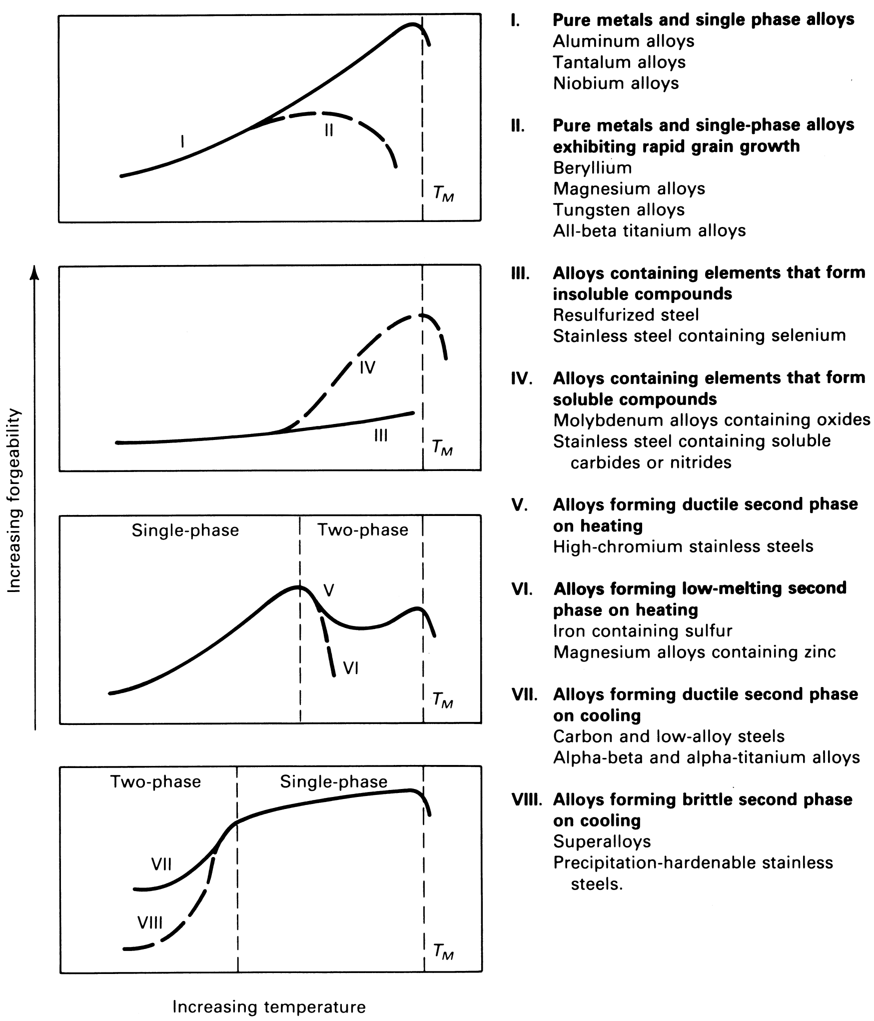


Fig. 35 Typical workability behaviors exhibited by different alloy systems. T_M = melting temperature. Source: Ref 111

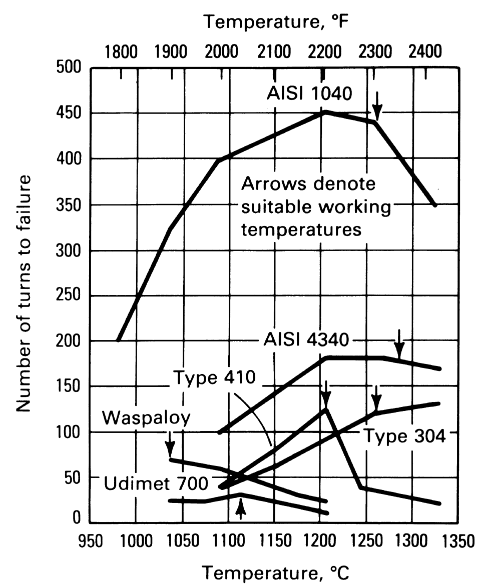


Fig. 36 Ductility determined in hot torsion tests. Source: Ref 112, 113

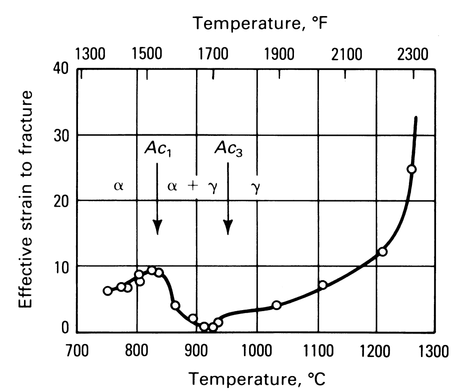


Fig. 37 Variation of ductility with temperature for a high-oxygen Swedish iron tested in torsion at an effective strain rate of 0.5 s^{-1}. Source: Ref 14

The nickel-base superalloys in Fig. 36 show inferior ductility compared to types 304 and 410 steels due to the generally high alloy content of the former and the formation of hard, second-phase particles at low to moderate hot working temperatures in these materials. These alloys exhibit group VIII behavior. Because the second phase is formed at relatively high temperatures, the superalloys must be worked close to the melting point. Hot torsion data in this range often show a decrease in ductility with temperature because of incipient melting problems.

Several other alloying effects in steels and aluminum alloys are illustrated in Fig. 38 to 40. Figure 38 shows the effects of carbon on the torsional ductility of plain-carbon steels and electrolytic iron-carbon alloys at 0.5 T_M. For both types of alloys, workability is controlled by ductile fracture due to cavitation at precipitates and inclusions.

For the plain-carbon steels, the ductility increases initially with carbon content because of a decrease in the number of oxide inclusions, which serve as void nucleation sites. Similarly, with carbon contents above 0.1%, ductility decreases with carbon content as a result of the increasing volume percent of carbides at which voids may also nucleate. The variation of ductility with carbon content is similar for the electrolytic alloys, although the level is generally higher. This trend is the result of a decrease in the volume fraction of manganese sulfide and oxide inclusions.

Quantitative analysis of the effect of sulfides on the workability of steels is shown in Fig. 39, in which the ductility of several rimmed steels is presented as a function of sulfide content. As inclusion content increases, the fracture strain decreases and thus workability during an actual forming operation is affected deleteriously.

Figure 40 gives the torsional ductility of pure aluminum and a series of aluminum-magnesium alloys as a function of temperature. At the testing temperatures used, magnesium is totally in solution. The observed trend is caused by increased amounts of grain-boundary sliding in the magnesium alloys as well as by a decrease in the kinetics of the dynamic softening processes. The latter leads to an increase in flow stress, thereby reducing the ability of the alloy to accommodate stress concentrations at inclusions and grain boundaries.

Effect of Strain Rate on Torsional Ductility

The effect of strain rate on workability should be considered when attempting to apply laboratory torsion measurements to metalworking processes carried out at much higher speeds. In many cases, the general dependence of ductility on temperature will be similar at different strain rates. However, the ductility peak and thus optimal working temperature may be shifted. An example of this effect is shown in Fig. 41 for the nickel-base superalloy Udimet 700. For this material, the ductility maximum is shifted to higher temperatures with increases in strain rate. This behavior occurs because the alloy is single phase at the temperatures under consideration, and the flow stress of single-phase alloys is a function of the Zener-Hollomon parameter ($Z = \dot{\varepsilon} \exp (Q/RT)$). Thus, as the strain rate is increased, similar flow and fracture processes should occur at correspondingly higher temperatures. As shown in Fig. 41, however, the peak ductility is not as large at high strain rates. This can be ascribed to the increased amount of deformation heating and the temperature increase at these rates, which raise the alloy to its incipient melting point.

The influence of strain rate on torsional ductility is also illustrated in Fig. 42(a) for type 304L stainless steel, in which the ductilities from tests conducted at von Mises surface effective strain rates of 0.01 and 10.0 s^{-1} are plotted. The values of γ_{sf} from the lower rate test are plotted versus the (constant) test temperature at which they were conducted. The high-strain-rate ductilities are plotted versus the estimated temperature at fracture. These estimated temperatures are equal to the nominal test temperature plus the adiabatic ΔT from the plots shown in Fig. 42(b). Ductility data for high-strain-rate tests performed at 20 °C (68 °F) and 200 °C (390 °F), in which flow localization occurred, are also shown. In these two instances, γ_{sf} was obtained from scribe-line measurements, and the ΔT versus γ plot was extrapolated to obtain the fracture temperature.

The low-strain-rate data for type 304L shown in Fig. 42(a) illustrate a classical dependence on temperature, that is, a modest ductility at cold working temperatures and a ductility minimum at warm working temperatures. The major effect of the higher strain rate is a translation of the lower-strain-rate data to higher temperatures. For example, the ductility minimum appears to be shifted.

A more detailed interpretation of low- and high strain-rate-data can be made through the Zener-Hollomon parameter, $Z = \dot{\varepsilon} \exp (Q/RT) = \{\dot{\gamma} \exp (Q/RT)\}/\sqrt{3}$. For this alloy,

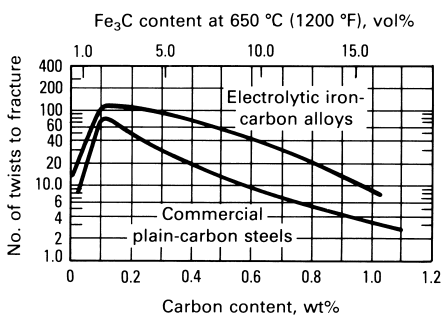


Fig. 38 Influence of carbon on the ductility of bcc iron tested in torsion at 0.03 s^{-1} and 650 °C (1200 °F), or one-half the absolute melting point. Source: Ref 21

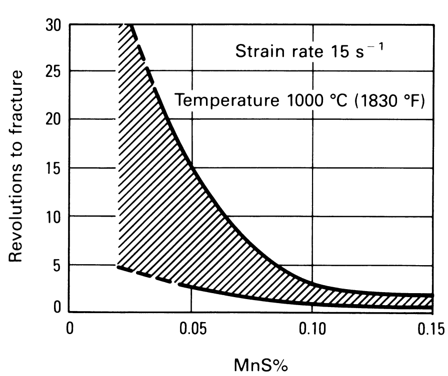


Fig. 39 Relation between manganese sulfide content and ductility in hot torsion tests on a variety of rimmed steels. Source: Ref 31

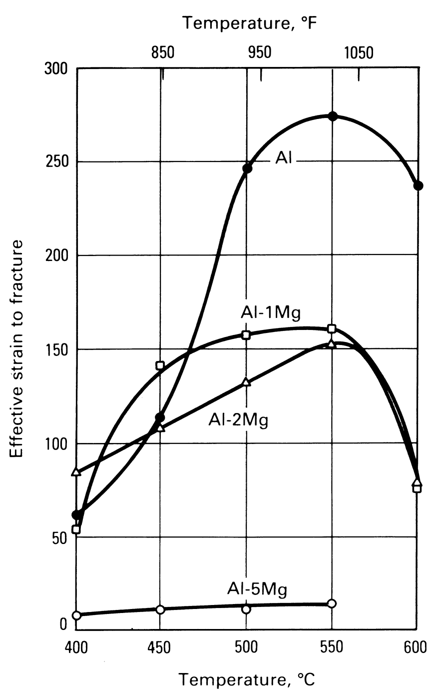


Fig. 40 Effect of the amount of magnesium in solid solution in aluminum on torsional ductility. Tested at an effective strain rate of 2.3 s^{-1}. With 5% Mg, ductility is severely reduced over the entire temperature range. Source: Ref 29

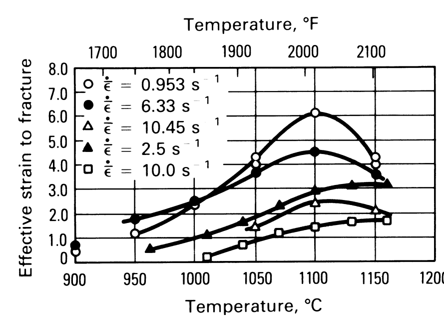


Fig. 41 Effect of test temperature on the torsional ductility of Udimet 700. Source: Ref 114

the steady-state flow stress at equal values of Z is nearly the same. Therefore, equal failure strains can be expected at fixed values of Z. This hypothesis may be checked by "translating" the low-strain-rate fracture locus $\gamma_{fa}(T_{fa})$ to $\gamma_{fb}(T_{fb})$ defined by $\gamma_{fb} = \gamma_{fa}$ and $\dot{\gamma}_b \exp(Q/RT) = \dot{\gamma}_a \exp/(Q/RT_{fa})$, where $\dot{\gamma}_b/\dot{\gamma}_a = 10/0.01 = 10^3$. For these type 304L stainless steel results, Q is taken to be the same as the activation energy from flow-stress data (98.2 kcal/mol, or 411 kJ/mol). The results of the calculation are shown in Fig. 42(a) and appear to be in fair agreement with the measured high-strain-rate fracture locus.

Nicholson et al. (Ref 115) corrected their ductility-temperature curves for deformation-heating effects in a similar fashion (Fig. 43).

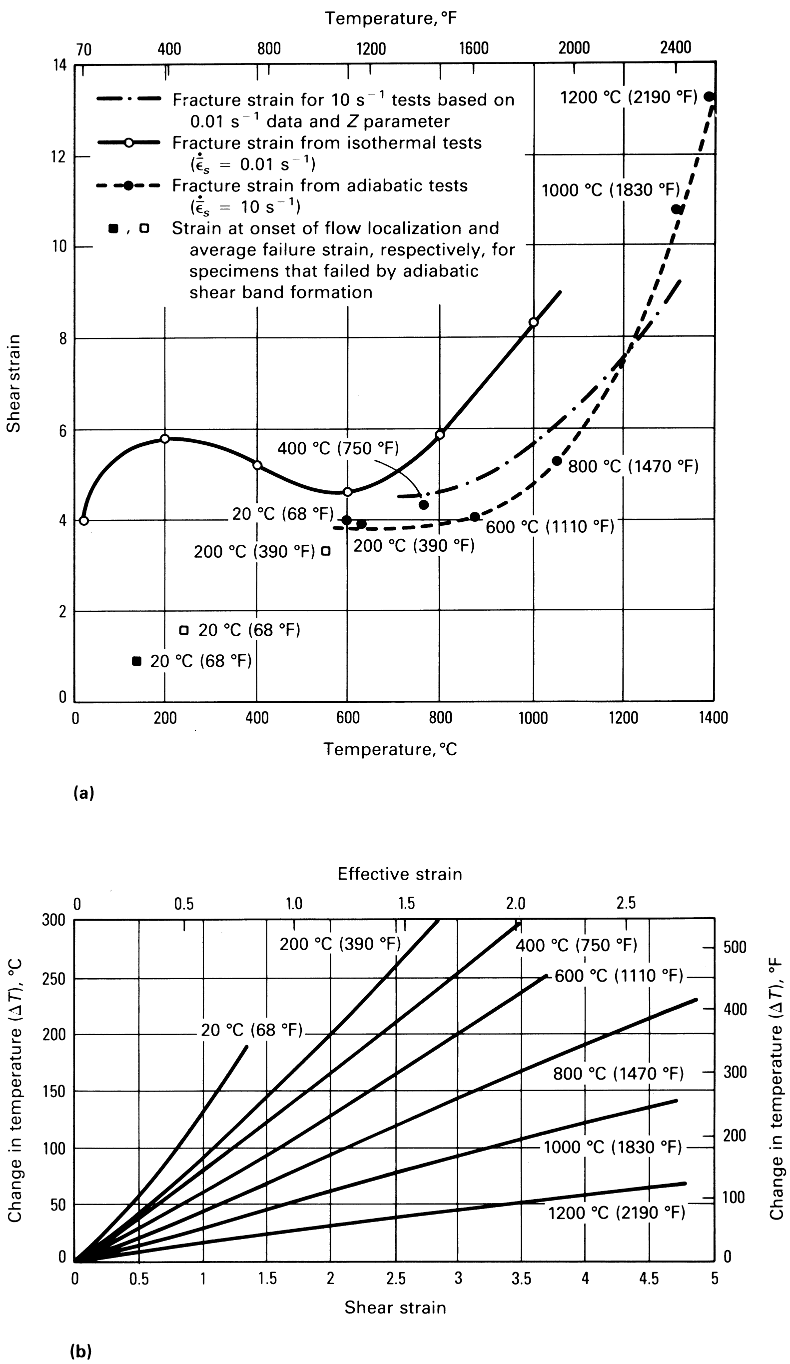


Fig. 42 (a) Fracture-strain data from type 304L austenitic stainless steel torsion tests and (b) estimated temperature changes during high-rate tests. Low-strain-rate ($\dot{\bar{\varepsilon}} = 0.01$ s^{-1}) data are plotted versus the actual test temperature. High-strain-rate (10.0 s^{-1}) data are plotted versus temperatures estimated from (b) a ΔT-γ plot and the nominal test temperature, which is shown beside each data point. Source: Ref 88

Correlation of Torsional Ductility Data to Other Workability Data

Because metalworking processes are not carried out under a state of pure torsional or shear loading, it is often necessary to convert the workability parameter measured in torsion to an index that is compatible with other deformation modes. Previously, attempts were made to correlate torsion, tension, and other types of data using the effective strain concept. Although a definite relationship exists between torsion and tension effective fracture strains (Fig. 44), such a method is incapable of explaining how free-surface fracture is avoided in homogeneous compression by preventing barreling. For fracture to occur, deformation must also involve tensile stresses to promote ductile fracture, wedge cracking, or some other failure mechanism.

One of the most successful hypotheses incorporating the effects of deformation and tensile stress is that proposed by Cockcroft and Latham (Ref 116). They postulated that fracture occurs after the maximum tensile stress (σ_T) carries out a fixed amount of work through the applied effective strain, or:

$$\int_0^{\bar{\varepsilon}_f} \sigma_T d\bar{\varepsilon} = \text{a constant } (C) \qquad \text{(Eq 24)}$$

The constant in the relation is a function of material, purity, and test temperature. In a torsion test, the maximum tensile stress occurs at 45° to the torsion axis and is equal to τ. As mentioned previously, $\bar{\varepsilon}$ for torsion is equal to $\gamma/\sqrt{3}$. If the material under consideration exhibits power-law strain hardening, $\tau = K'\gamma^n$, then C is equal to $(K'\gamma_{sf}^{n+1})/(\sqrt{3}(n+1))$, where γ_{sf} is the shear fracture strain. If the material does not harden in a power-law manner, the integration to evaluate C at fracture can be performed graphically.

In uniaxial tension, σ_T is the axial tensile stress, and $\bar{\varepsilon}$ is the axial strain. Upon necking, σ_T is higher than the effective flow stress by a correction factor that can be estimated from the work of Bridgman (Ref 117). This factor is on the order of 0 to 33% for true axial strains of 0 to 2.5 in the necked region. For a power-law-hardening material, $\sigma = K\varepsilon^n$. For tension, $C = (1 + \langle CF \rangle) \cdot \{K\varepsilon_f^{n+1}/(n+1)\}$, where $\langle CF \rangle$ is the average correction factor and ε_f is the axial fracture strain determined by reduction-in-area measurements and the constant volume assumption of plastic flow.

To apply the Cockcroft and Latham criterion to other deformation modes, the critical value of C must be known, and the maximum tensile stress and effective strain must be estimated. For arbitrary deformation paths, estimating these quantities can be very difficult and often requires sophisticated mathematical techniques, such as finite-element methods. Once σ_T and $\bar{\varepsilon}$ are known, however, numerical techniques can be used to estimate the value of the maximum work integral, which can then be compared with the critical value of C at frac-

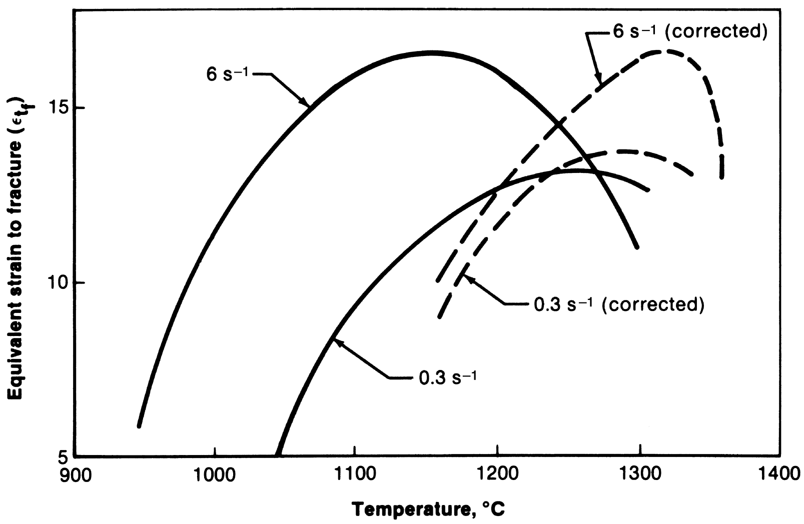


Fig. 43 Effect of strain rate on temperature dependence of fracture strain in torsion for type 321 stainless steel. The influence of adiabatic heating is shown by the corrected curves. After Nicholson et al. (Ref 115)

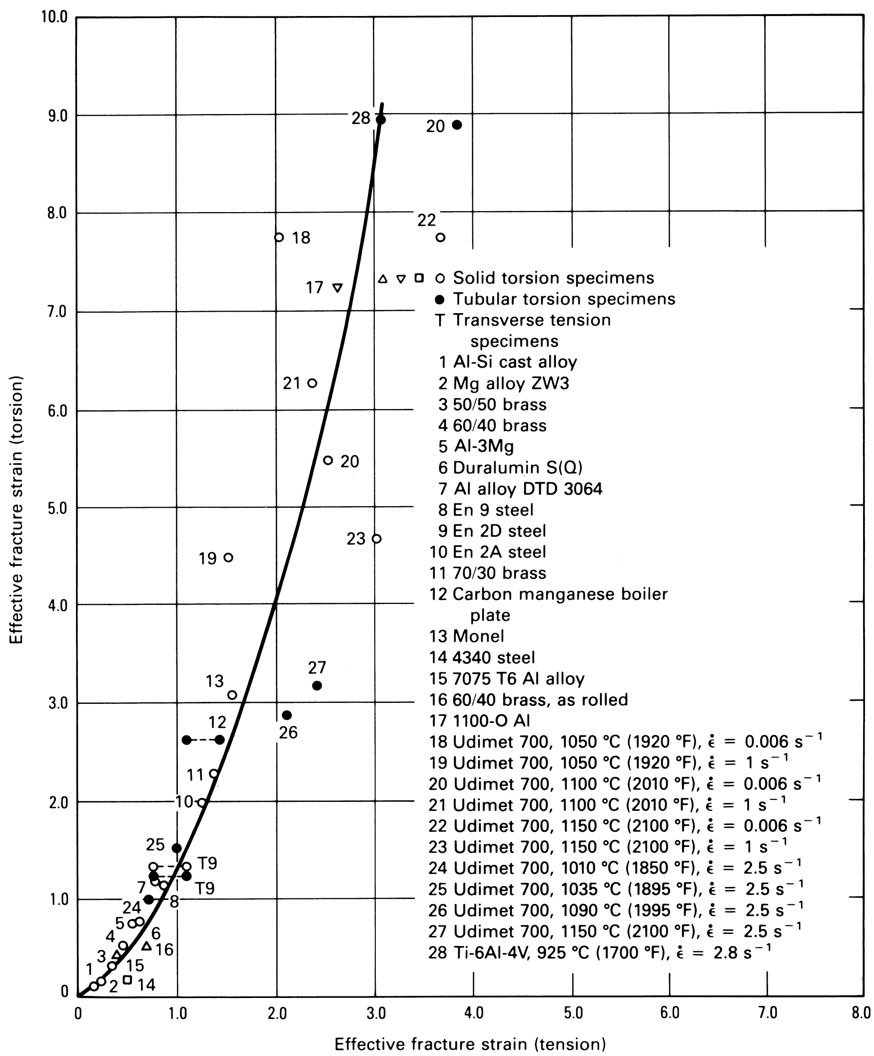


Fig. 44 Relation between effective fracture strain from tension and from torsion tests for several alloys. Source: Ref 114

ture determined by a simple workability test such as the torsion test.

The Cockcroft and Latham criterion has found its greatest success in correlating tensile and torsion fracture strains. The degree of success in these instances is judged by comparing the critical values of C established from flow and fracture results for the different tests. An example of such a comparison for type 304L stainless steel is given in Table 4. At 20 and 400 °C (68 and 750 °F), the material strain hardens in a power-law manner, enabling closed-form integration to obtain C values. At the two higher temperatures of 800 and 1000 °C (1470 and 1830 °F), a graphical integration procedure using flow curves (Fig. 29) was used. The integration procedure was straightforward for torsion, because deformation was uniform to fracture.

For tension, in which necking precludes obtaining flow stress data to fracture, the initial portion of the tensile curves was extrapolated to the required large strains using the compression and torsion curves as models. Results of the C parameter calculations for a given test temperature agreed within 10% except at 800 °C (1470 °F), where the two values differ by approximately 15%. The magnitudes of these differences are considered typical and acceptable, particularly because a gross continuum model was used for a mechanism that greatly depends on the presence of microstructural features such as inclusions. The same continuum model was successful over a wide range of temperatures, within which the fracture mechanism may vary significantly.

Attempts to include microstructural features and fracture mechanisms have met with less success than the Cockcroft and Latham formulation. One such attempt is that of Hoffmanner (Ref 118), who proposed a model that incorporates a factor describing the dependence of fracture strain on the magnitude of the normal stress perpendicular to the mechanical texture. Such a concept may have application in the analysis of fracture in torsion, during which the mechanical texture rotates with respect to the stress axes.

Because of the stress-state dependence of ductile failure, the axial stresses that sometimes develop during torsion may affect ductility. The ratio of the axial to the shear component of stress seldom exceeds 0.4, so the axial component has a limited influence on the effective stress. The axial stress, however, plays a significant role in determining the ductility measured by the shear or effective strain to failure. This has been clearly demonstrated by external application of longitudinal tension or compression during torsion (Ref 40, 45, 91, 105); data for Inconel 600 are shown in Fig. 45. When the ratio σ_A/τ is negative (compressive axial stress), the shear strain at fracture increases almost linearly with the magnitude of the ratio. In the regime of positive axial stresses, the curves show a marked upward curvature, which may, in part, reflect the development of geometric instabilities as the shear-to-normal stress ratio typical of tensile

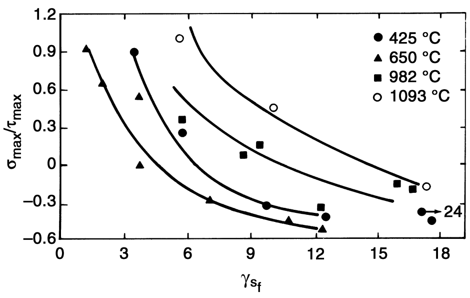


Fig. 45 Ratio of maximum axial stress to maximum shear stress versus values of shear strain at failure for Inconel 600. Source: Ref 45

Table 4 Cockcroft and Latham criterion for annealed type 304L fracture data

Test temperature, °C	Test temperature, °F	Deformation mode	Strain at fracture (torsional shear or tensile), γ_{sf} or $\bar{\varepsilon}_f$	Strength coefficient, G or K, MPa	Strength coefficient, G or K, ksi	Strain-hardening exponent, n	Average correction factor, ⟨CF⟩	C parameter, MPa	C parameter, ksi
20	68	Torsion	4.01	517	75	0.25	. . .	1360	197
		Tension	1.26	1172	170	0.352	0.075	1275	185
400	750	Torsion	5.17	310	45	0.137	. . .	1020	148
		Tension	1.51	772	112	0.307	0.105	1120	162
800	1470	Torsion	5.89	. . .	. . .	. . .	. . .	445	64.5
		Tension	1.94	. . .	. . .	. . .	0.12	525	76.1
1000	1830	Torsion	8.33	. . .	. . .	. . .	. . .	215	31.5
		Tension	2.75	. . .	. . .	. . .	0.14	235	34.1

Note: strain rate = 0.01 s^{-1}. Source: Ref 100

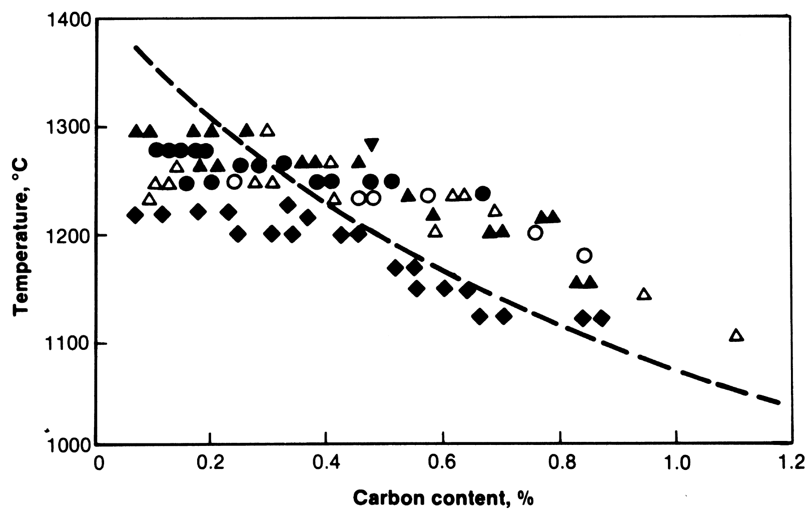


Fig. 46 Comparison of optimal ductility temperature determined in hot torsion tests at ~0.5 s^{-1} with operating temperatures for piercing, rolling, and forging of carbon steels. Dashed curve is for torsion; closed circles, Mannesmann tube mill; open circles, Stiefel tube mill; closed triangles, slabbing mill; open triangles, general rolling mill; closed diamonds, forging. Source: Ref 120

testing is approached. Although it was demonstrated that the σ_A/τ ratio influences initiation of cracks, the dominant influence is through the propagation phase. Dragon and his coworkers, (Ref 91, 105) have considered these and other observations and have argued that the influence of axial stress on ductility can be represented by:

$$\bar{\varepsilon}_T = \bar{\varepsilon}_0(1 + \phi\sigma_A\tau) \quad \text{(Eq 25)}$$

where $\bar{\varepsilon}_T$ is the total effective strain at failure on the application of an axial compressive stress σ_A, $\bar{\varepsilon}_0$ is the strain at fracture in the absence of an axial stress, and ϕ is a material constant that ranges from 2.5 to 3 for carbon steel, and from 10 to 20 for Inconel 600 at 900 °C (1650 °F), These authors suggest that $\bar{\varepsilon}_0$ is a measure of the "intrinsic" ductility of the material. Perhaps "ductility in simple shear" would be a more appropriate term, but the effects of the core constraints mentioned previously are not considered in this definition.

Correlation of Torsional Ductility Data with Working Practice

The hot torsion test has been used extensively to grade materials and assess optimal temperatures for particular working operations. Studies of the mechanisms that control hot ductility at high strains have met with reasonable success at relatively low temperatures for materials that undergo dynamic recrystallization (Ref 27, 89, 119). They have been less successful for materials that undergo only dynamic recovery during straining (Ref 89, 119). As yet, there remains no systematic basis for translating these basic concepts into a framework for workability predictions. On the other hand, measurement of hot ductility in torsion can successfully provide estimates of optimal working temperatures for several working operations. Zidek (Ref 120) has carried out a relatively complete survey of the relation between torsional ductility and hot working performance for carbon steel. His data are summarized in Fig. 46. These data confirm a reasonable correlation between the temperature of maximum ductility in torsion (dashed line in Fig. 46) and the optimal working temperature for a variety of practical working processes (data points in Fig. 46). The hot torsion test is shown to be particularly effective in predicting the optimal temperature for rotary piercing, which has been confirmed by other authors (Ref 3, 14). By contrast, this test tends to overestimate the ideal temperature for forging and extrusion. This discrepancy has been attributed in part to a propensity for deformation heating in torsion and piercing where high shear rates occur in a relatively small volume of material (Ref 109). The low-temperature limit for a particular working operation correlates with a specific total effective strain in torsion. Reynolds and Tegart (Ref 14), for example, showed that an effective strain of between 3 and 4 at $\dot{\bar{\varepsilon}} = 1$ s^{-1} is necessary in torsion testing to ensure successful extrusion, whereas $\bar{\varepsilon} = 18$ is required for successful rotary piercing.

Measuring Flow-Localization-Controlled Workability

Because torsion consists essentially of deformation in simple shear, torsion testing is frequently used to determine the material and process variables that contribute to the formation of shear bands during metalworking. Either solid or tubular specimens can be used for testing. When solid specimens are used, however, data analysis and interpretation are usually more complex than when tubular specimens are used. This difficulty arises because of the variation of strain, strain rate, and (at moderate to high strain rates) temperature across the specimen diameter.

Observation of Flow Localization in Torsion

The development of flow localization during torsion is usually detected by scribe-line measurements, variations in microstructure along the gage length, or torque-twist behavior. The first two of these techniques are unequivocal. Examples of flow localization detected by these two methods for the shear bands developed in tubular type 304L stainless steel torsion specimens twisted at $\dot{\bar{\varepsilon}}_s = 10$ s^{-1} at room temperature are shown in Fig. 47 and 48.

The use of the torque-twist behavior, on the other hand, to measure flow localization may be

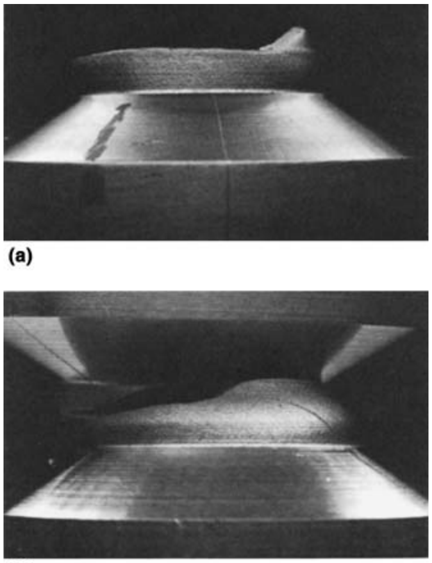


Fig. 47 Failed type 304L stainless steel torsion specimens from $\dot{\bar{\varepsilon}} = 10\ s^{-1}$ tests showing evidence of flow localization. (a) 20 °C (68 °F), average $\gamma_s = 1.3$. (b) 200 °C (390 °F), average $\gamma_s = 2.9$. Magnification: 2×. Source: Ref 88

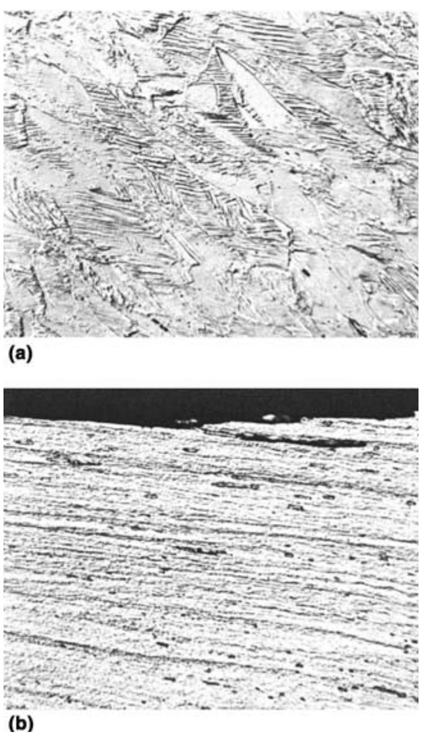


Fig. 48 Micrographs from room-temperature ($\dot{\bar{\varepsilon}} = 10\ s^{-1}$) torsion tests on type 304L stainless steel. (a) $\bar{\varepsilon} < 1.16$ (outside shear band). (b) $\bar{\varepsilon} > 1.16$ (inside shear band). Magnification: 310×. Source: Ref 88

misleading. Although a monotonically increasing torque-twist curve usually indicates homogeneous or nearly homogeneous flow, a curve that exhibits a torque maximum (often referred to as torque instability) and torque softening does not necessarily signify flow localization,

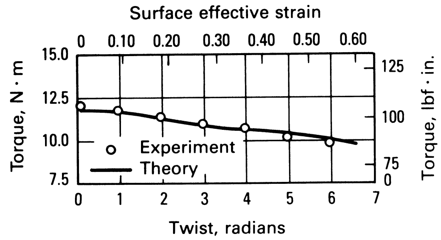


Fig. 49 Comparison of experimental and theoretical torque-twist curves for α + β (equiaxed alpha) microstructure Ti-6242Si hot torsion specimens. Tested at $\dot{\bar{\varepsilon}} = 0.9\ s^{-1}$; $T = 913$ °C (1675 °F). Source: Ref 52

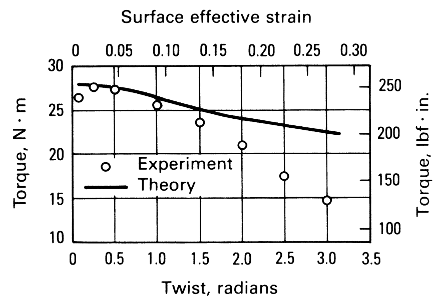


Fig. 50 Comparison of experimental and theoretical torque-twist curves for β (Widmanstätten alpha) microstructure Ti-6242Si hot torsion specimens. Tested at $\dot{\bar{\varepsilon}} = 0.9\ s^{-1}$; $T = 816$ °C (1500 °F). Source: Ref 52

particularly when the test alloy possesses a significant strain-rate sensitivity index.

The difficulties associated with the interpretation of torque-twist behavior are shown in Fig. 49 and 50, in which experimental data are presented for hot torsion tests on thick-wall tubular specimens of Ti-6Al-2Sn-4Zr-2Mo-0.1Si (Ti-6242Si) twisted at rates sufficiently high to minimize the effects of axial heat transfer. In Fig. 49, the alloy had a starting microstructure of equiaxed alpha (α + β microstructure); in Fig. 50, it had one of acicular, Widmanstätten alpha (β microstructure). The two torque-twist curves were similar in that they both exhibited torque maxima early in the deformation and subsequently decreasing torque with increasing twist. For the β microstructure, however, the softening rate was substantially greater. Scribe-line measurements revealed that the α + β specimen deformation was uniform, whereas the β specimen had undergone flow localization.

These results can be explained on the basis of a parameter that relates the degree of flow localization to the ratio of the torque-softening rate to strain-rate-sensitivity index. For the α + β microstructure test, the detrimental effects of torque softening were counterbalanced by a rather high strain-rate sensitivity, and localization was prevented in much the same way that it is during the tensile testing of superplastic materials. In contrast, for the β microstructure Ti-6242Si specimen, which also has a high strain-rate sensitivity, the larger amount of flow softening caused by the basic instability of the acicular microstructure was deduced to be the main reason for the localization behavior. For either case, however, a localization analysis that deals with the problem as a process rather than as an event is required to fully understand such observations.

Flow Localization Analyses

The main variables that control flow localization during torsion are material properties (mechanical and thermal), twist rate, and the presence of material or specimen imperfections. Of the material properties, strain- and strain-rate-hardening rates, temperature sensitivity of the flow stress, and thermal conductivity are the most important. Twist rate affects heat conduction and heat transfer. Material or torsion specimen imperfections offer sites for the initiation of flow localization. Such sites are particularly important when torsion is conducted at very low or very high rates, at which the axial temperature field is relatively uniform. Under these circumstances, temperature gradients and heat-transfer phenomena, which serve as a prime source of flow localization, are absent. The interpretation of torsion data under flow-localization conditions is discussed subsequently for two cases. In one, heat-transfer effects are not considered, while in the other their influence is taken into account during the analysis.

Analysis of Flow Localization in the Absence of Heat Transfer

Flow Localization Parameter. When heat-transfer effects are minimal, the study of flow localization in torsion aids understanding of the occurrence of shear bands in low-strain-rate isothermal metalworking operations, as well as those at very high strain rates. Examples of these processes include isothermal forging ($\dot{\bar{\varepsilon}} \cong 10^{-3}\ s^{-1}$) and high-energy-rate forming ($\dot{\bar{\varepsilon}} \cong 10^{3}\ s^{-1}$). In these cases, a defect is assumed to be the source of localization. For a torsion specimen, the defect may be a deficiency in radius (or wall thickness), strength coefficient, or another material property. The simplest imperfection to visualize is the geometric (or radius/wall thickness) defect, although the localization behavior is usually similar for similar sizes of geometric or material-property inhomogeneities.

When a torsion specimen with a geometric defect is twisted, variations in twist, twist rate, and temperature exist between the defect region and the nominally uniform region from the onset of deformation. Generally, all of these quantities will be higher in the defect area. These differences are generated to maintain torque equilibrium, so that the defect and uniform regions transmit an identical torque. They change as the deformation proceeds, depending on the material properties. Thus, localization, or the lack of it, must be viewed as a process. In contrast to the localization process and analysis, shear banding is often interpreted in terms of an instability condition. Such treatments are concerned only with material properties and with determining the twist, under a nominally homogeneous deformation field, at which the torque maximum ($dM = 0$) occurs. The results for Ti-6242Si discussed previously demonstrated that a torque maximum is not a sufficient condition for localization. Rapid localization cannot occur until after the instability condition is satisfied. However, if the material shows only a small amount of flow softening or has a large positive rate sensitivity, localization and the occurrence of shear banding in torsion will be minimal. Thus, the quantitative degree of localization cannot be predicted by the instability (torque maximum) analysis alone.

As its name implies, flow-localization analysis concerns variations of twist, twist rate, and temperature between a defect region and a nominally uniform region. During the entire strain-concentration process, the torque must be uniform along the axis, or $\delta M = 0$, where δ denotes a variation. In the present context, it is used to define variations of the various field quantities between the two regions of the torsion specimen. If $M = M(\theta, \dot{\theta}, T, R)$, where R is the specimen radius, the equilibrium condition $\delta M = 0$ may be used to obtain:

$$\frac{\delta \ln \dot{\theta}}{\delta \theta} = \frac{-\left\{ G + \left(\frac{\partial \ln M}{\partial \ln R}\right)\bigg|_{\theta,\dot{\theta}T} \frac{\delta \ln R}{\delta \theta} \right\}}{\left(\frac{\partial \ln M}{\partial \ln \dot{\theta}}\right)\bigg|_{\theta,T,R}} \quad \text{(Eq 26)}$$

where G denotes the normalized torque hardening (or softening) rate at a fixed $\dot{\theta}$:

$$G \equiv \left(\frac{d \ln M}{d\theta}\right)\bigg|_{\dot{\theta}} = \frac{\left(\frac{\partial \ln M}{\partial \theta}\right)\bigg|_{\dot{\theta},T,R} d\theta + \left(\frac{\partial \ln M}{\partial T}\right)\bigg|_{\theta,\dot{\theta},R} dT}{d\theta} \quad \text{(Eq 27)}$$

The derivation of Eq 26 and a detailed discussion of flow-localization analyses are presented in Ref 121. Once localization becomes noticeable, $\delta\theta \gg \delta \ln R$, and the second term in the braces in Eq 26 becomes negligible. Defining $(\delta \ln M/\delta \ln \dot{\theta})\big|_{\theta,T,R}$ as m^*, Eq 26 reduces to an expression for the torsional flow-localization parameter A:

$$A \equiv \frac{\delta \ln \dot{\theta}}{\delta \theta} = \frac{-G}{m^*} \quad \text{(Eq 28)}$$

The parameter A describes the variations of twist and twist rate that can be sustained under equilibrium conditions, as a function of the material properties through their influence on G and m^*. A material whose constant-strain-rate flow curve shows a large amount of flow softening will exhibit large negative G values. Similarly, a high strain-rate sensitivity, $m = \partial \ln \tau/\partial \ln \dot{\gamma}$, will result in large values of m^*.

A somewhat modified form of Eq 28 is useful in obtaining a quantitative estimate of the level of flow localization that can be expected during torsion testing. This form is obtained by substituting:

$$\dot{\theta} = \frac{(\sqrt{3})(L\dot{\bar{\varepsilon}})}{r_s} \text{ and } \theta = \frac{(\sqrt{3})(L\bar{\varepsilon})}{r_s} \quad \text{(Eq 29)}$$

where r_s and L are the outer radius and gage length of the specimen, respectively, into this equation to yield:

$$\left(\frac{\delta \ln \dot{\bar{\varepsilon}}}{\delta \varepsilon}\right) = -\frac{(\sqrt{3})(L)}{r_s} \frac{G}{m^*} \quad \text{(Eq 30)}$$

The left side of Eq 30 is known as the alpha (α) parameter. Much experimental and theoretical work has demonstrated that when α is equal to 5 or more, noticeable flow localization should be expected, either in torsion, compression, or some other deformation mode. When $G = 0$, which corresponds to the torque instability condition, $dM = 0$, α is equal to 0, and minimal localization is expected at this point. Beyond the torque instability, sufficiently large amounts of torque softening and/or low rate sensitivity are required to obtain values of α on the order of 5. Only when m^* ($\cong m$) is very small (as at cold working temperatures) will the onset of noticeable flow localization follow soon after the occurrence of the torque maximum.

Application of the Flow-Localization Parameter. An examination of Eq 30 indicates a dependence of the flow-localization parameter on specimen geometry, G, and m^*. As stated in Eq 27, G is the normalized torque softening rate under constant $\dot{\theta}$ conditions. This rate varies with θ, leading to the conclusion that α, and thus the rate of localization, can vary during torsion testing.

Difficulty arises, however, when attempting to use the measured M-θ curve from a torsion test in which localization has occurred, because the strain rate will have varied along the gage length even though the overall twist rate and average strain rate may have been held constant. Thus, it is often necessary to estimate G from the measured torque-twist behavior or to use some other mechanical test in which localization does not occur.

The values of G and m^* are readily estimated for thin-wall tube specimens. In this instance, m^* is identical to the strain-rate sensitivity index, m. G is obtained from the torque expression for this geometry:

$$M = 2\pi r_s^2 t\tau \quad \text{(Eq 31)}$$

Resulting in:

$$G = \frac{d \ln M}{d\theta} = \frac{d \ln \tau}{d\theta} = \frac{r_s}{L}\frac{d \ln \tau}{d\gamma} \quad \text{(Eq 32)}$$

Substituting this into Eq 30 and assuming $m^* = m$ yields:

$$\alpha = \left(\frac{\delta \ln \dot{\bar{\varepsilon}}}{\delta \bar{\varepsilon}}\right) = -\sqrt{3}\,\frac{\left.\frac{d \ln \tau}{d\gamma}\right|_{\dot{\gamma}}}{m} \quad \text{(Eq 33)}$$

or

$$\alpha = -\left(\frac{\left.\frac{d \ln \bar{\sigma}}{\delta \bar{\varepsilon}}\right|_{\dot{\bar{\varepsilon}}}}{m}\right) \quad \text{(Eq 34)}$$

in terms of von Mises effective stress and strain.

Equation 34 indicates that the α parameter for a thin-wall torsion-test specimen can be calculated directly from the normalized flow softening rate (at constant strain rate) and the strain-rate sensitivity index. Specimen geometry has no effect on the results. Furthermore, Eq 34 suggests that other mechanical tests may be useful in estimating values of $(d \ln \bar{\sigma}/d\bar{\varepsilon})|_{\dot{\bar{\varepsilon}}}$. In particular, tests in which flow localization can be avoided are preferable. The mechanical test in which localization occurs the most slowly is the uniaxial compression test due to the stabilizing effect of increases in cross-sectional area. To obtain values of the α parameter for thin-wall torsion tests, compression data from tests conducted at constant strain rates equivalent to those in torsion can be conducted to estimate the torque-softening rate and strain-rate sensitivity parameters.

When a thick-wall tube or solid-bar torsion test is to be analyzed, compression data are also useful in estimating the magnitude of the α parameter. Assuming that localization has been avoided in the compression tests, flow stresses as a function of strain, strain rate, and temperature can be converted to the equivalent shear stress τ as a function of γ, $\dot{\gamma}$, and T. For a given torsional $\dot{\theta}$, the shear stress can then be used via numerical integration to obtain the torque ($M = \int 2\pi r^2 dr$) required for the specific torsion specimen geometry. Note that this is the torque that would be required under uniform deformation conditions.

Calculations of M versus θ in such a case, which were based on the compression data, were performed for the thick-wall torsion specimens of the Ti-6242Si alloy discussed previously. The results (Fig. 51) demonstrate that the absolute magnitude of $G = (d \ln M/d\theta)\big|_{\dot{\theta}}$ was certainly much greater for the β microstructure than $\alpha + \beta$ microstructure. This behavior reflects the larger amount of flow softening that occurs in uniaxial compression at a variety of deformation rates.

An estimated constant-strain-rate M-θ curve is also plotted with each of the experimental results in Fig. 49 and 50. For the $\alpha + \beta$ microstructure, the predicted curve agrees quite well with the measured curve, confirming that the assumption of uniform deformation is valid. The maximum value of α over the twist range investigated is only 3.0. In contrast, the predicted M-θ curve for the microstructure in Fig. 50 lay substantially above the experimental results. This divergence, or the fact that deformation proceeded at torques below those required for uniform flow, is an indication of flow localization. In addition, the α parameter in this case is found to be as high as 14.3—much greater than the critical α of 5.

Application of the Flow-Localization Parameter to Other Deformation Modes. The rate of flow localization in torsion is related to the α parameter, as defined by Eq 30. Basically, this expression establishes that *marked* flow localization occurs only after a critical value of the ratio of the torque softening rate to the rate sensitivity of the torque is achieved.

Related metalforming research has demonstrated that similar parameters can be used to gage the rate of localization in isothermal uniaxial compression and plane-strain forging operations. In compression, α is defined as:

$$\alpha_c = -\left(\frac{\gamma' - 1}{m}\right) \qquad \text{(Eq 35)}$$

and is used to predict when unstable, localized bulging may occur. In isothermal plane-strain forging processes, an α defined by:

$$\alpha_p = \left(\frac{\gamma'}{m}\right) \qquad \text{(Eq 36)}$$

can be used to predict the formation of shear bands. In both cases, m is the strain-rate sensitivity and γ' is the normalized strain-hardening (or softening) rate at fixed strain rate:

$$\gamma' = \left(\frac{1}{\bar{\sigma}}\frac{d\bar{\sigma}}{d\bar{\varepsilon}}\right)\Bigg|_{\dot{\bar{\varepsilon}},T} = \frac{\left(\frac{\partial \ln \bar{\sigma}}{\partial \bar{\varepsilon}}\right)\Big|_{\dot{\bar{\varepsilon}}} d\bar{\varepsilon} + \left(\frac{\partial \ln \bar{\sigma}}{\partial T}\right)\Big|_{\bar{\varepsilon},\dot{\bar{\varepsilon}}} dT}{d\bar{\varepsilon}} \qquad \text{(Eq 37)}$$

and is readily determined from constant-strain-rate compression tests.

As in torsion, criteria based on a $\alpha_c = 5$ or $\alpha_p = 5$ are useful for predicting the occurrence of flow localization. Examples are given in Fig. 52 and 53 for the isothermal hot compression and isothermal plane-strain sidepressing, respectively, of Ti-6242Si. In each case, the α + β microstructure specimens, which developed low values of the α parameter (α_c or α_p), deformed uniformly. However, the β microstructure specimens in both cases had sufficiently high degrees of flow softening to promote α ≥ 5 and thus developed regions of nonuniform flow.

Effect of Temperature and Strain Rate on the α Parameter. The α parameter used to gage flow localization generally shows a sharp dependence on temperature and strain rate. These variations can be predicted, at least qualitatively, from the magnitudes of the terms that comprise it. For simplicity, only α_p, the flow-localization parameter for the occurrence of shear bands in plane-strain forging operations, is discussed here.

As defined by Eq 36, α_p depends on γ' and m. In turn, γ' is specified by Eq 37. At cold working temperatures, γ' is usually positive due to strain hardening. At high strain rates, however, the effect of thermal softening in Eq 37 may outweigh the strain-hardening effect (the first parenthetical term on the right side of the equation), resulting in an overall negative γ', or flow softening. Furthermore, at cold working temperatures, m is usually small (usually between 0 and 0.02). Therefore, small amounts of flow softening coupled with low m may be sufficient to generate α ≥ 5 and hence cause noticeable flow

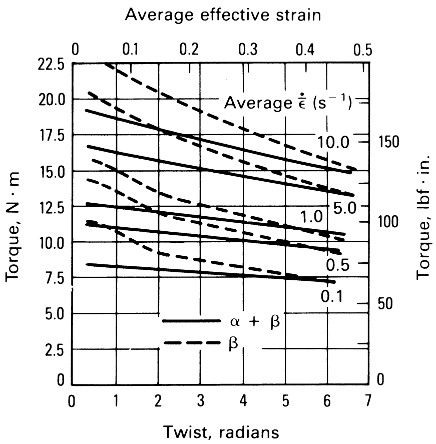


Fig. 51 Torque-twist curves for Ti-6242Si predicted from a numerical-deformation/heat-transfer simulation and measured compression flow-stress data. Results are for testing at 913 °C (1675 °F) and various average effective strain rates. Average effective strain rate = 0.6× surface effective strain rate for the tubular specimen design used. Source: Ref 52

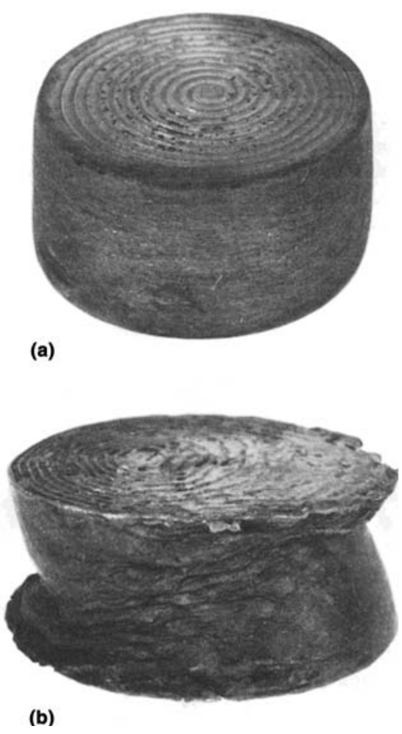


Fig. 52 Specimens of Ti-6242Si from isothermal hot compression tests. Tested at 913 °C (1675 °F); $\dot{\bar{\varepsilon}} = 2\ s^{-1}$. Starting microstructures were (a) α + β (equiaxed alpha). (b) β (Widmanstätten alpha). Magnification: 3×. Source: Ref 122

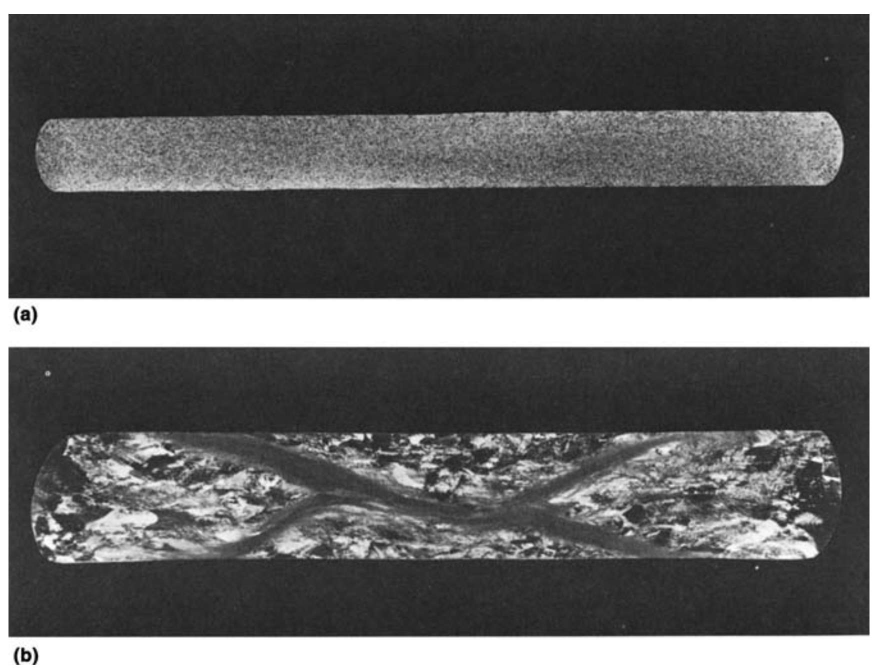


Fig. 53 Transverse metallographic sections of Ti-6242Si bars. (a) α + β (equiaxed alpha). (b) β (Widmanstätten alpha) microstructures. Isothermally sidepressed at 913 °C (1675 °F); $\dot{\bar{\varepsilon}} = 2\ s^{-1}$. Magnification: 5×. Source: Ref 52

localization. Shear bands generated at cold working temperatures at high strain rates are often called adiabatic shear bands.

At hot working temperatures, flow softening (negative γ') is prevalent not only as a result of thermal softening effects at high strain rates, but also because of microstructural softening at all deformation speeds. These microstructural effects, quantified by the first parenthetical term on the right side of Eq 37, are due to softening processes such as dynamic recovery, dynamic recrystallization, and the breakup of Widmanstätten microstructures to form equiaxed microstructures. During plane-strain forging, it is also possible for the strain path in crystallographically "hard" grains to change from homogeneous pure shear to localized simple shear due to texture-related flow softening at a grain scale. On the other hand, at hot working temperatures, m values tend to be substantially larger (usually between 0.1 and 0.3) than at cold working temperatures. Thus, substantial amounts of flow softening, much more than at lower temperatures, are required to produce marked flow localizations.

Analysis of Flow Localization in the Presence of Heat Transfer

At intermediate strain rates, heat-generation and heat-transfer effects must be taken into account to describe flow localization during torsion testing. Axial temperature (and hence flow stress) gradients are established as a result of uneven deformation heating caused by the presence of defects, as well as by the conduction of heat into the colder shoulders of the torsion specimen. In this case, the analysis involves the development of a torque-equilibrium equation as well as a relation to describe the heat-transfer aspects (Ref 121, 123, 124). In its most basic form, the heat-transfer relation considers axial heat conduction and deformation heat generation. In most applications, radiation, convection, and radial-heat-transfer effects can be neglected. The equilibrium and heat-transfer equations are solved incrementally, subject to the imposed boundary conditions (typically, a constant overall twist rate).

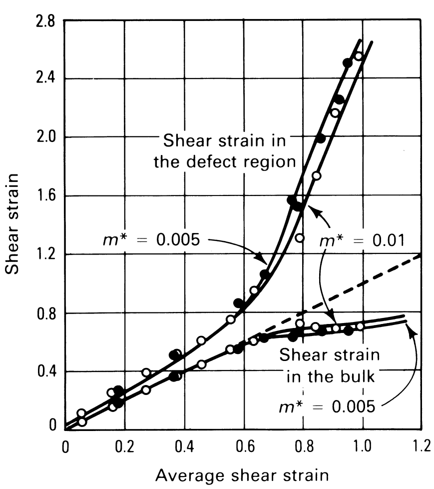


Fig. 54 Comparison of experimentally observed localization kinetics (data points) with simulation results (solid lines) in type 304 stainless steel specimens. Specimens had premachined radius defects at the center of the gage section and were tested in torsion at room temperature. The simulations were run with two rate sensitivities: $m^* = 0.01$ and $m^* = 0.005$, whose values bounded those measured in torsion tests on specimens without geometric defects. Average surface shear strain rate was approximately 0.05 s^{-1} in both experiments and simulations. Source: Ref 123

Calculations have been carried out to determine the effect of heat transfer on flow localization during the torsion of type 304 stainless steel torsion specimens at room temperature (Ref 123). The material coefficients required for the analysis were determined from low-speed tests at which localization does not occur. Using a specimen with a premachined 8% defect in radius at the center of the gage section, localization occurred during tests at $\dot{\gamma} \approx 0.05$ s^{-1}. The localization rate (measured using scribe lines) showed good agreement with the localization simulation based on material parameters, when the additional effect of geometry changes occurring during testing (because the specimen ends were not clamped) was taken into account (Fig. 54).

Microstructure Development During Deformation Processing

In previous sections of this chapter, the application of torsion testing to determine workability limited by excessive loads, fracture-controlled failure, and flow-localization-controlled failure has been discussed. Often, however, the ability to form a piece of metal into a particular shape with the available equipment comprises only the basic considerations of workability. It is usually important to control the microstructures that are developed, which ultimately determines the properties of the finished product and its suitability for further deformation processing, heat treatment, final machining, or service. The ability to accurately control test variables during torsion makes the technique attractive to establish the processing parameters that are required to produce the desired microstructures.

At cold and warm working temperatures, changes in microstructure are largely a distortion of the metal grains (observable at optical magnifications) and a process of dislocation multiplication and dynamic recovery, which is only detectable with the aid of transmission electron microscopy. Dynamic recovery consists of the regrouping of individual dislocations to form cells and subgrains. However, recovery, as well as many other microstructural transformations, occurs much more readily under hot working conditions. The hot torsion test is thus frequently used to detect the effects of deformation and deformation rate on these changes. Significant microstructural phenomena that can be studied with the hot torsion test include dynamic recovery and dynamic recrystallization, which are the main mechanisms controlling microstructural development for single-phase metals and alloys, and dynamic spheroidization and strain-induced precipitation in two-phase materials.

When microstructural observations are made, a particular type of cross section, which is neither axial nor transverse, should be selected. This section should be one in which the outer-diameter surface, or near-surface region of the round gage section, is examined. Axial-tangential sections reveal the true simple shear nature of the deformation.

Because the deformation varies along the radius of the gage section in torsion, the correlation of observed microstructural features to the measured shear stress, strain, and strain rate must be done with care. Barraclough et al. (Ref 35) have examined the problem in detail and have devised the concept of an "effective radius," a_m, for handling the difficulty of microstructure variation across the section. This radius is that whose local shear stress τ^* would generate the measured torque if that same shear stress were developed across the entire section and was derived as:

$$a_m = \left[\frac{3}{3+m'+n'} \times \frac{\left[a_2^{(3+m'+n')} - a_1^{(3+m'+n')}\right]}{a_2^3 - a_1^3} \right] \qquad \text{(Eq 38)}$$

in which a_2 and a_1 denote the outer and inner radii, respectively, of the torsion specimen, and m' and n' are material coefficients used to fit a constitutive equation:

$$\tau = K' \gamma^{n'} \dot{\gamma}^{m'} \qquad \text{(Eq 39)}$$

Dynamic Recovery and Recrystallization in Single-Phase Materials

In single-phase materials, the primary microstructural features are dislocations and grain boundaries. The torsion test is used to determine how these characteristics are affected by deformation temperature and deformation rate. Dynamic recovery at hot working temperatures leads to a reduction in the number of dislocations at a given strain, without noticeably affecting the gross deformation of the grains. By contrast, dynamic recrystallization is characterized by the motion of grain boundaries and annihilation of large numbers of dislocations in a single event, thereby producing new strain-free grains. Metals with high-stacking fault energies (e.g., pure aluminum and α-iron), in which climb of dislocations is easy because they are not dissociated, tend to soften primarily by dynamic-recovery processes. On the other hand, low stacking fault energy materials (e.g., γ-iron, copper, and nickel), in which dislocations are dissociated and thus able to climb and cross-slip only with difficulty, tend to store large reservoirs of strain energy at hot working temperatures. Dynamic recrystallization is initiated from these large stores of dislocations.

The flow curve from a torsion test on a metal in which dynamic recovery predominates at hot working temperatures has an appearance typified by iron tested in the α regime at 700 °C (1290 °F) (Fig. 55). Initially, strain hardening occurs, during which the rate of dislocation multiplication exceeds the rate of recovery. This is followed by a plateau in flow stress and steady-state flow. At this point, well-defined, equiaxed subgrains form whose walls (subboundaries) consist of fairly regular arrays of dislocations. These subgrains remain equiaxed during the remainder of the deformation process, even though the actual grain boundaries do not, as shown by the optical micrographs of Fig. 56. Figures 56(a) and (b) are macrographs of two different hot torsion specimens twisted a total of ten revolutions. The sample in Fig. 56(a) was given ten twists in the forward direction, while that in Fig. 56(b) was twisted forward for five turns and then given five turns of reverse twist. The grains in both specimens underwent a large degree of deformation, which increased from the center to the surface because of the strain variation inherent in the torsion of solid round bars. The large degree of deformation can be readily seen from the micrographs of the first test specimen, Fig. 56(c), in which the highly distorted nature of the twisted grains is evident at the bottom of the micrograph. The micrograph of the second test sample of identical geometry that was twisted five revolutions in one direction and then five in the reverse direction is shown in Fig. 56(d). It is apparent that the original equiaxed grain structure has been restored by this deformation schedule, even though these grains have received the same equivalent strain as those of Fig. 56(c).

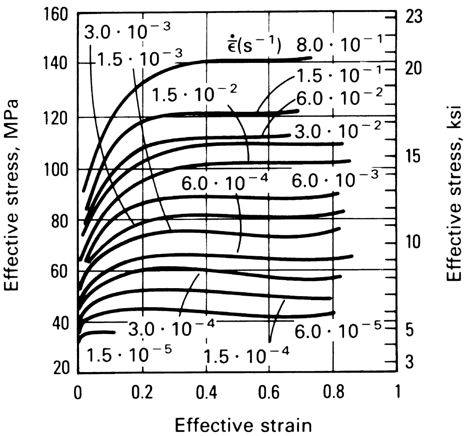


Fig. 55 Stress-strain curves for Armco iron. Strain-rate dependence of the flow stress at 700 °C (1290 °F), or 0.54× the absolute melting point, is evident. Data are from compression tests; torsion results exhibit similar behavior. Source: Ref 125

The torsional flow curves for materials that recrystallize dynamically at hot working temperatures are often quite different from those typical of metals in which dynamic recovery predominates. These curves may be used to determine the onset of dynamic recrystallization as a function of strain rate and temperature.

Stress-strain curves for nickel and carbon steel (in the single-phase austenite range) are shown in Fig. 57 and 58. The gross behavior consists of an initial strain-hardening stage, during which the dislocation density increases rapidly. At some point, a flow-stress maximum is achieved, after which marked flow softening occurs. The strain at the peak flow stress decreases with increasing temperature and decreasing strain rate. Such flow softening can be identified with the onset of dynamic recrystallization. At relatively low temperatures and/or high strain rates, the flow curve eventually achieves a steady state, indicative of ongoing (discontinu-

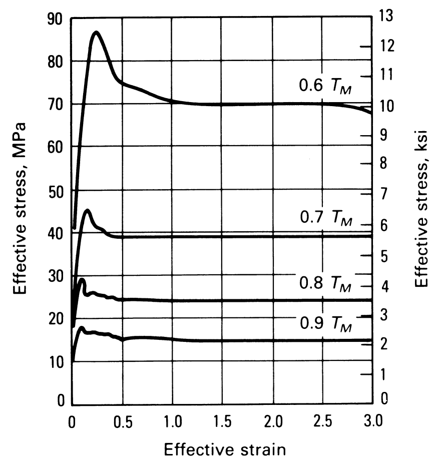


Fig. 57 Stress-strain curves derived from hot torsion data for nickel at an effective strain rate of 0.016 s^{-1}. The dependence of flow behavior on test temperature in a material that undergoes dynamic recrystallization is shown. Source: Ref 26

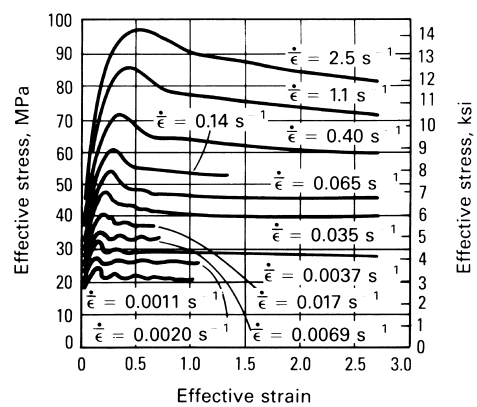


Fig. 58 Flow curves for 0.25% low-carbon steel in the austenitic state tested in torsion at 1100 °C (2010 °F). The strong influence of strain rate and a behavior indicative of dynamic recrystallization are shown. Source: Ref 39

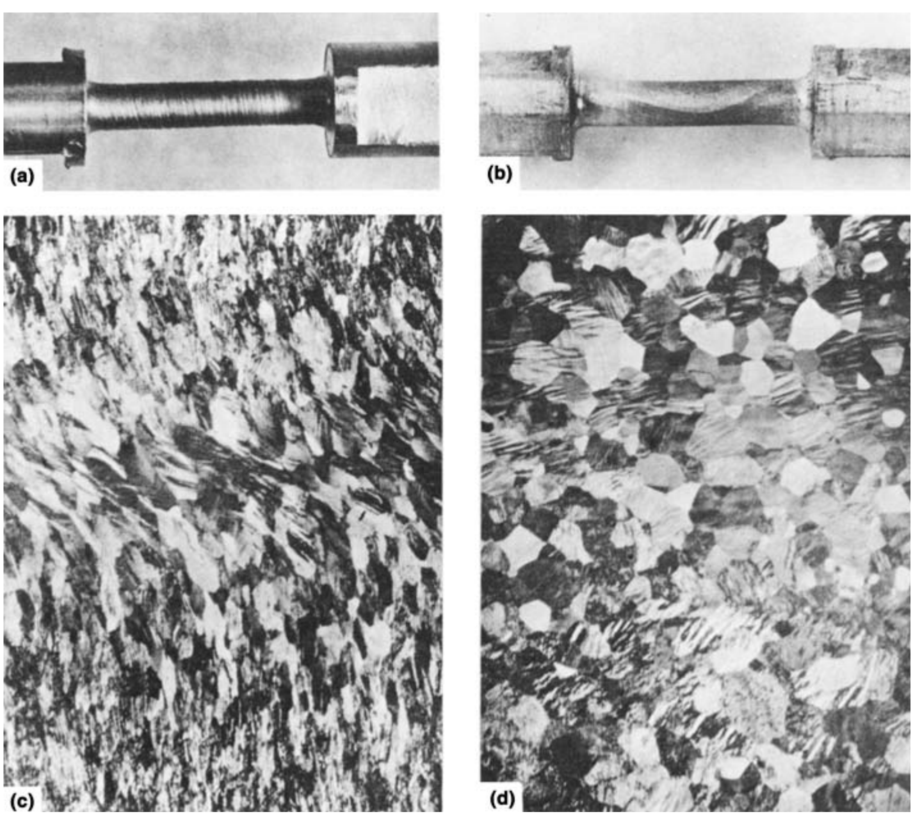


Fig. 56 Effect of twist reversal on specimen appearance and structure of superpurity aluminum after deformation at 400 °C (750 °F) and a strain rate of 2 s^{-1}. (a) 5 + 5 revolutions (2×). (b) 5 − 5 revolutions (2×). (c) 5 + 5 revolutions (25×). (d) 5 − 5 revolutions (25×). Source: Ref 54

ous) dynamic recrystallization. At low strain rates and/or high temperatures (as under conditions approaching those typical of creep deformation), the flow curve following the peak flow stress often exhibits cyclic hardening and softening.

The dynamically recrystallized grain size formed under equilibrium conditions is a function of the deformation temperature and strain rate. For example, the stable grain size in type 304L stainless steel twisted at several temperatures and strain rates is shown in Fig. 59, which illustrates that grain size increases with increasing temperature and decreasing strain rate. As with the flow stress and fracture behavior at hot working temperatures, the dependence on these two variables is best expressed in terms of the product $\dot{\varepsilon}$ exp (Q/RT), or the Zener-Hollomon parameter (Z). The strong correlation between the reciprocal of the average recrystallized grain size and Z is demonstrated for copper and nickel in Fig. 60.

A similar dependence between Z and the stable, dynamically recrystallized grain size, D_s, for a 0.16% C steel tested in the austenite regime is shown in Fig. 61. Also illustrated are the critical Z, or Z_c, values as a function of starting grain size, D_o; the values of Z_c represent the conditions associated with the transition between single-peak and multiple-peak (periodic) flow curves. This curve is labeled Z_c-D_o. In addition, a curve representing Z versus $2D_s$ is shown.

The similarity between the Z_c-D_o and Z-$2D_s$ curves establishes that the transition in flow behavior occurs when the equilibrium or stable, dynamically recrystallized size is equal to approximately one-half of the initial grain size. Thus, grain refinement leading to at least a halving of the starting grain size produces a single-peak flow curve, whereas grain coarsening (or refinement of less than one-half) results in cyclic flow curves (Ref 99).

These observations are useful in rationalizing the dependence of the shape of the torsional flow curve on temperature and strain rate. As mentioned previously, cyclic curves are most frequently observed at low strain rates and high temperatures, or the regime in which the stable, recrystallized grain size is large. Because all but coarse-grained materials undergo grain coarsening during torsion under these conditions, the flow curves are cyclic. These coarsening cycles continue to be observed until the recrystallized grain size attains the equilibrium value.

Thus, the torsion test can be very useful in establishing the occurrence of recovery or recrystallization during the hot working of single-phase materials. When dynamic recrystallization occurs, the shape of the flow curve may be used to determine the temperatures and strain rates at which refinement of the grain size, a characteristic important with respect to service properties, occurs. Moreover, it can be used to establish equivalent combinations of temperature and strain rate at which a given grain size is produced.

Development of Microstructure in Alloys with More Than One Phase

The torsion test has also been used to determine the effects of deformation parameters on the microstructures developed in two-phase and multiphase alloys. Torsion testing has the ability to impose large strains at rates up to and includ-

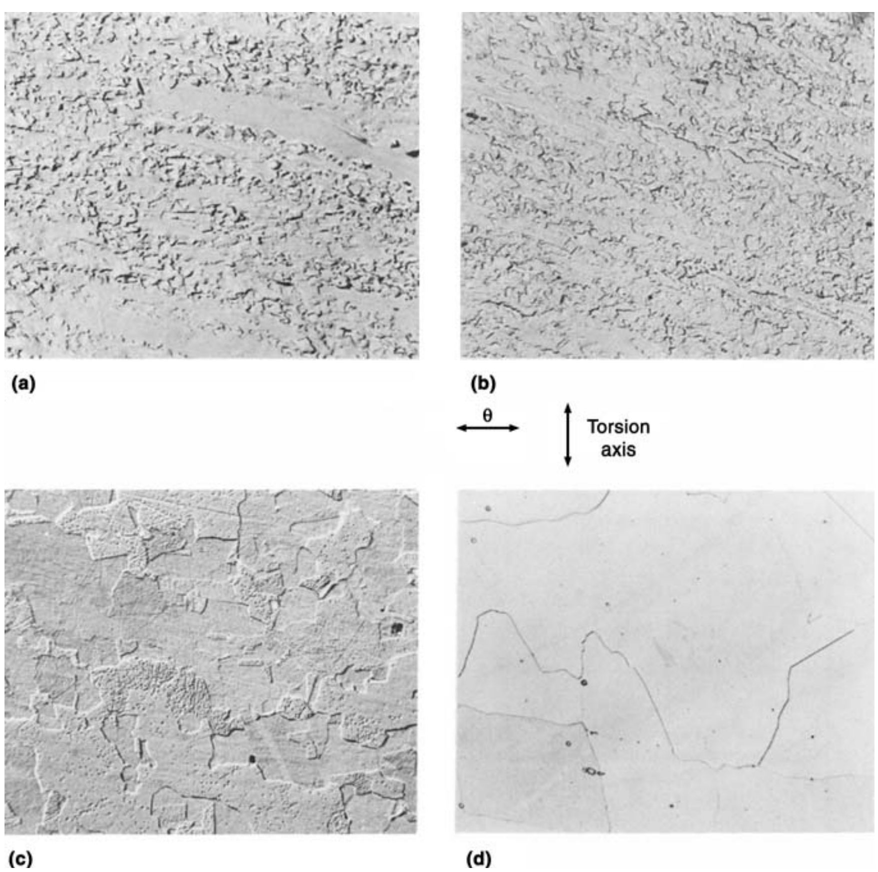


Fig. 59 Micrographs of type 304L stainless hot torsion specimens tested under various strain-rate/temperature conditions. (a) 0.01 s^{-1}, 800 °C (1470 °F) ($\bar{\varepsilon}$ = 1.99). (b) 10 s^{-1}, 800 °C (1470 °F) ($\bar{\varepsilon}$ = 3.81). (c) 0.01 s^{-1}, 1000 °C (1830 °F) ($\bar{\varepsilon}$ = 3.73). (d) 0.01 s^{-1}, 1200 °C (2190 °F) ($\bar{\varepsilon}$ = 4.64). Magnification: 350×. Source: Ref 88

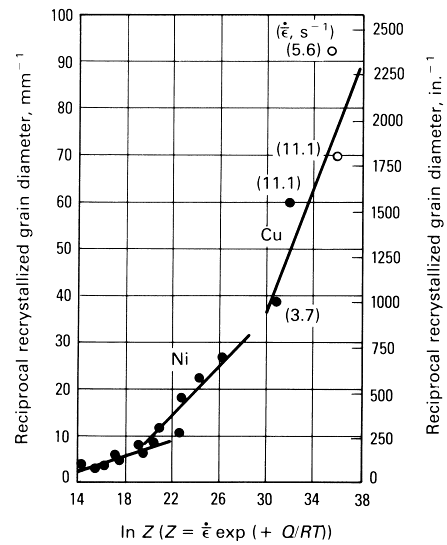


Fig. 60 Dynamically recrystallized grain sizes of copper and nickel as a function of the Zener-Hollomon parameter (Z). Source: Ref 126

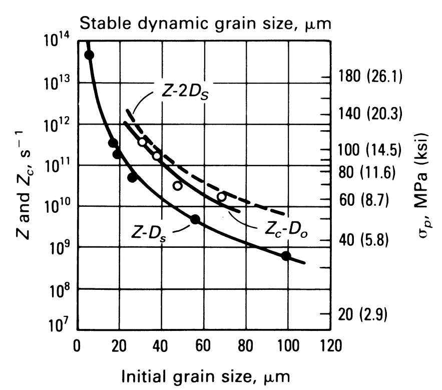


Fig. 61 Dependence of the critical parameter Z_c on initial austenite grain size D_o in a 0.16% C steel (open data points). The solid line fitted to the filled points is the Z-D_s relationship (D_s = stable grain size achieved during dynamic recrystallization). Note that the Z-$2D_s$ (broken line) and Z_c-D_o relations are nearly coincident. Source: Ref 99

ing those used in commercial metalworking operations. For example, torsion testing in two-phase alloys can be used to study flow softening and the breakup of unstable microstructures. These effects are especially strong in materials with lamellar or Widmanstätten phases, such as carbon steels and α-β titanium alloys.

In carbon steels, torsion may be used to determine the flow response and microstructural changes that occur in pearlite subject to deformation below the lower critical temperature. For steel of a eutectoid composition, large amounts of flow softening were measured in torsion, except when tested above the critical temperature of 742 °C (1368 °F) (Fig. 62). This softening was associated with the breakup and spheroidization of the cementite (Fe_3C) lamellae. As the temperature was lowered (or the strain rate was increased), the torsion results suggested that the rate of spheroidization increased. At the highest subcritical temperature (715 °C, or 1320 °F), the amount of softening was rather low, suggesting that microstructural changes were not as drastic as at lower temperatures, consisting primarily of the development of coarse cementite. This conclusion was supported by microstructural examination of the hot torsion specimens, which indicated a definite relationship among strain rate, temperature, and spacing of the spheroidal particles after a fixed amount of deformation (Fig. 63). Again, the strain rate and temperature dependence was expressed through the Zener-Hollomon parameter, $\dot{\varepsilon}\exp(Q/RT)$. When high strain rates or low temperatures were used, a fine, closely spaced dispersion was produced. On the other hand, low strain rates and high temperatures, which enhance diffusion processes, resulted in a large interparticle spacing, indicative of a coarse dispersion of cementite.

In a similar two-phase system, torsion has been used to study the modification of the Widmanstätten α microstructure of α-β titanium alloys subjected to deformation below the β-transus temperature. As with the eutectoid steel, tests of this type have shown that coarsening prevails at temperatures near the transus, whereas spheroidization results from deformation at high strain rates and lower subtransus temperatures.

The torsion test can also be used to study microstructural modification in multiphase alloys and to establish processing conditions based on this information. It is particularly useful for nickel-base superalloys. These materials are expensive and have a very limited working temperature range due to the presence of second phases (γ′, intermetallic carbides), which only go into solution at relatively high temperatures (if at all), and because of the problems of grain growth or incipient melting when deformation is performed near the melting point.

Figures 64 through 66 illustrate a typical application of the torsion test to study the development of microstructure in the nickel-base alloy Udimet 700. In the as-received condition (Fig. 64), the material had fine γ′, carbides, and borides, dispersed through the γ matrix, whose grain structure was not resolvable. After torsion at approximately 1060 °C (1940 °F), the structure consisted of a well-defined fine γ grain structure with carbides and borides situated at the grain boundaries Fig. 65(a).

This microstructure is very similar to that observed in Udimet 700 material extruded under almost identical conditions of strain, strain rate, and temperature. In addition, the microstructures developed in torsion and extrusion at the higher temperature of 1150 °C (2100 °F) (Fig. 66) closely resembled one another. These results and similar ones on other alloys establish the torsion

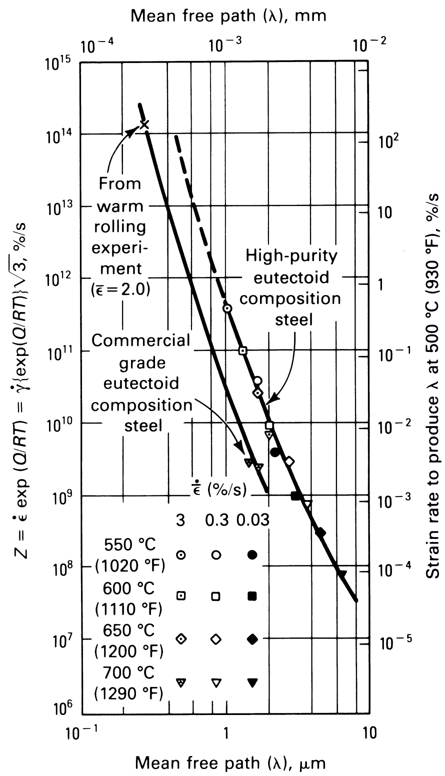


Fig. 63 The mean free path (λ) between spheroidite particles in hot-worked eutectoid steels as a function of the Zener-Hollomon parameter. The right ordinate is scaled to show the strain rate at 500 °C (930 °F), which would produce the indicated spacing after large plastic deformation. Data are from torsion, compression, and rolling experiments. Source: Ref 127

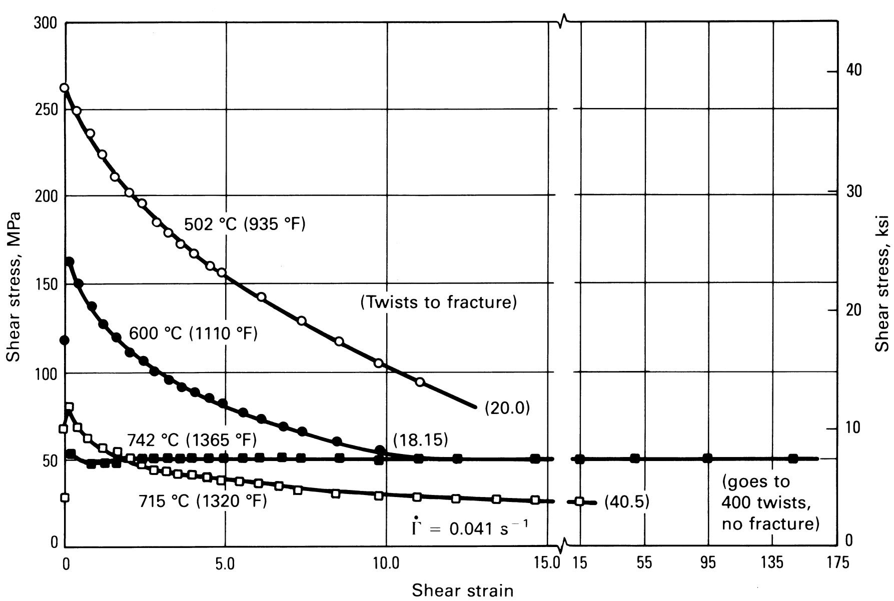


Fig. 62 Effect of test temperature on the torsional flow curve of a high-purity 0.8% C pearlitic iron-carbon alloy. Numbers in parentheses refer to the number of twists to fracture. The flow softening at the three lower temperatures can be attributed to pearlite spheroidization. Source: Ref 21

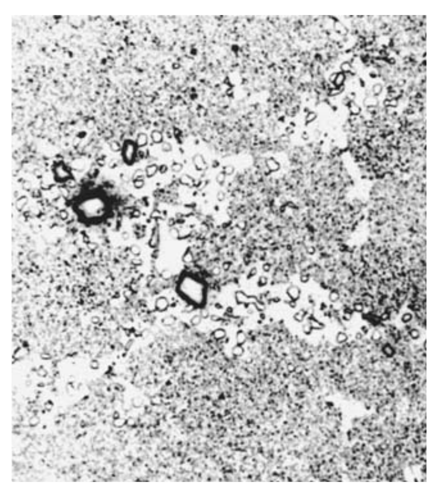


Fig. 64 Starting microstructure of Udimet 700 billet material used in torsion and extrusion studies. Magnification: 465×. Source: Ref 114

test as a valuable means of determining the microstructures that can be developed during actual metalforming processes.

Processing History Effects

Torsion testing describes the broad patterns of deformation, failure resistance, and microstructure development for a given alloy. In actual metalforming processes, however, the thermomechanical history of the workpiece is rarely so simple. Frequently, it will be preheated in a furnace and transferred to the processing equipment (forging press, rolling mill, etc.). During this dwell period, it will have cooled a certain amount. Also, the workpiece will chill when it is in contact with the tooling during conventional metalworking processes. This is in contrast to normal torsion tests in which the test specimen is heated to temperature, soaked for a period of time, and then twisted. Furthermore, during deformation processing, the strain rate is rarely constant, unlike that in the conventional torsion experiment.

With proper controls, the effects of temperature and strain-rate history on workability (flow-stress levels, fracture behavior, and microstructure development) can be assessed using the torsion test, provided means exist to replicate the thermal and/or deformation-rate history. The thermal effect of greatest importance is cooling during processing. Cooling histories are best controlled during testing through forced-air or argon convection around the specimen at rates that must be determined experimentally to obtain the desired results. Strain-rate histories are more readily controlled by interfacing closed-loop test systems with computers or function generators that provide the proper control signals representing the rotation rate-time dependence needed.

Figures 67 and 68 illustrate the effects of temperature and strain rate history during torsion on the flow stress of α-iron and copper, respectively. The α-iron was tested at a constant strain rate, but the specimen temperature was increased or decreased at a continuous rate of 50 °C/min (90 °F/min). The flow stresses from such tests were compared to isothermal stresses. If the specimen was heated during testing, the flow stress was higher than the isothermal stress. The reverse was true if the specimen was cooled. During either heating or cooling, the dislocation substructure does not change instantaneously. During the heating experiments, a lower-temperature, less highly recovered substructure is retained, giving rise to higher flow stresses than the isothermal tests. Similarly, during cooling, a softer, more highly recovered substructure leads to lower flow stresses than observed in isothermal experiments.

The effect of strain-rate history on the flow stress of copper is shown in Fig. 68. Under constant-strain-rate conditions, harder substructures were produced at higher strain rates. However, if the strain rate was increased or decreased during torsion testing, the inertia of the acquired substructure prevented changes in flow stress as

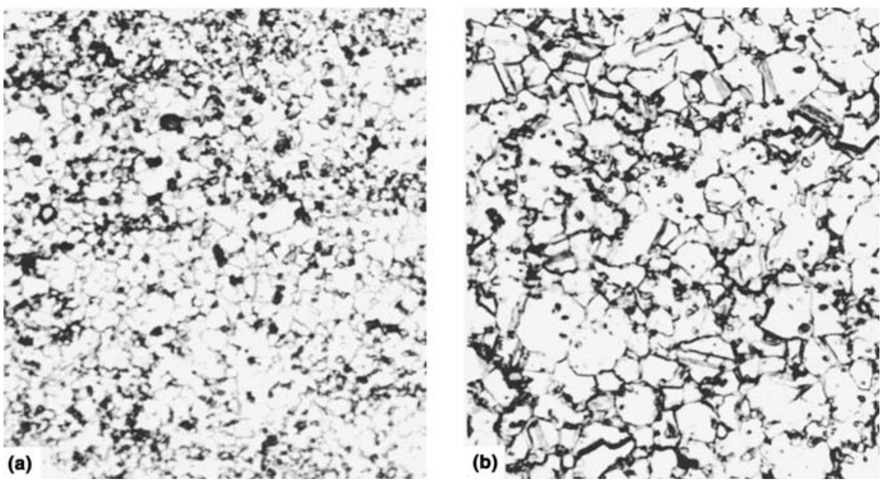


Fig. 65 Comparison of (a) torsion and (b) extrusion microstructures in Udimet 700. Deformed under nearly identical conditions of $\dot{\bar{\varepsilon}} = 7\ s^{-1}$, T = 1060 °C (1940 °F), and $\bar{\varepsilon} = 1.85$. Magnification 465×. Source: Ref 114

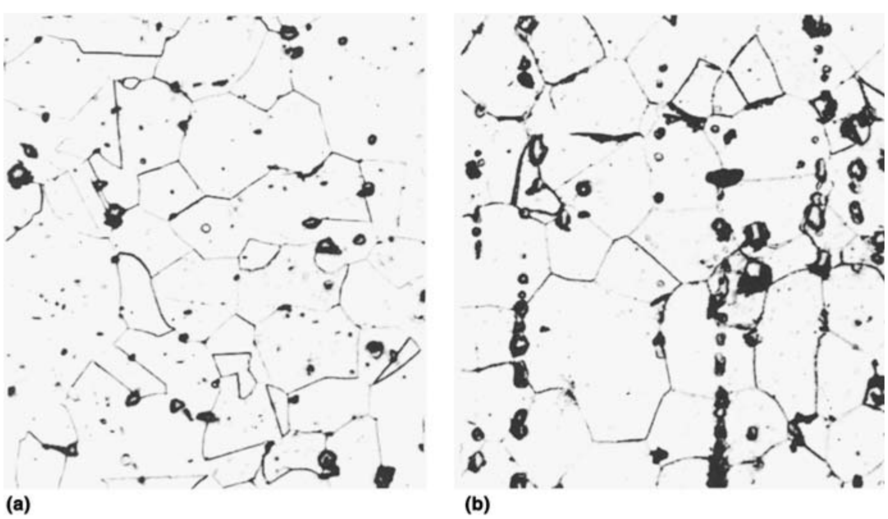


Fig. 66 Comparison of (a) torsion and (b) extrusion microstructures in Udimet 700. Deformed under nearly identical conditions of $\dot{\bar{\varepsilon}} = 4\ s^{-1}$, T = 1145 °C (2090 °F), and $\bar{\varepsilon} = 2.15$. Magnification: 465×. Source: Ref 114

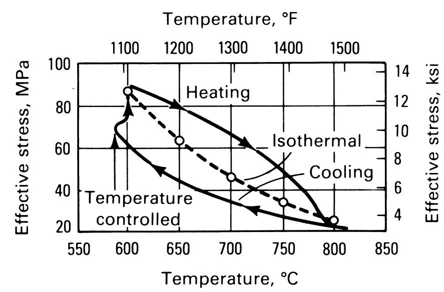


Fig. 67 Effect of continuous heating or cooling on the steady-state flow stress of vacuum-melted iron. Deformed in torsion at an effective strain rate of $1.5 \times 10^{-3}\ s^{-1}$. Source: Ref 128

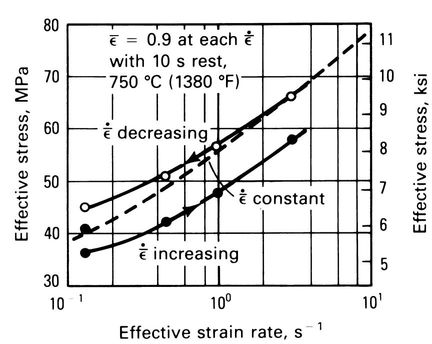


Fig. 68 Effect of increasing or decreasing strain rate on the flow stress of copper deformed in torsion at 750 °C (1380 °F). Source: Ref 129

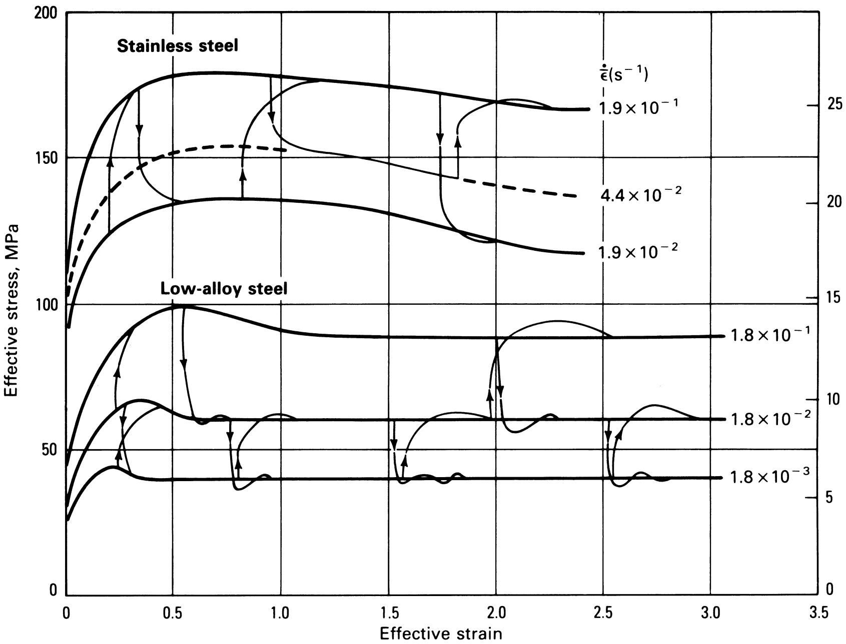


Fig. 69 Typical stress-strain curves for tests involving instantaneous changes in strain rate for an austenitic stainless steel and a ferritic low-alloy steel. Note that the rate-change strain-rate sensitivity is either lower (stainless steel) or higher (low-alloy steel) than that based on continuous (constant-strain-rate) torsion tests. Source: Ref 130

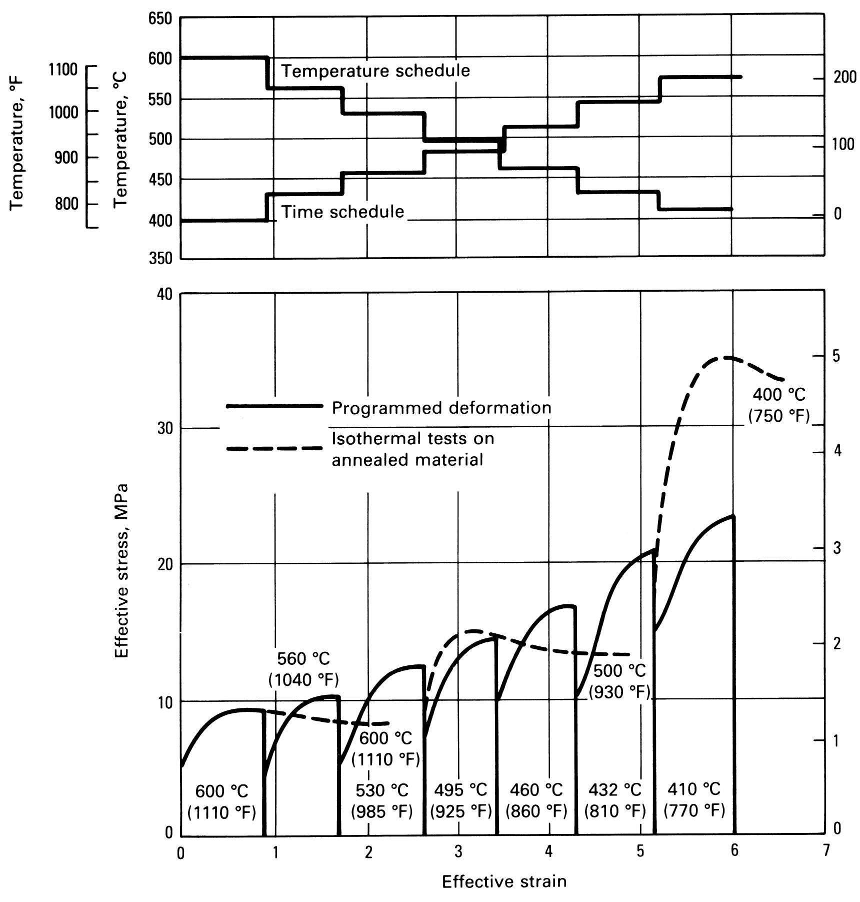


Fig. 70 Deformation-temperature-time schedule and resulting flow behavior of superpurity aluminum deformed in torsion at an effective strain rate of 2.3 s^{-1}. Source: Ref 54

large as those observed in a series of constant-strain-rate tests. Thus, strain-rate sensitivities measured in rate-change tests are often lower than those based on constant-strain-rate or so-called continuous flow curves in materials such as copper, aluminum, and austenitic stainless steels. In materials that exhibit dynamic strain aging (e.g., carbon steels at cold working temperatures), the relationship between the two rate-sensitivity parameters may be reversed, depending on the strain-rate regime and the kinetics of strain aging (Fig. 69).

The torsion test is especially useful to study the effects of history during deformation processes such as bar, plate, and sheet rolling (Ref 54, 131–145). For example, Fig. 70 shows the type of flow-stress behavior that might be expected for high-purity aluminum during rolling in which the temperature decreases continuously. As the temperature decreased, the flow stress increased, but not as much as would be expected based on isothermal measurements. This phenomenon was particularly evident once the temperature dropped below 450 °C (840 °F) and was a result of the retention of a soft high-temperature substructure.

Simulation of the rolling of high-strength low-alloy steels through the torsion test (Fig. 71–74) has been used to provide insight into how equipment requirements and final microstructures can be determined by this technique. In this instance, torsion tests were conducted in a servohydraulic test machine at a fixed strain rate. However, temperature was controlled to decrease continuously at a rate almost equivalent to that measured during actual production (Fig. 71). Under these conditions, the torsion flow stresses increased rapidly (Fig. 72). Using these data, the roll-separating force and rolling torque were estimated for the rolling schedule under study using standard formulas. These were compared to actual measurements made at the individual stands of the rolling facility. Agreement was fairly good (Fig. 73), indicating the usefulness of the torsion test in assessing the influence of processing parameters (rolling speed, temperatures, dwell time between stands, etc.) on equipment requirements. In addition, the final microstructure from the torsion simulation was almost identical to that from the production run (Fig. 74).

Other experiments using similar computer-controlled torsion setups have provided valuable information on the required loads and microstructures developed during rolling of a wide range of alloy steels.

The torsion test can also be used to simulate the effects of die chilling on workability in multiphase alloys. In particular, the effects of solutioning of second phases (during preheating) and subsequent reprecipitation during working because of chilling are readily simulated by the test. This is accomplished by preheating the torsion specimen to a high temperature and testing it on cooling. The torsional ductility obtained in such tests frequently may be used to determine a potential workability problem.

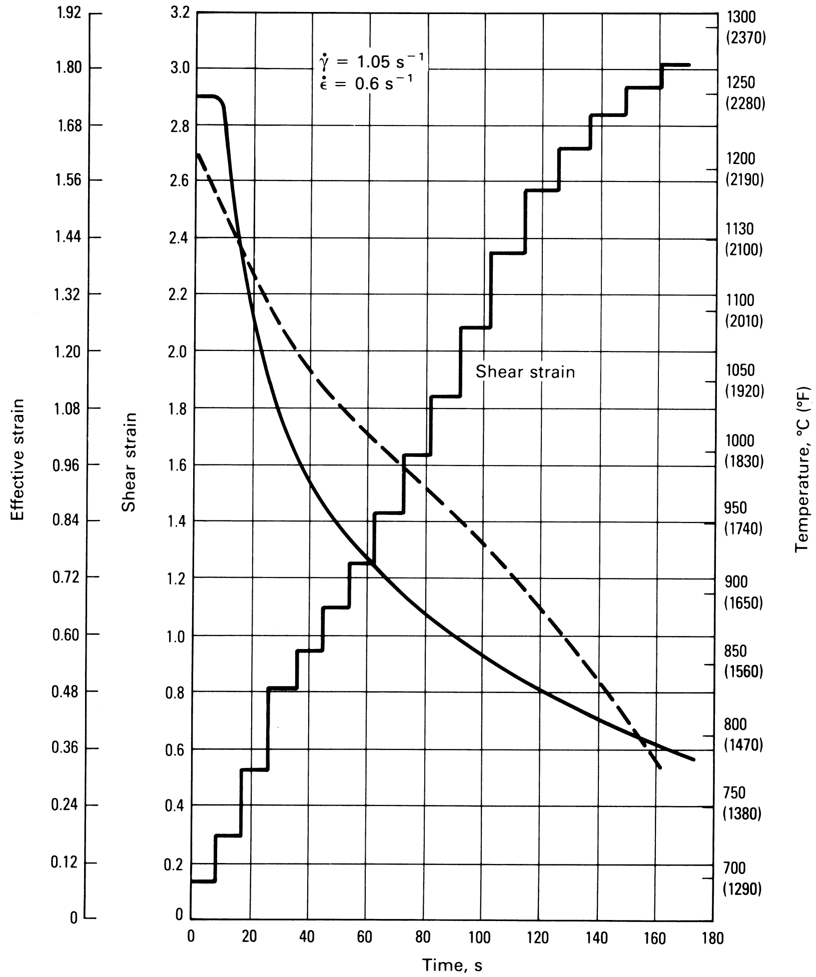


Fig. 71 Deformation-temperature-time sequence imposed during torsion testing of microalloyed steels. The temperature-time profile followed in a production plate mill (dashed curve) is compared with that experienced by the sample in the torsion machine (solid curve). Source: Ref 131

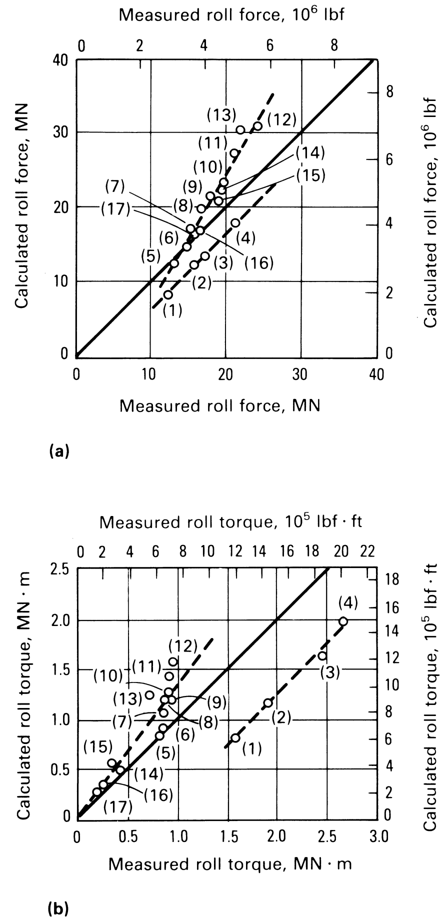


Fig. 73 Comparison of the predicted and measured (a) roll-separating forces and (b) roll torques associated with rolling of a niobium-vanadium microalloyed steel. The pass number is shown beside each data point. Source: Ref 131

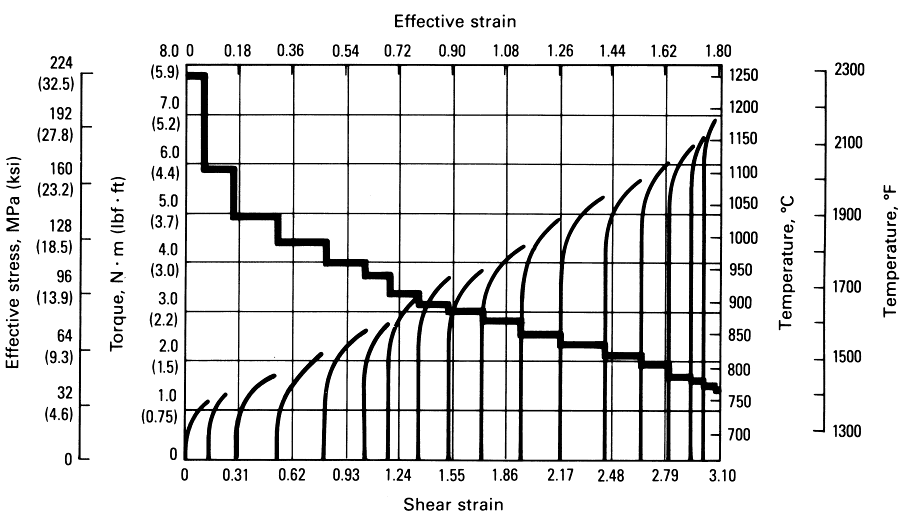


Fig. 72 Flow behavior for a niobium-vanadium microalloyed steel deformed in 17 passes in a torsion machine. The specimen temperatures are represented by the upper bold line. Source: Ref 131

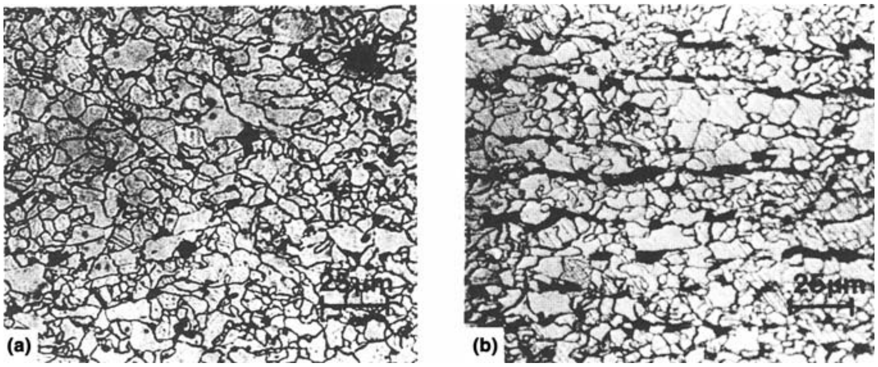


Fig. 74 Ferrite structure obtained in a niobium-vanadium microalloyed steel. (a) After 17 passes in the torsion machine. (b) After 17 passes in a production plate mill. Source: Ref 131

ACKNOWLEDGMENT

This article drew upon:

- S.L. Semiatin, G.D. Lahoti, and J.J. Jonas, "Application of the Torsion Test to Determine Workability," *Mechanical Testing,* Vol 8, *ASM Handbook,* American Society for Metals, 1985, p 154–184
- M.J. Luton, "Hot Torsion Testing," *Workability Test Techniques,* American Society for Metals, 1984, p 95–133

REFERENCES

1. C.L. Clark and J.J. Russ, *Trans. AIME,* Vol 167, 1946, p 736–748
2. H.K. Ihrig, *Trans. AIME,* Vol 167, 1946, p 749–790
3. D.E.R. Hughes, *J. Iron Steel Inst.,* Vol 170, 1952, p 214–220
4. J. Danvergne, M. Pelabon, and J. Ivernal, *Rev. Met.,* Vol 51, 1954, p 254–264
5. K. Tajima and K. Kugai, *Tetsu-to-Hagané (J. Iron Steel Inst. Jpn.),* Vol 42, 1956, p 980–985
6. W. Precht and J.R. Pickens, *Metall. Trans. A,* Vol 18A, 1987, p 1603–1611
7. M.P. Clode, *Alum. Ind.,* Vol 11, 1992, p 34–39
8. R. Raj, unpublished research, Cornell University, Ithaca, NY, 1981
9. W.A. Backofen, A.J. Shaler, and B.B. Hundy, *Trans. ASM,* Vol 46, 1954, p 655–680
10. H. Ormerod and W.J.McG. Tegart, *J. Inst. Met.,* Vol 89, 1960–61, p 94–96
11. D. Hardwick and W.J.McG. Tegart, *J. Inst. Met.,* Vol 90, 1961–62, p 17–20
12. D. Hardwick and W.J.McG. Tegart, *Mem. Sci. Rev. Met.,* Vol 58, 1961, p 869–880
13. J.L. Robbins, O.C. Shepard, and O.D. Sherby, *J. Iron Steel Inst.,* Vol 199, 1961, p 175–180
14. R.A. Reynolds and W.J.McG. Tegart, *J. Iron Steel Inst.,* Vol 200, 1962, p 1044–1059
15. F.E. White and C. Rossard, *Rev. Met.,* Vol 59, 1962, p 237–241
16. P.J. Regenet and H.P. Stuwe, *Z. Metallkde.,* Vol 54, 1963, p 275
17. H.P. Stuwe and H. Turck, *Z. Metallkde.,* Vol 55, 1964, p 699–703
18. H.P. Stuwe, *Z. Metallkde.,* Vol 56, 1965, p 633–642
19. C.M. Sellars and W.J.McG. Tegart, *Mem. Sci. Rev. Met.,* Vol 63, 1966, p 731–746
20. C. Rossard and A. LeBon, *Hart.-Tech. Mitt.,* Vol 24, 1967, p 221–225
21. J.L. Robbins, O.C. Shepard, and O.D. Sherby, *Trans. ASM,* Vol 60, 1967, p 205–216
22. J.L. Robbins, H. Wagenaar, O.C. Shepard, and O.D. Sherby, *J. Mater.,* Vol 2, 1967, p 271–299
23. C. Rossard, *Rev. Met.,* Vol 65, 1968, p 181–195
24. F.E. White, *Rev. Met.,* Vol 63, 1966, p 991–998
25. H.P. Stuwe, *Deformation under Hot Working Conditions,* Iron and Steel Institute, London, 1968, p 1–6
26. M.J. Luton and C.M. Sellars, *Acta Metall.,* Vol 17, 1969, p 1033–1043
27. M.J. Luton and W.J.McG. Tegart, *Met. Sci. J.,* Vol 3, 1969, p 142–146
28. J.P. Sah, G.J. Richardson, and C.M. Sellars, *J. Aust. Inst. Met.,* Vol 14, 1969, p 292–297
29. J. Cotner and W.J.McG. Tegart, *J. Inst. Met.,* Vol 97, 1969, p 73
30. G.A. Redfern and C.M. Sellars, *J. Iron Steel Inst.,* Vol 208, 1970, p 576–587
31. G. Mima and F. Inoko, *Trans. Iron Steel Inst. Jpn.,* Vol 10, 1970, p 216–221
32. F.A. Hodierne, *J. Inst. Met.,* Vol 91, 1962–63, p 267–273
33. C. Rossard and P. Blain, *Rev. Met.,* Vol 56, 1959, p 175–180
34. A. Johansen, B. Ronning, and N. Ryum, *Proc. Sixth Int. Conf. Aluminum Alloys,* Japan Institute of Light Metals, Tokyo, 1998, p 559–564
35. D.R. Barraclough, H.J. Whittaker, K.D. Nair, and C.M. Sellars, *J. Test. Eval.,* Vol 1, 1973, p 220–230
36. D.S. Fields and W.A. Backofen, *Proc. Am. Soc. Test. Mater.,* Vol 57, 1957, p 1259–1272
37. T. Morushima, *Tetsu-to-Hagané (J. Iron Steel Inst. Jpn.),* Vol 44, 1958, p 552–559, 660–668
38. C. Rossard and P. Blain, *Rev. Met.,* Vol 55, 1958, p 573–594, 595–598
39. C. Rossard, *Métaux-Corros.-Ind.,* Vol 35, 1960, p 102–115, 140–153, 190–205
40. G.E. Dieter, J.V. Mullins, and E. Shapiro, *Deformation under Hot Working Conditions,* Iron and Steel Institute, London, 1968, p 7–13
41. J.A. Bailey and S.L. Haas, *J. Mater.,* Vol 7, 1972, p 8–13
42. C. Rossard and P. Blain, *Rev. Met.,* Vol 59, 1962, p 223–236
43. C. Rossard and P. Blain, *Rev. Met.,* Vol 61, 1964, p 949–961
44. T.B. Vaughan and W.C. Aston, *Metallurgia,* Vol 471, 1969, p 39–44
45. E. Shapiro and G.E. Dieter, *Metall. Trans.,* Vol 1, 1970, p 1711–1719
46. G. Glover and C.M. Sellars, *Metall. Trans.,* Vol 4, 1973, p 756–775
47. H. Weiss, D.H. Skinner, and J.R. Everett, J. Phys. E, *Sci. Instrum.,* Vol 6, 1973, p 709–714
48. S. Fulop, K.C. Cadien, M.J. Luton, and H.J. McQueen, *J. Test. Eval.,* Vol 5, 1977, p 419–426
49. J.J. Mills, K. Nielsen, and W. Merriam, *Novel Techniques in Metal Deformation Testing,* TMS, Warrandale, PA, 1983, p 343–357
50. P. Choquet, A. LeBon, C. Rossard, C. Perdrix, and G. Joannes, *Proc. Int. Conf. Physical Metallurgy of Thermomechanical Processing of Steels and Other Metals (Thermec88),* Iron and Steel Institute of Japan, Tokyo, 1988, p 729–736
51. B.M. Perrett, *Mechanical Testing of Materials,* Institute of Metals and Materials Austalasia Ltd., Melbourne, Australia, 1994, p 41–44
52. S.L. Semiatin and G.D. Lahoti, *Metall. Trans. A,* Vol 12, 1981, p 1719–1728
53. G. Glover and C.M. Sellars, *Metall. Trans.,* Vol 3, 1972, p 2271–2280
54. M.M. Farag, C.M. Sellars, and W.J.McG. Tegart, *Deformation under Hot Working Conditions,* Iron and Steel Institute, London, 1968, p 60–67
55. E. Shapiro and G.E. Dieter, *Metall. Trans.,* Vol 2, 1971, p 1385–1391
56. A. Nadai, *Theory of Flow and Fracture,* Vol 1, 2nd ed., McGraw-Hill, 1950, p 349
57. P. Ludwik, *Elemente der Technologischen Mechanic,* Springer Verlag, Berlin, 1909, p 132
58. H.S. Kim, *Mater. Sci. Eng. A,* Vol A299, 2001, p 305–308
59. G.R. Canova, S. Shrivastava, J.J. Jonas, and C.G'Sell, *Formability of Metallic Materials—2000 A.D,* STP 753, ASTM, 1982, p 189–210.

60. G.R. Canova, Master's thesis, McGill University, Montreal, 1979
61. E.W. Hart, C.Y. Li, H. Yamada, and G.L. Wire, *Constitutive Equations in Plasticity,* A.S. Argon, Ed., MIT Press, 1975, p 149
62. A. Graber and K. Pohlandt, *Steel Res.,* Vol 61, 1990, p 212–218
63. S. Khoddam, Y.C. Lam, and P.F. Thomson, *Steel Res.,* Vol 66, 1995, p 45–49
64. S. Khoddam, Y.C. Lam, and P.F. Thomson, *J. Test. Eval.,* Vol 26, 1998, p 157–167
65. S. Khoddam, Y.C. Lam, and P.F. Thomson, *Steel Res.,* Vol 67, 1996, p 39–43
66. S. Khoddam, Y.C. Lam, and P.F. Thomson, *Proc. Int. Conf. Thermomechanical Processing of Steels and Other Metals (Thermec97),* T. Chandra and T. Sakai, Ed., TMS, Warrendale, PA, 1997, p 911–917
67. A. Gavrus, E. Massoni, and, J.-L. Chenot, *Steel Res.,* Vol 70, 1999, p 259–268
68. J.J. Jonas, F. Montheillet, and S. Shrivastava, *Scr. Metall.,* Vol 19, 1985, p 235–240
69. T. Spittel, M. Spittel, and J. Suchanek, *Neue Hütte,* Vol 31, 1986, p 233–238
70. M. Zhou and M.P. Clode, *Mater. Des.,* Vol 17, 1996, p 275–281
71. M. Zhou and M.P. Clode, *Mater. Sci. Technol.,* Vol 13, 1997, p 818–824
72. M. Zhou and M.P. Clode, *Finite Elements in Analysis and Design,* Vol 31, 1998, p 1–14
73. S. Khoddam, Y.C. Lam, and P.F. Thomson, *Mech. Mater.,* Vol 22, 1996, p 1–9
74. F. Montheillet, M. Cohen, and J.J. Jonas, *Acta Metall.,* Vol 32, 1984, p 2077–2089
75. F. Montheillet, P. Gilormini, and J.J. Jonas, *Acta Metall.,* Vol 33, 1985, p 705–717
76. L.S. Toth, J.J. Jonas, P. Gilormini, and B. Bacroix, *Int. J. Plast.,* Vol 6, 1990, p 83–108
77. K.W. Neale, L.S. Toth, and J.J. Jonas, *Int. J. Plast.,* Vol 6, 1990, p 45–61
78. L.S. Toth and J.J. Jonas, *Textures Microstruct.,* Vol 10, 1989, p 195–209
79. L.S. Toth, P. Szaszvari, I. Kovacs, and J.J. Jonas, *Mater. Sci. Technol.,* Vol 7, 1991, p 458–463
80. L.S. Toth and J.J. Jonas, *Scr. Metall. Mater.,* Vol 27, 1992, p 359–363
81. J.J. Jonas and L.S. Toth, *Scr. Metall. Mater.,* Vol 27, 1992, p 1575–1580
82. J.J. Jonas, *Inter. J. Mech. Sci.,* Vol 35, 1993, p 1065–1077
83. J. Baczynski and J.J. Jonas, *Acta Mater.,* Vol 44, 1996, p 4273–4288
84. L.S. Toth, *Thermomechanical Processing of Steel—J.J. Jonas Symposium,* S. Yue and E. Essadiqi, Ed., CIM, Montreal, Canada, 2000, p 109–120
85. J. Baczynski and J.J. Jonas, *Metall. Mater. Trans. A,* Vol 29A, 1998, p 447–462
86. R. von Mises, *Göttringer Nachrichten, Math.-Phys. K1,* 1913, p 582
87. S. Shrivastava, J.J. Jonas, and G. Canova, *J. Mech. Phys. Solids,* Vol 30, 1982, p 75–90
88. S.L. Semiatin and J.H. Holbrook, *Metall. Trans. A,* Vol 14, 1983, p 2091–2099
89. W.J.McG. Tegart, in *Ductility,* American Society for Metals, 1968, p 133
90. T.L.F. Muller, G.J. Richardson, and H.P. Stuwe, *Deformation under Hot Working Conditions,* Iron and Steel Institute, London, 1968, p 131
91. I. Dragon, M.Sci. thesis, University of Aston in Birmingham, 1967
92. R. Hill, *The Mathematical Theory of Plasticity,* Oxford University Press, 1950, p 346
93. P. Ludwik and R. Schen, *Stahl Eisen,* Vol 45, 1925, p 373
94. C. Zener and J.H. Hollomon, *Trans. ASM,* Vol 33, 1944, p 163
95. A Nadai and E.A. Davis, *J. Appl. Phys.,* Vol 8, 1937, p 213
96. D.J. Lloyd, *Met. Sci.,* Vol 14, 1980, p 193–198
97. C.M. Sellars, *Philos. Trans. R. Soc. (London) A,* Vol A288, 1978, p 147
98. A. Eichinger, *Festigkeits Theoretische Untersuchungenn* in *Hanbuch der Werkstoffprufung,* 2nd ed., Springer Verlag, Berlin, 1955
99. T. Sakai and J.J. Jonas, *Acta Metall.,* Vol 32, 1984, p 189–209
100. S.L. Semiatin, J.H. Holbrook, and M.C. Mataya, *Metall. Trans. A,* Vol 16A, 1985, p 145–148
101. E. Aernoudt and J.G. Sevillano, *J. Iron Steel Inst.,* Vol 211, 1973, p 718
102. F.E. White and C. Rossard, *Deformation under Hot Working Conditions,* Iron and Steel Institute, London, 1968, p 14–20
103. P. Choquet, *J. Nucl. Mater.,* Vol 52, 1975, p 34–44
104. J. Kortmann, *Neue Hütte,* Vol 2, 1977, p 371–375
105. I. Dragon, C.G. Radn, and C. Vaida, *Metallurgia (Bucharest),* Vol 30, 1978, p 439–443
106. T. Nishumura, H. Miguro, and S. Kikuchi, *Denki Seiko (Electric Furn. Steel),* Vol 49, 1978, p 187–195
107. H.J. McQueen, J. Sankar, and S. Fulop, *Mechanical Behavior of Materials,* Cambridge, England, Vol 2, 1979, p 675–684
108. F.H. Hammad, G.A. Hassan, and M.M. Dougol, *Aluminum,* Vol 55, 1979, p 457–461
109. C.M. Sellars and W.J.McG. Tegart, *Inter. Met. Rev.,* Vol 17, 1972, p 1–24
110. J. Coupry, *Rev. Alum.,* Vol 451, 1976, p 259–270
111. A.M. Sabroff, F.W. Boulger, and H.J. Henning, *Forging Materials and Practices,* Rheinhold, 1968
112. "Evaluating the Forgeability of Steel," The Timken Company, Canton, OH, 1974
113. H.J. Henning and F.W. Boulger, *Mechanical Working of Steel I,* P.H. Smith, Ed., Gordon and Breach, 1964, p 107
114. C.M. Young and O.D. Sherby, *Metal Forming—Interrelation between Theory and Practice,* A.L. Hoffmanner, Ed., Plenum Press, 1971
115. A. Nicholson, D. Smith, and P. Shaw, *Deformation under Hot Working Conditions,* Iron and Steel Institute, London, 1968, p 161
116. M.G. Cockroft and D.J. Latham, *J. Inst. Met.,* Vol 96, 1968, p 33–39
117. P.W. Bridgman, *Studies in Large Plastic Flow and Fracture,* McGraw-Hill, 1952, p 9
118. A.L. Hoffmanner, *Metal Forming—Interrelation Between Theory and Practice,* A.L. Hoffmanner, Ed., Plenum Press, 1971, p 349–391
119. A. Gittins and C.M. Sellars, *Met. Sci. J.,* Vol 6, 1972, p 118
120. M. Zidek, *Sbornik VSB Ostrave (Hutnicka),* Vol 11, 1965, p 687
121. S.L. Semiatin and J.J. Jonas, *Formability and Workability of Metals: Plastic Instability and Flow Localization,* American Society for Metals, 1984
122. S.L. Semiatin, G.D. Lahoti, and T. Altan, *Process Modelling—Fundamentals and Applications to Metals,* T. Altan et al., Ed., American Society for Metals, 1980, p 387–408
123. E. Rauch, G.R. Canova, J.J. Jonas, and S.L. Semiatin, *Acta Metall.,* Vol 33, 1985, p 465–476
124. E. Rauch, M. Eng. thesis, McGill University, Montreal, Quebec, Canada, 1983
125. J.-P.A. Immarigeon and J.J. Jonas, *Acta Metall.,* Vol 22, 1974, p 1235–1247
126. H.J. McQueen and J.J. Jonas, *Treatise on Material Science and Technology,* Vol 6, *Plastic Deformation of Materials,* R.J. Arsenault, Ed., Academic Press, 1975, p 393–493
127. O.D. Sherby, M.J. Harrigan, L. Chamagne, and C. Sauve, *Trans. ASM,* Vol 62, 1969, p 575
128. G. Glover, Ph.D. thesis, University of Sheffield, 1969
129. R. Bromley, Ph.D. thesis, University of Sheffield, 1969
130. D.R. Barraclough and C.M. Sellars, *Mechanical Properties at High Rates of Strain,* Institute of Physics Conference Series No. 21, J. Harding, Ed., Institute of Physics, London, 1974, p 111
131. I. Weiss, J.J. Jonas, and G.E. Ruddle, *Process Modelling Tools,* J.F. Thomas, Jr., Ed., American Society for Metals, 1981, p 97–125
132. N.D. Ryan, *The Science and Technology of Flat Rolling,* B. Fazan et al., Ed., IRSID, Mazieres-Le-Metz, France, 1987, p 17.1–17.9
133. N.D. Ryan and H.J. McQueen, *Mater. Sci. Technol.,* Vol 7, 1991, p 818–826
134. N.D. Ryan and H.J. McQueen, *J. Mater. Proc. Technol.,* Vol 36, 1993, p 103–123
135. S. Yue, F. Boratto, and J.J. Jonas, *Proc. Conf. on Hot- and Cold-Rolled Sheet*

Steels, R. Pradhan and G. Ludkovsky, Ed., TMS, Warrendale, PA, 1988, p 349–359

136. F.H. Samuel, R. Barbosa, F. Boratto, S. Yue, and J.J. Jonas, *Proc. Int. Conf. Physical Metallurgy of Thermomechanical Processing of Steels and Other Metals (Thermec88),* Iron and Steel Institute of Japan, Tokyo, 1988, p 721–728
137. F.H. Samuel, S. Yue, J.J. Jonas, and B.A. Zbinden, *ISIJ Int.,* Vol 29, 1989, p 878–886
138. L.N. Pussegoda, S. Yue, and J.J. Jonas, *Metall. Trans. A,* Vol 21A, 1990, p 153–164
139. L.N. Pussegoda and J.J. Jonas, *ISIJ Int.,* Vol 31, 1991, p 278–288
140. A. Najafi-Zadeh, J.J. Jonas, and S. Yue, *Metall. Trans. A,* Vol 23A, 1992, p 2607–2617
141. J.J. Jonas and C.M. Sellars, *Iron Steelmaker,* Vol 19, 1992, p 67–71
142. P.R. Cetlin, S. Yue, and J.J. Jonas, *ISIJ Int.,* Vol 33, 1993, p 488–497
143. T.M. Maccagno, S. Yue, J.J. Jonas, and K. Dyck, *Metall. Trans. A,* Vol 24A, 1993, p 1589–1596
144. T.M. Maccagno, J.J. Jonas, S. Yue, B.J. McCrady, R. Slobodian, and D. Deeks, *ISIJ Int.,* Vol 34, 1994, p 917–922
145. L.P. Karjalainen, P. Kantanen, T.M. Maccagno, and J.J. Jonas, *Proc. Int. Conf. Thermomechanical Processing of Steels and Other Metals (Thermec97),* T. Chandra and T. Sakai, Ed., TMS, Warrendale, PA, 1997, p 819–825

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 122-127
DOI:10.1361/hwpd2003p122

Chapter 9

Hot Working Simulation by Hot Torsion Testing

Joseph R. Pickens, Concurrent Technologies Corporation

PROCESS PARAMETERS for commercial hot working processes such as rolling, extrusion, and forging are normally based on prior experience with similar alloys. If the parameters are successful in producing material that meets quality control and property targets at economic production rates, little additional effort is generally extended to optimize the process. Production equipment and personnel are often not available for the experimentation required to perform an optimization study, and the cost of processing (then possibly scrapping) many large ingots or billets is excessive.

In addition, alloy developers are often faced with the dilemma of selecting hot working parameters to produce experimental material that may be expensive and in short supply. They are primarily concerned with producing material in the desired product form without losing material due to cracking or loading the process equipment excessively. Conservative hot working parameter selection may result in microstructures and properties that do not reach the ultimate potential of the experimental material.

One solution to these problems is to perform a scaled-down physical simulation of the hot working process by deforming relatively small amounts of material under carefully controlled conditions. The investigator can thereby use specimens weighing hundreds of grams as opposed to production ingots weighing hundreds or thousands of kilograms. Testing is carried out at temperatures and strain rates in the range of commercial hot working practice. The deformation parameters used and output data generated during the test can be related to commercial processes through analytical relationships. By using hot working physical simulation, one investigator can quickly generate the data necessary to improve or optimize a production hot working schedule. In addition, the data generated can be used as input for process modeling, increasing the accuracy of computer simulations of the process.

Primary input variables in hot working physical simulations are the deformation temperature, T_{def}, and strain rate, $\dot{\varepsilon}$. The strain, ε, to which the simulation experiment is performed depends on the process being simulated and the amount of deformation occurring in the process. The most useful output variables are flow stress, σ_o, and strain to failure, ε_f. Basically, σ_o is a measure of the resistance of a material to deformation, and ε_f is a measure of the deformation limits of a material.

Types of Hot Working Simulation Tests

The capabilities of various hot working simulation tests are summarized in Table 1 (Ref 1). These test methods are described in more detail in preceding chapters, and so the following is just a brief summary of hot tension, compression, and torsion testing used in the assessment of workability.

Hot tension testing can be used to simulate hot working operations at low strain rates. The peak flow stress, or stress for steady-state deformation, gives a measure of the material deformation resistance. A limitation of elevated-temperature tensile testing is necking, which prematurely limits or obscures the strain to which the specimen may be deformed. Elevated-temperature tensile testing is often used to assess superplasticity by determining the uniform deformation up to necking.

Uniaxial compression of cylindrical specimens is another test by which the flow stress can be measured accurately and reproducibly. However, uniaxial compression specimens undergo barreling due to friction at the contact surfaces, which limits uniform deformation to a true strain of about 0.8. A more important limitation is that the test does not provide a practical measure of working limits.

The cam plastometer is a hot working simulator in which two flat, opposing platens impact a flat, rectangular specimen. The platens are moved together rapidly by energy stored in a rotating flywheel acting through a cam, which produces strain rates as high as 500 s^{-1}. Cam plastometers are expensive and also have the disadvantage that the working limits of the material are not measured.

Hot torsion testing involves subjecting a dog-bone-shaped specimen to torsion on a machine that can be loosely described as an instrumented lathe. One end of the specimen is prevented from rotating, while the other is subjected to torque provided by a motor. A tube furnace is used to heat the specimen in place. Angular displacement is monitored from the drive-motor grip system and converted to torsion strain. The fixed end of the specimen is connected to a transducer that measures torque, which is converted to torsion stress.

Table 1 Comparison of hot working simulators

Simulators	$\dot{\varepsilon}$, max	ε, max	Working limit measurement	Temperature control	Flow stress measurement	Ease of microexam	Ease of quench	Limiting instabilities
Tension	10^{-1}	0.3	No	Good	Good	Poor	Fair	Necking
Uniaxial compression	10^{+1}	0.8	No	Fair	Good	Good	Fair	Barreling
Cam plastometer	10^{+2}	2	No	Fair	Good	Good	Fair	Friction
Torsion	10^{+3}	20	Yes	Good	Fair	Fair	Good	None

Note: Modified from Ref 1. max, maximum

Sophisticated hot torsion machines use dual axis transducers on the fixed end to measure both torque and axial load. Changes in axial load occur during torsion if the material undergoes recrystallization or other internal material transformations.

Hot Torsion Testing Practice

This section illustrates the use of hot torsion testing for optimization of hot working processes. Billets of an experimental aluminum alloy have been homogenized by a conservative temperature-time schedule and are ready for trial processing. What temperature and strain rate should be used to maximize the probability that useful product can be extruded?

A transverse slice from one experimental billet is cut, and torsion specimens of the geometry shown in Fig. 1 (Ref 2) are machined in a chordal orientation. This ensures that the shear strain during torsion testing is in the same orientation as the dominant shear deformation during extrusion. Fifteen specimens are deformed, one at each combination of five working temperatures and three torsional strain rates. These parameters are based on previous plant experience for similar alloys. During each test, the torque, M, is measured as a function of angular displacement.

As described in more detail in the preceding chapter, the shear strain in torsion (γ) increases linearly, starting from zero at the axis and increasing linearly to a maximum value at the surface. The maximum shear strain is:

$$\gamma = R\theta/L \quad \text{(Eq 1)}$$

where R is the radius of the test specimen, L is the length of the gage section, and θ is the angle of rotation of the movable end relative to the fixed end.

Because the shear strain at all points is beyond the elastic limit, the shear stress at all points across the cross section is in the plastic range, and, neglecting strain hardening or strain-rate hardening, the shear stress is uniform, except at the central axis. From equilibrium, the shear stress is related to the torque and specimen geometry by:

$$\tau = 3M/2\pi(R_o^3 - R_i^3) \quad \text{(Eq 2)}$$

where τ is the shear stress, M is the torque, R_o is the outside radius, and R_i is the inside radius. (For a solid test specimen, $R_i = 0$.) For materials that strain harden or strain-rate harden, the radial variation in strain and strain rate leads to a modification of Eq 2 for solid specimens (Ref 3):

$$\tau = (3 + m + n)\, M/2\pi\, R_o^3 \quad \text{(Eq 3)}$$

where m is the strain-rate hardening exponent of the material, and n is the strain-hardening exponent.

For the 15 tests performed, calculations using Eq 1 and 2 or 3 are plotted in Fig. 2 to determine τ_o, the shear yield strength, and γ_f, the shear strain to failure.

For isotropic materials, the shear yield strength and strain-to-failure values can be converted to equivalent tensile stress and strain through the relationships (Ref 4):

$$\sigma_o = \tau\sqrt{3} \quad \text{(Eq 4)}$$

$$\varepsilon_f = \gamma_f\sqrt{3} \quad \text{(Eq 5)}$$

where σ_o is tensile yield strength (or flow stress), and ε_f is tensile strain at failure or fracture. For this hypothetical example, the output variables, equivalent tensile flow stress, or σ_o, and equivalent tensile strain to failure, or ε_f, are plotted versus T_{def} for the various strain rates, as shown in Fig. 3.

Interpreting the results, clearly the flow stress decreases significantly and monotonically with T_{def}; however, at the highest strain rate, $\dot{\varepsilon}_3$, the flow stress approaches zero at the highest T_{def}. Metallographic evaluation was carried out in this sample, which showed that the vanishing flow stress resulted from hot shortness, that is, localized melting. Also, ε_f shows a peak at all strain rates at the intermediate deformation temperature, T_2.

Thus, the extrusion preheat temperature selected is slightly below T_2, where ε_f is near its maximum. The highest strain rate, $\dot{\bar{\varepsilon}}_3$, is avoided, because it may lead to hot shortness at T_1. To generate the desired strain rate, the appropriate billet and extrusion process parameters can be obtained from (Ref 4):

$$\dot{\bar{\varepsilon}} = 6VD_b^2 \ln R \tan\alpha/(D_b^3 - D_e^3) \quad \text{(Eq 6)}$$

where $\dot{\bar{\varepsilon}}$ is the mean equivalent strain rate in extrusion; D_b is the extrusion chamber diameter; D_e is the extrusion product diameter; α is the

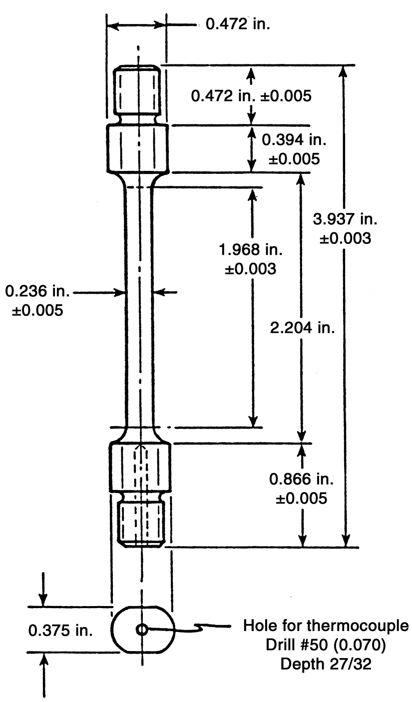


Fig. 1 Typical torsion specimen

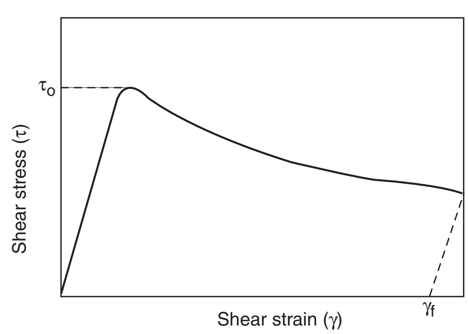


Fig. 2 Typical torsion stress-strain curve at elevated temperatures

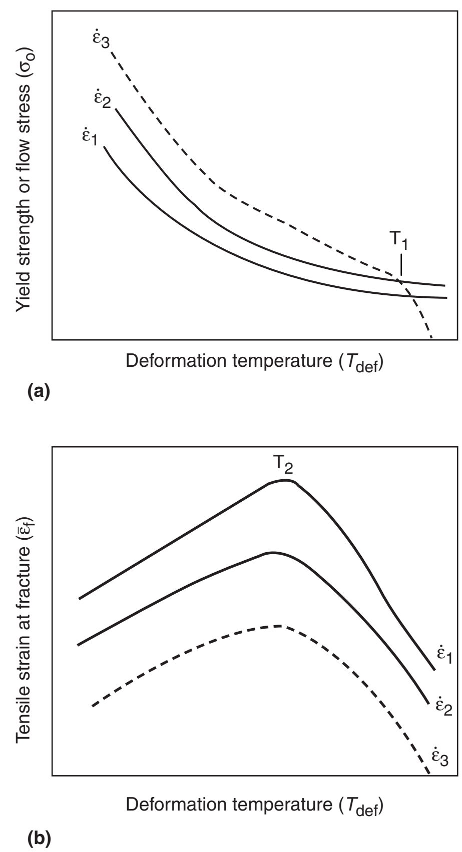


Fig. 3 Hypothetical dependence of (a) flow stress and (b) strain to failure on deformation temperature at three strain rates where $\dot{\varepsilon}_1 < \dot{\varepsilon}_2 < \dot{\varepsilon}_3$

semidie angle of extrusion die; V is the extrusion ram speed; and R is the extrusion ratio = billet area/extruded cross-section ratio = D_b^2/D_e^2.

This example illustrates the details in application of hot torsion testing. Following are real-world examples of the use of this approach.

Hot Torsion Application Examples

Example 1: Extrusion of Al-Mg Alloy. An extrusion fabricator considered extruding a proprietary aluminum-magnesium (5*xxx*) alloy that is very similar in composition to an alloy that is usually hot rolled. From rolling practice, it is known that this alloy is sensitive to homogenization schedule, and four candidate schedules are to be considered. The alloy has high magnesium content, which is well known to produce high flow stress. The production supervisor was concerned that the high flow stress might stall the press, and removing a stalled billet from the container wastes time. The press was to be taken out of production temporarily for experimental extrusion of several billets of the alloy, and the supervisor did not want the press to be unavailable for production any longer than necessary.

To solve this problem, torsion specimens were machined from billets that were subjected previously to the various homogenization treatments. The specimens were tested at temperatures spanning the expected extrusion range of 340 to 400 °C (650 to 750 °F) and at a strain rate typical of commercial extrusion. Some specimens were tested at very high strain rates to simulate the more rapid deformation on the surface of the extrusion billet. Such high strain rates may lead to hot shortness.

The results show that slight differences in flow stress exist (Fig. 4), but for the limited number of specimens tested, the differences between treatments are not statistically significant. However, the ε_f data (Fig. 5) show a clear, statistically significant superiority for treatment 3. While variations in homogenization treatments do not significantly alter the resistance of the material to deformation, they do have a large effect on the limits to which the material can be hot worked. The most advantageous homogenization schedule (schedule 3) is identified without taking the extrusion press out of production.

Note also that a few tests were run at 425 °C (800 °F), beyond the hot working range expected by plant metallurgists, and hot ductility (ε_f) was still increasing. Based on the hot torsion results, the production supervisor attempted to extrude this alloy with exit temperatures higher than those previously thought permissible. Plant trials confirmed that this was indeed possible.

Example 2: Determining the Effect of Alloy Variation Within Specification. Many metal producers have internal versions of standard alloys, and the versions are used for specific applications, depending on user requirements. The composition specifications for commercial alloys are generally wide and allow the producer to have narrower composition "sweet spots" for specific end-use applications. For example, an alloy may have alloying element composition in the upper part of the allowable range to enable service in strength-critical applications, such as extrusions under heavy compressive loading. Alternatively, for corrosion-critical applications, the alloying element content may be toward the lower end of the allowable range but just high enough to reliably meet minimum strength requirements. Such changes in composition, which are within specification but often not highlighted to the user, can have dramatic effects on hot working behavior.

For example, a vertically integrated extrusion fabricator wanted to make a highly alloyed version of an Al-Zn-Mg-Cu alloy for weight-critical applications where high strength could allow down-gaging. The standard version of this alloy (7*xxx*A, hereafter) has relatively low alloying content in the allowable composition range to minimize stress-corrosion-cracking susceptibility. The fabricator wanted to know what the effect of the increased alloying content would be on extrusion breakout pressure at various ram speeds. One "log" was cast for 7*xxx*A and another for a more highly alloyed variant (7*xxx*B, hereafter). Extrusion billets were machined from each log. Torsion specimens were machined

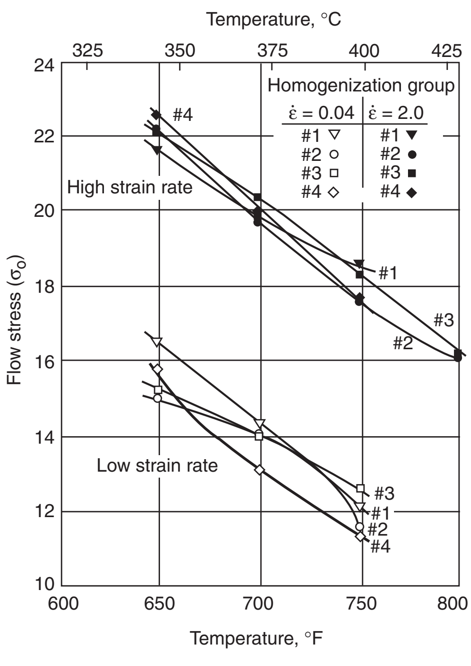


Fig. 4 Flow stress as a function of deformation temperature for an aluminum-magnesium alloy homogenized by different schedules (#1, #2, #3, #4)

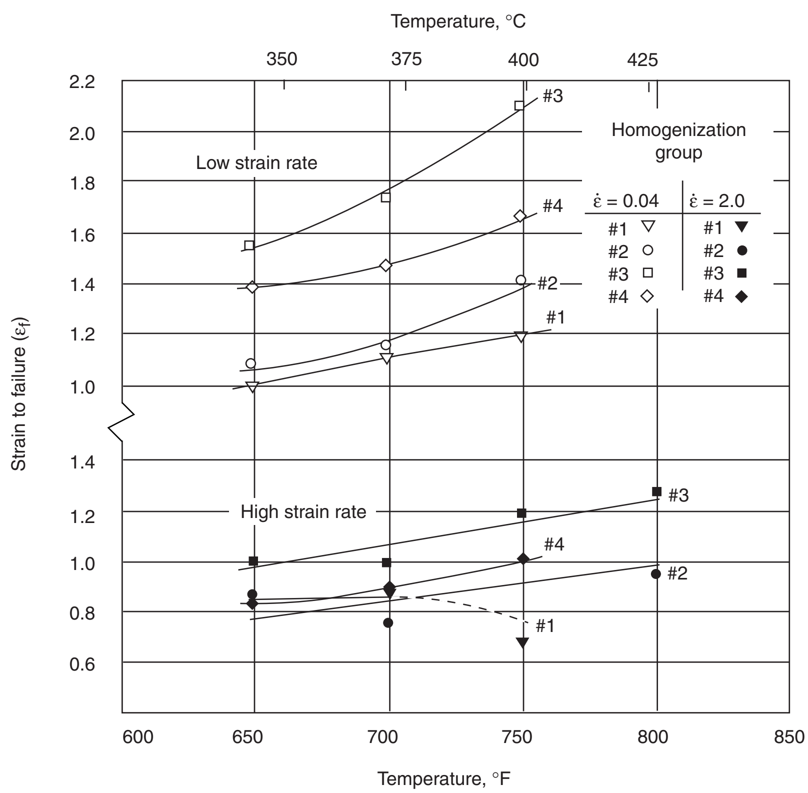


Fig. 5 Strain to failure as a function of deformation temperature for an aluminum-magnesium alloy homogenized by different schedules

from each log, with the long axis of the specimens parallel to the casting direction. Hot torsion testing was performed at strain rates corresponding to the mean-equivalent strain rate for the extruded shape in question, based on ram speeds that have been successful for 7*xxx*A. These strain rates were calculated from Eq 6. The standard billet preheat temperature of 315 °C (600 °F) for 7*xxx*A was used.

Flow stress increased monotonically with increasing strain rate for each alloy, as expected (Fig. 6). However, the flow stress for 7*xxx*B was significantly higher than that for 7*xxx*A, which was directly related to the increased alloying content. No difference in strain to failure was discerned. Subsequent extrusion of 7*xxx*B at 315 °C (600 °F) did indeed stall the press. Successful extrusion was conducted at higher temperatures.

Example 3: Torsion Test in a Cost-Benefit Analysis of EMC Ingot. An aluminum company considered investing in electromagnetic casting (EMC) equipment, with the hope of increasing casting recovery rates and reducing production costs. A slice of an EMC ingot of aluminum-magnesium alloy 5182 was purchased from a company that was casting the alloy by EMC. Visual inspection of the ingot clearly showed superior surface finish relative to conventional direct-chill-cast (DC) ingots. In addition, metallographic evaluation revealed that the EMC microstructure was finer than that of the DC ingot. As the aluminum company was performing the cost-benefit analysis to determine whether to invest in EMC equipment, uncertainties were raised, such as: Could EMC reduce homogenization costs? Would the finer microstructure of the EMC material lead to reduced edge cracking and result in higher mill recovery? The company did not want to invest in purchasing several ~10,000 kg (22,000 lb) ingots to answer these questions by a plant trial. However, investing in a small hot working simulation study was deemed worthwhile (Ref 5).

Torsion specimens were machined from the 5182 EMC ingot and a slice of a conventional 5182 DC ingot—each in the as-cast condition. Unfortunately, the DC ingot supplied was significantly lower in magnesium content than the EMC ingot (Table 2), and magnesium is well known to be a major determiner of flow stress in aluminum alloys. Specimens were homogenized for different times (from 3 to 9 h) at 523 °C (975 °F), and others were homogenized for 6 h at temperatures ranging from 482 to 548 °C (900 to 1020 °F). A representative strain rate of 1.2 s^{-1} was calculated using commercial rolling parameters and the following equation (Ref 6):

$$\bar{\dot{\varepsilon}} = (2\pi/60)N\left[\sqrt{(R'/H_2)\cdot(1-r)/r}\right]\cdot \ln[1/(1-r)] \quad \text{(Eq 7)}$$

where r is the reduction = $(H_1 - H_2)/H_1$; H_1 is the entry thickness; H_2 is the exit thickness; N is the angular speed of rolls in revolutions/min; and R' is the deformed roll radius.

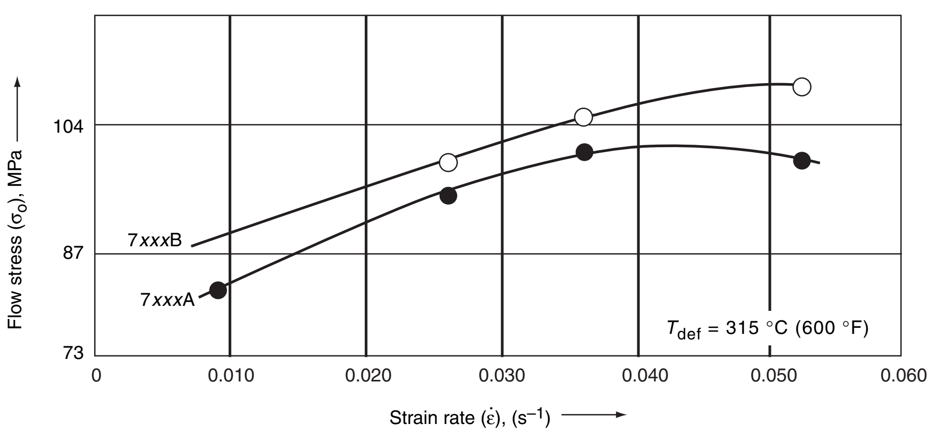


Fig. 6 Flow stress for two versions of a 7*xxx* aluminum alloy as a function of strain rate at T_{def} = 315 °C (600 °F)

Table 2 Composition of 5182 DC and 5182 EMC ingots (balance aluminum)

	Composition, wt%								
Ingot	Si	Fe	Cu	Mn	Cr	Zn	Ti	Na	Mg
EMC	0.133	0.214	0.032	0.383	0.010	0.014	0.009	0.0002	4.39
DC	0.100	0.223	0.017	0.330	0.035	0.012	0.010	0.0001	3.87

Note: EMC, electromagnetic cast; DC, direct chill cast

Flow stress for both 5182 DC and 5182 EMC were essentially unaffected by homogenization time for the times investigated (Fig. 7). In addition, strain to failure for 5182 DC was largely unaffected by homogenization time at 523 °C (975 °F) (Fig. 7). However, 5182 EMC displayed a dramatic increase in strain to failure, resulting in a 30% increase in hot ductility relative to 5182 DC. Similar results were observed at other homogenization temperatures. Based on this inexpensive study, the company believed that an optimization study could lead to reduced homogenization costs for 5182 EMC, and, furthermore, edge cracking would likely be lower for the EMC material. As an added benefit to this simulation, plant metallurgists evaluated the data and then questioned whether the existing homogenization practice for 5182 DC was accomplishing anything, so they initiated a hot torsion investigation to improve the homogenization practice.

Example 4: Determining Advantageous Rolling Temperatures. An aluminum company was uncertain as to the preferred rolling temperature for high-volume aluminum-magnesium alloy 5052. The conventional practice was to use the standard homogenization practice and begin rolling at the homogenization temperature of 549 °C (1020 °F) as the ingot was quickly transferred to the reversing mill from the homogenization furnace. To identify a more advantageous rolling temperature, a hot torsion investigation was initiated (Ref 7). Torsion specimens were machined from various regions of a cross section of a commercial-scale ingot and were homogenized by the standard practice (which was previously determined by hot torsion testing). The strain rate calculated from Eq 7, based on typical commercial rolling parameters, was 1.2 s^{-1} for the first pass on the reversing mill. Specimens were deformed at temperatures from 425 to 560 °C (800 to 1040 °F).

Results showed that the flow stress decreases monotonically with increasing deformation temperature (Fig. 8). However, a surprising peak in strain to failure was observed at approximately 500 °C (930 °F). The plant metallurgists had assumed that hot ductility increased monotonically with increasing deformation temperature until the onset of hot shortness. This surprising result suggested a revision in the plant rolling practice. The temperature of the homogenization furnace could be lowered toward the end of the treatment so that the ingots could begin rolling at a temperature slightly higher than the peak in Fig. 8. Thus, as the ingot cooled during hot rolling, it would be deformed at temperatures around the strain-to-failure peak; that is, the amount of cooling during a typical 20-pass reversing mill operation was measured, and that temperature change was placed symmetrically around the 500 °C (930 °F) rolling temperature, where the peak in strain to failure is observed. These predictions were confirmed in a plant trial with 20 production ingots.

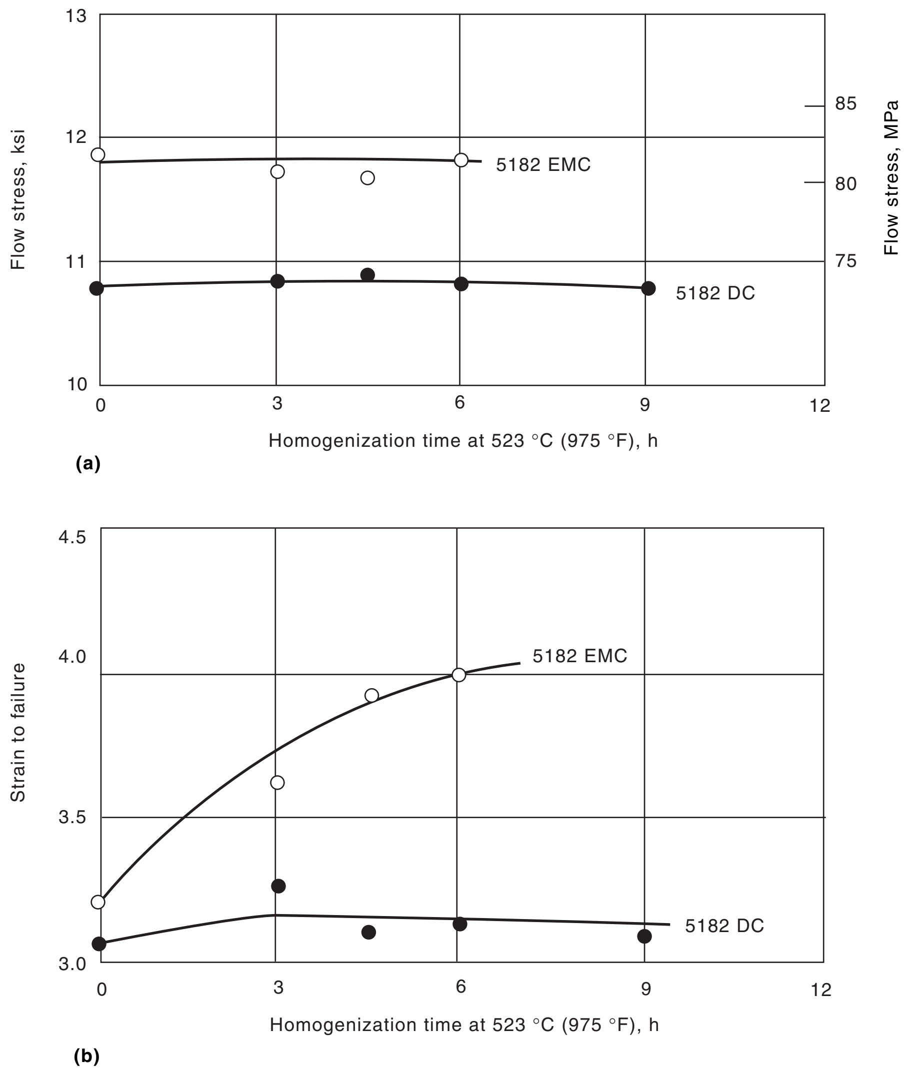


Fig. 7 (a) Flow stress and (b) strain to failure as a function of homogenization time at 523 °C (975 °F). Deformation temperature is 493 °C (920 °F).

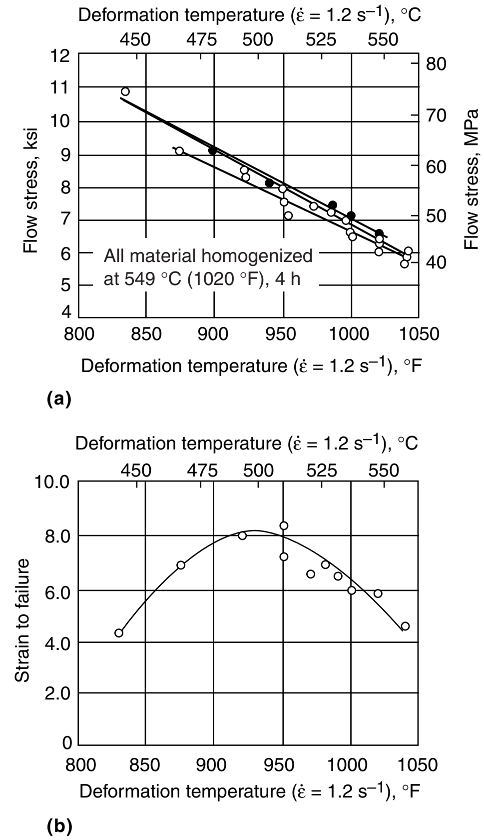


Fig. 8 (a) Flow stress and (b) strain to failure as a function of deformation

Example 5: Determining Commercial Rolling Parameters. For army applications, a stronger aluminum-base armor alloy has been sought for many years as a replacement for Al-4.4wt%Mg alloy 5083. One obvious way to increase strength of 5083-type alloys is to increase the magnesium content, but this raises the issue of stress-corrosion-cracking susceptibility (when sensitized by exposure to warm temperatures for long periods) and decreased workability. In fact, the propensity for edge cracking of aluminum-magnesium alloys during hot rolling generally increases with magnesium content and impurity levels.

The army was evaluating a candidate aluminum-magnesium alloy with 8.5 wt% Mg, and funds were not available to produce the several 10,000 kg (22,000 lb) ingots needed to develop commercial rolling parameters for the alloy. A hot torsion investigation was initiated (Ref 8) to determine whether an enhanced-purity version of the alloy would have superior edge-cracking resistance relative to a standard-purity version, to develop a viable homogenization schedule, to identify an optimal deformation temperature, and to elucidate the effect of strain rate on edge-cracking resistance. The strain to failure in hot torsion was used as an assessment of resistance to edge cracking.

Lab-scale ingots were cast at the two impurity levels, and torsion specimens were machined from locations equidistant from the surfaces of the ingots to minimize effects of segregation in the ingots. Specimens were deformed at a typical commercial strain rate for alloy 5083 over a wide range of deformation temperatures. In the as-cast condition, that is, not homogenized, the flow stress decreased monotonically with deformation temperature, and there was essentially no difference in flow stress for the two impurity levels (Fig. 9). However, strain to failure peaked at approximately 315 °C (600 °F), and, importantly, the enhanced-purity version displayed an ~40% improvement in strain to failure (Fig. 9).

The decision was made to develop a two-stage homogenization schedule, and the first step, 16 h at 427 °C (800 °F), was selected based on differential thermal analysis to reduce coarse composition gradients in this overalloyed material. This first step increased strain to failure for both purity levels (Fig. 10). A fixed, 8 h second homogenization step was performed at temperatures from 470 to 505 °C (880 to 940 °F), and the enhanced-purity material displayed further improvements in strain to failure. However, the standard-purity version actually displayed a surprising decrease in strain to failure with homogenization temperature (Fig. 10). This decrease was accompanied by a change in fracture mode to quasi-cleavage, which may have been caused by solid metal induced embrittlement by the impurities (Ref 8). This surprising decrease in hot ductility and edge-cracking resistance was found efficiently and inexpensively by hot torsion simulation of hot rolling.

The homogenization study was repeated with specimens deformed at a strain rate an order of magnitude lower than typical for 5083 (0.1 s^{-1}

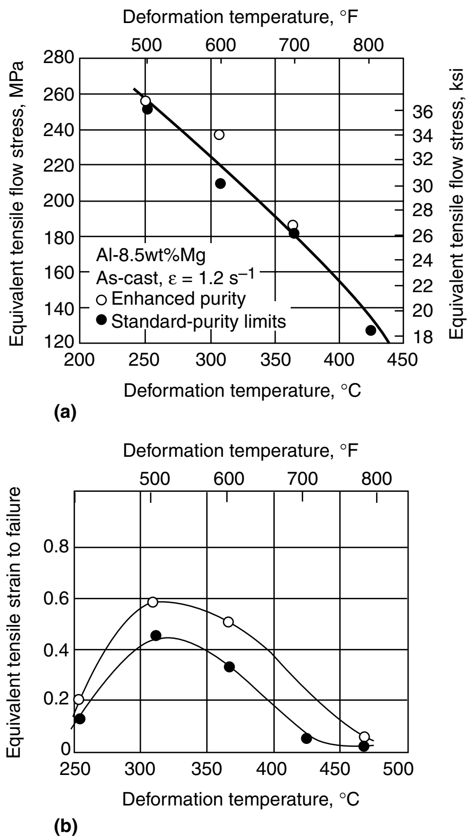


Fig. 9 (a) Equivalent tensile flow stress and (b) strain to failure of as-cast specimens deformed by hot torsion at various temperatures

instead of 1.2 s^{-1}). Strain to failure improved dramatically for both purity levels, with the enhanced-purity material displaying a relatively high strain to failure of 2.7 at the lower strain rate.

The lab-scale ingots were homogenized and rolled, with visual edge cracks correlating with the strain-to-failure values. A homogenization and a rolling schedule for enhanced-purity material were recommended to the U.S. Army, emphasizing that the lower strain rate would produce acceptably low edge cracking. It was recommended that the U.S. Army evaluate the effects of this slow strain rate on overall cost.

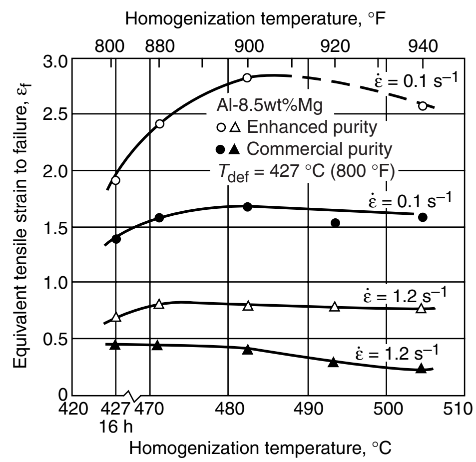


Fig. 10 Alloy purity and strain-rate effects on equivalent tensile strain to failure for the second homogenization step of 8 h at different temperatures following a first homogenization of 427 °C (800 °F) for 16 h

REFERENCES

1. H.J. McQueen and J.J. Jonas, Hot Workability Testing Techniques, *Metal Forming: Interrelation Between Theory and Practice*, 1971, p 93–428
2. C. Rossard, IRSID, Iron and Steel Institute, France
3. D.S. Fields and W.A. Backofen, *Proc. ASTM*, Vol 57, 1957, p 1259
4. G.E. Dieter, *Mechanical Metallurgy*, 2nd ed., McGraw-Hill, Inc., 1976, p 385, 652
5. J.R. Pickens, W. Precht, and J.J. Mills, Hot Rolling Simulation of Electromagnetically Cast and Direct-Chill Cast 5182 Aluminum Alloy by Hot Torsion Testing, *Proceedings of Mater., 30th Sagamore Army Research Conference: Innovations in Materials Processing*, Aug 1983, G. Bruggeman and V. Weiss, Ed., Plenum Press, 1985, p 101–116
6. A. Gittins, J.R. Everett, and W.J.M. Tegart, *Metalwork. Technol.*, Vol 4, 1977, p 377–383
7. W. Precht and J.R. Pickens, A Study of the Hot Working Behavior of Al-Mg Alloy 5052 by Hot Torsion Testing, *Metall. Trans. A,* Vol 18, Sept 1987, p 1603–1611
8. J.R. Pickens and F.H. Heubaum, "A Study of the Hot Workability of Al 8.5 wt. Pct. Mg Alloys for Armor Plate Applications," final report on contract DAAG46-85-C-0034, Report MTL TR 89-3, Jan 1989

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 128-135
DOI:10.1361/hwpd2003p128

Chapter 10

Thermomechanical Testing

Stéphane Guillard and Koushik Ray, Concurrent Technologies Corporation

THERMOMECHANICAL PROPERTIES TESTS, as covered in this chapter, are used to gain insight into the causes of problems that arise during a given thermomechanical process. Material data can be divided into two categories: mechanical properties and thermophysical properties. Tests needed to obtain the mechanical properties for most process modeling endeavors are described in previous chapters. Because this chapter deals with thermomechanical properties testing, it is important to mention at least some of the tests needed to obtain properties, such as specific heat, coefficient of thermal expansion, thermal conductivity/diffusivity, and density, all of which can significantly affect the behavior of a material during processing. Descriptions of these tests are given in the section "Thermophysical Properties Tests" in this chapter.

Successful outcomes from thermomechanical testing require proper test-selection, appropriate experiment design to maximize efficiency in obtaining test results, and using those results in an effective manner. Table 1 provides information about the most common processes, but the potential number of process and problem combinations is extremely high. Therefore, the intention here is not to cover every possible situation, but rather to present through examples how to approach solving typical problems associated with thermomechanical processes. A series of examples, each using the same general approach, is also included. They illustrate how the process, the problem, and previous knowledge influence the way the test was selected, conducted, and analyzed.

To get information about a specific thermomechanical process, the testing must resemble the process. Table 1 highlights how different processes correspond to various types of mechanical tests. The state of stress experienced during the actual metalworking operation should be present during the test. This correlation is the most direct way of ensuring that phenomena observed during the tests can be expected during the actual process.

One key element is to determine what parameters will be considered as possible influences on the phenomenon being observed. For example, when grain size during a metalworking process is the chief concern of the process designer, the parameters likely to influence grain size should be considered. Some parameters are easily contemplated, such as initial grain size, temperature, strain, and deformation rate. Other parameters are less obvious, such as prior thermomechanical history or frictional effects. The numbered examples provided in this chapter demonstrate how significant parameters were selected for specific tests.

Table 1 Guide to selecting a thermomechanical process test based on problem encountered and process used

Process	Problem	Test
Rolling	Surface cracks, edge cracking, barreling, center bursts	Upset, plane-strain indentation, partial width indentation, hot tensile (σ_f, A_f)
Forging	Hot shortness, center bursts, triple-point cracks/fracture, grain-boundary cavitation/fracture, shear bands/fracture	Wedge forging, side pressing, isothermal tensile or compressive, temperature-cycling tensile
Closed-die forging	Laps or buckles, die fill	Wedge forging, side pressing, isothermal tensile or compressive, temperature-cycling tensile
Extrusion, wire drawing	Center bursts, fracture	Hot tensile (σ_f, A_f)
Wire/rod drawing	Surface defects: fir-tree cracks, splitting, crows' feet, orange peel; center bursts (due to inhomogeneous deformation)	Boundary layer lubrication studies
Sheet metal forming	Tearing (localized necking), wrinkling and earring, Lüders bands, orange peel, fracture	Limiting dome height, tensile (R, n values), punch (forming-limit diagram)

Note: Besides thermomechanical testing, a number of tools can be useful in solving the type of problems listed above. Among them are reviewing tool and die design rules, using finite-element analysis, physical modeling, and metallography, including optical and scanning electron microscopes.

Typical Types of Problems

Essentially three types of problems can occur during a thermomechanical process, and each type requires the use of a specific test or analysis tool. First, the part might not get formed, or the final geometry might differ from expectations. An example would be a lack of complete die filling during a forging operation. This type of problem is usually solved through redesign of the die, die series, or preform, or by adjusting process variables (e.g., temperature, strain rate). Such design issues are outside the scope of this chapter; however, physical modeling, using wax or plasticine, can be one effective way of exploring such problems.

Recently, this type of issue has been tackled very effectively through the use of process modeling by finite-element-based software codes, such as DEFORM (Scientific Forming Technologies Corp.) or MSC.MARC (MacNeal-Schwendler Corp.), or by recently developed finite-volume-based software, such as MSC. SUPERFORGE (MacNeal-Schwendler Corp.). Independent of the type of analysis used, the accuracy of the results provided by these codes not only depends on how well the process model matches the actual process but also on how accurate and reliable the material data are.

A second type of problem commonly experienced arises when the desired geometry is obtained, but cracking occurs or, more generally, some type of defect is observed. Examples include edge cracking during rolling operations, central bursting during extrusion, cavitation during superplastic forming of aluminum alloys, or even cracking of the tooling itself. In these instances, combining the use of damage criterion (Ref 1–5) with process modeling generally allows enough insight to generate a solution that usually consists of redesigning dies/pass schedules or adjusting the processing parameters. Another approach, which involves extensive testing, consists of determining a map showing the occurrence of damage as a function of various processing parameters. The well-known forming-limit diagram (Ref 6, 7) is used to avoid failure through proper design of strain paths (Fig. 1). Another example of this approach is the strain-rate sensitivity (m) map (Fig. 2).

Although Fig. 2 does not specifically identify regions where damage will occur, it does highlight regions of high strain-rate sensitivity. These regions occur where damage is less likely because "necks" tend to push deformation to other, nonnecking areas, thus diffusing the necking mechanism. A strain-rate sensitivity about 0.2 to 0.3 is considered quite satisfactory for most typical forming operations where necking is a concern.

Of course, avoiding failure at strains much larger than those experienced during forming operations can be achieved with materials having very high strain-rate sensitivity. For example, *m* reaches values above 0.5 and as high as 0.9 in so-called superplastic materials, such as Ti-6Al-4V or aluminum 7475. Interestingly, the failure of 7475 under these conditions is not typically through necking but through cavitation. The cavitation, which is the nucleation and growth of tiny cavities, probably occurs at the intersection of surfaces of grain-boundary sliding or cooperative grain-boundary sliding within the material (Ref 8).

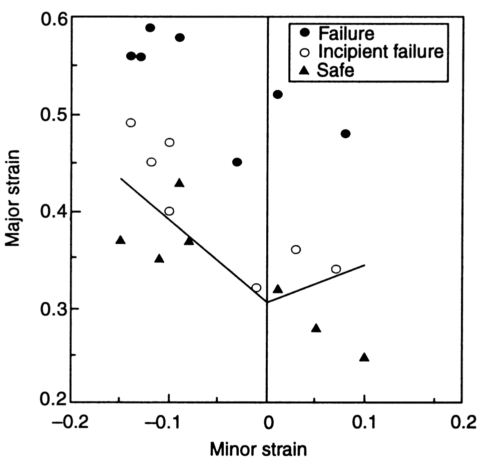


Fig. 1 Forming limit diagram for 301 stainless steel in the ¼ hard condition. Samples were machined in the rolling direction. Source: Ref 7

Consequently, although a given processing map might seem suitable to design how a material should be processed under specific conditions, it might not correspond to the controlling mechanism throughout the planned processing range of the material. In this particular situation, a cavitation map would be used in conjunction with the strain-rate sensitivity map to ensure safe processing throughout the intended strain range. Therefore, limitations in using processing maps show how carefully they must be utilized. Indeed, the various phenomena likely to occur must be known to determine what processing map or combination thereof should be considered.

Physical modeling can also be used to solve defect problems in thermomechanical processing. However, it is only helpful when an analog exists, such as plasticine for hot steel, and when the damage is not caused by a microstructural-level phenomenon but rather by a stress- or strain-state effect. For example, central bursting during extrusion can be physically modeled because it is caused by tensile stresses along the center of the product. Cavitation in aluminum cannot be physically modeled because it is caused by accommodation of a microstructural phenomenon—grain-boundary sliding.

A third type of problem relates to poor properties of the product itself, which arise from inadequate microstructure. The link between properties and microstructure is generally well understood. For example, small-grained microstructures are known to favor room-temperature tensile strength and fatigue crack resistance. Large grains, on the other hand, increase high-temperature creep resistance. As a result, inadequate properties can generally be traced back to inadequate microstructure.

Figure 3 relates the overall approach for solving thermomechanical processing problems to microstructural evolution. It highlights test choice as a function of both the thermomechanical process used and the problem experienced. This chapter addresses which tests should be conducted, how they should be conducted, and how they should be interpreted to gain sufficient insight into the microstructural evolution taking place in a material during thermomechanical processing. Once combined with other tools at the process designer's discretion, this information can be used so that the process yields the desired properties in the final product.

Thermophysical Properties Tests

In addition to thermomechanical behavior, thermophysical properties are also of particular interest when designing or optimizing a metalworking process. Thermophysical properties include specific heat, coefficient of thermal expansion, thermal conductivity/diffusivity, and density. In addition, interfacial heat-transfer coefficients between workpiece and tooling is important.

Specific Heat. Enthalpy is the overall heat (or energy) at constant pressure. The average specific heat, C_p, over a temperature range at constant pressure is defined as the change in enthalpy divided by the temperature change:

$$C_p = d(H_T - H_{273})_p/dT \qquad \text{(Eq 1)}$$

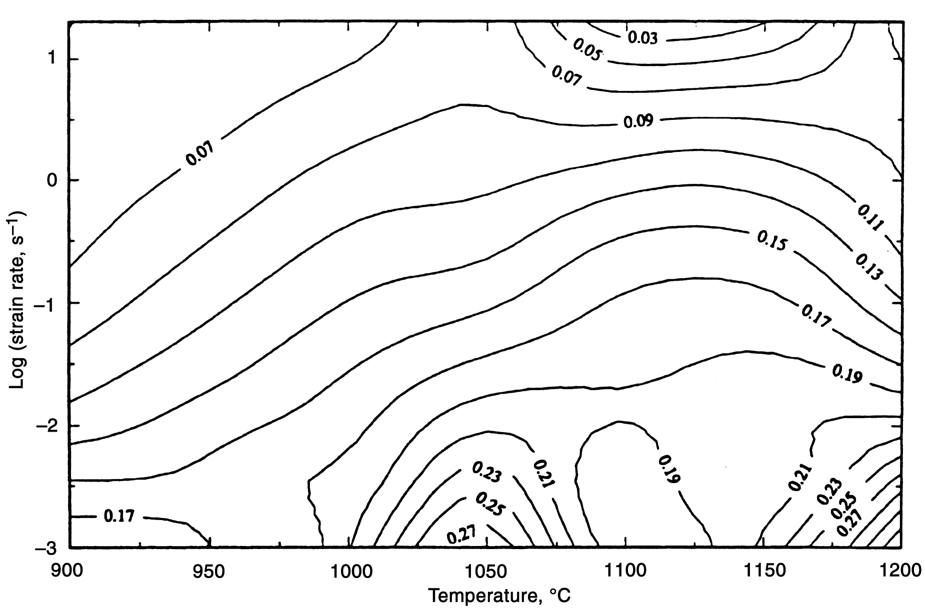


Fig. 2 Map depicting the evolution of strain-rate sensitivity as a function of temperature and strain rate for CF-8M stainless steel. Source: Ref 7

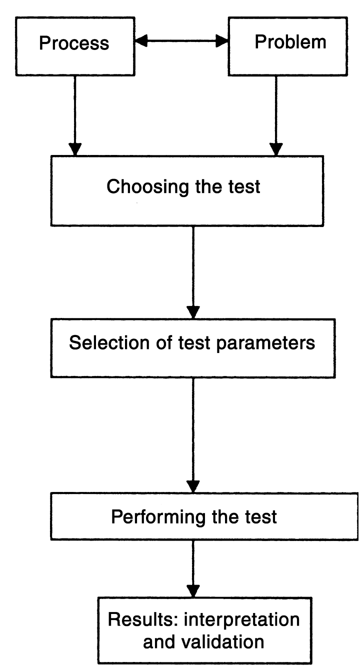


Fig. 3 Overall approach for solving problems encountered during thermomechanical processing

where H_T is the enthalpy at a given temperature, H_{273} is the enthalpy at 0 °C, and T is the temperature. The specific heat influences the rate of cooling and heating of the workpiece and how much the temperature changes with local heating by deformation and friction.

The two most common methods of measuring specific heat are differential thermal analysis (DTA) and differential scanning calorimetry (DSC). Figure 4 shows a DTA cell in which a specimen and a reference are heated by a single heating element while each temperature is monitored. These temperature curves can be used to derive the specific heat of the sample, but not with great accuracy.

The DSC cell (Fig. 5) has separate heaters for the sample and the reference material, with the temperatures of each monitored and controlled. The amount of energy required to force both to follow the same heating, or cooling, curve is measured, which provides a direct, accurate indication of the specific heat.

Thermal Expansion and Density. The coefficient of thermal expansion (CTE) is a measurement of strain resulting from a temperature change. Instruments that measure CTE are based on a precisely controlled and highly uniform temperature furnace combined with a system for measuring the change in length of the specimen or a change relative to a reference material. Available temperature ranges are usually only limited by the availability of stable reference materials. Temperatures can range up to 2000 °C for some systems.

The three most common methods of measuring CTE are through the use of a quartz dilatometer, a vitreous silica dilatometer, or an interferometer. The quartz dilatometer (Fig. 6) uses a quartz tube as a reference for the expansion of the specimen. Quartz is a crystalline form of silica that has an extremely low and well-defined thermal expansion coefficient. The specimen is contained in the quartz tube and pressed against a quartz push rod whose motion is monitored by a linear variable differential transformer (LVDT). The specimen temperature is monitored by a thermocouple maintained in an inert gas environment. The ratio of specimen strain to temperature change gives the CTE.

A vitreous silica dilatometer essentially uses glassy silica as the reference material. Current systems use optical encoders or optical interferometers to measure the displacement. However, the overall principle does not change. The specimen is positioned in a highly uniform and accurately controlled furnace, and the extension is measured and compensated for the small expansion of the silica.

Density is obtained quite easily as a function of temperature once the CTE as a function of temperature is known, based on conservation of mass.

Thermal conductivity and thermal diffusivity define the ability of the workpiece or tooling to conduct or spread heat, respectively. The difference between the two comes from the heat capacity and density of the material:

$$K = D\rho\, C_p \qquad \text{(Eq 2)}$$

where K is the thermal conductivity, D is the thermal diffusivity, and ρ is the density. Therefore, materials with the highest diffusivity have high conductivity and low heat capacity and density.

During metalworking, the thermal conductivity of both the workpiece and the tooling can be important. Two common methods for measuring these properties are the heat flow meter and the guarded plate method. These techniques essentially force heat to flow through both the specimen and a known conductivity reference material. Because lateral heat flow is prevented through the use of various guards, the amount of heat flowing through the sample and the reference material is the same. Therefore, measuring the temperature gradient through the specimen provides an indication of how easily heat flows through it. For example, a small temperature difference means high conductivity.

The laser flash technique, a newer method to measure thermal diffusivity, utilizes a transient or sudden change method as opposed to a steady-state method. The sample surface is blackened by carbon to give it an emissivity resembling a black body. It is then suspended in a furnace maintained at a known temperature. A sudden flash from a laser heats one side of the sample. The emission of infrared radiation from the other side is monitored, and the emission curve is analyzed to determine the heat diffusion rate. Essentially, the diffusivity is the measurement of the emitted infrared heat curve by using the time, $t_{1/2}$, to reach half of the maximum temperature. A heat loss parameter, w, is a geometric constant calculated for the specimen geometry. The specimen thickness, L, is also used:

$$D = wL/t_{1/2} \qquad \text{(Eq 3)}$$

Figure 7 shows a schematic of a laser flash experimental setup.

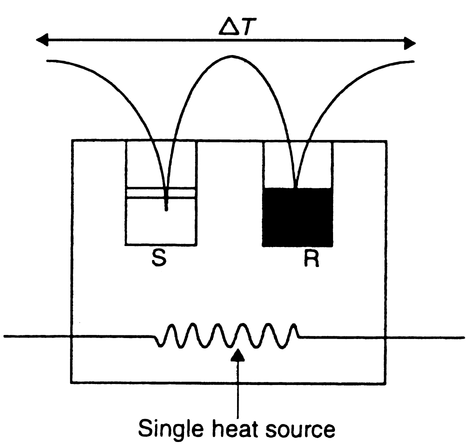


Fig. 4 Differential thermal analysis cell. S, specimen; R, reference; *T*, temperature

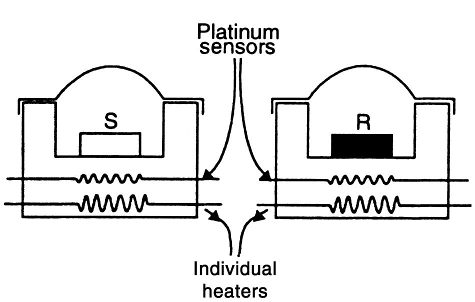


Fig. 5 Differential scanning calorimetry cell. S, specimen; R, reference

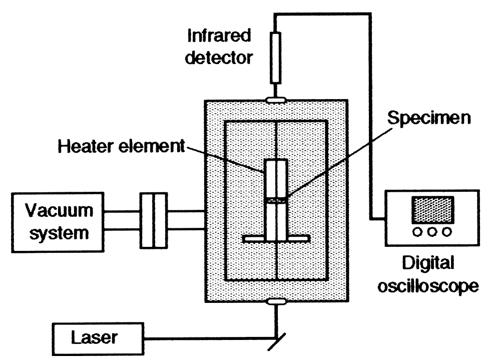


Fig. 7 Laser flash system

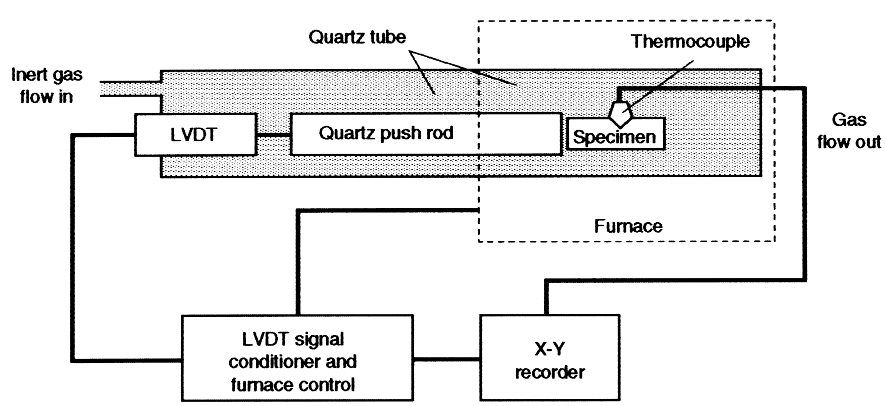


Fig. 6 Quartz dilatometer. LVDT, linear variable differential transformer

Testing for the Heat-Transfer Coefficient. Deformation processing often involves heat transfer, and one of the difficulties in modeling is determining the input value of the heat-transfer coefficient. The most commonly used method of determining the heat-transfer coefficient is the inverse method. In this method, the temperatures of the workpiece and tooling are measured at different locations as a function of time. Subsequently, the process is modeled with trial values of the heat-transfer coefficient to match the temperature measured in the experiments. The trial-and-error method for determining the heat-transfer coefficient can be programmed to loop the simulations until an optimal value of the coefficient is determined (see the section "Testing for the Heat-Transfer Coefficient" in the article "Testing for Deformation Modeling" in Volume 8 of *ASM Handbook* for a discussion of this for rolling operations).

Designing Thermomechanical Tests

The testing should be designed to maximize the chances of deriving useful information. The key element is to identify the mathematical form used to express the influence of the processing parameters and other parameters on the variable of interest. For example, when looking at the recrystallization behavior of a material, the following can be assumed to be correct:

$$X = 1 - \exp[-0.693(t/t_{0.5})^n] \quad \text{(Eq 4)}$$

where X is the percentage of recrystallization, t is the dwell time, $t_{0.5}$ is the time needed to obtain 50% recrystallization at that temperature and strain rate, and n is a material constant. Equation 4 requires tests for determining the effect of temperature, strain, and strain rate on $t_{0.5}$.

Factorial Design Testing. Knowing the mathematical form that the results should take is not a prerequisite. For example, techniques such as response surface determination can be used to obtain polynomial expressions that do not contain physical meaning, but are often sufficient for material and process modeling applications (e.g., example 3 in this chapter). A simple, yet effective, technique is called factorial design of experiments.

Often, the process parameters that can affect the quality of the product are known, but the relative importance is not clear. A sensitivity analysis is then needed, and factorial design of experiments aids in understanding the effects of the process parameters (now called factors) in the most efficient manner. The method of factorial design is best demonstrated by example.

Obtaining a superior microstructure in a material through optimization of the factors (temperature, strain rate, strain) is desired. The measure of superior microstructure will be in the form of tensile strength measurements made from samples that have undergone different processing parameter combinations. A schedule of experiments is set up in which the upper and lower bounds of the factors are chosen (Table 2).

Table 2 Example of lower and upper bounds selected for three factors considered to optimize a process

Factor	Lower bound (−)	Upper bound (+)
1. Strain, m/m	0.2	1.5
2. Strain rate, m/m·s	0.1	10
3. Temperature, °C	900	1150

Table 3 2^3 factorial design of experiments

Test	Factor 1	Factor 2	Factor 3
1	+1	+1	+1
2	+1	+1	−1
3	+1	−1	−1
4	+1	−1	+1
5	−1	−1	+1
6	−1	+1	+1
7	−1	−1	+1
8	−1	+1	+1

Note: For simplification, the lower bound of each variable is represented by −1 and the upper bound by +1. Refer to Table 2 for the appropriate values.

These upper and lower bounds are the practical and realistic limits within which the value of the parameter is expected to lie. Tensile test measurements are made on samples processed under the combination of factors listed in Table 3. The two values of each factor combine with the two values of every other factor, thus leading to 2^n tests (n = number of factors). Following this method, the factors, and more importantly the combination of factors, having the greatest influence on the quality of the product (i.e., tensile strength) are known.

Variations on this method exist, such as the fractional factorial method. A number of excellent reference books should be consulted prior to using any of these techniques (Ref 9–11). More recently, various software packages have been developed and are available commercially to speed up both the design of experiments and the analysis of results. These software packages include Design-Expert (STAT-EASE Inc.) and ECHIP Inc. (Hockessin, DE).

Improving Testing Efficiency. It has been assumed that one data point, such as a grain size value, is obtained from each sample. However, samples themselves can be designed so that many data points are obtained from each sample. The wedge-test specimen, because of its shape, allows the metallurgist to evaluate the effect of uniformly varying amounts of strain on the microstructural evolution of a material in a single test (Fig. 8). The relevant limits of strain can be fixed by designing the geometry of the wedge appropriately (Ref 12).

An important practical application of this test helps the process engineer determine the optimal reduction-anneal sequence for obtaining a recrystallized structure after breakdown of a cast ingot. In this instance, wedge specimens large

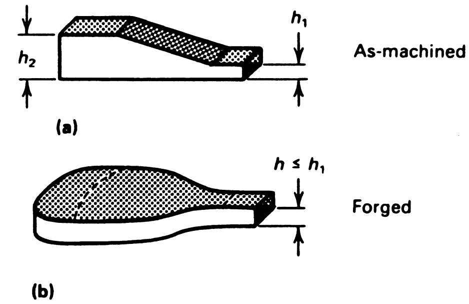


Fig. 8 Wedge test specimen. (a) As-machined specimen. (b) Specimen after deformation. Source: Ref 12

enough to be representative of the behavior of the cast material are machined from the cast structure whose microstructural evolution properties are to be determined. Typically, 12 to 15 grains across the diameter of the specimen is considered the minimum size to obtain a representative behavior. The wedge specimens are heated to different preheat temperatures. These specimens are then deformed to one specified height reduction and annealed at different temperatures for various time lengths to obtain a recrystallized microstructure. Finally, the specimens are quenched. Through microstructural examination of each quenched specimen, the process engineer can evaluate how strain affects the recrystallization behavior. Therefore, for one combination of annealing temperature and hold time, the engineer needs only one test to determine the ideal strain to input into the material. It is possible to obtain the entire range of strain values from only one sample. From the collection of samples, it is also possible to determine how the strain effect varies with the other processing parameters and to specify how much strain should be input into the material depending on the various processing parameters.

Obtaining, Analyzing, and Using Thermomechanical Testing Results

When attention has been paid to choosing the right test and carefully designing the test matrix, analyzing the results becomes a simple task. Indeed, the form the results should take is usually determined beforehand, and this step then essentially consists of characterizing the samples. The principal tool is undoubtedly the optical microscope. In most situations, the microstructural features of interest to the process engineer are the grain size, percent recrystallization, amount and distribution of second phase, and texture. These features directly influence the properties of the product. Image analysis can be very helpful at this stage, especially when a large number of samples is involved.

Use of the Results. At this stage, the process engineer should be able to express the relationship between parameters that can be controlled and variables that influence, in a known fashion, the properties of the product. The most effective approach is then to combine this relationship with a numerical model of the process, such as a finite-element model. This combination allows a "what if" scenario to be considered, with the microstructure of the product as one of the aspects to optimize.

Examples of Thermomechanical Testing: Design, Experiment, and Analysis

Four examples illustrate how the various considerations in testing are successfully used to solve practical thermomechanical processing problems.

Example 1: Superplastic Forming of Aluminum Alloy 7475. This first example illustrates how a biaxial test (closely resembling the intended process) allows the process engineer to verify the viability of superplastic forming for a sheet of aluminum alloy 7475. In superplastic forming, a key factor is to keep the grain size under control to within a small level (less than 15 μm), knowing the behavior of grain-size growth during the process. Before committing funds to making tooling and actually forming a component, the process designer decided to verify that the grain size would not be a problem. To do so, the material was tested by superplastically forming a double-well pan, which is a fairly standard test item for superplastic forming (Fig. 9).

This test was chosen as opposed to a standard uniaxial tensile test because the stress state needed to be biaxial to accurately reflect what happens in the aluminum sheet during actual forming. Following deformation, samples were cut from the pan at various locations and prepared for metallographic examination. The grain size was measured and remained almost constant at approximately 11.5 μm for deformation up to strains between 1 and 1.4. This result is in contrast to other findings that show increasing grain size with increasing strain. The explanation is two-fold. First, the other studies were performed using uniaxial tensile tests, which do not accurately reflect the biaxial state of stress present during superplastic forming. Second, the other studies considered a starting material that had equiaxed grains. The process engineer had noticed that the starting material had slightly elongated grains, which could influence the results.

As a consequence of conducting tests more directly applicable to the process, the process engineer was able to ascertain that grain growth would not be a problem in manufacturing the component. Had the specific tests not been run, the component might have been determined a very risky candidate for superplastic forming. The process engineer might have either increased the level of process control or selected another manufacturing method, both of which likely would have resulted in increased manufacturing costs.

Example 2: Cogging of Nickel-Base Superalloy 718. Another example of how a test should be designed to closely fit the thermomechanical process can be seen in how compression tests were designed by various researchers to depict phenomena occurring during the ingot breakdown by cogging of superalloy ingots. Cogging consists of a succession of open-die forging-type hits along the length of an ingot to break down the casting structure and promote recrystallization and grain refinement. As in other processes, the resulting microstructure is strongly influenced by the way the ingot is processed. In particular, the amount of strain received during each hit, the number of hits, and the dwell time between hits affect microstructural evolution. The process engineer must optimize all these variables to produce the most desired microstructure. In addition, because of the high costs involved in trying something new during an actual cogging operation, most operators only apply small changes to a process known to work at least reasonably well.

Therefore, obtaining relevant information from laboratory-scale testing is the key to finding out what works. Although laboratory-scale work is performed to simplify real-life situations, oversimplification is detrimental. When interested in cogging or radial forging to break down cast ingots, performing single-hit laboratory-scale tests, for example, is unlikely to yield appropriate results. For this reason, double-hit and even four-hit tests have been considered by process designers (Ref 13, 14). Even a factor as simple as load is unlikely to be represented accurately if a single-hit test is used to represent the full range of strain expected during the operation, when the total strain is, in fact, obtained from four successive hits (Fig. 10). In this situation, using the results from a single-hit test would be accurate for the first hit, would lead to overprediction of the load during most of the second hit, and would lead to underestimation of the maximum loads for the third and fourth hits. The process designer could not even conclude whether the test represents an upper or a lower bound.

Adding to the confusion, the previous observation is itself a function of the processing temperature. Indeed, Fig. 11, which represents the same comparison between single-hit and four-

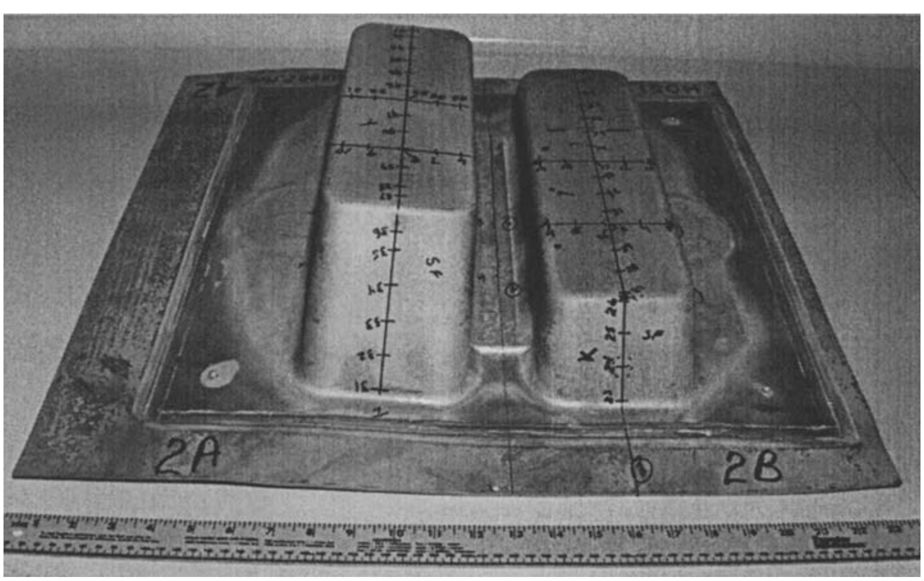


Fig. 9 Double-well pan superplastically formed using aluminum alloy 7475

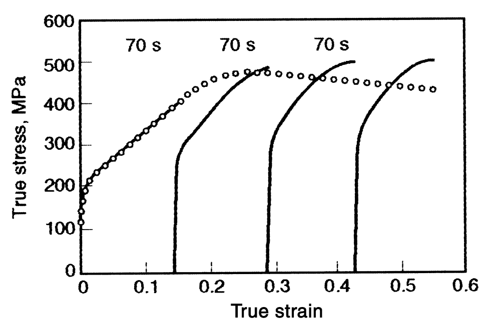


Fig. 10 Stress-strain curves comparing single-hit (open circles) and multiple-hit (continuous lines) tests for wrought alloy 718 at 950 °C. Source: Ref 13

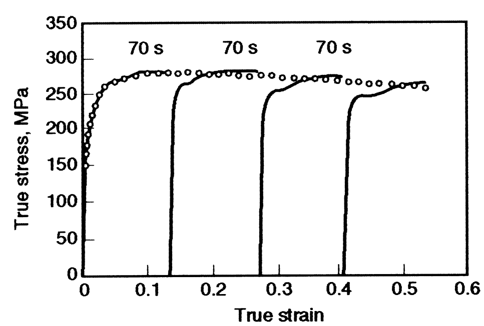


Fig. 11 Stress-strain curves comparing single-hit (open circles) and multiple-hit (continuous lines) tests for wrought alloy 718 at 1050 °C. Source: Ref 13

hits tests but at 100 °C higher than in Fig. 10, reveals that at 1050 °C using the single-hit test results to predict the load would yield acceptable results. However, because of the poor results obtained at 950 °C, the process engineer would probably not trust the results at 1050 °C.

The uncertainty is not limited to loads. Microstructurally, single-hit tests yield increased levels of recrystallization, which is explained by the fact that they favor adiabatic heating and, therefore, higher temperatures. Also, extensive deformation is applied without much chance for recovery (Ref 13). In a multiple-hit test, dwell times are available for recovery to occur, thereby reducing the driving force for recrystallization. Whatever phenomena might occur during the process, and especially when those phenomena are unknown, testing must closely match the actual process in terms of temperature, stress state, strain rates, strain paths, lubrication, or prior thermomechanical history of the material.

Example 3: Processing of Near-Gamma Titanium Aluminide. This example illustrates how statistical design of experiments can be used to maximize the amount of information obtained from very few tests. The challenge consisted of quantitatively determining the influence of processing on percent spheroidization (Ref 15). The processing parameters selected were temperature, strain, and strain rate during the deformation of the material. The type of experimental design selected was composite rotatable design (Ref 11), which is just a more advanced version of factorial design. However, the composite rotatable design allows the determination of a response surface equation. Therefore, if percent recrystallization could be measured on samples having been deformed under conditions of temperature, strain (ε) and strain rate ($\dot{\varepsilon}$) recommended by the design, then a polynomial equation could be obtained:

$$X = a\varepsilon + b\dot{\varepsilon} + cT + d\varepsilon^2 + e\dot{\varepsilon}^2 + fT^2 + g\varepsilon T + h\dot{\varepsilon}T + i\varepsilon\dot{\varepsilon} \quad \text{(Eq 5)}$$

Table 4 illustrates the levels of the three factors for a central composite rotatable design. As the +1 values in Table 3 show the normalized levels for the extrema values of the parameters for the trials, the +1 values in Table 4 represent normalized values of corresponding parameters. However, in a central composite rotatable design, the factorial design trials are only a subset of all the trials. So-called star trials, with values exceeding the extrema values, are also included (represented in Table 4 by normalized values of −1.68, 1.68, and the square values). Ultimate explanation for these values is based on group theory. The additional star trials make the composite rotatable design of experiments a more robust method than factorial design.

Compression tests were conducted for all conditions. Due to friction at the sample/platen interfaces, the samples exhibited some amount of barreling. As a result, the strain across the midplane of the sample, perpendicular to the compression axis, was not uniform. Because strain was being studied for its influence on percent spheroidization, it was critical to know what strain specific portions of the samples had experienced. In fact, because the central rotatable design called for specific strains to be considered, finite-element modeling was used to predict what areas of the samples experienced those specific strains. Figure 12 shows where a micrograph frame had to be taken to be sure that the material had experienced a strain of 0.70 ± 0.05.

Using an image-analysis system, the percent spheroidization was measured for all samples. The results were input into a "treatment matrix" where the coefficients corresponding to temperature, strain, strain rate, and the combinations were determined. Although it is not the intent of this chapter to provide details about any of these methods, the central composite rotatable design does allow determination of the significance of the various effects so that those with little importance can be deleted from the final mathematical expression (note that log $\dot{\varepsilon}$ was used instead of $\dot{\varepsilon}$ itself, for convenience only):

$$X = 3268.63 + 7.017 \log \dot{\varepsilon} - 4.7605T + 56.84\varepsilon + 0.001776T^2 - 12.52\varepsilon^2 \quad \text{(Eq 6)}$$

where T is temperature in Kelvin. Equation 6 is a response surface equation, and contour plots of that surface can be drawn. For example, Fig. 13 depicts the influence of strain and strain rate on percent spheroidization at 1330 K. The process engineer can use this plot, or others at various temperatures, to determine the optimal processing regime for this specific material if it is known that maximizing the percent spheroidization is beneficial to the properties of the product. A full derivation of the central composite rotatable design for this specific example, which also includes consideration of the spheroidized grain size, is provided in Ref 15.

Table 4 Central composite rotatable design of experiments for 3 factors

Test	T	ε	$\dot{\varepsilon}$	$\dot{\varepsilon}^2$	ε^2	$\dot{\varepsilon}^2$	$T\varepsilon$	$T\dot{\varepsilon}$	ε
1	−1	−1	−1	+1	+1	+1	+1	+1	+1
2	−1	−1	+1	+1	+1	+1	+1	−1	−1
3	−1	+1	−1	+1	+1	+1	−1	+1	−1
4	−1	+1	+1	+1	+1	+1	−1	+1	−1
5	+1	−1	−1	+1	+1	+1	−1	−1	+1
6	+1	−1	+1	+1	+1	+1	−1	+1	+1
7	+1	+1	−1	+1	+1	+1	+1	−1	−1
8	+1	+1	+1	+1	+1	+1	+1	+1	−1
9	−1.68	0	0	+2.83	0	0	0	0	0
10	+1.68	0	0	+2.83	0	0	0	0	0
11	0	−1.68	0	0	+2.83	0	0	0	0
12	0	+1.68	0	0	+2.83	0	0	0	0
13	0	0	−1.68	0	0	+2.83	0	0	0
14	0	0	+1.68	0	0	+2.83	0	0	0
15–20	0	0	0	0	0	0	0	0	0

Note: The values 1.68 and 2.83 are chosen to yield special properties that facilitate the analysis of the results. Only the first three columns are used to define the values of the processing factors during the experiments. The other columns provide information used to determine the values of the parameters corresponding to the respective effects and interactions at the top of each column.

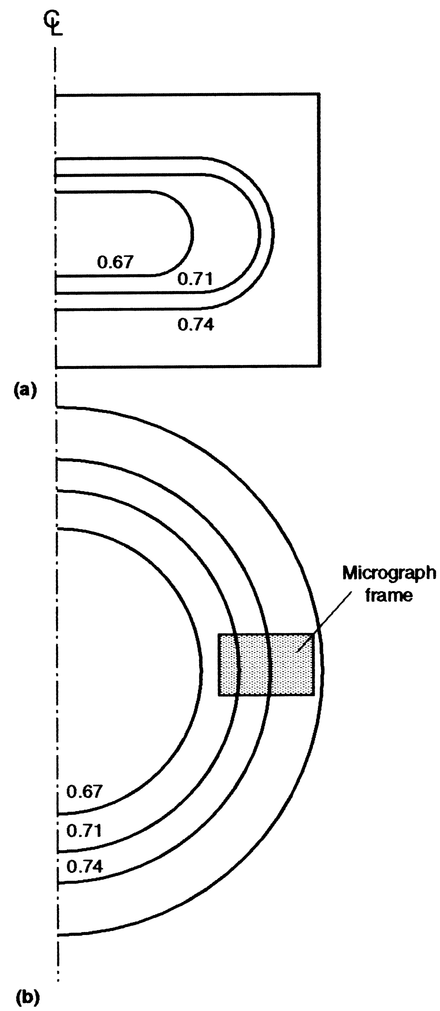


Fig. 12 Strain contours obtained after upsetting a compression sample in a median plane (a) parallel to and (b) perpendicular to the compression axis. Note the micrograph frame positioned to ensure observation of microstructure resulting from imposition of 0.7 strain.

Example 4: Using Damage Modeling to Eliminate Defects During Ingot Side Pressing. The objective in calculating damage in a metalforming operation is to determine regions where failure is most likely to occur. Because failure usually is caused by accumulation of tensile strain, the damage number should reflect the accumulation of tensile strain at a point along the path of the evolution during metalforming. Damage evaluation considers more than strain accumulation. It also considers the effective acting stress that might have assuaged or aggravated the accumulation of the strain, which makes the number a more realistic reflection of how "damaging" the strain has been to the material. A damage value can be determined:

$$\int_0^{\varepsilon_f}\left(\frac{\sigma^*}{\bar{\sigma}}\right)d\bar{\varepsilon}_p = C \quad \text{(Eq 7)}$$

where ε_f is the fracture strain, σ^* is the maximum tensile stress, σ is the equivalent stress, and ε_p is the equivalent strain.

There are two limitations when using the damage parameter. First, the numbers are purely a relative measure of the probability of damage and cannot be used to determine with certainty whether damage will occur. Secondly, there might be other modes of damage not incorporated in the damage parameter, such as what might cause the material to fail at otherwise-predicted low damage parameter values. The knowledge, then, to be gleaned from damage predictions is where the most critical areas during thermomechanical processing are and how could an adjustment of the processing parameters reduce the probability of failure.

The concept of using damage values as potential indicators of areas of concern can be shown by damage modeling a side-pressing operation. K-Monel has limited ductility at warm working temperatures. Hence, there is a need for optimizing the processing parameters, such as die design, temperature, and ram speed; otherwise, cracking might occur. Because of the high cost involved with full-scale trials to validate optimal parameters, physical modeling with plasticine can be used. In one instance, the plasticine ingot cracked during the side-pressing operation. The cracking is shown in Fig. 14 as voids adjacent to the inserted hard beads. The challenge was to optimize the parameters in such a way to prevent the cracking. The process engineer constructed a finite-element analysis model and plotted the damage values associated with it during the forming operation (Fig. 15). Cracking corresponded to the region with the highest damage value in the model, which is expected. Next, the processing parameters were modified in various ways to lower the damage contour values. By judicious changes in the platen design, the process engineer was able to lower the contours (Fig. 16). Incorporating the suggested changes in the

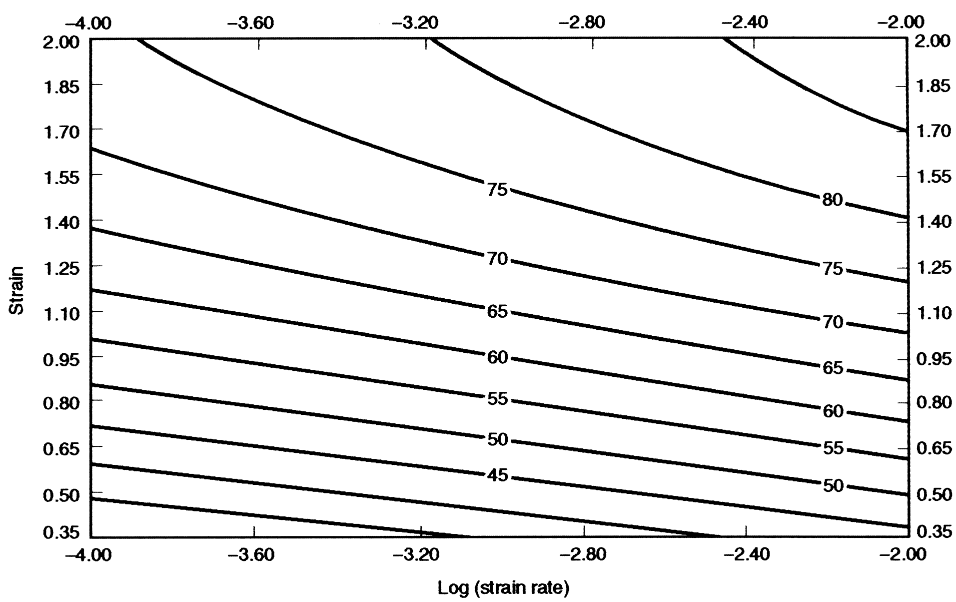


Fig. 13 Effect of strain and strain rate on percent spheroidization of Ti-49Al-2V at 1330 K. Source: Ref 15

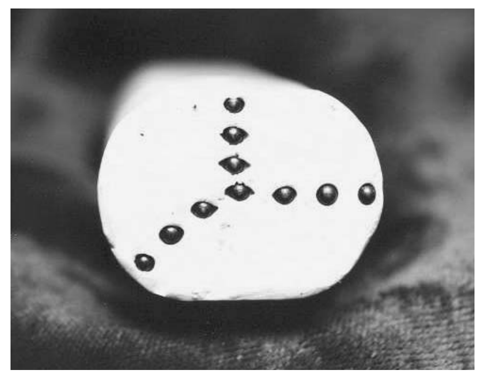


Fig. 14 Plasticine ingot representing side pressing of a K-Monel ingot. Hard beads have been embedded to illustrate the presence of tensile stresses at the center of the ingot. Damage seen is directly related to the tensile stresses.

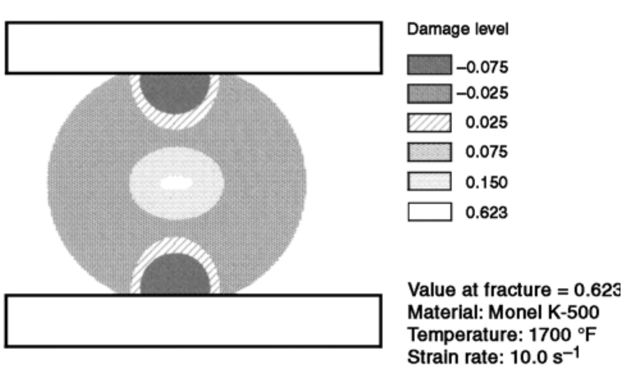


Fig. 15 Model showing damage values corresponding to the process used to manufacture the part shown in Fig. 14

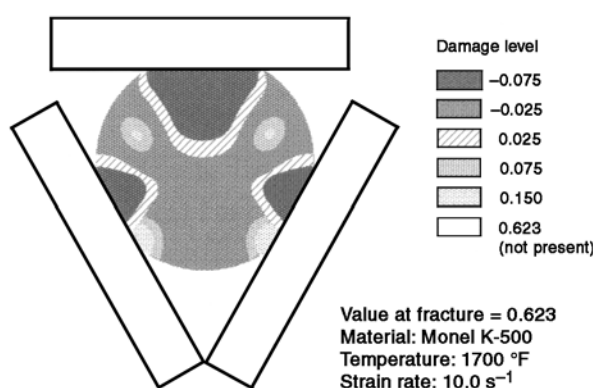


Fig. 16 Model showing damage values obtained after modification to the platen used to manufacture the part shown in Fig. 14

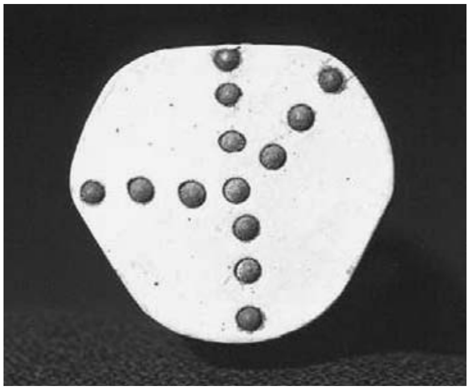


Fig. 17 Successfully side-pressed plasticine ingot after platen modification suggested by model shown in Fig. 16. Note that the plasticine around the hard beads does not flow in this instance, suggesting that damage has been minimized.

plasticine model obtained defect-free parts (Fig. 17). The damage modeling exercise, in this case coupled with physical modeling for further cost minimization, helped avoid the cost of running in-plant experiments and prototypes.

REFERENCES

1. M.G. Cockcroft and D.J. Latham, Ductility and the Workability of Metals, *J. Inst. Met.,* Vol 96, 1968, p 33–39
2. M. Oyane, Criteria of Ductile Fracture Strain, *Bull. Jpn. Soc. Mech. Eng.,* Vol 15, 1972, p 1507–1513
3. A.M. Freudenthal, *The Inelastic Behaviour of Solids,* John Wiley & Sons, 1950
4. D. Zhao, J.P. Bandstra, and H.A. Kuhn, A New Fracture Criterion for Fracture Prediction in Metalworking Processes, *Concurrent Engineering Approach to Materials Processing,* S.N. Dwivedi, A.J. Paul, and F.R. Dax, Ed., TMS, 1992, p 107–119
5. J.P. Bandstra, "3D Extension of Kuhn Surface Fracture Criterion," CTC/JPB-M0469-95, CTC Memorandum, Concurrent Technologies Corporation, 1995
6. S.S. Hecker, *Sheet Met. Ind.,* Vol 52, 1975, p 671–675
7. *Atlas of Formability* Bulletins, National Center for Excellence in Metalworking Technology, Concurrent Technologies Corporation, p 15
8. A.H. Chokshi, A.K. Mukherjee, and T.G. Langdon, *Mater. Sci. Eng. Res.,* Vol 10, 1993, p 237–274
9. C.R. Hicks, *Fundamental Concepts in the Design of Experiments,* Holt, Rinehart, and Winston, 1973
10. W.G. Hunter and J.S. Hunter, *Statistics of Experimenters,* John Wiley & Sons, 1978
11. W.J. Diamond, *Practical Experiment Design for Engineers and Scientists,* Van Nostrand Reinhold, 1989
12. S.L. Semiatin, *Workability Testing Techniques,* George Dieter, Ed., American Society of Metals, 1984, p 202
13. M.C. Mataya, Simulating Microstructural Evolution during the Hot Working of Alloy 718, *JOM,* Jan 1999, p 18–26
14. D. Zhao, S. Guillard, and A.T. Male, High Temperature Deformation Behavior of Cast Alloy 718, *Superalloys 718, 625, 706 and Various Derivatives,* E.A. Loria, Ed., TMS, 1997, p 193–204
15. S. Guillard, "High Temperature Micro-Morphological Stability of the (α_2 + γ) Lamellar Structure in Titanium Aluminides," Ph.D. dissertation, Clemson University, 1994

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 139-171
DOI:10.1361/hwpd2003p139

Chapter 11

Design for Deformation Processes

PRODUCT FORMS may be shaped by a wide variety of metalworking methods, and the success or failure of a bulk working process depends on the interaction of many variables, such as process conditions (strain rate, temperature, and shape) and metallurgical conditions (e.g., flow stress, workability, and microstructure). There are four major design considerations in selecting a deformation process:

- The workpiece material and its flow-stress behavior
- The fracture behavior of the material and the effects of temperature, stress state, and strain rate on ductility and fracture (i.e., workability)
- The desired microstructure needed to produce an acceptable product and a determination of which process should be used to produce this microstructure
- The added constraints of available equipment and economics in addition to flow stress, forming, and part-performance considerations

The fourth consideration usually dominates the other considerations, sometimes to the detriment of the material being worked. It is also necessary to consider the effect of friction at the tool/workpiece interface, as well as the ability of the tool to withstand the loads and temperatures required to deform the workpiece into the desired geometry.

Proper design of deformation processes involves many decisions. The design includes both part-related and process-related decisions, such as those listed in Table 1. This division or categorization of decisions is somewhat arbitrary, as processing and part decisions can be interrelated. Preform (initial workpiece) geometry, the deformation temperature, amount of force and strain rate, the friction conditions, and the metallurgical condition of the workpiece are all interrelated factors in the design. Nonetheless, Table 1 identifies some major factors in metalworking design. These decisions are needed to accentuate the advantages of deformation processing (Table 2), while still preventing problems of inferior parts or poor equipment utilization (Table 3). It is also interesting to note that many of the advantages of deformation can become disadvantages if the process is poorly designed and/or poorly executed.

This chapter briefly introduces the basic design concepts of bulk deformation processes that are described in more detail in subsequent chapters. This chapter also introduces reasons behind the selection of a deformation process as a production method for producing a part or product form. Product form may be produced by a wide variety of methods, and the basic goal of a deformation process is the same as other manufacturing processes; that is, a desired form of a particular material is sought with certain mechanical and/or physical characteristics at a minimum cost. What are the characteristics of deformation processing that are desirable for designers? What are the advantages of bulk forming as opposed to achieving the shape by solidification, by molding of powder, or by machining? These general questions are discussed briefly in this chapter. For a more complete discussion of the equipment and methods used for deformation processes, see Ref 1 to 4 and subsequent chapters in this book.

Table 1 Design decisions associated with deformation processes

Part-related decisions
Part or product material selection
Geometry and dimensions to be produced
Required properties (mechanical, physical, and metallurgical)
Process-related decisions
Equipment selection (type, rate, and load requirements)
Starting material geometry (plate, bar, sheet, etc.)
Workpiece temperature and tooling temperature
Orientation of part during deformation step(s)
Location of flash or scrap loss
Number of deformation steps
Lubrication and method of application
Starting microstructure and control of microstructure during forging sequence (preheat practice and intermediate heating steps, if any)

Table 2 Advantages of deformation processing

Improved internal quality due to compressive deformation
Uniform grain structure
Elimination of casting porosity
Breakup of macrosegregation patterns
Beneficial grain-flow pattern for improved part performance
Improved toughness due to grain flow and fibering
Improved fatigue resistance due to grain-flow pattern
Controlled surface quality
Burnished surface can have improved fatigue resistance due to quality of as-forged surface
High throughput due to potentially high rates of forming
Ability to produce a net-shape or near-net-shape part

Why Use Deformation Processes?

Cost, dimensions and tolerances, surface finish, throughput, available equipment, and part-performance requirements dictate the material and process selection for production. The simplest practice that still achieves the desired product form should be used. Deformation offers many advantages, especially in terms of microstructural benefits, but these processes also have disadvantages, the main ones being the cost of equipment and tooling.

Advantages and Disadvantages of Deformation Processes. Discussions of process advantages must be approached carefully because most final parts are generally subjected to more than one type of manufacturing process. For example, a screw-machined part has been first cast, hot rolled to bar stock, and possibly cold rolled or drawn prior to screw machining. A forging may have been cast, hot rolled to bar form, cropped into a billet, forged through multiple stations, and then finish machined. There are steps needed to produce the starting material (ingot or cast shape), intermediate steps needed to shape the material into a manageable interim form (bar, plate, tube, sheet, wire), and then steps needed to make the final part. Both processes to make interim product forms or stock and processes to make parts are included here.

The objective of hot forging or hot rolling of cast materials is to refine the structure that results from solidification. To alter the inhomogeneous structure due to solidification and to pro-

Table 3 Potential disadvantages of deformation processing

Fracture-related problems
Internal bursts or chevron cracks
Cracks on free surfaces
Cracks on die contacted surfaces
Metal-flow-related problems
End grain and poor surface performance
Inhomogeneous grain size
Shear bands and locally weakened structures
Cold shuts, folds, and laps
Flow-through defect
Control, material selection, and utilization problems
Underfill, part distortion, and poor dimensional control
Tool overload and breakage
Excessive tool wear
High initial investment due to equipment cost
Poor material utilization and high scrap loss

duce a more workable microstructure, cast ingots and continuously cast slabs and blooms are typically hot worked into interim product forms, that is, plate, bars, tubes, or sheet. Large deformation in combination with heat is very effective for refining the microstructure of a metal, breaking up macrosegregation patterns, collapsing and sealing porosity, and refining the grain size.

Many design decisions are required in order to take advantage of the benefits of deformation processing while avoiding potential problems of flow-related defects: fracture or poor microstructure. Some of the advantages of a part produced by deformation are listed in Table 2. In addition, the process can be tailored to achieve tight control of dimensions for mass production, and typically some net surfaces can be achieved. While the goal is to achieve a net shape, it is rare that a totally net shape is produced by bulk deformation processes, and some machining is typically needed to produce a usable part. Sheet-forming processes, however, often result in net functional surfaces. The production rate for many deformation processes can be high, so that high-volume production requirements can be met with efficient machinery utilization.

Disadvantages of deformation processes are listed in Table 3. It is interesting to note that many of the advantages of deformation also show up on the list of disadvantages. If the deformation process is poorly designed and/or poorly executed, the sought-after advantages will not be realized, and instead, an inferior part will be produced. The categorization in Table 3 is somewhat arbitrary because metal flow, fracture, die wear, and tool stresses are so interlinked. The decisions that the designer must make concerning the preform or initial workpiece geometry, the deformation temperature, amount of force and forging speed, the friction conditions, and the metallurgical condition of the workpiece are all interrelated. Decisions about the deformation process (e.g., Table 1) must be made to accentuate the advantages listed in Table 2 and to overcome or avoid disadvantages listed in Table 3.

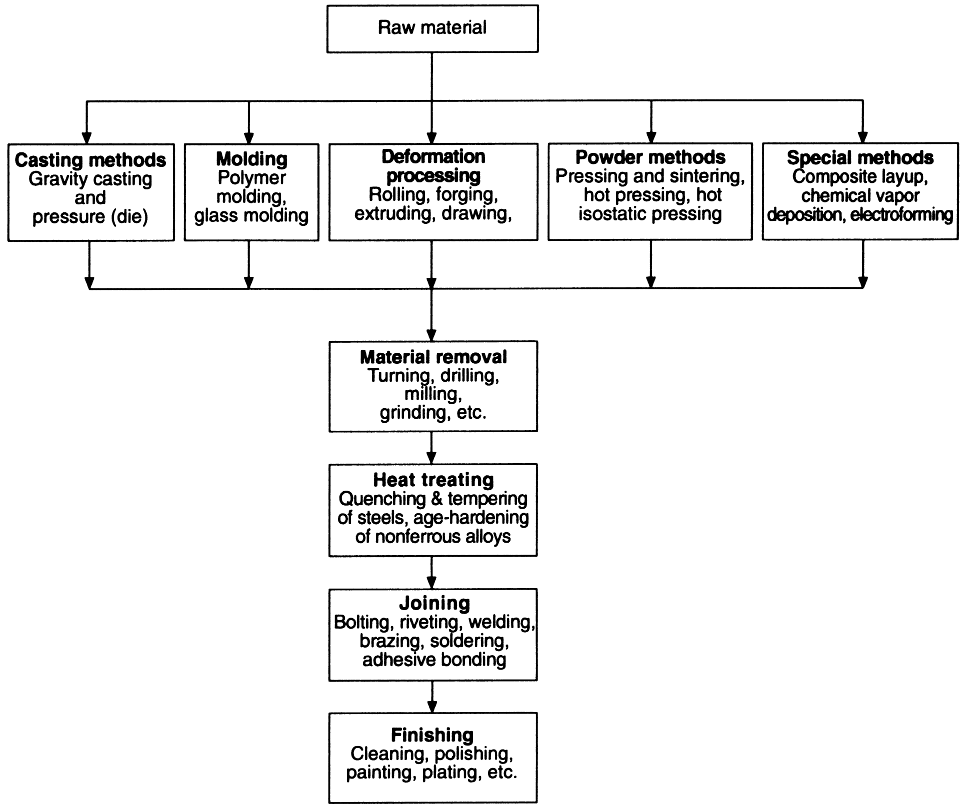


Fig. 1 The nine classes of manufacturing processes. The first row contains the primary forming (shaping) processes. The processes in the vertical column below are the secondary forming and finishing processes. Source: Ref 6

Characteristics of Manufacturing Processes (from Ref 5)

Figure 1 (Ref 6) gives a breakdown of manufacturing processes into nine broad classes, and the goal is to select processes that maximize quality and minimize the cost of the part. Selection depends on the basic characteristics of manufacturing processes in terms of:

- Material factors
- Size and shape factors
- Process factors

These factors are described briefly in rather elementary form, although the intent is to provide sufficient detail for identification of likely processes at the conceptual or embodiment stages of a design. Further details on the selection of manufacturing processes are described in Ref 7. Process-selection charts introduced by Ashby (Ref 6) are also useful in comparing manufacturing processes (e.g., casting, metalworking, polymer processing, power fabrication, and machining) during the conceptual stage of design.

Much attention has also been directed in recent years to developing design rules and computer methods for enhancing manufacturability through proper design. This approach, sometimes referred to as design for manufacturing (DFM), integrates product and process concepts to ensure ease of manufacture by matching product and process requirements. The DFM approach examines both product and process design in an integrated way. While DFM methods have not given much attention to the materials selection, the ideas of DFM are important to consider when selecting a material in relationship to possible manufacturing processes. The product (including material selection) and the process are designed together to ensure effective production. Typical objectives of DFM, as outlined by Bralla (Ref 8) include:

- Minimize total number of parts.
- Standardize design of products, subassemblies, components, modules, and individual parts.
- Use readily processed materials with two important caveats: (1) never compromise the quality of a part by selecting a more readily processed material and (2) it is the final part cost that counts, not the initial material cost.
- Fit the design to the manufacturing process. This argues for a process-first, materials-second approach to materials selection. For the many applications where performance is not paramount or difficult to achieve, this is the approach to use.
- Design each part to be easy to make. The inherent capabilities and constraints of each manufacturing process must be known and considered by the design team. With this knowledge, the designer should specify tolerances easily met by the process and try to minimize the need for fixturing and secondary operations.

Additional information on DFM can be found in Ref 9 to 11.

Materials Aspects

The selection of a material must be closely coupled with the selection of a manufacturing process. This is not an easy task, as there are many processes that can produce the same part. In a very general sense, the selection of the material determines a range of processes that can be used to process parts from the material. Table 4 shows the manufacturing methods used most frequently with different metals and plastics

Table 4 Compatibility between materials and manufacturing processes

Process	Cast iron	Carbon steel	Alloy steel	Stainless steel	Aluminum and aluminum alloys	Copper and copper alloys	Zinc and zinc alloys	Magnesium and magnesium alloys	Titanium and titanium alloys	Nickel and nickel alloys	Refractory metals	Thermoplastics	Thermoset plastics
Casting/molding													
Sand casting	•	•	•	•	•	•	—	•	—	•	—	X	X
Investment casting	—	•	•	•	•	•	—	—	—	•	—	X	X
Die casting	X	X	X	X	•	—	•	•	X	X	X	X	X
Injection molding	X	X	X	X	X	X	X	X	X	X	X	•	—
Structural foam molding	X	X	X	X	X	X	X	X	X	X	X	•	X
Blow molding (extrusion)	X	X	X	X	X	X	X	X	X	X	X	•	X
Blow molding (injection)	X	X	X	X	X	X	X	X	X	X	X	•	X
Rotational molding	X	X	X	X	X	X	X	X	X	X	X	•	X
Forging/bulk forming													
Impact extrusion	X	•	•	—	•	•	•	—	X	X	X	X	X
Cold heading	X	•	•	•	•	•	—	—	X	—	X	X	X
Closed-die forging	X	•	•	•	•	•	X	•	•	—	—	X	X
Pressing and sintering (P/M)	X	•	•	•	•	•	X	•	—	•	•	X	X
Hot extrusion	X	•	—	—	•	•	X	•	—	—	—	X	X
Rotary swaging	X	•	•	•	•	—	—	•	X	•	•	X	X
Machining													
Machining from stock	•	•	•	•	•	•	•	•	—	—	—	—	—
Electrochemical machining	•	•	•	•	—	—	—	—	•	•	—	X	X
Electrical discharge machining (EDM)	X	•	•	•	•	•	—	—	—	•	—	X	X
Wire EDM	X	•	•	•	•	•	—	—	—	•	—	•	X
Forming													
Sheet metal forming	X	•	•	•	•	•	—	—	—	—	X	X	X
Thermoforming	X	X	X	X	X	X	X	X	X	X	X	•	X
Metal spinning	X	•	—	•	•	•	•	—	—	—	—	X	X

•, Normal practice; —, less-common practice; X, not applicable. Source: Adapted from Ref 9

(Ref 9). The melting point of a material and level of deformation resistance (hardness) and ductility determine these relationships. Low-melting-point metals can be used with any of a large number of casting processes, but as the melting point of the material rises, the number of available processes becomes limited. Similarly, yield strength, or hardness, determines deformation limits, and forging and rolling loads are related to yield strength. Some materials are too brittle to be plastically deformed; others are too reactive to be cast or have poor weldability.

The ultimate criterion for selection of materials is the cost to produce a quality part. For selection of the best material for producing a part the following factors must be considered:

- Material composition: grade of alloy or plastic
- Form of material: bar, tube, wire, strip, plate, powder, and so forth
- Size: dimensions and tolerances
- Heat treated condition
- Directionality of mechanical properties (anisotropy)
- Surface finish
- Quality level: control of impurities, inclusions, and microstructure
- Quantity: volume of production (batch size)
- Ease of manufacture: workability, weldability, castability, and so forth
- Ease of recyclability
- Cost of material

Cost is the basis to which the other factors are reduced to reach a decision on materials selection. Cost of material may be a dominant factor, as the least expensive material consistent with design objectives and required properties is selected. However, as previously noted, the important objective is to minimize final part cost. A more expensive material may result in a less expensive product because it can employ a more economical manufacturing process.

The act of selecting a material is a form of engineering decision making. It is important to consider a spectrum of materials during conceptual stage of design and then narrow the list during embodiment or configuration design. The materials-selection decision is ultimately a trade-off between performance and cost. In many materials-selection problems, one property stands out as the most dominant service requirement. In this case, a reasonable selection criterion may be a method based on cost per unit property. Cost per unit property is a type of performance index for material cost. Processing factors also influence cost, and trade-offs may be needed in deciding whether to use material A with process B versus material C with manufacturing process D. For geometrically simple parts, such as straight shafts, the most economical raw material form and method of manufacture are readily apparent. As the shape of the part becomes more complex, the applicability of two or more forms and methods of manufacture add complexity to the selection process. The trade-offs depend on the materials-selection factors and processing factors.

Part Size, Shape, and Tolerances

Each process has associated with it a range of part shapes and sizes that can be produced. Thus, the first decision in selecting a process is its capability of producing parts of the required size and shape. The overall guide should be to select a primary process that makes the part as near to final shape as possible (near-net-shape forming) without requiring additional secondary processes such as machining or grinding. Sometimes the form of the starting material is important. For example, a hollow shaft can be made best by starting with a tube rather than a solid bar.

Maximum size often is controlled by equipment considerations. Overall size of the part is expressed by volume, projected area, or weight. For example, Fig. 2 plots the typical regions of part size and complexity for deformation processing, casting, and fabrication operations. In

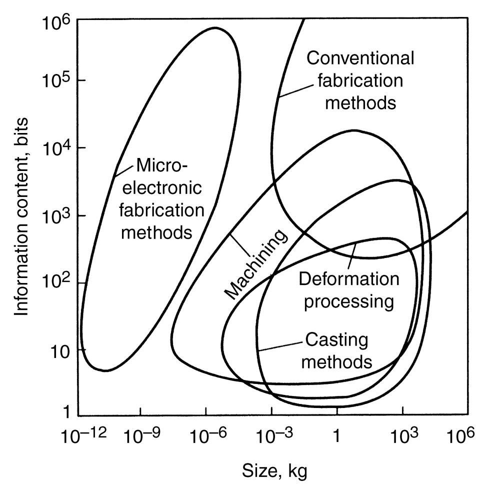


Fig. 2 Ashby process-selection chart for the relationship between part size and complexity of the shape. Source: Ref 6

this case, part complexity is expressed in terms of the information content of the part, that is, the number of independent dimensions that must be specified to describe the shape. Machining adds precision and allows production of parts with a greater range of complexity (Fig. 2). Simple shapes require only a few bits of information; for example, the casting for an engine block might have 10^3 bits of information, but after machining, the complexity increases both by adding new dimensional features and improving the precision (or dimensional tolerances).

Part Shape and Complexity. Shape is an essential feature of all manufactured parts; complexity of this shape often determines what processes can be considered for making it. In the most general sense, increasing complexity narrows the range of processes and increases cost. A cardinal rule of design is, therefore, to keep the shape as simple as possible. This rule may, however, be broken if a more complex shape allows consolidation of several parts and/or elimination of one or more manufacturing steps.

The classification of shapes is an active field of research in manufacturing, but there is no universally accepted system. In general, shape is often characterized by aspect ratio, the surface-to-volume ratio, or the web thickness-to-depth ratio. Overall complexity of a shape is influenced by a variety of factors such as:

- Requirements to achieve a minimum section thickness
- Lack of symmetry
- Surface details (e.g., the smallest radius of curvature at a corner is a surface detail that influences complexity)
- The presence of undercuts and reentrant angles
- The presence of internal cavities

Figure 3 (Ref 12) is one example of some basic shapes with increasing spatial complexity. Another scheme for shape classification is shown in Fig. 4, where the complexity or difficulty of making the shape increases down and to the right.

Group technology (GT) is a term that refers to the classification of shapes suitable for processing by the same techniques. Group technology uses a classification and coding system to identify and understand part similarities and to establish parameters for action. Because similar shapes tend to be produced by similar processing methods, the aim of GT is to classify similarity of parts for more effective design and manufacturing (Ref 13). Benefits include:

- Ability to quickly generate designs for similar-shaped parts
- Control of part proliferation and redundant designs
- Standardization of design features such as chamfers, corner radii, and tolerances
- Grouping of machine tools into more productive units
- Development of jigs and fixtures that can accommodate different members of a part family
- Development of computer-based process-selection methods

Manufacturing engineers use GT to decide on more efficient ways to increase system flexibility by streamlining information flow, reducing

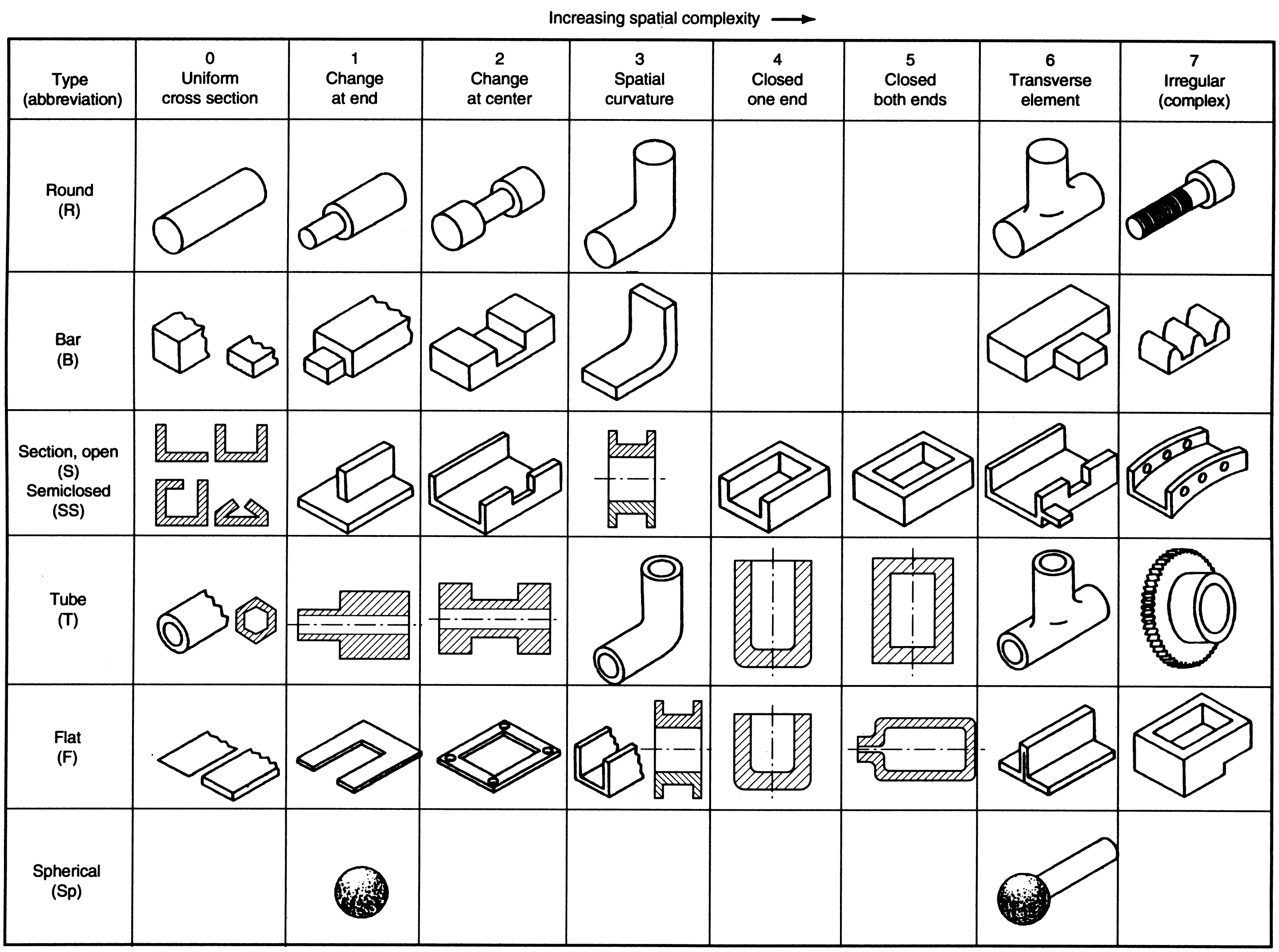


Fig. 3 The choice of possible manufacturing methods is aided by classifying shapes according to their geometric features. Source: Ref 12

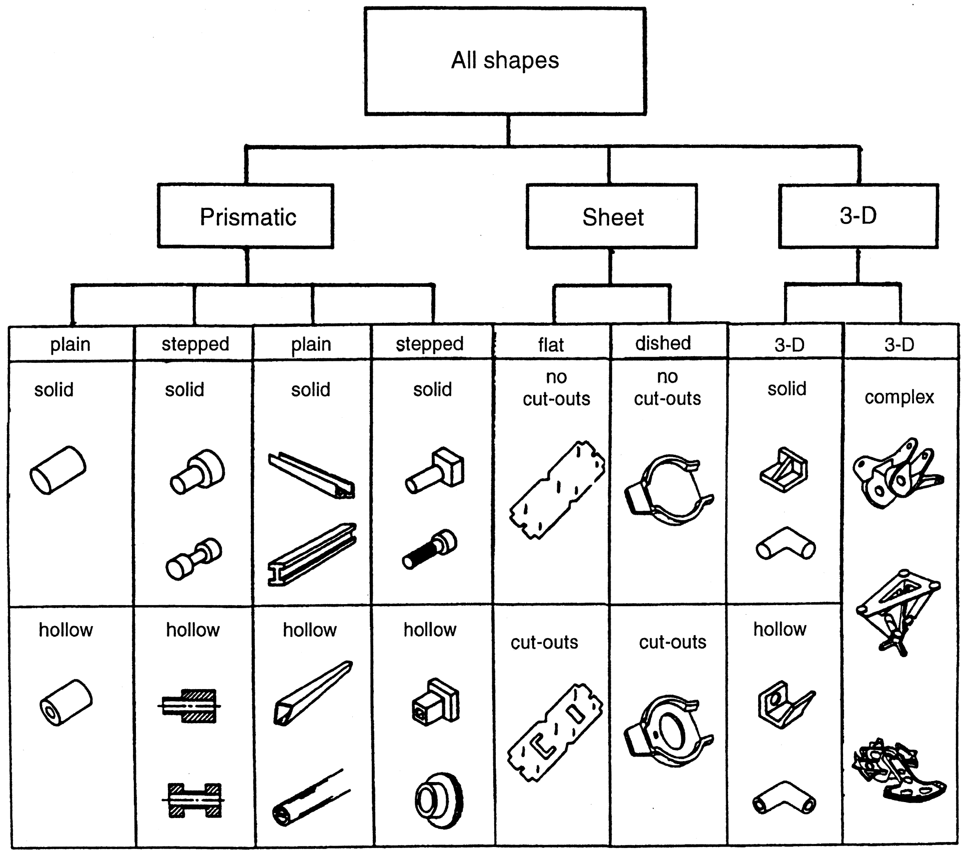


Fig. 4 Classification of shapes by complexity and process manufacturing difficulty where the complexity or difficulty of making the shape increases down and to the right. 3-D, three-dimensional. Source: Ref 6

setup time and floor space requirements, and standardizing procedures for batch-type production. Design engineers use GT to reduce design time and effort as well as part and tooling proliferation. With increasing emphasis on flexible and integrated manufacturing, GT is also an effective first step in structuring and building an integrated database. Standardized process planning, accurate cost estimation, efficient purchasing, and assessment of the impact of material costs are benefits that are often realized. A more complete discussion of GT, especially as it applies to manufacturing, is given in Ref 14.

The method of GT for part coding and classification began with metal-removal processes, but coding and classification schemes have been developed for metalforming operations (e.g., Ref 15). Coding details are based on part geometry, tolerance level, surface finish, material, heat treatment, and defect level. Hybrid component classification systems have also been developed where part families are defined by codes of both geometrical and technological factors of a part (Ref 16). This allows for more effective planning, estimating, and flexible manufacturing.

Dimensional Tolerances. Part complexity also depends on the dimensional tolerances and the roughness of the surface finish. Tolerance is the degree of deviation from ideal that is permitted in the dimensions of a part, and it is closely related to the surface finish. Each manufacturing process has the capability of producing a part with a certain range of tolerance and surface finish (Fig. 5). Every manufacturing process has standard or commercial tolerances that are generally well known to manufacturing process engineers, although standard tolerance values are not always explicitly defined in print for designers.

Manufacturing costs increase with tighter dimensional tolerances. The manufacturing tolerances are those that can be produced with the normal attention paid to process control and inspection. Part designs based on standard manufacturing tolerances will be the least expensive to produce. Moreover, the cost of manufacturing also depends on the number of tolerances that must be critically controlled. Controlling one or two critical dimensions, unless they are of an especially difficult type or extremely tight, is often relatively easy to do if the other dimensions of the part do not need special control.

At the configuration stage of design for parts, even before exact dimensions are assigned, designers can and should think ahead about both dimensional tolerances and geometric tolerances (e.g., flatness) that will be needed to achieve the desired functionality. Geometric tolerances refer to such issues as flatness, straightness, parallelism, perpendicularity, and so forth. These issues are discussed in other references such as ASME standard Y 14.5M and the *Handbook of Geometric Tolerancing* (Ref 17). In general, configuration design should seek to reduce the number of geometric and dimensional tolerances and the stringency or tightness of the tolerance requirements. By way of example, the following are some general tolerance guidelines for aluminum extrusion:

- Avoid tight straightness tolerances on long, thin-wall sections. Add small ribs to help achieve straightness.
- Angles of unsupported wall sections are difficult to control precisely. Hence, tolerances on size dimensions involving long legs are difficult to control. If necessary for tolerance control, support long legs with webs connecting to other parts of the extrusion. (This may create a more costly hollow shape, but if tolerances are critical, it may be necessary.)
- Using uniform wall thickness throughout the cross section makes all dimensions easier to control.
- Rounding inside corners makes angles easier to control than when corners are sharp.
- The dimensions of symmetrical shapes are generally easier to control than nonsymmetrical shapes.

In forgings, shrinkage and warpage contribute largely to the difficulties of controlling tolerances. Forged parts with simple shapes and common proportions can help minimize the effects of shrinkage and warpage. It is also important to provide for well-rounded corners, fillets, and edges. Thin sections, such as webs, should be kept to a minimum in number and should be made as thick as possible.

Process Factors

Key manufacturing process factors include:

- *Cycle time:* Once the process has been set up and is operating properly, cycle time is the time required to produce one unit. The inverse of cycle time is production rate.
- *Quality:* A widely inclusive term describing fitness for use. It may include dimensional tolerance, freedom from defects, and performance properties.
- *Flexibility:* The ease of adapting a process to produce different products or variations of the same product. Product customization is becoming more important, and so this characteristic has gained importance. Flexibility is influenced greatly by the time to change and set up tooling.
- *Materials utilization:* The amount of processed material used in the final product. Machining operations can generate 60 to 80% scrap. Net-shape forging and injection molding are at the other extreme. As materials costs become the greater part of the product cost, materials utilization becomes of greater importance.
- *Operating costs and equipment utilization:* Operating costs involve both the capital cost (plant infrastructure, machinery, and

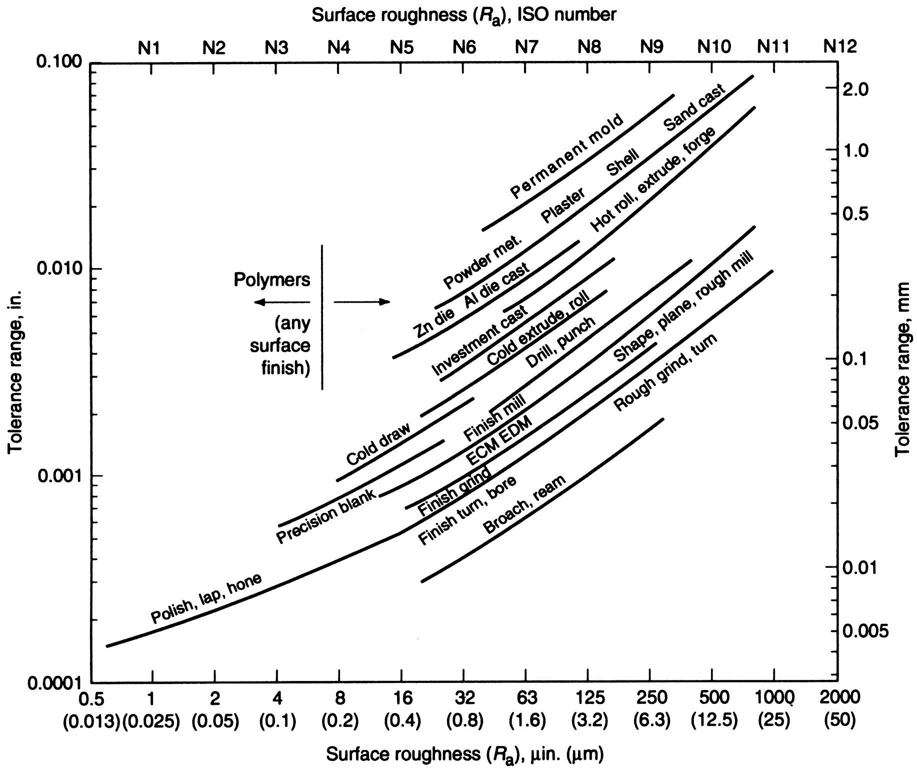


Fig. 5 Approximate values of surface roughness and tolerance on dimensions typically obtained with different manufacturing processes. ECM, electrochemical machining; EDM, electrical discharge machining. Polymers are different from metals and ceramics in that they can be processed to a very high surface smoothness, but tight tolerances are seldom possible because of internal stresses left by molding and creep at service temperatures. Source: Ref 12

tooling) and the labor costs of setting up and running the process. Process selection often is constrained by the available equipment, particularly if an alternate process is costly to install. In any manufacturing situation, one available decision is to outsource the production to a qualified subcontractor.

A rating system for evaluating these five process characteristics is given in Table 5. This is applied to rating the most common manufacturing processes in Table 6.

Production Volume. An important practical consideration is the quantity of parts required. For each process, there is a minimum batch size below which it is not economical to go because of costs of tooling, fixtures, and equipment. Thus, process choice is heavily influenced by the total number of parts to be made and by the required rate of production (i.e., the number of parts produced in a given time period). In general, a larger production quantity justifies greater investment in dies, equipment, and automation, whereas small quantities are often made with more labor input.

The total number of parts is frequently produced in lots. Lot size used to be determined to a large extent by total quantity, with lot size chosen to provide a supply for several days or weeks. Some standard components such as fasteners are still produced in large lots by hard automation. However, the spread of just-in-time delivery schedules and the increasing use of quick die-changing techniques and flexible automation have contributed to shrinking lot sizes even in mass production; the limit is reached when a single part constitutes a lot.

It is not feasible to state hard rules of what the economical production quantity and lot size is. Much depends on the equipment, degree of automation, and process control available in a given plant. Also related to part cost is the production rate or the cycle time, that is, the time required to produce one part. The most commonly used manufacturing processes are evaluated with respect to these characteristics in Table 7 (Ref 20). These relative values are meant to serve only as general guidelines.

Categories of Deformation Processes

Dieter (Ref 21) categorizes deformation processes into five broad classes:

- *Direct-compression processes:* Force is applied directly to the surface of the workpiece and material flow is normal to the application of the compressive force; examples are open-die forging and rolling.
- *Indirect-compression processes:* Deformation is imposed by compressive loads generated as the workpiece is pushed or pulled through a converging die. The direction of the external load applied to the workpiece is in the direction of workpiece motion. Examples include extrusion, wire drawing, and deep drawing.
- *Tension-based processes:* Tensile loading is developed in the workpiece to cause thinning, with stretch forming being a primary example.
- *Bending processes:* A bending moment is applied to cause a geometry change, the deformation being limited to the local region of the bend. Sheet bending, rod bending and coiling, and plate bending are example processes.
- *Shearing processes:* Metal deformation is highly localized in a workpiece as offset blades moving in opposite directions generate a plane of intense shear to intentionally cause a shear failure. Hole punching, plate shearing, blanking, and slitting are examples of shearing processes.

Other terminology is also recognized in the industry. Bulk forming processes are processes that have large volumes of material participating in the deformation and may be termed *three-dimensional processes.* The starting material in bulk deformation processes is a slab, ingot, billet, and so forth, produced by casting into stationary molds or by continuous-casting techniques. Primary deformation processes such as hot rolling, tube piercing, extrusion, and open-die forging are then used for converting the cast structure. The product may be suitable for immediate application, but in many cases it serves as the starting material for

Table 5 Scale for rating manufacturing processes

Rating	Cycle time	Quality	Flexibility	Materials utilization	Operating costs
1	>15 min	Poor quality, average reliability	Changeover extremely difficult	Waste > 100% of finished component	Substantial machine and tooling costs
2	5 to 15 min	Average quality	Slow changeover	Waste 50 to 100%	Tooling and machines costly
3	1 to 5 min	Average to good quality	Average changeover and setup time	Waste 10 to 50%	Tooling and machines relatively inexpensive
4	20 s to 1 min	Good to excellent quality	Fast changeover	Waste < 10% finished part	Tooling costs low/little equipment
5	<20 s	Excellent quality	No setup time	No appreciable waste	No setup costs

Rating scheme: 1, poorest; 5, best. Source: Ref 18

Table 6 Rating of characteristics for common manufacturing processes

Process	Shape	Cycle time	Flexibility	Material utilization	Quality	Equipment tooling costs
Casting						
Sand casting	3D	2	5	2	2	1
Evaporative foam	3D	1	5	2	2	4
Investment casting	3D	2	4	4	4	3
Permanent mold casting	3D	4	2	2	3	2
Pressure die casting	3D solid	5	1	4	2	1
Squeeze casting	3D	3	1	5	4	1
Centrifugal casting	3D hollow	2	3	5	3	3
Injection molding	3D	4	1	4	3	1
Reaction injection molding (RIM)	3D	3	2	4	2	2
Compression molding	3D	3	4	4	2	3
Rotational molding	3D hollow	2	4	5	2	4
Monomer casting, contact molding	3D	1	4	4	2	4
Forming						
Forging, open die	3D solid	2	4	3	2	2
Forging, hot closed die	3D solid	4	1	3	3	2
Sheet metal forming	3D	3	1	3	4	1
Rolling	2D	5	3	4	3	2
Extrusion	2D	5	3	4	3	2
Superplastic forming	3D	1	1	5	4	1
Thermoforming	3D	3	2	3	2	3
Blow molding	3D hollow	4	2	4	4	2
Pressing and sintering	3D solid	2	2	5	2	2
Isostatic pressing	3D	1	3	5	2	1
Slip casting	3D	1	5	5	2	4
Machining						
Single-point cutting	3D	2	5	1	5	5
Multiple-point cut	3D	3	5	1	5	4
Grinding	3D	2	5	1	5	4
Electrical discharge machining	3D	1	4	1	5	1
Joining						
Fusion welding	All	2	5	5	2	4
Brazing/soldering	All	2	5	5	3	4
Adhesive bonding	All	2	5	5	3	5
Fasteners	3D	4	5	4	4	5
Surface treatment						
Shot peening	All	2	5	5	4	5
Surface hardening	All	2	4	5	4	4
CVD/PVD	All	1	5	5	4	3

Rating scheme: 1, poorest; 5, best. Ratings from Ref. 19

another deformation process, the so-called secondary deformation processes, such as drawing and hot and cold forging.

Much primary working is conducted at elevated temperature, in the hot-working temperature range (typically above one-half the melting point on the absolute temperature scale) where dynamic recovery and recrystallization ensure low flow stress and high workability. Cold working produces better surface finish, can impart higher strength, and allows thinner products, but does so at the expense of higher forces. Warm working in an intermediate temperature range can combine some of the benefits of hot and cold working. Many processes are capable of producing net-shape or near-net-shape products, and Table 8 refers to these. Limits of size and web thickness are often set by the pressure increase due to friction at high width-to-thickness ratios (Fig. 6) (Ref 12).

Rolling is an important primary working process, practiced in highly specialized, high-productivity, capital-intensive plants. The products of flat hot rolling are relatively thick plate, sheet, and strip, typically down to 1.6 mm (0.060 in.). Cold rolling is capable of producing the thinnest of all metal products (strip and foil, to a few microns or a fraction of a thousandth of an inch) to very tight tolerances and controlled surface finish. Most of these products are further processed in sheet metalworking operations. Hot-rolled bar and wire serve as starting materials for other bulk deformation processes or machining. Shapes are mostly hot rolled or hot extruded for structural applications, although cold-rolled precision shapes fill an important niche. All these processes yield products of two-dimensional configuration, often in very long lengths, and are capable of producing high-quality (close-tolerance) products at very high rates and low cost. Closer to final application are three-dimensional products of hot and cold ring rolling (rotating rings of jet engines, ball-bearing races), forge rolling for complex shapes

Table 7 Manufacturing processes and their attributes

Process	Surface roughness	Dimensional accuracy	Complexity	Production rate	Production run	Relative cost	Size (projected area)
Pressure die casting	L	H	H	H/M	H	H	M/L
Centrifugal casting	M	M	M	L	M/L	H/M	H/M/L
Compression molding	L	H	M	H/M	H/M	H/M	H/M/L
Injection molding	L	H	H	H/M	H/M	H/M/L	M/L
Sand casting	H	M	M	L	H/M/L	H/M/L	H/M/L
Shell mold casting	L	H	H	H/M	H/M	H/M	M/L
Investment casting	L	H	H	L	H/M/L	H/M	M/L
Single point cutting	L	H	M	H/M/L	H/M/L	H/M/L	H/M/L
Milling	L	H	H	M/L	H/M/L	H/M/L	H/M/L
Grinding	L	H	M	L	M/L	H/M	M/L
Electrical discharge machining	L	H	H	L	L	H	M/L
Blow molding	M	M	M	H/M	H/M	H/M/L	M/L
Sheet metal working	L	H	H	H/M	H/M	H/M/L	L
Forging	M	M	M	H/M	H/M	H/M	H/M/L
Rolling	L	M	H	H	H	H/M	H/M
Extrusion	L	H	H	H/M	H/M	H/M	M/L
Powder metallurgy	L	H	H	H/M	H	H/M	L
Key:							
H	>250	<0.005	High	>100	>5000	High	>0.5
M	>63 and <250	>0.005 and <0.05	Medium	>10 and <100	>100 and <5000	Medium	>0.02 and <0.5
L	<63	>0.05	Low	<10	<100	Low	<0.02
Units	µin.	in		Parts/h	Parts		m^2

Source: Ref 20

Table 8 General characteristics of bulk deformation processes

Characteristic	Hot forging: Open die	Hot forging: Impression	Hot extrusion	Cold forging, extrusion	Shape drawing	Shape rolling	Transverse rolling
Part							
Material (wrought)	All	All	All	All	All	All	All
Shape(a)	RO–3; B; T1, 2; F0; Sp6	R; B; S; T1, 2, 4; (T6, 7); Sp	R; B; S ;SS; T1, 4; Sp	Same as hot	R0; B0; S0; T0	R0; B0; S0	R1–2; T1–2; Sp
Size, kg	0.1–200,000	0.01–100	1–500	0.001–50	10–1000	10–1000	0.001–10
Minimum section, mm	5	3	1	(0.005) 1	0.1	0.5	1
Minimum hole diameter, mm	(10) 20	10	20	(1) 5	0.1		
Surface detail(b)	E	C	B–C	A–B	A	A–B	A–C
Cost							
Equipment(b)	A–D	A–B	A–B	A–C	B–D	A–C	A–C
Die(b)	F	B–C	C–D	A–B	C–D	A–C	A–C
Labor(b)	A	B–D	B–C	C–E	C–E	C–E	C–E
Finishing(b)	A	B–C	C–D	D–E	E	E	D–E
Production							
Operator skill(b)	A	B–C	C–E	C–E	D–E	B	B–C
Lead time	Hours	Weeks	Days–weeks	Weeks	Days	Weeks	Weeks–months
Rates, pieces/machine	1–50 per h	10–300 per h	10–100 per h	100–10,000 per h	10–2000 m/min	20–500 m/min	100–1000 per h
Minimum quantity or length (m)	1	100–1000	1–10	1000–100,000	1000 m	50,000 m	1000–10,000

(a) From Fig. 3. (b) Comparative ratings with A indicating the highest value, E the lowest. Values in parenthesis may be obtained by special techniques. Source: Ref 12

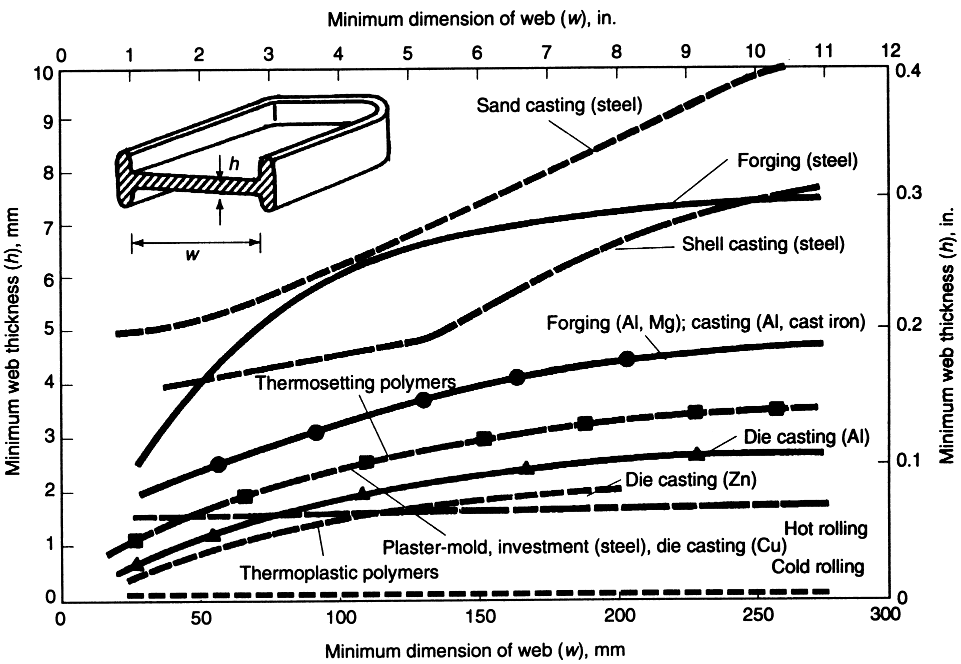


Fig. 6 For a given process, the minimum web thickness increases with the distance over which material must move. Source: Ref 12

(cutlery), gear rolling, and cross (transverse) rolling for shapes of axial symmetry (preforms for forging double-ended wrenches). Seamless tubes are made by hot piercing or hot extrusion and are further reduced by special rolling processes for immediate use (for example, in oil fields) and for further working, including cold reduction and drawing.

Drawing. A rolled preform is drawn through a converging gap, usually cold. Round bars and wires as well as two-dimensional sections of often complex cross-sectional shape are drawn through stationary dies or, less frequently, rollers. Tubes may be simply drawn through a die (tube sinking), but more often wall-thickness reduction is achieved by drawing over a plug or a bar (also called mandrel, from which the tube is released by cross rolling). Die costs are relatively low and production rates can be high, making for low-cost, high-precision products of wide size and wall-thickness range.

Extrusion. A billet held in a container is pushed against a die (direct or forward extrusion), or the die penetrates the billet (indirect or reverse extrusion), either hot or cold. Friction over the container and die surfaces has a decisive influence on material flow and die pressures.

Only lower-melting alloys, including aluminum alloys, can be extruded hot without a lubricant, and then deformation occurs with shearing the alloy at the die-to-billet interface. Two-dimensional shapes of great complexity and thin walls (e.g., architectural extrusions) can be made at a relatively low die cost. In the absence of a lubricant, divided material strands can be reunited and welded in a bridge-type die, allowing extrusions with one or more closed cavities, in a very wide size range (multihole tubes are extruded with walls as thin as 0.25 mm, or 0.010 in.). Hot extrusion of higher-melting alloys requires a high-temperature lubricant, often a glass; shapes are more limited and wall thickness is greater. The purpose may be that of producing a two-dimensional semimanufactured product or, if part of the billet is retained at the starting diameter, that of making near-net-shape three-dimensional products such as large valves and fasteners. The skill of the die designer is called upon to ensure a sound product. Extrusion is the preferred way for making bar stock from difficult-to-work high-temperature alloys.

Cold extrusion is always conducted with an effective lubricant (often with the lubricant superimposed on or reacted with a conversion coating on the metal) and is normally used to create near-net-shape products such as fasteners, automotive components, and so forth. Impact extrusion is a term employed for the cold extrusion of thin-walled products such as toothpaste tubes.

Forging. The aim in most forging operations is that of producing a part as close to the finished shape as possible. Open-die forging uses simple tools in a programmed sequence of basic operations (upsetting, drawing out), mostly in the hot-working temperature range, and the products (ranging from the one-off products of the blacksmith to huge turbine rotors) usually require finishing by machining. Rotary forging and swaging on special-purpose machines produce parts of axial symmetry to much tighter tolerances (axles, gun barrels).

Hot impression-die forging (sometimes termed *closed-die forging*) shapes the part between two die halves; thus productivity is increased, albeit at the expense of higher die costs. Excess metal is allowed to escape in the flash; thus pressure is kept within safe limits while die filling is ensured. More complex shapes, thinner walls, and thinner webs may necessitate forging in a sequence of die cavities, as for connecting rods and crankshafts. Die design calls for a thorough knowledge of material flow and is greatly aided by computer models and expert systems. At high width-to-thickness ratios, friction sets a limit to minimum web thickness (Fig. 6) that decreases with effective lubrication. In true closed-die forging the material is trapped in the die cavity.

With dies heated to or close to forging temperature (isothermal or hot-die forging), cooling is prevented and thin walls and webs can be produced, provided the die material is stronger than the workpiece material at the temperatures and strain rates ($\dot{\varepsilon}$) prevailing in the process. This is relatively easy for aluminum alloys (airframe parts); low press speeds (low $\dot{\varepsilon}$) help to keep stresses and forces low. Titanium alloys and superalloys can be forged in the superplastic state (jet-engine fan blades and turbine disks).

The sequence of operations can be carried out by moving the heated end of a bar through the die cavities in an upsetter, achieving high production rates. Mechanized transfer between cavities in conventional presses is also possible. In all impression-die forging, die design calls for considerable knowledge and die cost can be high (Table 8), but the product often has superior properties because material flow can be directed to give the best orientation of the structure relative to loading direction in the service of the part.

Cold forging is related to cold extrusion and, when a complex shape is to be formed in a single step, requires special lubricants, often with a conversion coating, as in making spark-plug bodies. Alternatively, the shape is developed by moving the bar or slug through a sequence of cavities, using a liquid lubricant. Cold forging is often combined with cold extrusion. It is the preferred process for mass producing near-net-shape parts such as bolts, nuts, rivets, and many automotive and appliance components.

Table 9 Values for the work-hardening exponent and strength coefficient for selected metals

Metal	Condition	Work-hardening exponent (n)	Strength coefficient (K) MPa	ksi
0.05% C steel	Annealed	0.26	531	77
4340 steel	Annealed	0.15	641	93
0.6% C steel	Quenched and tempered 540 °C (1000 °F)	0.10	1572	228
0.6% C steel	Quenched and tempered 700 °C (1300 °F)	0.19	1227	178
Copper	Annealed	0.54	317	46
70/30 brass	Annealed	0.49	896	130

Source: Ref 21

Cold Working

Cold working results in a deformed, unrecrystallized grain structure with the grains being elongated in the direction of metal flow. Deformation begins in the grain interiors when the critical resolved shear stress of the material is exceeded. At cold-working temperatures, the grain boundaries are more resistant to deformation, so workpieces with fine grains and a large amount of grain-boundary area are stronger than coarse-grained material of the same alloy. This dependence of yield strength on grain size and the amount of grain-boundary area in a material is captured by the Hall-Petch equation:

$$\sigma_y = \sigma_i + k/\sqrt{D} \qquad \text{(Eq 1)}$$

where σ_y is the yield stress of polycrystalline metal, σ_i is the stress related to resistance of dislocation motion within a grain, k is the parameter relating grain-boundary-hardening effect, and D is the average grain diameter. Equation 1 indicates that the yield strength of an alloy increases as the grain size becomes finer, where yield strength is the initial flow stress.

Once deformation is initiated, the moving dislocations interact with each other and with the grain boundaries to make continued yielding more difficult. This is work hardening, and a further feature of cold forming is that work-hardening effects continue to build with continued deformation. An empirical relationship for cold working between flow stress and plastic strain is:

$$\sigma_0 = K \cdot \varepsilon^n \qquad \text{(Eq 2)}$$

where σ_0 is the flow stress, K is the strength coefficient (stress when ε = 1.0), ε is the plastic strain, and n is the work-hardening exponent. From Eq 2, a high strength coefficient indicates a high initial resistance to plastic flow. Metals with a high K require large machines for deformation. Work hardening is a measure of how the resistance to plastic flow increases as the metal is deformed. Typically, n has values of 0.1 to 0.5 for cold working, with 0 being a perfectly plastic metal (no work hardening). A metal with a high work-hardening exponent but a low strength coefficient will achieve a high strength level after a large amount of deformation. Copper, brasses, and low-carbon steels are typical examples of metals that are cold worked to produce improved hardness and strength in the formed part. Table 9 contains some values of K and n for these metals (Ref 21). For steels, K increases with carbon content, while n generally decreases. Both copper and brass have a much higher work-hardening exponent than steel. Both K and n are affected not only by chemistry, but also by prior history and the microstructure. This is shown in Fig. 7 (Ref 22) for the work-hardening exponent for a variety of steels and microstructures.

Over the range of strain rates at which cold-deformation processes are conducted (0.1–100/s), the sensitivity to strain rate for most metals is low. Strain level rather than strain rate controls the flow stress, in addition to the initial strength coefficient K. Grain growth is not a factor in cold working. However, grain flow and a change in grain-aspect ratio is very much a factor. As the grains distort, a well-defined grain-flow pattern is developed due to grain-boundary alignment. Nonmetallic inclusions may also participate and further define a definite directionality in the microstructure and mechanical properties due to this mechanical fibering (see Chapter 2, "Bulk Workability of Metals," Fig. 22). Extremely deformed microstructures, as are present in cold-rolled sheet products, may also show alignment of crystallographic planes or texture, as well as grain-boundary alignment. The result is anisotropic behavior of the deformed material, either in service or in subsequent deformation steps. The designer must be

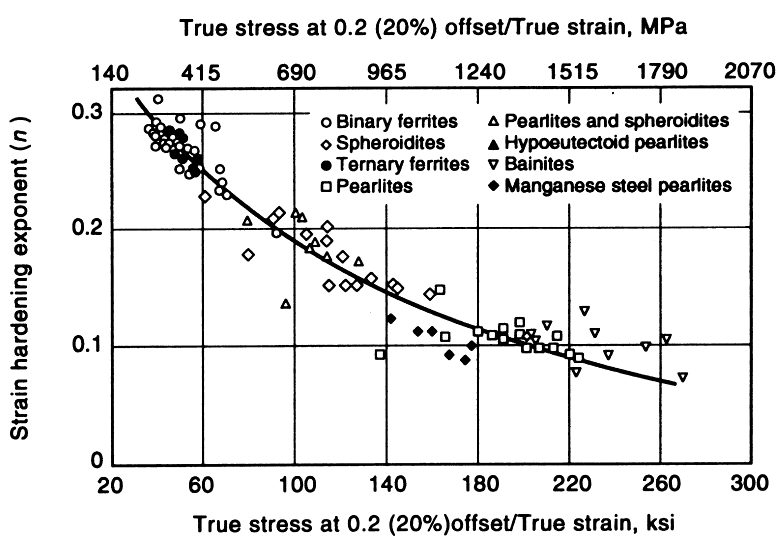


Fig. 7 Relationship between yield strength and the strain-hardening exponent (n) for a variety of steel microstructures. Source: Ref 22

aware of the effects of microstructural features such as fibering or preferred orientation on mechanical properties and the relationship of microstructural alignment and performance stresses. This is especially critical in cases where fatigue and fracture toughness are design issues (for more details, see Chapter 3, "Evolution of Microstructure during Hot Working").

Hot Working

Hot working takes place roughly above a homologous temperature of 0.5 T_M, with typical hot-working temperatures being 70 to 80% of the absolute melting temperature. At these temperatures, there is a high amount of internal energy available, and a number of deformation mechanisms, in addition to slip, are also available. These additional mechanisms include power-law creep mechanisms such as dislocation glide and climb and diffusional flow such as diffusion of vacancies and boundary motions. Hot working also allows static or dynamic recovery and recrystallization of the microstructure. Dynamic recovery and dynamic recrystallization occur during deformation, while static recovery and recrystallization occur after deformation while the workpiece is still hot. Dynamic recovery and recrystallization ensure low flow stress and high workability.

At high temperature, work hardening is low and the flow-stress curve becomes very different from that of cold deformation. Because the recovery processes take time, flow stress (σ_0) during hot working is a function of strain rate:

$$\sigma_0 = C\dot{\varepsilon}^m \qquad \text{(Eq 3)}$$

where C is the strength coefficient (decreasing with increasing temperature) and m is the strain-rate-sensitivity exponent. A high m value means that an incipient neck becomes stronger and spreads to neighboring material, allowing more deformation in tension. In some very fine-grained metals the value of m may reach 0.4 or 0.5, but only at very low strain rates and in a limited temperature range, and then superplastic deformation is possible to large strains and with low stresses.

The stacking-fault energy relates to the dislocation structure of the crystal. Low stacking-fault energy results in wide stacking faults that have a relatively high resistance to thermally activated mechanisms, and these metals strain harden rapidly. Metals with high stacking-fault energy have narrower stacking faults, the dislocations are more mobile, and as a result the rate of work hardening is low. Metals with low stacking-fault energies include brass and austenitic stainless steels, and metals having high stacking-fault energy include aluminum and nickel alloys. The effect of hot deformation on changes in grain structure depends on stacking-fault energies and the level of deformation (Fig. 8).

Dynamic recovery occurs in all cases with high or low stacking-fault energies for both moderate deformation (Fig. 8a) and high strain

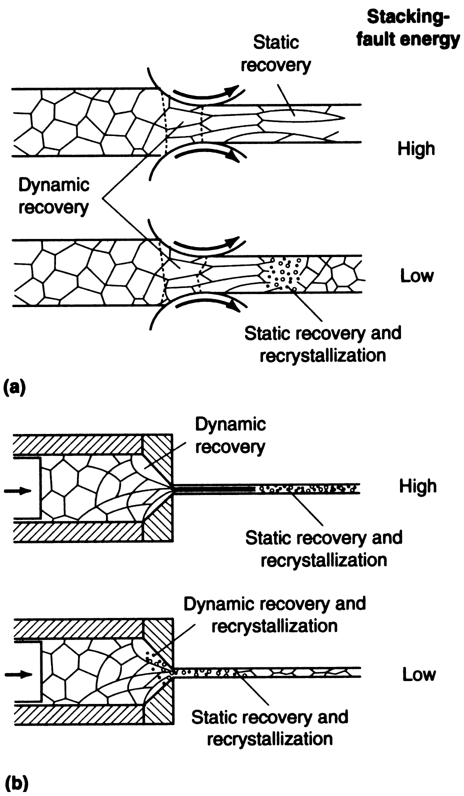


Fig. 8 Hot-working effects on microstructure. (a) Rolling with a thickness strain of 50%. (b) Extrusion with a strain of 99%. Source: after Ref 3

(Fig. 8b). Dynamic recovery occurs when there is sufficient atomic mobility to balance or nearly balance work hardening. That is, dislocations are sufficiently active to move in response to local stresses associated with dislocation tangles and forests, the presence of second phases, and other local stress concentrations. A metal that is undergoing dynamic recovery during hot working will exhibit negligible work hardening, with most low-carbon and low-alloy steels being primary examples. Figure 9(a) shows a flow stress-strain curve for a metal that dynamically recovers during hot working.

However, a high deformation level is required to produce recrystallization in metals with high stacking-fault energy, while metals with low stacking-fault energy can recrystallize at a lower level of deformation. Recrystallization occurs if a critical level of strain energy is achieved so that a new set of grains forms. If recrystallization occurs during hot deformation, the result is flow softening as shown in Fig. 9(b) (Ref 3). Examples of metals that may flow soften during hot working include nickel-aluminum bronze, commercial-purity titanium, super α_2 titanium aluminide, Ti-6Al-4V, and 7075 aluminum. Whether or not recrystallization and flow softening occur is a function of the working temperature and strain rate and the local deformation level. In practice, most conventional hot-working processes are too fast for dynamic recrystallization to occur. Flow softening can lead to strain localization and highly nonuniform mi-

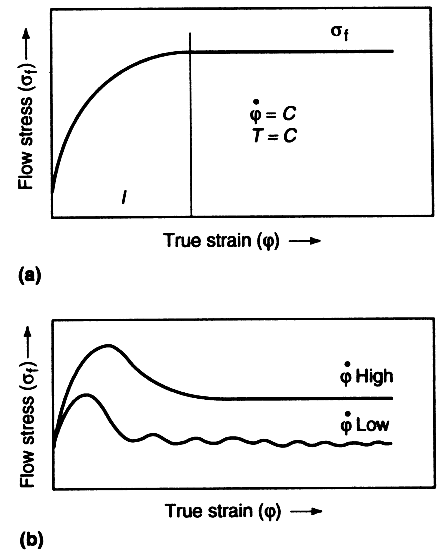


Fig. 9 Flow-stress curves representative of (a) dynamic recovery during hot working and (b) dynamic recovery and dynamic recrystallization. $\dot{\varphi}$, shear-strain rate; T, temperature; C, constant. Source: Ref 3

crostructures, and in extreme cases, shear cracking. This is mainly true in cases where this phenomenon occurs over a narrow range of temperatures or strain rates.

Recovery and recrystallization are thermally activated, and therefore these mechanisms are very dependent on temperature. They are also very dependent on the level of deformation because strain is indicative of a growing volume fraction of defects in the crystal structure and greater internal energy or stress within the grains. The benefit of recovery is that work-hardening effects can be minimized by allowing atomic rearrangement to reduce the internal stress within the grain. This is why hot-working processes can accomplish large deformation levels while maintaining relatively low working loads. For processes such as hot rolling, extrusion, and forging, the time in the deformation zone is short. Grain refinement is accomplished by static recrystallization after hot working. A high level of hot deformation followed by a hold time at an elevated temperature causes static recovery and recrystallization to result in a fine grain size. This may occur in hot rolling, where there is time between roll passes, or after hot forging where the workpiece slowly cools in a bin. Metallurgical specifications more frequently include grain size limits, and therefore it is becoming more critical to control workpiece temperature, deformation rate, the amount of deformation per working step, and the time between steps in order to control the microstructure of the deformed workpiece.

The other microstructural phenomenon that can occur during hot-working processes is grain growth. The natural drive for a polycrystalline material is to minimize internal energy, and because grain boundaries are regions of higher internal energy, grain growth is a way for nature to minimize energy by minimizing the grain-boundary content. Grain growth is also thermally driven, and because hot-working processes may hold a workpiece at a high temperature for a long time, grain growth can occur. In fact, in an extended hot-working process such as ingot breakdown rolling, a cyclic history of grain deformation, recrystallization, and growth is established for each deformation step. The ability to put work into the grains at a level sufficient to cause recrystallization is the reason that fine grains can be developed from a coarse-grained structure by hot working. Hot-working processes must balance recovery and recrystallization against grain growth in order to be effective in refining large-grained microstructures or in homogenizing microstructures of mixed grain sizes. For the designer, this is important because grain size has such a pronounced effect on mechanical performance of the part.

At temperatures above the equicohesive temperature, the grain interiors are more resistant to deformation than the grain boundaries, and the grain boundaries can sustain deformation. If a very fine grain size can be achieved and maintained during deformation at a low strain rate, grain-boundary sliding can occur. This mechanism, in combination with other thermally activated deformation mechanisms, is used successfully to deform very-fine-grained metals to large deformation levels, and this is termed superplastic behavior. Superplastic behavior is not a primary feature in most hot-working processes because of the required low deformation rates and extremely fine grain size. Creep forming, hot-die forging, isothermal forging and sizing, and isothermal rolling are processes that rely in part on grain-boundary sliding and other thermally activated deformation mechanisms.

Some general observations concerning recovery, recrystallization, and grain growth include:

- If the metal has been previously deformed, grain growth is preceded by recovery and by recrystallization.
- The driving force for recrystallization exceeds that of grain growth; thus recrystallization can occur at lower temperatures than grain growth.
- If the workpiece is stress free, grain growth begins without recovery or recrystallization occurring, but the temperature must be relatively high.
- The rate of grain growth is a function of the starting grain size, grain shape, and, most importantly, temperature. The driving force for grain growth is the minimization of internal energy, with the amount of grain-boundary area representing internal energy. Therefore, a fine-grained workpiece will experience a higher rate of grain growth than a coarse-grained workpiece will at the same temperature.
- As the workpiece is held at temperature, a saturation grain size will be reached, and holding the workpiece for a longer time will not result in further appreciable grain growth.
- Grain growth is retarded by the presence of stable second-phase particles, as these tend to hinder grain-boundary movement. Aluminum-killed steels have finer grain size than nonkilled steels because of the presence of AlN precipitates that lock up grain boundaries. Other common second phases that can be used to control grain size in iron and other metals include thermally and chemically stable carbides, nitrides, and oxides.

More details on the microstructural changes during hot working are described in Chapter 3, "Evolution of Microstructure during Hot Working."

Forgeability of Alloys

Workability is the ability to shape metals by bulk deformation, and various techniques are used to evaluate workability, as described in previous chapters of this book. For example, ductility as measured by the tensile test is a good way to compare the deformation resistance of different materials. However, ductility alone is insufficient in the evaluation of workability, because workability may involve several factors besides fracture susceptibility of a material. For example, workability may be defined in more general terms by a variety of factors, such as the generation of a rough surface finish or the inability to achieve a required tolerance on a critical dimension. From this practical point of view, a workability problem occurs whenever the produced part is unacceptable and must be scrapped or reworked.

Workability is a complex characteristic that depends on the specific stress-state conditions of a particular metalworking process. For example, the favorable compressive stresses generated during extrusion allow a high deformation capacity in comparison to other deformation processes (see Fig. 2 in Chapter 20, "Extrusion"). Therefore, it is possible to extrude some metals that can only be slightly deformed by other methods. Another comparison of workability for various processes is shown in Fig. 10 (Ref 23) for free-surface cracking that can occur as a workpiece expands due to a deformation process. The strain path that is experienced at a free surface is a function of friction and deformation-zone geometry, where a steeper strain path is observed for either higher friction or lower workpiece aspect ratio (which for upsetting is the height/diameter ratio of the workpiece). For forging processes, this would include cracking on the exposed surfaces during upsetting and cracking on surfaces at the leading edge of localized extrusion during forging. For rolling processes, this may be edge cracking of rolled slabs, plates, or rings. In these processes, the uniaxial compression test can be used to construct workability limits to evaluate materials for free-surface cracking (e.g., see Chapter 5 "Cold Upset Testing"). Localized stress and strain conditions on workability also are described in Chapter 12, "Workability Theory and Application in Bulk Forming Processes."

Different metals have different workability limits, and the limits change with chemistry, grain size, temperature, second-phase content, and possibly with strain rate. The remainder of this chapter briefly reviews the forgeability of common alloys. These brief summaries of alloy forgeability are only meant for the purpose of broad comparison. Typical forging temperatures and lubrication are also discussed. Lubrication variations cause local velocity gradients and secondary tensile stresses in the plane of the flowing surface. Even though there is compression normal to the workpiece surface, the local in-plane tension may be high enough to cause cracking. Once a small crack forms, the tension is relieved so the crack does not propagate. These cracks usually do not propagate deeply into the workpiece, but can result in unacceptable surface quality. Cracking on surfaces in contact with a die is a common problem, and it can lead to unacceptable parts in net-shape forging or unacceptable machining depths if that surface is to be finish machined.

Carbon and Alloy Steels

Carbon and alloy steels are by far the most commonly forged materials and are readily forged into a wide variety of shapes using hot-, warm-, or cold-forging processes and standard equipment. Despite the large number of available compositions, all of the materials in this category exhibit essentially similar forging characteristics. Exceptions to this are steels containing free-machining additives such as sulfides; these materials are more difficult to forge than are non-free-machining grades.

One common means of measuring the forgeability of steels is the hot-twist test. As the name implies, this test involves twisting of heated bar specimens to fracture at a number of different temperatures selected to cover the possible hot-working temperature range of the test material. The number of twists to fracture, as well as the torque required to maintain twisting at a constant rate, are reported. The temperature at which the number of twists is the greatest, if such a maximum exists, is assumed to be the optimal hot-working temperature of the test material.

Selection of forging temperatures for carbon and alloy steels is based on carbon content, alloy composition, the temperature range for optimal plasticity, and the amount of reduction required to forge the workpiece. Of these factors, carbon content has the most influence on upper-limit forging temperatures. Table 10 lists the typical hot-forging temperatures for a variety of carbon and alloy steels; it can be seen that, in general, forging temperatures decrease with increasing carbon and alloy content.

The hot forging of carbon and alloy steels into intricate shapes is rarely limited by forgeability aspects with the exception of the free-machining grades mentioned previously. Section thickness,

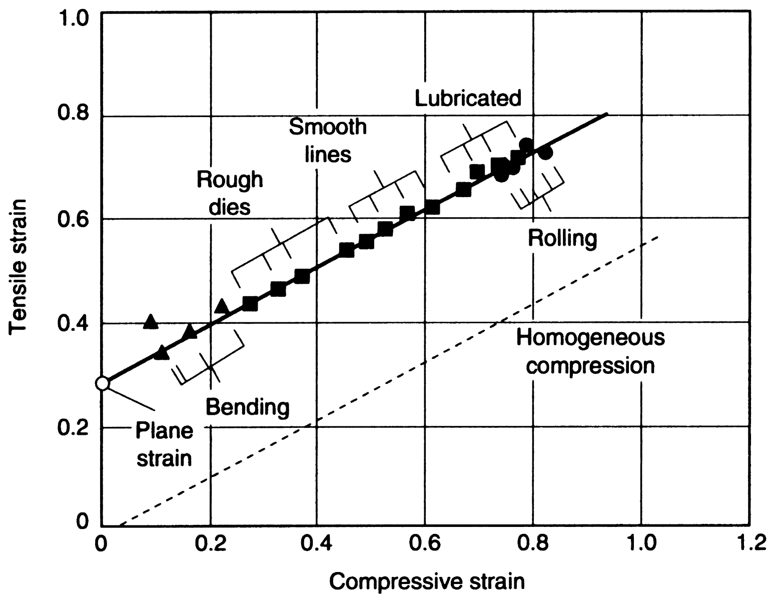


Fig. 10 Workability limit for upsetting, bending, and rolling with varying aspect ratios and friction conditions. Source: Ref 23

Table 10 Typical forging temperatures for various carbon and alloy steels

Steel	Major alloying elements	Typical forging temperature °C	°F
Carbon steels			
1010	...	1315	2400
1015	...	1315	2400
1020	...	1290	2350
1030	...	1290	2350
1040	...	1260	2300
1050	...	1260	2300
1060	...	1180	2160
1070	...	1150	2100
1080	...	1205	2200
1095	...	1175	2150
Alloy steels			
4130	Chromium, molybdenum	1205	2200
4140	Chromium, molybdenum	1230	2250
4320	Nickel, chromium, molybdenum	1230	2250
4340	Nickel, chromium, molybdenum	1290	2350
4615	Nickel, molybdenum	1205	2200
5160	Chromium	1205	2200
6150	Chromium, vanadium	1215	2220
8620	Nickel, chromium, molybdenum	1230	2250
9310	Nickel, chromium, molybdenum	1230	2250

Source: Ref 24

shape complexity, and forging size are limited primarily by the cooling that occurs when the heated workpiece comes into contact with the cold dies. For this reason, equipment that has relatively short die contact times, such as hammers, is often preferred for forging intricate shapes in steel. Generally, the hot forgeability of carbon and alloy steels improves as deformation rate increases. The improvement in workability has been primarily attributed to the increased heat of deformation generated at high deformation rates.

Forging Lubricants (Ref 25). For many years, oil-graphite mixtures were the most commonly used lubricants for forging carbon and alloy steels. Recent advances in lubricant technology, however, have resulted in new types of lubricants, including water/graphite mixtures and water-base synthetic lubricants. Each of the commonly used lubricants has advantages and limitations (Table 11) that must be balanced against process requirements.

Lubricant selection for forging is based on several factors, including forging temperature, die temperature, forging equipment, method of lubricant application, complexity of the part being forged, and environmental and safety considerations. At normal hot-forging temperatures for carbon and alloy steels, water-base graphite lubricants are used almost exclusively, although some hammer shops may still employ oil-based graphite.

The most common warm-forming temperature range for carbon and alloy steels is 540 to 820 °C (1000 to 1500 °F). Because of the severity of forging conditions at these temperatures, billet coatings are often used in conjunction with die lubricants. The billet coatings used include graphite in a fluid carrier or water-based coatings used in conjunction with phosphate conversion coating of the workpiece.

For still lower forging temperatures (less than about 400 °C, or 750 °F), molybdenum disulfide has a greater load-carrying capacity than does graphite. Molybdenum disulfide can either be applied in solid form or dispersed in a fluid carrier.

Microalloyed Steel Forgings. Microalloying—the use of small amounts of elements such as vanadium and niobium to strengthen steels—has been in practice since the 1960s to control the microstructure and properties of low-carbon steels. Most of the early developments were related to plate and sheet products in which microalloy precipitation, controlled rolling, and modern steelmaking technology combined to increase strength significantly relative to that of low-carbon steels.

The application of microalloying technology to forging steels has lagged behind that of flat-rolled products because of the different property requirements and thermomechanical processing of forging steels. Forging steels are commonly used in applications in which high strength, fatigue resistance, and wear resistance are re-

Table 11 Advantages and limitations of the principal lubricants used in the hot forging of steels

Type of lubricant	Advantages	Limitations
Water-base micro-graphite	Eliminates smoke and fire; provides die cooling; is easily extended with water	Must be applied by spraying for best results
Water-base synthetic	Eliminates smoke and fire; is cleaner than oils or water-base graphite; aids die cooling; is easily diluted, and needs no agitation after initial mixing; reduces clogging of spray equipment; does not transfer dark pigment to part	Must be sprayed; lacks the lubricity of graphite for severe forging operations
Oil-base graphite	Fluid film lends itself to either spray or swab application; has good performance over a wide temperature range (up to 540 °C, or 1000°F)	Generates smoke, fire, and noxious odors; explosive nature may shorten die life; has potentially serious health and safety implications for workers

Source: Ref 25

quired. These requirements are most often filled by medium-carbon steels. Thus, the development of microalloyed forging steels has centered around grades containing 0.30 to 0.50% C.

The metallurgical fundamentals described previously were first applied to forgings in the early 1970s. A West German composition, 49MnVS3 (nominal composition: 0.47C-0.20Si-0.75Mn-0.060S-0.10V), was used successfully for automotive connecting rods. The steel was typical of the first generation of microalloy steels, with a medium carbon content (0.35–0.50% C) and additional strengthening through vanadium carbonitride precipitation. The parts were subjected to accelerated air cooling directly from the forging temperature. The AISI grade 1541 microalloy steel with either niobium or vanadium has been used in the United States for similar automotive parts for many years.

The driving force behind the development of microalloyed forging steels has been the need to reduce manufacturing costs. The objective of microalloy forging steels is to obtain the enhanced mechanical properties of hot-formed steel parts while simultaneously eliminating the need for heat treating of the steel. Elimination of the heat treating operation reduces energy consumption and processing time as well as the material inventories resulting from intermediate processing steps. This is accomplished by means of a simplified thermomechanical treatment (that is, a controlled cooling following hot forging) that achieves the desired properties without the separate quenching and tempering treatments required by conventional carbon and alloy steels.

Recent advances in titanium-treated and direct-quenched microalloy steels provide new opportunities for the hot forger to produce tough, high-strength parts without special forging practices. Product evaluations of these microalloy steels indicate that they are comparable to conventional quenched-and-tempered steels. Warm forging continues to make steady progress as a cost-effective, precision-manufacturing technique because it significantly reduces machining costs. Microalloy steels austenitized at 1040 °C (1900 °F), cooled to a warm forging temperature of 925 °C (1700 °F), forged, and cooled by air or water (depending on composition), will produce a range of physical properties. The resulting cost savings has the potential to improve the competitive edge that forging has over other manufacturing techniques.

First-generation microalloy forging steels generally have ferrite-pearlite microstructures, tensile strengths above 760 MPa (110 ksi), and yield strengths in excess of 540 MPa (78 ksi). The room-temperature Charpy V-notch toughness of first-generation forgings is typically 7 to 14 J (5 to 10 ft · lbf), ambient. It became apparent that toughness would have to be significantly improved to realize the full potential of microalloy steel forgings.

Second-generation microalloy forging steels were introduced in the mid-1980s. These are typified by the West German grade 26MnSiVS7 (nominal composition: 0.26C-0.70Si-1.50Mn-0.040S-0.10V-0.02Ti). The carbon content of these steels was reduced to between 0.10 and 0.30%. They are produced with either a ferrite-pearlite microstructure or an acicular-ferrite structure. The latter results from the suppression of pearlite transformation products by an addition of about 0.10% Mo.

Titanium additions have also been made to these steels to improve impact toughness even further. Titanium-treated microalloy steels are currently in production in the United States, Germany, and Japan. The resistance to grain coarsening imparted by titanium nitride precipitation increases the toughness of the forgings.

One of the primary concerns of any steel user is the consistency of finished-part properties. Heat treatment has successfully addressed this concern, and a method must be found to ensure the consistency of finished-part properties for microalloy steels. One disadvantage of ferrite-pearlite microalloy steels is that the finished strength and hardness are functions of the cooling rate. Cooling rate can vary because of either process changes or part geometry. The ultimate strength of first- and second-generation microalloy steels is adequate for many engineering applications, but these steels do not achieve the toughness of conventional quenched-and-tempered alloys under normal hot-forging conditions.

Third-generation microalloy forging steels went into commercial production in the United States in 1989. This generation of microalloy forging steels has five to six times the toughness at −30 °C (−20 °F) and twice the yield strength of second-generation materials. No special forging practices are required except for the use of a water-cooling system.

These steels differ from their predecessors in that they are direct quenched from the forging temperature to produce microstructures of lath martensite with uniformly distributed temper carbides. Without subsequent heat treatment, these materials achieve properties, including toughness, similar to those of standard quenched-and-tempered steels. The metallurgical principles behind this development are based on:

- Niobium additions sufficient to exceed the solubility limit at the forging temperature, so that undissolved Nb(CN) retards the recrystallization and grain growth of austenite during forging, trimming, and entry into the quenchant
- Composition control to ensure that the martensite finish temperature is above 200 °C (400 °F)
- A fast cold-water quench performed on a moving conveyor through a spray chamber or by other appropriate equipment

The relatively high martensite finish temperature, combined with the mass effect of a forging, results in an autotempered microstructure with excellent toughness.

Stainless Steels

Forging Methods. Open-die, closed-die, upset and roll forging, and ring rolling are among the methods used to forge stainless steel. As in the forging of other metals, two of these methods are sometimes used in sequence to produce a desired shape. Roll forging can be used to forge specific products, such as tapered shafts. It is also used as a stock-gathering operation prior to forging in closed dies. Ring rolling is used to produce some ringlike parts from stainless steel at lower cost than by closed-die forging. The techniques used are essentially the same as those for the ring rolling of carbon or alloy steel.

Stainless steels, based on forging pressure and load requirements, are considerably more difficult to forge than carbon or low-alloy steels, primarily because of the greater strength of stainless steels at elevated temperatures and the limitations on the maximum temperatures at which stainless steels can be forged without incurring microstructural damage. Forging-load requirements and forgeability vary widely among stainless steels of different types and compositions; the most difficult alloys to forge are those with the greatest strength at elevated temperatures.

Open-die forging (hand forging) is often used for smaller quantities for which the cost of closed dies cannot be justified and in cases in which delivery requirements dictate shortened lead times. Generally, products include round bars, blanks, hubs, disks, thick-wall rings, and square or rectangular blocks or slabs in virtually all stainless grades. Forged stainless steel round bar can also be produced to close tolerances on radial forge machines. Although massive forgings are normally associated with open-die forging, most stainless steel open-die forgings are produced in the range of 10 to 900 kg (25 to 2000 lb).

Closed-die forging is extensively applied to stainless steel in order to produce blocker-type, conventional, and close-tolerance forgings. Selection from these closed-die types invariably depends on quantity and the cost of the finished part.

The relative forgeability characteristics of stainless steels can be most easily depicted through examples of closed-die forgings. Stainless steels of the 300 and 400 series can be forged into any of the hypothetical parts illustrated in Fig. 11. However, the forging of stainless steel into shapes equivalent to part 3 in severity may be prohibited by shortened die life (20 to 35% of that obtained in forging such a shape from carbon or low-alloy steel) and by the resulting high cost. For a given shape, die life is shorter in forging stainless steel than in forging carbon or low-alloy steel.

Forgings of mild severity, such as part 1 in Fig. 11, can be produced economically from any stainless steel with a single heating and about five blows. Forgings approximating the severity of part 2 can be produced from any stainless steel with a single heating and about ten blows. For any type of stainless steel, die life in the forging of part 1 will be about twice that in the forging of part 2.

Part 3 represents the maximum severity for forging all stainless steels and especially those with high strength at elevated temperature, namely, types 309, 310, 314, 316, 317, 321, and 347. Straight-chromium types 403, 405, 410, 416, 420, 430, 431, and 440 are the easiest to forge into a severe shape such as part 3 (although type 440, because of its high carbon content, would be the least practical). Types 201, 301, 302, 303, and 304 are intermediate between the two previous groups.

Upset forging is sometimes the only suitable forging process when a large amount of stock is needed in a specific location of the workpiece. For many applications, hot upset forging is used as a preforming operation to reduce the number of operations, to save metal, or both when the forgings are to be completed in closed dies.

The rules that apply to the hot upset forging of carbon and alloy steels are also applicable to stainless steel; that is, the unsupported length should never be more than two times the diameter (or, for a square, the distance across flats) for single-blow upsetting. Beyond this length, the unsupported stock may buckle or bend, forcing metal to one side and preventing the formation of a concentric forging. Exceeding this limitation also causes grain flow to be erratic and nonuniform around the axis of the forging and encourages splitting of the upset on its outside edges. The size of an upset produced in one blow also should not exceed two diameters (or, for a square, two times the distance across flats). This varies to some extent, depending on the thickness of the upset. For extremely thin upsets, the maximum size may be only two diameters, or even less.

Forgings of the severity represented by hypothetical parts 4, 5, and 6 in Fig. 12 can be hot upset in one blow from any stainless steel. However, the conditions are similar to those encountered in hot-die forging. First, with a stainless steel, die wear in the upsetting of part 6 will be several times as great as in the upsetting of part 4. Second, die wear for the forming of any shape will increase as the elevated-temperature strength of the alloy increases. Therefore, type 410, with about the lowest strength at high temperature, would be the most economical stainless steel for forming any of the parts, particularly part 6. Conversely, type 310 would be the least economical.

The austenitic stainless steels are more difficult to forge than the straight-chromium types, but are less susceptible to surface defects. Most of the austenitic stainless steels can be forged over a wide range of temperatures above 930 °C (1700 °F), and because they do not undergo major phase transformation at elevated temperature, they can be forged at higher temperatures than the martensitic types (Table 12). Exceptions to the above statements occur when the composition of the austenitic stainless steel promotes the formation of δ-ferrite, as in the case of the 309S, 310S, or 314 grades. At temperatures above 1100 °C (2000 °F), these steels, depending on their composition, may form appreciable amounts of δ-ferrite. Delta-ferrite formation adversely affects forgeability, and compensation for the amount of ferrite present can be accomplished with forging-temperature restrictions.

Equally important restrictions in forging the austenitic stainless steels apply to the finishing temperatures. All but the stabilized types (321, 347, 348) and the extra-low-carbon types should be finished at temperatures above the sensitizing range (~815 to 480 °C, or 1500 to 900 °F) and cooled rapidly from 870 °C (1600 °F) to a black heat. The highly alloyed grades, such as 309, 310, and 314, are also limited with regard to finishing temperature, because of their susceptibility at lower temperatures to hot tearing and σ formation. A final annealing by cooling rapidly from about 1065 °C (1950 °F) is generally advised for nonstabilized austenitic stainless steel forgings in order to retain the chromium carbides in solid solution.

Finishing temperatures for austenitic stainless steels become more critical where section sizes increase and ultrasonic testing requirements are specified. During ultrasonic examination, coarse-grain austenitic stainless steels frequently display sweep noise that can be excessive due to a coarse-grain microstructure. The degree of sound attenuation normally increases with section size and may become too great to permit detection of discontinuities. Careful control of forging conditions, including final forge reductions of at least 5%, can assist in the improvement of ultrasonic penetrability.

Martensitic stainless steels have high hardenability to the extent that they are generally air hardened. Therefore, precautions must be taken in cooling forgings of martensitic steels, especially those with high carbon content, in order to prevent cracking. The martensitic alloys are generally cooled slowly to about 590 °C (1100 °F), either by burying in an insulating medium or by temperature equalizing in a furnace. Direct water sprays, such as might be employed to cool dies, should be avoided, because they would cause cracking of the forging. Forgings of the martensitic steels are often tempered in order to soften them for machining. They are later quench hardened and tempered.

Maximum forging temperatures for these steels are low enough to avoid the formation of δ-ferrite. If δ-ferrite stringers are present at forging temperatures, cracking is likely to occur. Delta-ferrite usually forms at temperatures from 1095 to 1260 °C (2000 to 2300 °F). Care must be exercised so as not to exceed this temperature during forging and to avoid rapid metal movement that might result in local overheating. Surface decarburization, which promotes ferrite formation, must be minimized.

The δ-ferrite formation temperature decreases with increasing chromium content, and small amounts of δ-ferrite reduce forgeability significantly. As the δ-ferrite increases above about 15%, forgeability improves gradually until the structure becomes entirely ferritic. Finishing temperatures are limited by the allotropic transformation, which begins near 815 °C (1500 °F). However, forging of these steels is usually stopped at about 925 °C (1700 °F), because the metal is difficult to deform at lower temperatures.

Sulfur or selenium can be added to type 410 to improve machinability. These elements can cause forging problems, particularly when they form surface stringers that open and form cracks. This can sometimes be overcome by adjusting the forging temperature or the procedure. With sulfur additions, it may be impossible to eliminate all cracking of this type. Therefore, selenium additions are preferred.

Ferritic Stainless Steels. The ferritic straight-chromium stainless steels exhibit virtually no increase in hardness upon quenching. They will work harden during forging; the degree of work hardening depends on the temperature and the amount of metal flow. Cooling from the forging temperature is not critical.

The ferritic stainless steels have a broad range of forgeability, which is restricted somewhat at higher temperature because of grain growth and structural weakness but is closely restricted in finishing temperature only for type 405. Type 405 requires special consideration because of the grain-boundary weakness resulting from the development of a small amount of austenite. The other ferritic stainless steels are commonly finished at any temperature down to 705 °C (1300

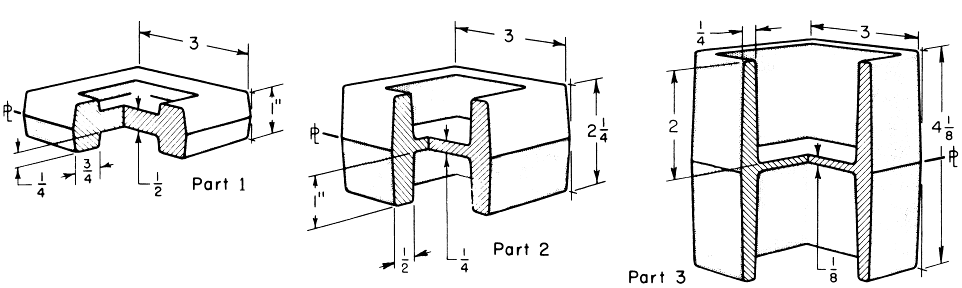


Fig. 11 Three degrees of forging severity. Dimensions are given in inches.

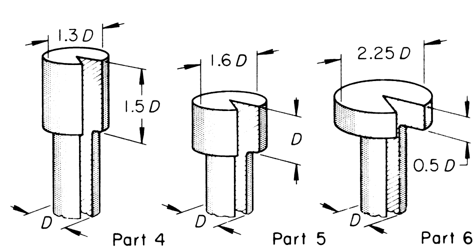


Fig. 12 Three degrees of upsetting severity

Table 12 Typical compositions and forging temperature ranges of high-temperature alloys

	Typical composition, %						Temperature	
Alloy	C	Cr	Ni	Mo	Co	Other	°C	°F
More difficult to hot work								
Carpenter 41	0.09	19.0	Bal	10.0	11.0	3.1 Ti, 1.5 Al, 0.005 B	1040–1175	1900–2145
Pyromet 718	0.10	18.0	55.0	3.0	. . .	1.3 Ti, 0.6 Al, 5.0 Nb	925–1120	1700–2050
M252	0.15	18.0	38.0	3.2	20.0	2.8 Ti, 0.2 Al	980–1175	1800–2145
Waspaloy	0.07	19.8	Bal	4.5	13.5	3.0 T1, 1.4 Al, 0.005 B	1010–1175	1850–2145
Pyromet 860	0.1	14.0	45.0	6.0	4.0	3.0 Ti, 1.3 Al, 0.01 B	1010–1120	1850–2050
Carpenter 901	0.05	12.5	42.5	6.0	. . .	2.7 Ti, 0.2 Al, 0.015 B	1010–1120	1850–2050
N155	0.12	21.0	20.0	3.0	19.5	2.4 W, 1.2 Nb, 0.13 N	1040–1150	1900–2100
V57	0.05	15.0	27.0	1.3	. . .	3.0 Ti, 0.2 Al, 0.01 B, 0.3 V	955–1095	1750–2000
A-286	0.05	15.0	25.0	1.3	. . .	2.1 Ti, 0.2 Al, 0.004 B, 0.3 V	925–1120	1700–2050
Carpenter 20Cb-3	0.05	20.0	34.0	2.5	. . .	3.5 Cu	980–1230	1800–2245
Pyromet 355	0.12	15.5	4.5	3.0	. . .	0.10 N	925–1150	1700–2100
Type 440F	1.0	17.0	. . .	0.5	. . .	0.15 Se	925–1150	1700–2100
Type 440C	1.0	17.0	. . .	0.5	. . .	. . .	925–1150	1700–2100
19-9DL/19DX	0.32	18.5	9.0	1.5	. . .	1.4 W plus Nb or Ti	870–1150	1600–2100
Types 347 and 348	0.05	18.0	11.0	. . .	. . .	0.07 Nb	925–1230	1700–2245
Type 321	0.05	18.0	10.0	. . .	. . .	0.40 Ti	925–1260	1700–2300
AMS 5700	0.45	14.0	14.0	. . .	. . .	2.5 W	870–1120	1600–2050
Type 440B	0.85	17.0	. . .	0.5	. . .	. . .	925–1175	1700–2145
Type 440A	0.70	17.0	. . .	0.5	. . .	. . .	925–1200	1700–2200
Type 310	0.15	25.0	20.0	. . .	. . .	. . .	980–1175	1800–2145
Type 310S	0.05	25.0	20.0	. . .	. . .	. . .	980–1175	1800–2145
17-4 PH	0.07	17.0	4.0	. . .	. . .	3.0–3.5 Cu, 0.3 Nb + Ta	1095–1175	2000–2145
15-5 PH	0.07	15.0	5.0	. . .	. . .	3.5 Cu, 0.3 Nb + Ta	1095–1175	2000–2145
13-8 Mo	0.05	13.0	8	2.25	. . .	0.90–1.35 Al	1095–1175	2000–2145
Type 317	0.05	19.0	13.0	3.5	. . .	. . .	925–1260	1700–2300
Type 316L	0.02	17.0	12.0	2.5	. . .	. . .	925–1260	1700–2300
Type 316	0.05	17.0	12.0	2.5	. . .	. . .	925–1260	1700–2300
Type 309S	0.05	23.0	14.0	. . .	. . .	. . .	980–1175	1800–2145
Type 309	0.10	23.0	14.0	. . .	. . .	. . .	980–1175	1800–2145
Type 303	0.08	18.0	9.0	. . .	. . .	0.30 S	925–1260	1700–2300
Type 303Se	0.08	18.0	9.0	. . .	. . .	0.30 Se	925–1260	1700–2300
Type 305	0.05	18.0	12.0	. . .	. . .	. . .	925–1260	1700–2300
Easier to hot work								
Types 302 and 304	0.05	18.0	9.0	. . .	. . .	. . .	925–1260	1700–2300
UNS S21800	0.06	17	8.5	. . .	. . .	8.0 Mn, 0.12 N	1095–1175	2000–2145
No. 10	0.05	16.0	18.0	. . .	. . .	. . .	925–1230	1700–2245
Lapelloy	0.30	11.5	0.30	2.8	. . .	0.3 V	1040–1150	1900–2100
Lapelloy C	0.20	11.5	0.40	2.8	. . .	2.0 Cu, 0.08 N	1040–1150	1900–2100
636	0.23	12.0	0.8	1.0	. . .	0.3 V, 1.0 W	1040–1175	1900–2145
H46	0.17	12.0	0.5	0.8	. . .	0.4 Nb, 0.07 N, 0.3 V	1010–1175	1850–2145
AMS 5616 (Greek Ascoloy)	0.17	13.0	2.0	0.2	. . .	3.0 W	955–1175	1750–2145
Type 431	0.16	16.0	2.0	. . .	. . .	. . .	900–1200	1650–2200
Type 414	0.12	12.5	1.8	. . .	. . .	. . .	900–1200	1650–2200
Type 420F	0.35	13.0	. . .	. . .	. . .	0.2 S	900–1200	1650–2200
Type 420	0.35	13.0	. . .	. . .	. . .	. . .	900–1200	1650–2200
Pyromet 600	0.08	16.0	74.0	. . .	. . .	8.0 Fe	870–1150	1600–2100
Type 416	0.1	13.0	. . .	. . .	. . .	0.3 S	925–1230	1700–2245
Type 410	0.1	12.5	. . .	. . .	. . .	. . .	900–1200	1650–2200
Type 404	0.04	11.5	1.8	. . .	. . .	. . .	900–1150	1650–2100
Type 501	0.2	5.0	. . .	0.5	. . .	. . .	980–1200	1800–2200
Type 502	0.05	5.0	. . .	0.5	. . .	. . .	980–1200	1800–2200
HiMark 300	0.02	. . .	18.0	4.8	9.0	0.7 Ti, 0.1 Al	815–1260	1500–2300
HiMark 250	0.02	. . .	18.0	4.8	7.5	0.4 Ti, 0.1 Al	815–1260	1500–2300
Carpenter 7-Mo (Type 329)	0.08	28.0	5.8	1.6	. . .	. . .	925–1095	1700–2000
Type 446	0.1	25.0	. . .	. . .	. . .	. . .	900–1120	1650–2050
Type 443	0.1	21.0	. . .	. . .	. . .	1.0 Cu	900–1120	1650–2050
Type 430F	0.08	17.0	. . .	. . .	. . .	0.3 S	815–1150	1500–2100
Type 430	0.06	17.0	. . .	. . .	. . .	. . .	815–1120	1500–2050

Source: Ref 26

°F). For type 446, the final 10% reduction should be made below 870 °C (1600 °F) to achieve grain refinement and room-temperature ductility. Annealing after forging is recommended for ferritic steels.

Precipitation-Hardening Stainless Steels. The semiaustenitic and martensitic precipitation-hardening stainless steels can be heat treated to high hardness through a combination of martensite transformation and precipitation. They are the most difficult to forge and will crack if temperature schedules are not accurately maintained. The forging range is narrow, and the steel must be reheated if the temperature falls below 980 °C (1800 °F). They have the least plasticity (greatest stiffness) at forging temperature of any of the classes and are subject to grain growth and δ-ferrite formation. Heavier equipment and a greater number of blows are required to achieve metal flow equivalent to that of the other types.

During trimming, the forgings must be kept hot enough to prevent the formation of flash-line cracks. To prevent these cracks, it is often necessary to reheat the forgings slightly between the finish-forging and trimming operations. Cooling, especially the cooling of the martensitic grades, must be controlled to avoid cracking.

Die Lubrication. Dies should be lubricated before each blow. For forging in shallow impressions, a spray of colloidal graphite in kerosene or in low-viscosity mineral oil is usually adequate. Ordinarily, dies are sprayed manually, but in press forging, automatic sprays timed with the press stroke are sometimes used. For deeper cavities, however, it is often necessary to use a supplemental spray (usually manual) to reach the deep areas of the cavity or to swab the cavity with a conventional forging oil. Forging oils are usually mixtures of oil and graphite; the oil should be free of lead and sulfur. Forging oils are often purchased as greases and are then diluted with mineral oil to the desired viscosity. Any volatile lubricant should be used sparingly. With even a slight excess, vapor explosions are likely, and greater amounts can cause explosions that will eject the workpiece, possibly injuring personnel.

Glass is sometimes used as a lubricant or billet coating in press forging. The glass is applied by dipping the heated forging in molten glass or by sprinkling the forging with glass frit. Glass is an excellent lubricant, but its viscosity must be compatible with the forging temperature used. For optimal results, the viscosity of the glass should be maintained at 10 Pa · s (100 cP). Therefore, when different forging temperatures are used, a variety of glass compositions must be stocked. Another disadvantage of glass is that it will accumulate in deep cavities, solidify, and impair metal flow. Therefore, the use of glass is generally confined to shallow forgings that require maximum lateral flow.

Superalloys

Because superalloys are designed to resist deformation at high temperatures, it is not surprising that they are very difficult to hot work; ductility is limited, and the flow stress is high. Further, any alloying addition that improves the service qualities usually decreases workability. Moreover, machining of heat-resistant alloys is difficult and expensive and can sometimes amount to 40% of the cost of production. Therefore, the use of computer-aided design and modeling are particularly significant in the forging of heat-resistant alloys, because of the premium placed on near-net-shape manufacturing. In addition, the thermomechanical conditions of forging are often designed to impart desired properties and controlled microstructures. The complexity of these demands makes computers very relevant in analyzing and simulating the forging of heat-resistant alloys.

Methods (adapted from Ref 27). Forged superalloy components are produced by:

- Die forging
- Upsetting
- Extrusion forging
- Roll forging
- Swaging (or versions using proprietary rotary forging machines)
- Ring rolling
- Two or more of these methods used in sequence

The die-forging categories can be further subdivided as:

- Open-die forging
- Open-die forgings formed with the aid of plugs and rings to impart certain shapes
- Closed-die blocker-type forgings
- Closed-die finish forgings

Which type of method is used depends on complexity of shape and tolerances required. For example, closed-die finish forgings have much thinner ribs and webs, tighter radii, and closer tolerances than blocker forgings (see Table 13 for some conventional superalloys). Closed-die forging is widely used for forging heat-resistant alloys. The procedures are generally different from those used for similar shapes from carbon or low-alloy steels in that preforms (produced by open-die forging, upsetting, rolling, or extrusion) are used to a greater extent for the closed-die forging of heat-resistant alloys than for steel.

Regardless of the method used, the forging of superalloys generally should be done as part of total thermomechanical processing. In other words, shaping should not be the only factor in forging. Work energy can be introduced and managed via temperature and deformation controls to impart the most useful or desired design qualities in a component. This discussion assumes that forging intends to create both shape and properties.

In some cases, forgings are deliberately processed for better tensile properties, stress-rupture behavior, creep strength, or low-cycle fatigue life. Therefore, the objectives for the forging cycle may be:

- Uniform grain refinement
- Control of second-phase morphology
- Controlled grain flow
- Structurally sound components

Fine grain size is not necessarily a desired outcome for all alloys. The objectives are to optimize properties (as defined by the component specification) in all sections of the component. This may require generating a grain size that is within a defined size range. The soundness and uniformity of the forging also must be ensured. Most forgings are inspected by ultrasonic testing, macroetch, and mechanical test of integral coupons. In order to impart optimal work during each stage, it may even be necessary to include redundant work if work penetration in the subsequent processing sequence is not likely to be uniform. The forging process today does not operate as a stand-alone function as it did at the start of the superalloy age.

Controlling the deformation process. Recrystallization must be achieved in each operation to obtain the desired grain size and flow characteristics in a forged superalloy. Recrystallization also helps to eliminate the grain- and twin-boundary carbides that tend to develop during static heating or cooling. Nonuniform distribution of inhomogeneities will likely lead to problems. Up to 80% of metal reduction is accompanied by recrystallization, usually completed over falling temperatures; the remaining 20% can be as warm work at lower temperatures for additional strengthening. However, the range of applications for superalloy forgings is diverse, and in some circumstances, the aim of the forging process may be to produce a duplex, not a single, grain size in the finished component.

During the latter quarter of the 20th Century, a trend developed to lower the strain rate and to heat the dies. Faster strain rates lead to frictional heat buildup, nonuniform recrystallization, and metallurgical instabilities, and they are also likely to cause radial-type ruptures, especially in high γ' alloys. Superalloys can be forged by a variety of methods, and two or more of these methods are often used in sequence. A particular outcome of lower strain rate was the introduction of isothermal superplastic forging/forming or, at least, isothermal forging (see the discussion of "Isothermal Forging" in the section on "Nickel-Base Alloys" in this article).

Die heating. Dies are always heated for the forging of heat-resistant alloys. The heating is usually done with various types of burners, although embedded elements are sometimes used. Optimal die temperature for conventional hot forging varies from 150 to 260 °C (300 to 500 °F); the lubricant used is an important limitation on maximum die temperature. Die temperature is controlled by the use of temperature-sensitive crayons or surface pyrometers.

Lubrication. Dies should be lubricated before each forging. For shallow impressions, a spray of colloidal graphite in water or in mineral oil is usually adequate. Dies are usually sprayed manually, although some installations include automatic sprays that are timed with the press stroke. Deeper cavities, however, may require the use of a supplemental spray (usually manually controlled) to ensure coverage of all surfaces, or they can be swabbed with a conventional forging oil. These oils are readily available as proprietary compounds.

Cooling. Specific cooling procedures are rarely, if ever, needed after the forging of heat-resistant alloys. If forging temperatures are correctly maintained, the forgings can be cooled in still air, after which they will be in suitable condition for heat treating.

Forgeability Ratings. Table 14 lists the most commonly forged heat-resistant alloys, and their forging temperatures and forgeability ratings. Generally, these alloys can be grouped into two categories: solid-solution-strengthened alloys and precipitation strengthened alloys. The latter group is much more difficult to forge than the former. These alloys are usually worked with the precipitates dissolved, but the higher concentration of dissolved alloying elements (40–50% total) gives rise to higher flow stress, higher recrystallization temperature, and lower solidus temperature, thus narrowing the useful temperature range for hot forming.

Iron-base superalloys evolved from austenitic stainless steels and are based on the principle of combining a closed-packed face-centered cubic (fcc) matrix with (in most cases) both solid-solution hardening and precipitate-forming elements. The austenitic matrix is based on nickel and iron, with at least 25% Ni needed to stabilize the fcc phase. Other alloying elements, such as chromium, partition primarily to the austenite for solid-solution hardening. The strengthening precipitates are primarily ordered intermetallics, such as Ni_3Al (or γ') Ni_3Ti (or η), and Ni_3Nb (or γ''), although carbides and carbonitrides may also be present. Elements that partition to grain boundaries, such as boron and zirconium, perform a function similar to that which occurs in nickel-

Table 13 Design guides for some conventional superalloy forgings

Alloy	Type of forging	Web thickness, min mm	Web thickness, min in.	Rib width, min. mm	Rib width, min. in.	Thickness tolerance mm	Thickness tolerance in.	Corner radii, min. mm	Corner radii, min. in.	Fillet radii, min. mm	Fillet radii, min. in.
A-286, Inco 901, Hastelloy X, Waspaloy, Udimet 630, TD-nickel(a)	Blocker	19.1–31.8	0.75–1.25	19.1–25.4	0.75–1.00	4.6–6.4	0.18–0.25	15.8	0.62	19.1–31.8	0.75–1.25
	Finish	12.7–25.4	0.50–1.00	15.8–19.8	0.62–0.78	3.0–4.6	0.12–0.18	12.7	0.50	15.8–25.4	0.62–1.00
Inco 718, René 41, X-1900(a)	Blocker	25.4–38.1	1.00–1.50	25.4–31.8	1.00–1.25	5.1–6.4	0.20–0.25	19.1	0.75	25.4–50.8	1.00–2.00
	Finish	19.1–31.8	0.75–1.25	19.8–25.4	0.78–1.00	3.8–5.1	0.15–0.20	15.8	0.62	19.1–38.1	0.75–1.50
Astrology, B-1900(a)	Blocker	38.1–63.5	1.50–2.50	31.8–38.1	1.25–1.50	6.4–7.6	0.25–0.30	25.4	1.00	31.8–63.5	1.25–2.50
	Finish	25.4–38.1	1.00–1.50	25.4–31.8	1.00–1.25	4.6–6.4	0.18–0.25	19.1	0.75	25.4–50.8	1.00–2.00

Note: For forgings over 258,064 mm^2 (400 in.^{2}) in plan area. For forgings of 64,516–258,064 mm^2 (100–400 in.^{2}) plan area, design allowables can be reduced 25%. For forgings under 64,516 mm^2 (100 in.^{2}), design allowables can be reduced 50%. Recommended draft angles are 5–7°. Machining allowance for finish forgings is 3.81–6.35 mm (0.15–0.25 in.). Some shapes can require higher minimum allowables than shown above. (a) Based on limited data. Source: Ref 27 base alloys; that is, grain-boundary fracture is suppressed under creep-rupture conditions, resulting in significant increases in rupture life.

The inclusion content of the alloys has a significant effect on their forgeability. Alloys containing titanium and aluminum can develop nitride and carbonitride segregation, which later appears as stringers in wrought bars and affects forgeability. This type of segregation has been almost completely eliminated through the use of vacuum melting. Precipitation-hardening iron-base alloys are electric-furnace or vacuum-induction melted and then vacuum-arc or electroslag remelted. Double-vacuum-induction melting may be employed when critical applications are involved.

Alloy A-286 (UNS S66286) is among the most forgeable of the heat-resistant alloys (Table 14), and its forgeability approaches that of AISI type 304 stainless steel. For example, Fig. 13(a) (Ref 28) plots the effect of temperature on the forging pressure required for a 10% upset reduction of various steels, where the forging pressures for A-286 and AISI type 304 converge at 1100 °C (2000 °F). However, alloy A-286 requires more power and more frequent reheating. Forging temperatures range from 980 to 1175 °C (1800 to 2150 °F). Reductions of at least 15% must be used under 980 °C (1800 °F) to prevent formation of coarse grains on solution treating (Ref 29).

Alloy 556 (UNS R30556, or N-155) is an iron-rich superalloy containing about 20 wt% Co, 20 wt% Ni, and 20 wt% Cr along with additions of molybdenum, tungsten, and niobium for improved solid-solution strengthening. It is single phase (austenitic) and strengthened primarily by work hardening. Maximum starting temperature is 1230 °C (2250 °F), and minimum finishing temperature is 955 °C (1750 °F). Hot-cold working may be done down to 760 °C (1400 °F), but reduction should exceed 10% below 980 °C (1800 °F) to prevent formation of coarse grains on solution treating (Ref 30).

Alloy 800 (UNS N08800) is workable, both hot and cold. The major part of the forging should be done between 1010 and 1230 °C (1850 and 2250 °F) metal temperature. Light working without tensile or bending stresses could be continued as low as 870 °C (1600 °F), but no work should be attempted between 650 and 870 °C (1200 and 1600 °F) because of the susceptibility to cracking within that range. Exposure at high temperatures to sulfurous atmospheres or to other sources of sulfur must be avoided. The furnace atmosphere for heating the material should be slightly reducing with approximately 2% CO. The rate of cooling is not critical with respect to thermal cracking, but the alloy is susceptible to carbide precipitation between 535 and 760 °C (1000 and 1400 °F).

The material is cold worked by much the same practice that is used for stainless steel. However, it work hardens to a slightly less degree than does stainless steel.

Nickel-base Superalloys (see also "Nickel-Base Alloys"). As shown in Table 14, all but one of the nickel-base superalloys are less forgeable than the iron-base alloys. Astroloy (UNS N13017) and alloy U-700 are the two most difficult to forge nickel-base alloys. For a given percentage of upset reduction at a forging temperature of 1095 °C (2000 °F), these alloys require about twice the specific energy needed for the iron-base A-286. In the forgeability ratings listed in Table 14, Astroloy and U-700 alloys have about one-fifth the forgeability of Alloy 600 (UNS N06600). However, these ratings reflect only a relative ability to withstand deformation without failure; they do not indicate the energy or pressure needed for forging, nor can the ratings be related to low-alloy steels and other alloys that are considerably more forgeable.

The forging of nickel-base alloys requires close control over metallurgical and operational conditions. Particular attention must be given to control of the work metal temperature. Recording usually is required for data on transfer time, soaking time, finishing temperature, and percentage of reduction. Critical parts are usually numbered, and precise records are kept. These records are useful in determining the cause of defective forgings, and they permit metallurgical analysis so that defects can be avoided in future products.

Nickel-base alloys are sensitive to minor variations in composition, which can cause large variations in forgeability, grain size, and final properties. In one case, wide heat-to-heat variations in

Table 14 Forging temperatures and forgeability ratings for heat-resistant alloys

Alloy	UNS designation	Forging temperature(a): Upset and breakdown, °C	Upset and breakdown, °F	Finish forging, °C	Finish forging, °F	Forgeability rating(b)
Iron-base alloys						
A-286	S66286	1095	2000	1035	1900	1
Alloy 556	R30556	1175	2150	1175	2150	3
Alloy 800	N08800	1150	2100	1035	1900	1
U-57		1095	2000	1035	1900	1
16-25-6		1095	2000	1095	2000	1
Nickel-base alloys						
R-235	...	1205	2200	1205	2200	3
Astroloy	N13017	1120	2050	1120	2050	5
Hastelloy W	W80004	1205	2200	1035	1900	4
Hastelloy X	N06002	1175	2150	1175	2150	3
Alloy 214	...	1160	2125	1040	1900	3
Alloy 230	...	1205	2200	1205	2200	3
Inconel 600	N06600	1150	2100	1035	1900	1
Inconel 700	...	1120	2050	1105	2025	4
Inconel 718	N07718	1095	2000	1035	1900	2
Inconel X-750	N07750	1175	2150	1120	2050	2
Inconel 751	N07751	1150	2100	1150	2100	3
Incoloy 901	N09901	1150	2100	1095	2000	2
M-252	N07252	1150	2100	1095	2000	3
Rene 41	N07041	1150	2100	1120	2050	3–4
U-500	N07500	1175	2150	1175	2150	3–4
U-700	...	1120	2050	1120	2050	5
Waspaloy	N07001	1160	2125	1040	1900	3
Cobalt-base alloys						
J-1570	...	1175	2150	1175	2150	2
J-1650	...	1150	2100	1150	2100	2
S-816	R30816	1150	2100	1150	2100	4
HS-25 (L-605)	...	1230	2250	1230	2250	3
Haynes 188	R30188	1205	2200	1175	2150	3

(a) Lower temperatures are often used for specific forgings to conform to appropriate specifications or to achieve structural uniformity.
(b) Based on the considerations stated in text. 1, most forgeable; 5, least forgeable.

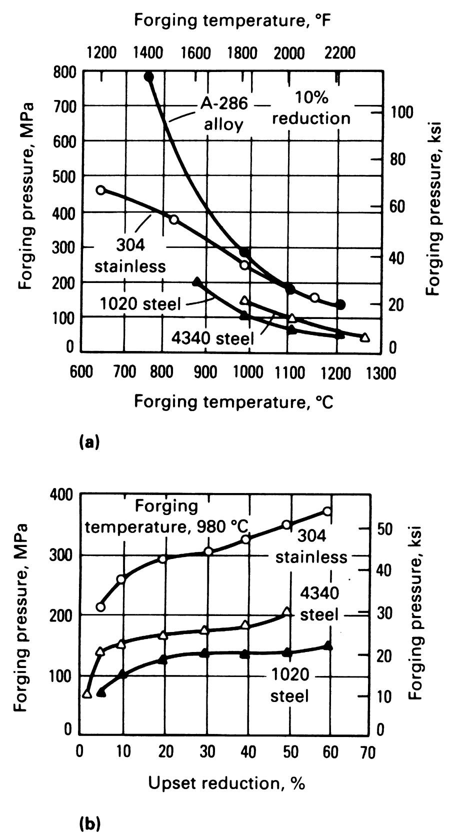


Fig. 13 Forging pressure required for upsetting versus (a) forging temperature and (b) percentage of upset reduction. Source: Ref 28

grain size occurred in parts forged from alloy 901 (UNS N09901) in the same sets of dies. For some parts, optimal forging temperatures had to be determined for each incoming heat of material by making sample forgings and examining them after heat treatment for variations in grain size and other properties. Improved ingot metallurgy is making the forging operation more consistent and easier to monitor, and wide variations in product structure are less frequent than in the past.

In the forging of nickel-base alloys, the forging techniques developed for one shape usually must be modified when another shape is forged from the same alloy; therefore, development time is often necessary for establishing suitable forging and heat treating cycles. This is especially true for stronger, more advanced alloys such as Waspaloy and Astroloy.

Cobalt-Base Alloys. Many of the cobalt-base alloys cannot be successfully forged because they ordinarily contain more carbon than the iron-base alloys and therefore greater quantities of hard carbides, which impair forgeability. The cobalt-base alloys listed in Table 14 are forgeable. The strength of these alloys at elevated temperatures, including the temperatures at which they are forged, is considerably higher than that for iron-base alloys; consequently, the pressures required in forging them are several times greater than those for the iron-base alloys.

Even when forged at its maximum forging temperature, Alloy-25 work hardens; therefore, forging pressure must be increased with greater reductions. Accordingly, this alloy generally requires frequent reheating during forging to promote recrystallization and to lower the forging pressure for subsequent steps.

Forging conditions (temperature and reduction) have a significant effect on the grain size of cobalt-base alloys. Because low ductility, notch brittleness, and low fatigue strength are associated with coarse grains, close control of forging and of final heat treatment is important. Cobalt-base alloys are susceptible to grain growth when heated above about 1175 °C (2150 °F). They heat slowly and require a long soaking time for temperature uniformity. Forging temperatures and reductions, therefore, depend on the forging operation and the part design.

The alloys are usually forged with small reductions in initial breakdown operations. The reductions are selected to impart sufficient strain to the metal so that recrystallization (and usually grain refinement) will occur during subsequent reheating. Because the cross section of a partly forged section has been reduced, less time is required to reach temperature uniformity in reheating. Consequently, because reheating time is shorter, the reheating temperature may sometimes be increased 30 to 85 °C (50 to 150 °F) above the initial forging temperature without harmful effects. However, if the part receives only small reductions in subsequent forging steps, forging should be continued at the lower temperatures. These small reductions, in turn, must be in excess of about 5 to 15% to avoid abnormal grain growth during subsequent annealing. The forging temperatures given in Table 14 are usually satisfactory.

Refractory Metals (Ref 31)

Refractory metals are forged from as-cast ingots or from billets that have been previously broken down by forging or extrusion. Forgeability depends to some extent on the method used to work the ingot into a billet. The forging characteristics of refractory metals and alloys are listed in Table 15.

Niobium and Niobium Alloys. Niobium and several of its alloys, notably, Nb-1Zr and Nb-33Ta-1Zr, can be forged directly from as-cast ingot. Most impression-die forging experience, however, has been with unalloyed niobium. Unalloyed niobium, Nb-1Zr, and Nb-33Ta-1Zr

Table 15 Forging characteristics of refractory metals and alloys

Metal or alloy	Approximate solidus temperature		Recrystallization temperature, minimum		Hot-working temperature, minimum(a)		Forging temperature		Forgeability
	°C	°F	°C	°F	°C	°F	°C	°F	
Niobium and niobium alloys									
99.2% Nb	2470	4475	1040	1900	815	1500	20–1095	70–2000	Excellent
Nb-1Zr	2400	4350	1040	1900	1150	2100	20–1260	70–2300	Excellent
Nb-33Ta-1Zr	2520	4570	1205	2200	1315	2400	1040–1480	1900–2700	Good
Nb-28Ta-10W-1Zr	2590	4695	1230	2250	1315	2400	1260–1370(b)	2300–2500	Good(b)
Nb-10Ti-10Mo-0.1C	2260	4100	1205	2200	1370	2500	1040–1480	1900–2700	Moderate
Nb-10W-1Zr-0.1C	2595	4700	1150	2100	1205	2200	1095–1205(b)	2000–2200	Moderate(b)
Nb-10W-2.5Zr	...	...	1150	2100	1260	2300	1205–1425(b)	2200–2600	Good(b)
Nb-15W-5Mo-1Zr	2480	4500	1425	2600	1650	3000	1315–1650	2400–3000	Fair
Nb-10Ta-10W	2600	4710	1150	2100	1315	2400	925–1205	1700–2200	Good
Nb-5V-5Mo-1Zr	2370	4300	1150	2100	1315	2400	1205–1650	2200–3000	Moderate(b)
Nb-10W-10Hf-0.1Y	...	...	1095	2000	1205	2200	1095–1650(b)	2000–3000	Good(b)
Nb-30Ti-20W	>2760	>5000	1260	2300	1150	2100	1150–1260	2100–2300	Good
Tantalum and tantalum alloys									
99.8% Ta	2995	5425	1095	2000	1315	2400	20–1095(b)	70–2000	Excellent(b)
Ta-10W	3035	5495	1315	2400	1650	3000	980–1260(b)	1800–2300	Good(b)
Ta-12.5W	3050	5520	1510	2750	>1650	>3000	>1095(b)	>2000	Good(b)
Ta-30Nb-7.5V	2425	4400	1150	2200	1540	2800	1150–1315(b)	2200–2400	Good(b)
Ta-8W-2Hf	2980	5400	1540	2800	>1650	>3000	>1095(b)	>2000	Good(b)
Ta-10Hf-5W	2990	5420	1315	2400	1650	3000	>1095(b)	>2000	Good(b)
Ta-2.5W	>2760	>5000	1260	2300	1150	2100	20–1150	70–2100	Excellent
Molybdenum and molybdenum alloys									
Unalloyed Mo	2610	4730	1150	2100	1315	2400	1040–1315	1900–2400	Good
Mo-0.5Ti	2595	4700	1315	2400	1480	2700	1150–1425	2100–2600	Good-fair
Mo-0.5Ti-0.08Zr	2595	4700	1425	2600	1650	3000	1205–1480	2200–2700	Good
Mo-25W-0.1Zr	2650	4800	1425	2600	1650	3000	1040–1315	1900–2400	Fair
Mo-30W	2650	4800	1260	2300	1370	2500	1150–1315	2100–2400	Fair
Tungsten and tungsten alloys									
Unalloyed W	3410	6170	1370–1595	2500–2900	...	...	1205–1650	2200–3000	...
W-1ThO_2	3410	6170	1595–1650	2900–3000	...	...	1315–1925	2400–3500	...
W-2ThO_2	3410	6170	1650–1760	3000–3200	...	...	1315–1370	2400–2500	...
W-2Mo	3385	6125	1540–1650	2800–3000	...	...	1150–1370	2200–2500	...
W-15Mo	3300	5970	1480–1595	2700–2900	...	...	1095–1370	2000–2500	...
W-26Re	3120	5650	>1870	>3400	...	...	>1480	>2700	...
W-0.5Nb	3405	6160	1705–1870	3100–3400	...	...	1205–1650	2200–3000	...

(a) Minimum hot-working temperature is the lowest forging temperature at which alloys begin to recrystallize during forging. (b) Based on breakdown forging and rolling experience can be cold worked. Other alloys, such as Nb-15W-5Mo-1Zr, generally require initial hot working by extrusion to break down the coarse grain structure of as-cast ingots before finish forging.

The billets are usually heated in a gas furnace using a slightly oxidizing atmosphere. Niobium alloys tend to flow laterally during forging. This results in excessive flash that must be trimmed from forgings. Niobium and its alloys can be protected from oxidation during hot working by dipping the billets in an Al-10Cr-2Si coating at 815 °C (1500 °F), then diffusing the coating in an inert atmosphere at 1040 °C (1900 °F). The resulting coating is about 0.05 to 0.1 mm (2 to 4 mils) thick and provides protection from atmospheric contamination at temperatures to 1425 °C (2600 °F). Glass frit coatings can also be applied to the workpiece before heating in a gas-fired furnace.

Molybdenum and Molybdenum Alloys. The forging behavior of molybdenum and molybdenum alloys depends on the preparation of the billet. Billets prepared by pressing and sintering can be forged directly. Large billets are open-die forged or extruded before closed-die forging. Arc-cast billets are usually brittle in tension; they cannot be forged before extruding, except at extremely high temperatures. A minimum extrusion ratio for adequate forgeability is 4 to 1.

Workpieces subjected to large reductions usually exhibit anisotropy and will recrystallize at lower temperatures than parts given less reduction. Forging temperature and reduction must be carefully controlled to avoid premature recrystallization in service and the resulting loss in strength.

Gas- or oil-fired furnaces can be used to heat molybdenum and its alloys to approximately 1370 °C (2500 °F). Induction heating is required for higher forging temperatures. Above 760 °C (1400 °F), molybdenum forms a liquid oxide that volatilizes rapidly enough that surface contamination is rarely a problem. If metal losses are excessive, protective atmospheres such as argon, carbon monoxide, or hydrogen can be used during heating. The liquid oxide formed during heating also serves as a lubricant. Glass coatings are also used; in addition to providing lubrication, glass coatings reduce heat losses during forging. Molybdenum disulfide and colloidal graphite are suitable lubricants for small forgings.

Tantalum and Tantalum Alloys. Unalloyed tantalum and most of the single-phase alloys listed in Table 15 can be forged directly from cast ingots. However, breakdown operations are usually required in order to avoid laps, wrinkles, internal cracks, and other forging defects. The breakdown temperature is 1095 to 1315 °C (2000 to 2400 °F). After about 50% reduction, the forging temperature may be permitted to drop slightly below 1095 °C (2000 °F). Billets produced by powder-metallurgy techniques do not lend themselves to direct forging and must be subjected to breakdown.

Most of the forging experience to date has been with the Ta-10W alloy. Billets are heated to 1150 to 1205 °C (2100 to 2200 °F) in gas-fired furnaces using an oxidizing atmosphere. Breakdown forging below 980 °C (1800 °F) or continued working below 815 °C (1500 °F) can cause internal cracking. Forgeability of the tantalum alloys decreases sharply as tungsten content exceeds about 12.5%. Interstitial elements such as carbon, oxygen, and nitrogen also have a deleterious effect on forgeability.

Two types of coatings—glasses and aluminides—have been successfully used to protect tantalum from oxidation during forging. A 0.076 mm (3 mil) thick coating of aluminum has provided protection for the Ta-10W alloy when it was heated in air at 1370 °C (2500 °F) for 30 min. Glass coatings are generally preferred for their lubricating properties. Various borosilicate glasses are available that can be used for forging operations carried out in the range of 1095 to 1315 °C (2000 to 2400 °F).

Tungsten and Tungsten Alloys. Tungsten-base materials, like the other refractory alloy systems, can be classified into two broad groups: unalloyed tungsten and solid-solution or dispersion-strengthened alloys. These classifications are convenient because they group the alloys in terms of metallurgical behavior and applicable consolidation methods. Solid-solution alloys and unalloyed tungsten can be produced by powder-metallurgy or conventional melting techniques; dispersion-strengthened alloys can be produced only by powder-metallurgy methods.

The forgeability of tungsten alloys, like that of molybdenum alloys, is dependent on the consolidation technique used. Billet density, grain size, and interstitial content all affect forgeability. Metallurgical principles in the forging of tungsten are much the same as those for molybdenum. Tungsten is usually forged in the hot/cold-working temperature range, in which hardness and strength increase with increasing reductions. Both systems exhibit increasing forgeability with decreasing grain size.

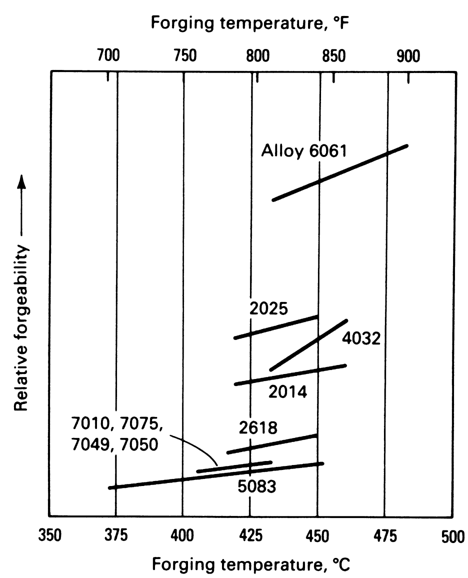


Fig. 14 Forgeability and forging temperatures of various aluminum alloys. There are wrought aluminum alloys, such as 1100 and 3003, whose forgeability would be rated significantly above those presented; however, these alloys have limited application in forging because they cannot be strengthened by heat treatment.

Tungsten requires considerably higher forging pressures than molybdenum; therefore, in-process annealing is often necessary to reduce the load requirements for subsequent forging operations. Because the need for lateral support during forging is greater for tungsten than for molybdenum, the design of preliminary forging tools is more critical. This is especially true for pressed-and-sintered billets, which have some porosity and are less than theoretical density.

Tungsten oxide, which becomes molten and volatilizes at forging temperatures, serves as an effective lubricant in the forging of tungsten. Mixtures of graphite and molybdenum disulfide are also used. Sprayed on the dies, these films provide lubricity and facilitate removal of the part from the dies. Glass coatings are also used, but they can accumulate in the dies and interfere with complete die filling.

Aluminum Alloys (Ref 32)

Compared to the nickel/cobalt-base alloys and titanium alloys, aluminum alloys are considerably more forgeable, particularly in conventional forging-process technology, in which dies are heated to 540 °C (1000 °F) or less. Figure 14 illustrates the relative forgeability of ten aluminum alloys that constitute the bulk of aluminum alloy forging production. This forgeability rating is principally based on the deformation per unit of energy absorbed in the range of forging temperatures typically employed for the alloys in question. Also considered in this index is the difficulty of achieving specific degrees of severity in deformation as well as the cracking tendency of the alloy under forging-process conditions. As shown in Fig. 14, there is considerable variation in the effect of temperature on forgeability of aluminum alloys. The high-silicon alloy 4032 shows the greatest effect, while the high-strength Al-Zn-Mg-Cu 7*xxx* alloys display the least effect.

The 15 aluminum alloys that are most commonly forged, as well as recommended temperature ranges, are listed in Table 16. All of these alloys are generally forged to the same reduction, although some alloys may require more forging power and/or more forging operations than others. The forging temperature range for

Table 16 Recommended forging temperature ranges for aluminum alloys

Aluminum alloy	Forging temperature range °C	°F
1100	315–405	600–760
2014	420–460	785–860
2025	420–450	785–840
2219	425–470	800–880
2618	410–455	770–850
3003	315–405	600–760
4032	415–460	780–860
5083	405–460	760–860
6061	430–480	810–900
7010	370–440	700–820
7039	380–440	720–820
7049	360–440	680–820
7050	360–440	680–820
7075	380–440	720–820
7079	405–455	760–850

most alloys is relatively narrow (generally <55 °C, or 100 °F), and for no alloy does the range extend beyond 85 °C (155 °F). Obtaining and maintaining proper metal temperatures in the forging of aluminum alloys is critical to the success of the forging process. Die temperature and deformation rates play key roles in the actual forging temperature achieved. Unlike some forging processes for carbon and alloy steels, the dies used in virtually all hot-forging processes for aluminum alloys are heated in order to facilitate the forging process. Therefore, die temperature is another critical element in the forgeability and forging-process optimization of this alloy class.

Table 17 summarizes the die-temperature ranges typically used for several aluminum forging processes. The criticality of die temperature in the optimization of the process depends on the forging equipment being employed, the alloy being forged, and the severity of the deformation or the sophistication of the forging design. For slower deformation processes, such as hydraulic press forging, die temperature frequently controls the actual metal temperature during deformation, and, in fact, aluminum alloys forged in hydraulic presses are isothermally forged; that is, metal and dies are at the same temperature during deformation. Therefore, the die temperatures employed for hydraulic press forging exceed those typical of more rapid deformation processes, such as hammers and mechanical presses. Both remote and on-press die-heating systems are used in the forging of aluminum alloys.

Forging Methods. Aluminum alloys are produced by all of the current forging methods available, including open-die (or hand) forging, closed-die forging, upsetting, roll forging, orbital (rotary) forging, spin forging, mandrel forging, ring rolling, and extrusion. Selection of the optimal forging method for a given forging shape is based on the desired forged shape, the sophistication of the forged-shape design, and cost. In many cases, two or more forging methods are combined in order to achieve the desired forging shape and microstructure. For example, open-die forging frequently precedes closed-die forging to prework the alloy (especially when cast ingot forging stock is used) and to preshape (or preform) the metal to conform to the subsequent closed dies and to conserve input metal.

Table 17 Die temperature ranges for the forging of aluminum alloys

Forging process/equipment	Die temperature °C	Die temperature °F
Open-die forging		
Ring rolling	95–205	200–400
Mandrel forging	95–205	200–400
Closed-die forging		
Hammers	95–150	200–300
Upsetters	150–260	300–500
Mechanical presses	150–260	300–500
Screw presses	150–260	300–500
Orbital (rotary) forging	150–260	300–500
Spin forging	150–315	200–600
Roll forging	95–205	200–400
Hydraulic presses	315–430	600–800

Aluminum alloy forgings are produced on a wide variety of forging equipment. Aluminum alloy forgings are produced on the full spectrum of forging equipment, ranging from hammers and presses to specialized forging machines. Selection of forging equipment for a given forging shape and type is based on the capabilities of the equipment, forging design sophistication, desired forging process, and cost. Deformation or strain rate is also a critical element in the successful forging of given alloy. The deformation or strain rate imparted by equipment varies considerably, ranging from very fast (for example, >10 s^{-1} on equipment such as hammers, mechanical presses, and high-energy-rate machines) to relatively slow (for example, <0.1 s^{-1} on equipment such as hydraulic presses). Although aluminum alloys are generally not considered to be as sensitive to strain rate as other materials (such as titanium and nickel/cobalt-base superalloys), selection of the strain rate in a given forging process or differences in deformation rates inherent in various types of equipment affect the forging-pressure requirements, the severity of deformation possible, and therefore the sophistication of the forging part that can be produced.

Open-die forging is used frequently to produce small quantities of aluminum alloy forgings when the construction of expensive closed dies is not justified or when such quantities are needed during the prototype fabrication stages of a forging application. The quantity that warrants the use of closed dies varies considerably, depending on the size and shape of the forging and on the application for the part. However, open-die forging is by no means confined to small or prototype quantities, and in some cases it may be the most cost-effective method of aluminum forging manufacture. For example, as many as 2000 pieces of biscuit forgings have been produced in open dies when it was desired to obtain the properties of a forging, but closed dies did not provide sufficient economic benefits.

Open-die forgings in aluminum alloys can be produced to a wide variety of shapes, ranging from simple rounds, squares, or rectangles to very complex contoured forgings. In the past, the complexity and tolerances of the open-die forging of aluminum and other materials depended on the skill of the press operator; however, with the advent of programmable computer-controlled open-die forging presses, it is possible to produce such shapes to overall thickness/width tolerance bands of 1.27 mm (0.050 in.). Because the open-die forging of aluminum alloys is also frequently implemented to produce preforms for closed-die forgings, these state-of-the-art forging machines also provide very precise preform shapes, improving the dimensional consistency and tolerances of the resulting closed-die forging and reducing closed-die forging cost through further input material conservation.

Closed-Die Forging. Most aluminum alloy forgings are produced in closed dies. The four types of aluminum forgings shaped in closed dies are blocker-type (finish forging only), conventional (block-and-finish forging or finish forging only), high-definition (near net shape), and precision (no draft, net shape). These closed-die forging types are illustrated in Fig. 15.

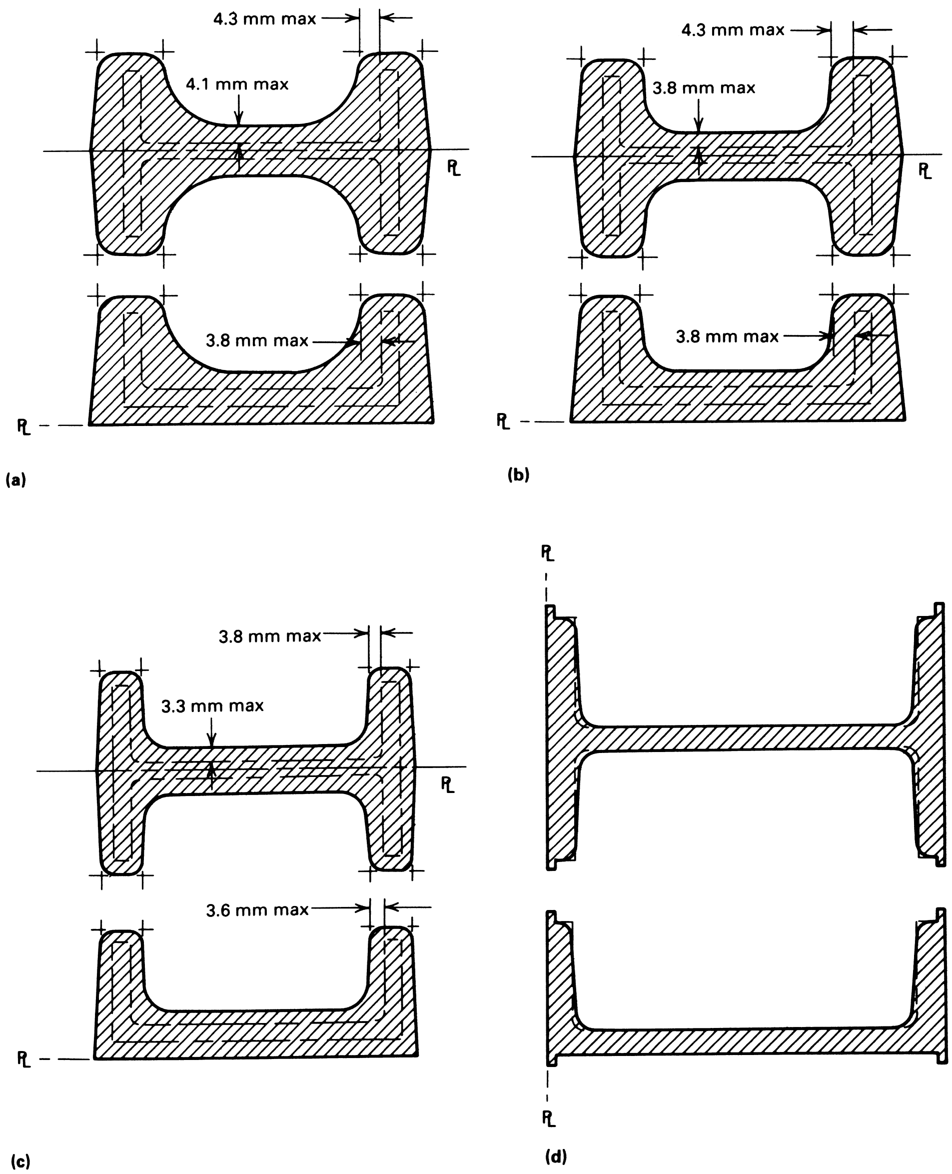
Blocker-type forgings (Fig. 15a) are produced in relatively inexpensive, single sets of dies. In dimensions and forged details, they are less refined and require more machining than conventional or high-definition closed-die forgings. A blocker-type forging costs less than a comparable conventional or high-definition forging, but it requires more machining.

Conventional closed-die forgings (Fig. 15b) are the most common type of aluminum forging. They are produced with either a single set of finish dies or with block-and-finish dies, depending on the design criteria. These forgings have less machine stock and tighter tolerances than blocker-type forgings, but require additional cost (both for the dies and for fabrication) to produce.

High-definition, near-net-shape closed-die forgings (Fig. 15c) are a result of improved forging equipment and process control. They offer forging design and tolerance enhancement over conventional or blocker-type forgings to affect further reduction in machining costs. High-definition forgings are produced with multiple die sets, consisting of one or more blocker dies and finish dies, and are frequently used with some as-forged surfaces remaining unmachined by the purchaser.

Precision forgings (Fig. 15d) represent net-shape products that require no subsequent machining. Net-shape aluminum forgings are produced in two-piece, three-piece through-die, and/or multiple-segment wrap-die systems to very restricted design and tolerances necessary for assembly. Precision aluminum forgings are produced with very thin ribs and webs; sharp corner and fillet radii; undercuts, backdraft, and/or contours; and, frequently, multiple parting planes that may optimize grain-flow characteristics. Design and tolerance criteria for precision aluminum forgings have been established to provide a finished product suitable for assembly or further fabrication. Precision aluminum forgings do not necessarily conform to the tolerances provided by machining of other product forms; however, as outlined in Table 18, design and tolerance criteria are highly refined in comparison to other aluminum alloy forging types and are suitable for the intended application of the product without subsequent machining by the purchaser.

Precision aluminum forgings are typically produced on hydraulic presses, although in some cases mechanical and/or screw presses have been effectively employed. Although many precision aluminum forgings have been produced on small-to-intermediate hydraulic presses with capacities in the range of 9 to 900 kN (1 to 100 tonf), the size of precision parts demanded by users has increased, and so heavy hydraulic presses in the range of 135 to 310,000 kN (15 to 35,000 tonf) have been added or upgraded to



Characteristic	Tolerance, mm (in.) Blocker-type	Conventional	High-definition	Precision
Die closure	+2.3, −1.5 (+0.09, −0.06)	+1.5, −0.8 (+0.06, −0.03)	+1.25, −0.5 (+0.05, −0.02)	+0.8, −0.25 (+0.03, −0.01)
Mismatch	0.5 (0.02)	0.5 (0.02)	0.25 (0.01)	0.38 (0.015)
Straightness	0.8 (0.03)	0.8 (0.03)	0.5 (0.02)	0.4 (0.016)
Flash extension	3 (0.12)	1.5 (0.06)	0.8 (0.03)	0.8 (0.03)
Length and width	±0.8 (±0.03)	±0.8 (±0.03)	±0.8 (±0.03)	+0.5, −0.25 (+0.02, −0.01)
Draft angles	5°	5°	3°	1°

Fig. 15 Types of aluminum closed-die forgings and tolerances for each. (a) Blocker-type. (b) Conventional. (c) High-definition. (d) Precision

produce this product. Forging-process criteria for precision aluminum forgings are similar to those for other aluminum alloy forging types, although the metal and die temperatures used are usually controlled to near the upper limits of the temperature ranges outlined in Tables 16 and 17 to enhance producibility and to minimize forging pressures. As with other aluminum forging processes, die lubrication also is a critical element in precision aluminum forging. The die lubricants employed, although of the same generic graphite/mineral-oil formulations used for other aluminum forging processes, frequently use other organic and inorganic compounds tailored to the process demands.

Upset forging can be accomplished in specialized forging equipment called upsetters (a form of mechanical press) or high-speed multiple-station formers and is frequently used to produce forging shapes that are characterized by surfaces of revolution, such as bolts, valves, gears, bearings, and pistons. Upset forging may be the sole process used for the shape, such as pistons, or it can be used as a preliminary operation to reduce the number of impressions, to reduce die wear, or to save metal when the products are finished in closed dies. Wheel and gear forgings are typical products for which upsetting is advantageously used in conjunction with closed-die forging. As a rule, in the upset forging of aluminum alloys, the unsupported length of forgings must not exceed three diameters for a round shape or three times the diagonal of the cross section for a rectangular shape.

Table 18 Design and tolerance criteria for aluminum precision forgings

Characteristics	Tolerance
Draft outside	0° +30′, −0
Draft inside	1° +30′, −0
Corner radii	1.5 ± 0.75 mm (0.060 ± 0.030 in.)
Fillet radii	3.3 ± 0.75 mm (0.130 ± 0.030 in.)
Contour	±0.38 mm (±0.015 in.)
Straightness	0.4 mm in 254 mm (0.016 in. in 10 in.)
Minimum web thickness(a)	2.3 mm (0.090 in.)
Minimum rib thickness	2.3 mm (0.090 in.)
Length/width tolerance	+0.5 mm, −0.25 mm (+0.020 in., −0.010 in.)
Die closure tolerance	+0.75, −0.25 mm (+0.030, −0.010 in.)
Mismatch tolerance	0.38 mm (0.015 in.)
Flash extension	0.75 mm (0.030 in.)

(a) Web thicknesses as small as 1.5 mm (0.060 in.) have been produced in certain forging designs

Roll forging can be used as a preliminary preform operation to reduce metal input or to reduce the number of closed-die operations. In roll forging, the metal is formed between moving rolls, either or both containing a die cavity, and is most often used for parts, such as connecting rods, where volume is high and relatively restricted cross-sectional variations typify the part.

Orbital (rotary) forging is a variant of closed-die mechanical or hydraulic press forging in which one or both of the dies is caused to rotate, usually at an angle, leading to the incremental deformation of the workpiece. Orbital forging is used to produce parts with surfaces of revolution with both hot and cold aluminum alloy forging processes, and it provides highly refined close-tolerance final shapes.

Spin forging, a relatively new aluminum alloy forging technique, combines closed-die forging and computer numerically controlled (CNC) spin forgers to achieve close-tolerance axisymmetric hollow shapes. Because spin forging is accomplished over a mandrel, inside-diameter contours are typically produced to net shape, requiring no subsequent machining. Outside-diameter contours can be produced net or with very little subsequent machining and to much tighter out-of-round and concentricity tolerances than competing forging techniques, such as forward or reverse extrusion (see below), resulting in material savings. Parts with both ends open, one end closed, or both ends closed can also be produced.

Ring rolling is also used for aluminum alloys to produce annular shapes. The procedure used to ring roll aluminum alloys is essentially the same as that used for steel. Both rectangular and contoured cross section rolled rings, with or without subsequent machining by the forger, are produced in many aluminum alloys. The temperatures employed for the ring rolling of aluminum alloys are quite similar to those for other forging processes, although care must be taken to maintain metal temperature. The deformation achieved in the ring rolling of aluminum typically results in the predominant grain flow in the tangential or circumferential orientation. If predominant grain flow is desired in other directions, such as axial or radial, other ringmaking processes, such as hollow-biscuit open-die forgings, mandrel forging, or reverse/forward extrusion, can be employed. The economy of ring rolling in aluminum alloys depends on the volume, size, and contour of the forging. For some ring parts, it may be more economical to produce the shape by mandrel forging or to cut rings from hollow extruded cylinders. Both techniques are discussed below.

Mandrel forging is used in aluminum alloys to produce axisymmetric, relatively simple, hollow ring or cylindrical shapes, in which the metal is incrementally forged, usually on a hammer or hydraulic press, over a mandrel. In the incremental forging process, the wall thickness of the preform is reduced, and this deformation enlarges the diameter of the piece. The mandrel forging of aluminum has been found to be economical for relatively low-volume part fabrication and/or in the fabrication of very large ring shapes (up to 3.3 m, or 130 in., in diameter). With control of the working history of the input material and the mandrel-forging process, mandrel-forged rings can be produced with either circumferential or axial-predominant grain orientations.

Reverse or forward extrusion, a variant of closed-die forging for aluminum, can be used to produce hollow, axisymmetric shapes in aluminum alloys with both ends open or with one end closed. Extrusion also frequently plays an important role in the closed-die forging of aluminum alloy parts other than hollow shapes (such as wheels). More information on extrusion of aluminum is available in Chapter 20, "Extrusion."

Heating and Temperature Control. Metal temperature is a critical element in the aluminum forging process. Aluminum alloys form a very tenacious oxide coating upon heating. The formation of this coating is self-limiting; therefore, aluminum alloys do not scale to the same extent that steel does. However, most aluminum alloys are susceptible to hydrogen pickup during reheating operations such that reheating equipment and practices are also critical elements of forging-process control. In the open-die forging of aluminum alloys, it is generally desirable to have billets near the high side of the forging temperature range when forging begins and to finish the forging as quickly as possible before the temperature drops excessively. Open-die forging and multiple-die closed-die forging of aluminum alloys are frequently conducted without reheating as long as critical metal temperatures can be maintained.

Aluminum alloys have a relatively narrow temperature range for forging. Therefore, careful control of the temperature in preheating is important. The heating equipment should have pyrometric controls that can maintain ±5 °C (±10 °F). Continuous furnaces used to preheat aluminum typically have three zones: preheat, high heat, and discharge. Most furnaces are equipped with recording/controlling instruments and are frequently surveyed for temperature uniformity in a manner similar to that used for solution treatment and aging furnaces. Heated aluminum alloy billets are usually temperature checked by using either contact or noncontact pyrometry based on dual-wavelength infrared systems. This latter technology, although sensitive to emissivity, has been successfully incorporated into the fully automated temperature-verification systems used in automated high-volume aluminum forging processes to provide significantly enhanced temperature control and process repeatability.

As previously noted, the methods of die heating include remote die-heating systems and on-press die-heating systems. Remote die-heating systems are usually gas-fired die heaters capable of slowly heating the die blocks. These systems are used to preheat dies to the desired temperature prior to assembly into the forging equipment. On-press die-heating systems range from relatively rudimentary systems to highly engineered systems designed to maintain very tight die-temperature tolerances. On-press die-heating systems include gas-fired equipment, induction-heating equipment, and/or resistance-heating equipment. In addition, presses used for the precision forging of aluminum alloys frequently have bolsters that can be heated or cooled as necessary.

State-of-the-art on-press aluminum die-heating equipment can hold die-temperature tolerances within ±15 °C (±25 °F) or better. Specific on-press die-heating systems vary with the forging equipment used, the size of the dies, the forging process, and the type of forging produced.

Die lubrication is another critical element in the aluminum forging process and is the subject of major engineering and developmental emphasis, both in terms of the lubricants themselves and the lubricant-application systems. The lubricants used in aluminum alloy forging are subject to severe service demands. They must be capable of modifying the surface of the die to achieve the desired reduction in friction, withstand the high die and metal temperatures and pressures employed, and yet leave the forging surfaces and forging geometry unaffected. Lubricant formulations are typically highly proprietary and are developed either by the lubricant manufacturers or by the forgers themselves. Lubricant composition varies with the demands of the forging process used and the forging type. The major active element in aluminum alloy forging lubricants is graphite; however, other organic and inorganic compounds are added to colloidal suspensions in order to achieve the desired results. Carriers for aluminum alloy forging lubricants vary from mineral spirits to mineral oils to water.

Lubricant application is typically achieved by spraying the lubricant onto the dies while the latter are assembled in the press; however, in some cases, lubricants are applied to forging stock prior to reheating or just prior to forging. Several pressurized-air or airless spraying systems are employed, and with high-volume highly automated aluminum forging processes, lubricant application is also automated by single- or multiple-axis robots. Lubricant can be applied with or without heating. State-of-the-art lubricant-application systems have the capability of applying very precise patterns or amounts of lubricant under fully automated conditions such that the forging processes are optimized and repeatable.

Copper and Copper Alloys

Copper and copper alloy forgings offer a number of advantages over parts produced by other processes, including high strength as a result of working, closer tolerances than competing processes such as sand casting, and modest overall cost. The most forgeable copper alloy, forging brass (alloy C37700), can be forged into a given shape with substantially less force than that required to forge the same shape from low-carbon steel. A less forgeable copper alloy, such as an aluminum bronze, can be forged with approximately the same force as that required for low-carbon steel. Copper and copper alloy forgings, particularly brass forgings, are used in valves, fittings, refrigeration components, and other high-pressure liquid and gas-handling applications. High-strength bronze forgings find application as mechanical parts such as gears, bearings, and hydraulic pumps.

Copper C10200 and the copper alloys most commonly forged are listed in Table 19. They comprise at least 90% of all commercially produced copper alloy forgings. Forging brass, the least difficult alloy to forge, has been assigned an arbitrary forgeability rating of 100. Some copper alloys cannot be forged to any significant degree because they will crack. Leaded copper-zinc alloys, such as architectural bronze, which may contain more than 2.5% Pb, are seldom recommended for hot forging. Although lead content improves metal flow, it promotes cracking in those areas of a forging, particularly deep-extruded areas, that are not completely supported by, or enclosed in, the dies. This does not mean that the lead-containing alloys cannot be forged, but rather that the design of the forging may have to be modified to avoid cracking.

Table 19 Relative forgeability ratings of commonly forged copper alloys

Ratings are in terms of the most forgeable alloy, forging brass (C37700).

Alloy	Nominal composition	Relative forgeability(a), %
C10200	99.95 min Cu	65
C10400	Cu-0.027 Ag	65
C11000	99.9 min Cu	65
C11300	Cu-0.027 Ag + O	65
C14500	Cu-0.65Te-0.008P	65
C18200	Cu-0.10 Fe-0.90Cr-0.10 Si-0.05Pb	80
C37700	Cu-38Zn-2 Pb	100
C46400	Cu-39.2Zn-0.8Sn	90
C48200	Cu-38Zn-0.8Sn-0.7Pb	90
C48500	Cu-37.5Zn-1.8 Pb-0.7Sn	90
C62300	Cu-10Al-3Fe	75
C63000	Cu-10Al-5Ni-3Fe	75
C63200	Cu-9Al-5Ni-4Fe	70
C64200	Cu-7Al-1.8Si	80
C65500	Cu-3Si	40
C67500	Cu-39Zn-1.4Fe-1Si-0Mn	80

(a) Takes into consideration such factors as pressure, die wear, and hot plasticity

The solubility of lead in β-brass at forging temperatures is about 2% maximum, but lead is insoluble in β-brass at all temperatures. Consequently, although a lead content of up to 2.5% is permissible in Cu-40Zn α-β brasses, lead in excess of 0.10% in a Cu-30Zn α-brass will contribute to catastrophic cracking. Other copper alloys, such as the copper-nickels, can be forged only with greater difficulty and at higher cost. The copper-nickels, primarily because of their higher forging temperatures, are sometimes heated in a controlled atmosphere, thus complicating the process. The silicon bronzes, because of their high forging temperatures and their compositions, cause more rapid die deterioration than the common forging alloys.

Heating Practices. Optimal forging temperature ranges for ten alloys are given in Table 20. Atmosphere protection during billet heating is not required for most alloys, especially when forging temperatures are below 705 °C (1300 °F). For temperatures toward the top of the range in Table 20, a protective atmosphere is desirable and is sometimes required. An exothermic atmosphere is usually the least costly, and it is satisfactory for heating copper alloys at temperatures above 705 °C (1300 °F).

Gas-fired furnaces are almost always used, and furnace design is seldom critical. Open-fired conveyor chain or belt types are those most commonly used. Any type of temperature control that can maintain temperature within ±5 °C (±10 °F) is suitable. As billets are discharged, a periodic check with a prod-type pyrometer should be made. This permits a quick comparison of billet temperature with furnace temperature.

Heating Time. The time at temperature is critical for all copper alloys, although to varying degrees among the different alloys. For forging brass (alloy C37700), the time is least critical, but for aluminum bronze, naval brass, and copper, it is most critical. Time in excess of that required to bring the billet uniformly to forging temperature is detrimental, because it causes grain growth and increases the amount of scale.

Table 20 Recommended forging temperature ranges for copper alloys

Alloy	Temperature range °C	Temperature range °F
C12200	730–845	1350–1550
C18200	650–760	1200–1400
C37700	650–760	1200–1400
C46400	595–705	1100–1300
C62400	705–815	1300–1500
C64200	730–900	1350–1650
C67000	595–705	1100–1300
C67300	595–730	1100–1350
C67400	595–730	1100–1350
C67500	595–705	1100–1300

Reheating Practice. When forging in hammers, all of the impressions are usually made in one pair of dies, and reheating is rarely required. In press forging, particularly in high-production applications, blocking is often done separately, followed by trimming before the forging is completed. The operations are likely to be performed in different presses; therefore, the partially completed forging is reheated to the temperature originally used.

Heating of Dies. Dies are always heated for forging copper and copper alloys, although because of the good forgeability of copper alloys, die temperature is generally less critical than for forging aluminum. Dies are seldom preheated in ovens. Heating is usually accomplished by ring burners. Optimal die temperatures vary from 150 to 315 °C (300 to 600 °F), depending on the forging temperature of the specific alloy. For alloys having low forging temperatures, a die temperature of 150 °C (300 °F) is sufficient. Die temperature is increased to as much as 315 °C (600 °F) for the alloys having the highest forging temperatures shown in Table 20.

Dies designed for forging copper or copper alloys usually differ from those designed for forging the same shapes from steel, as follows:

- The draft angle can be decreased for forging copper (3° max and often less than 3°).
- The die cavity is usually machined to dimensions that are 0.005 in./in. less than those for forging steels.
- The die cavity is usually polished to a better surface finish for forging copper and copper alloys.

Die materials and hardnesses selected for forging copper alloys depend on part configuration (forging severity) and number of parts to be produced. Figure 16 illustrates the forging severities of parts listed in Table 21. Whether the dies are made entirely from a hot-work steel such as H11 or H12 or whether or not inserts are used depends largely on the size of the die. Common practice is to make the inserts from a hot-work steel and to press them into rings or holders made from a low-alloy die block steel (Table 21) or L6 tool steel. Hardness of the ring or holder is seldom critical; a range of 341 to 375 HB is typical.

Die Lubrication. Dies should be lubricated before each forging operation. A spray of colloidal graphite and water is usually adequate. Many installations include a spray that operates automatically, timed with the press stroke. However, the spray is often inadequate for deep cavities and is supplemented by swabbing with a conventional forging oil.

Forging Methods. Most copper alloy forgings are produced in closed dies. The sequence of operations is the same as that used for forging a similar shape from steel, that is, fullering, blocking, and finishing, as required. However, it is estimated that 90% of the forgings produced from forging brass are forged completely in one or two blows in a finishing die. The starting slugs or blanks are usually cut from extruded bars or tubes to eliminate the blocking operation. Excessive flash is produced, but it is easily trimmed and remelted. In the forging of parts of mild to medium severity, in plants where remelting facilities are available, cutting slugs from bars or tubes is usually the least expensive approach. However, in plants that do not remelt their scrap, the flash must be sold as scrap, and it is sometimes more economical to use blocking.

Cylindrical slugs are sometimes partially flattened before forging to promote better flow and consequently better filling of an impression. This can usually be done at room temperature between flat dies in a hammer or a press. A rectangular slug is occasionally obtained by extrud-

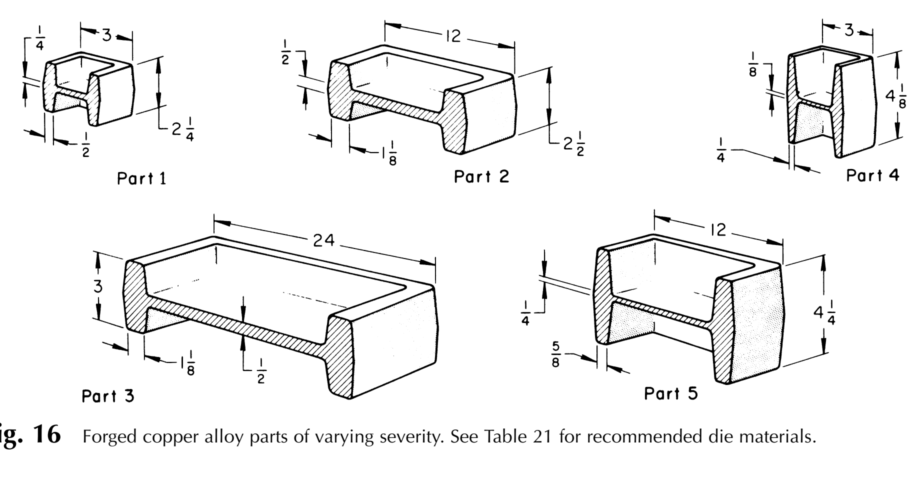


Fig. 16 Forged copper alloy parts of varying severity. See Table 21 for recommended die materials.

Table 21 Recommended die materials for the forging of copper alloys

Part configurations of varying severity are shown in Fig. 16.

	Total quantity to be forged			
	100–10,000		≥10,000	
Maximum severity	Die material	Hardness, HB	Die material	Hardness, HB
Hammer forging				
Part 1	H11	405–433	H12	405–448
	6G, 6F2	341–375		
Part 2	6G, 6F2	341–375	6G, 6F2	341–375
			H12(a)	405–448
Part 3	6G, 6F2	269–293	6G, 6F2	302–331
Part 4	H11	405–433	H11	405–433
Part 5	6G, 6F2	302–331	6G, 6F2(b)	302–331
Press forging				
Part 1	H12	477–514	H12	477–514
	6G, 6F2	341–375		
Part 2	6G, 6F2	341–375	H12	477–514
Part 3	Part normally is not press forged from copper alloys			
Part 4	H11	405–433	6G, 6F2(c)	341–375
Part 5	6G, 6F2	341–375	H12	477–514

(a) Recommended for long runs—for example, 50,000 pieces. (b) With either steel, use H12 insert at 405–448 HB. (c) With either steel, use H12 insert at 429–448 HB

ing rectangular-section bar stock and sawing slugs from it.

Upset forging is used less frequently for copper alloys than for steels, primarily because copper alloys are so easily extruded. A part having a long shaftlike section and a larger-diameter head can often be made at less cost by extruding the smaller cross section from a larger one than by starting with a small cross section and upsetting to obtain the head.

In the upsetting of copper alloys, the same rule applies for maximum unsupported length as is used for steels, that is, not more than three times stock diameter. For the forging of brass, single-blow upsetting as severe as 3 to 1 (upset three times starting diameter) is considered reasonable. In practice, however, upsets of this severity are rare. The degree of allowable upset for other copper alloys is somewhat less than that for forging brass, generally in proportion to forgeability (Table 19). In most designs, the amount of upset can be reduced by using slugs cut from specially shaped extrusions or by using one or more blocking impressions in the forging sequence.

Ring rolling is sometimes used as a means of saving material when producing ring gears or similar ringlike parts. Temperatures are the same as those for forging the same alloy in closed dies. Cost usually governs the minimum practical size for ring rolling. Most rings up to 300 mm (12 in.) in outside diameter are more economically produced in closed dies. However, if the face width is less than about 25 mm (1 in.) it is often less expensive to produce rings no larger than 200 mm (8 in.) in outside diameter by the rolling technique. The alloy being forged is also a factor in selecting ring rolling or closed-die forging. For example, alloys such as beryllium copper that are difficult to forge are better adapted to ring rolling. For these alloys, ring rolling is sometimes used for sizes smaller than the minimum practical for the more easily forged alloys.

Magnesium Alloys

Magnesium and its alloys, like other ductile metals, can be formed by any of the bulk working processes, including forging, extrusion, and rolling. Because of its hexagonal crystal structure, magnesium has relatively low cold workability. Magnesium alloys also are good conductors of heat; therefore, dies must be heated to temperatures not much lower than those used to heat the stock (Table 22).

The forgeability of magnesium alloys depends on three factors: the solidus temperature of the alloy, the deformation rate, and the grain size. Only forging-grade billet or bar stock should be used in order to ensure good workability. This type of product has been conditioned and inspected to eliminate surface defects that could open during forging, and it has been homogenized by the supplier to ensure good forgeability. Table 22 lists commonly forged magnesium alloys along with their forging temperatures.

Magnesium alloys are often forged within 55 °C (100 °F) of their solidus temperature. An exception is the high-zinc alloy ZK-60, which sometimes contains small amounts of the low-melting eutectic that forms during ingot solidification. Forging of this alloy above about 315 °C (600 °F)—the melting point of the eutectic—can cause severe rupturing. This problem can be minimized by holding the cast ingot for extended periods at an elevated temperature to redissolve the eutectic and to restore a higher solidus temperature.

Hydraulic presses or slow-action mechanical presses are the most commonly used machines for the open-die and closed-die forging of magnesium alloys. In these machines, magnesium alloys can be forged with small corners and fillets and with thin web or panel sections. Corner radii of 1.6 mm ($^1/_{16}$ in.), fillet radii of 4.8 mm ($^3/_{16}$ in.), and panels or webs 3.2 mm ($^1/_8$ in.) thick are not uncommon. The draft angles required for extraction of the forgings from the dies can be held to 3° or less.

Magnesium alloys are seldom hammer forged or forged in a rapid-action press because they will crack unless exacting procedures are used. Alloys ZK60A, AZ31B, and HM21A are more easily forged by these methods than AZ80A, which is extremely difficult to forge. Cracking can occur also in moderately severe, unsupported bending. Magnesium alloys generally flow laterally rather than longitudinally. This characteristic must be considered in the design of tools.

An important factor in the forging of magnesium alloys is to refine the grain size. Alloys that are subject to rapid grain growth at forging temperatures (AZ31B, AZ61A, and AZ80A) are generally forged at successively lower temperatures for each operation. Common practice is to reduce the temperature about 15 to 20 °C (25 to 35 °F) after each step. For parts containing regions that receive only small reductions, all forging is often done at the lowest practical temperature to permit strain hardening. Grain growth in ZK60A and HM21A is slow at forging temperatures, and there is little risk of extensive grain growth.

Magnesium alloy forgings are water quenched directly from the forging operation to

Table 22 Recommended forging temperature ranges for magnesium alloys

	Recommended forging temperature(a)			
	Workplace		Forging dies	
Alloy	°C	°F	°C	°F
Commercial alloys				
ZK21A	300–370	575–700	260–315	500–600
AZ61A	315–370	600–700	290–345	550–650
AZ31B	290–345	550–650	260–315	500–600
High-strength alloys				
ZK60A	290–385	550–725	205–290	400–550
AZ80A	290–400	550–750	205–290	400–550
Elevated-temperature alloys				
HM21A	400–525	750–975	370–425	700–800
EK31A	370–480	700–900	345–400	650–750
Special alloys				
ZEA42A	290–370	550–700	300–345	575–650
ZE62	300–345	575–675	300–345	575–675
QE22A	345–385	650–725	315–370	600–700

(a) The strain-hardening alloys must be processed on a declining temperature scale within the given range to preclude recrystallization prevent further recrystallization and grain growth. With some of the age-hardening alloys, the quench retains the hardening constituents in solution so that they are available for precipitation during subsequent aging treatments.

Nickel-Base Alloys (Ref 33)

Most nickel-base alloys are stronger and stiffer than steel, but alloy 200 (UNS N02200) and alloy 400 (UNS N04400) are softer than many steels. For example, Table 23 lists the pressures developed in the roll gap at 20% reduction in hot rolling for five nickel-base alloys and two steels at four hot-working temperatures. Higher pressures indicate greater resistance to hot deformation. Sufficiently powerful equipment is of particular importance when forging alloys 800 (UNS N08800), 600 (UNS N06600), 625 (UNS N06625), and the precipitation-hardenable alloys such as 718 (UNS N07718) and X-750 (UNS N07750). These alloys were specifically developed to resist deformation at elevated temperatures.

General Guidelines for the Breakdown of Nickel-Base Alloys (Ref 34). Because of their high alloy content and generally narrow working temperature range, nickel-base alloys must be converted from cast ingots with care. Initial breakdown operations are generally conducted well above the γ' solvus temperature, with subsequent deformation completed below it but still high enough to avoid excessive warm working and an unrecrystallized microstructure. The original cast structure must be completely refined during breakdown, that is, before final forging, particularly when substantial levels of reduction not imposed during closed-die forging.

Good heat-retention practice during ingot breakdown is an important factor in obtaining a desirable billet microstructure. Rapid transfer of the ingot from the furnace to the forging press, as well as the use of such techniques as reheating during breakdown, is necessary to promote sufficient recrystallization during each forging pass. In addition, it has been found that diffusion of precipitation-hardening elements is associated with recrystallization during ingot conversion. Mechanical factors such as cycling speed (which affects heat losses), reduction, length of pass, die design, and press capacity all influence the degree of work penetration through the billet cross section and therefore the rate of ingot conversion.

General Guidelines for the Finish Forging of Nickel-Base Alloys. Figure 17 shows the temperature ranges for the safe forging of 12 nickel-base alloys. Use of the lower part of the temperature range may be required for the development of specific mechanical properties.

Closed-die forging of nickel-base alloys is generally done below the γ' solvus temperature in order to avoid excessive grain growth. Approximately 80% of the reduction is scheduled in the recrystallization temperature range, with the remaining 20% done at lower temperatures to introduce a certain amount of warm work for improved mechanical properties. Preheating of all tools and dies to about 260 °C (500 °F) is recommended to avoid chilling the metal during working.

Forging Rate. A very rapid rate of forging often causes heat buildup (due to friction and deformation heating), a nonuniform recrystallized grain size, and mechanical-property variations. Susceptibility to free-surface ruptures also increases with forging rate (and forging temperature). Therefore, slow strain rates are typically used during the initial closed-die reductions of such alloys as Astroloy (UNS N13017) and René 95 (Ni-14Cr-8Co-3.5Mo-3.5W-3.5Nb-3.5Al-2.5Ti). With proper selection of starting stock and forging temperature, however, the forging rate is less critical. For example, some Astroloy turbine components are currently hammer forged.

Forging Reduction. A sufficient amount of recrystallization is necessary in each of a series of closed-die forging operations to achieve the desired grain size and to reduce the effects of the continuous grain-boundary or twin-boundary carbide networks that develop during heating and cooling. This condition contributes more to mechanical-property and other problems than any other single factor. Poor weldability, low-cycle fatigue, and stress-rupture properties are associated with continuous grain-boundary carbide networks. Heat treatment can do very little to correct this problem without creating equally undesirable mechanical-property problems when higher solution-treatment temperatures are used. All portions of a part must receive some hot work after the final heating operation in order to achieve uniform mechanical properties.

In open-die forging, a series of moderate reduction passes along the entire length of the forging is preferred. In working a square section into a round, the piece should be worked down in the square form until it approaches the final size. It should then be converted to an oversize octagon before finishing into the round. Billet corners that will be in contact with dies should be chamfered rather than left square. The work should be lifted away from the dies occasionally to permit relief of local cold areas.

Other Considerations. The precipitation-hardenable nickel alloys are subject to thermal cracking. Therefore, localized heating is not recommended. The entire part should be heated to the forging temperature.

If any ruptures appear on the surface of the metal during hot working, they must be removed at once, either by hot grinding or by cooling the work and cold overhauling. If the ruptures are not removed, they may extend into the body of the part.

For sections equal to or larger than 400 mm (16 in.) square, precautions should be taken in heating precipitation-hardenable alloys. They should be charged into a furnace at 870 °C (1600 °F) or colder and brought up to forging temperature at a controlled rate of 40 °C/h (100 °F/h).

Cooling after Forging. The rate of cooling after forging is not critical for alloys 200, 400, and 625. Alloys K-500 (UNS N05500) and 301 (UNS N03301) should be water quenched from forging temperatures to avoid the excessive hardening and cracking that could occur if they were cooled slowly through the age-hardening range and to maintain good response to subsequent aging. Alloy 825 (UNS N08825) should be cooled at a rate equal to or faster than air cooling.

Alloys 800 and 600 are subject to carbide precipitation during heating in or slow cooling through the temperature range of 540 to 760 °C (1000 to 1400 °F). If sensitization is likely to prove disadvantageous in the end use, parts made of these alloys should be water quenched or cooled rapidly in air.

The precipitation-hardenable alloys should, in general, be cooled in air after forging. Water quenching is not recommended, because of the possibility of thermal cracking, which can occur during subsequent heating for further forging or heat treating.

Forging Practice for Specific Alloys. The following practices are used in the forging of nickel-base alloys. However, as noted in the section on "Superalloys," state-of-the-art forging practices for nickel-base alloys depend on thermomechanical process design for the development of desirable properties and microstructures. Variations from these procedures may be necessary for some specialized applications (see the sections "Thermomechanical Processing" and "Isother-mal Forging" in this article).

Alloy 200 should be charged to a hot furnace, withdrawn as soon as the desired temperature has been reached, and worked rapidly. The recommended range of forging temperatures is 650

Table 23 Hot-forming pressures for several nickel-base alloys

Pressures developed in the hot forming of 1020 steel and AISI type 302 stainless steel are shown for comparison.

		Pressure developed at working temperature(a)							
		870 °C (1800 °F)		1040 °C (1900 °F)		1095 °C (2000 °F)		1150 °C (2100 °F)	
Alloy	UNS No.	MPa	ksi	MPa	ksi	MPa	ksi	MPa	ksi
400	N04400	124	18	106	15.3	83	12	68	9.8
600	N06600	281	40.8	239	34.6	195	28.3	154	22.3
625	N06625	463	67.2	379	55	297	43	214	31
718	N07718	437	63.3	385	55.8	333	48.3	283	41
X-750	N07750	335	48.6	299	43.3	265	38.4	230	33.3
1020 steel	G10200	154	22.4	126	18.3	99	14.3	71	10.3
Type 302 stainless steel	S30200	192	27.8	168	24.3	148	21.4	124	18

(a) Pressure developed in the roll gap at 20% reduction in hot rolling

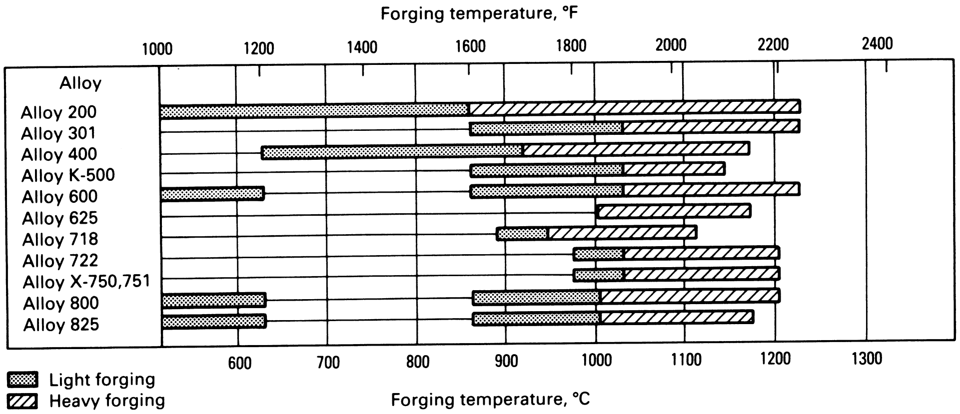


Fig. 17 Forging temperature ranges for 12 nickel-base alloys

to 1230 °C (1200 to 2250 °F). Because the metal stiffens rapidly when cooled to about 870 °C (1650 °F), all heavy work and hot bending should be done above that temperature. High mechanical properties can be produced by working lightly below 650 °C (1200 °F). The best range for hot bending is 870 to 1230 °C (1600 to 2250 °F).

Alloy 301 has an optimal forging range of 1065 to 1230 °C (1900 to 2250 °F). Light finishing work can be done down to 870 °C (1600 °F). Finer grain size is produced in forgings by using 1175 °C (2150 °F) for the final reheat temperature and by taking at least 30% reduction of area in the last forging operation.

After hot working, the alloy should be quenched from a temperature of 790 °C (1450 °F) or above. Quenching retains the strain hardening imparted by the forging operation and produces better response to subsequent age hardening. Quenching in water containing about 2 vol% alcohol results in less surface oxidation.

Material that must be cooled prior to subsequent hot working should also be quenched. Slow cooling may cause age hardening, which sets up stresses in the workpiece that can cause cracking during subsequent reheating.

Alloy 400 has a maximum heating temperature for forging of 1175 °C (2150 °F). Prolonged soaking at the working temperature is detrimental. If a delay occurs during processing, the furnace temperature should be reduced to 1040 °C (1900 °F) and not brought to 1175 °C (2150 °F) until operations are resumed.

The recommended metal temperature for heavy reductions is 925 to 1175 °C (1700 to 2150 °F). Light reductions may be taken at temperatures down to 650 °C (1200 °F). Working at the lower temperatures produces higher mechanical properties and smaller grain size.

A controlled forging procedure is necessary to meet the requirements of some specifications for forged hot-finished parts. Both the amount of reduction and the finishing temperature must be controlled in order to develop the desired properties.

One procedure for producing forgings to such specifications consists of taking a 30 to 35% reduction after the final reheat. This is done by:

- Reheat.
- Forge to a section having about 5% larger area than the final shape (take at least 25% reduction).
- Cool to 705 °C (1300 °F).
- Finish to size (5% reduction).

High-tensile forgings, as described in certain military specifications, also require a minimum of 30 to 35% reduction after the last reheat. This is taken in the following manner:

- Reheat.
- Forge to a section having an area about 25% larger than the final shape (take about 5% reduction).
- Cool to 705 °C (1300 °F).
- Finish to size (25% reduction).
- Grain refinement is achieved by using a temperature of 1095 °C (2000 °F) for the final reheat and by increasing the amount of reduction taken after the last reheat.

Alloy K-500 has a maximum recommended heating temperature for forging of 1150 °C (2100 °F). Metal should be charged into a hot furnace and withdrawn when uniformly heated. Prolonged soaking at this temperature is harmful. If a delay occurs such that the material would be subject to prolonged soaking, the temperature should be reduced to or held at 1040 °C (1900 °F) until shortly before working is to begin, then brought to 1150 °C (2100 °F). When the piece is uniformly heated, it should be withdrawn. In the event of a long delay, the work should be removed from the furnace and water quenched.

The forging range is 870 to 1150 °C (1600 to 2100 °F). Heavy work is best done between 1040 and 1150 °C (1900 and 2100 °F), and working below 870 °C (1600 °F) is not recommended. To produce finer grain in forgings, 1095 °C (2000 °F) should be used for the final reheat temperature, and at least 30% reduction of area should be taken in the last forging operation.

When forging has been completed or when it is necessary to allow alloy K-500 to cool before further hot working, it should not be allowed to cool in air, but should be quenched from a temperature of 790 °C (1450 °F) or higher. If the piece is allowed to cool slowly, it will age harden to some extent, and stress will be set up that may lead to thermal splitting or tearing during subsequent reheating. In addition, quenched material has better response to age hardening because more of the age-hardening constituent is retained in solution.

Alloy 600 has a normal forging range of 870 to 1230 °C (1600 to 2250 °F). Heavy hot work should be done in the range from 1040 to 1230 °C (1900 to 2250 °F). Light working can be continued down to 870 °C (1600 °F). Generally, forging should not be done between 650 and 870 °C (1200 and 1600 °F) because of the low ductility of the alloy in this temperature range. Judicious working at a temperature below 650 °C (1200 °F) will develop higher tensile properties.

The rate of cooling after forging is not critical with respect to thermal cracking. However, alloy 600 is subject to carbide precipitation in the range between 540 and 760 °C (1000 and 1400 °F), and if subsequent use dictates freedom from sensitization, the part should be rapidly cooled through this temperature range.

Alloy 625 should be heated in a furnace held at 1175 °C (2150 °F) but no higher. The work should be brought as close to this temperature as conditions permit. Forging is done from this temperature down to 1010 °C (1850 °F); below 1010 °C (1850 °F) the metal is stiff and hard to move, and attempts to forge it may cause hammer splits at the colder areas. The work should be returned to the furnace and reheated to 1175 °C (2150 °F) whenever its temperature drops below 1010 °C (1850 °F). To guard against duplex grain structure, the work should be given uniform reductions. For open-die work, final reductions of a minimum of 20% are recommended.

Alloy 718 is strong and offers considerable resistance to deformation during forging. The forces required for hot deformation are somewhat higher than those employed for alloy X-750. Alloy 718 is forged in the range from 900 to 1120 °C (1650 to 2050 °F). In the last operation, the metal should be worked uniformly with a gradually decreasing temperature, finishing with some light reduction below 955 °C (1750 °F). In heating for forging, the material should be brought up to temperature, allowed to soak a short time to ensure uniformity, and withdrawn.

Alloy 718 should be given uniform reductions in order to avoid duplex grain structure. Final reductions of 20% minimum should be used for open-die work, and 10% minimum for closed-die work. Parts should generally be air cooled from the forging temperature, rather than water quenched.

Alloy 706 (UNS N09706) is similar to alloy 718, except that alloy 706 is more readily fabricated, particularly by machining. Forging should be done using the same procedures and temperatures as those for alloy 718.

Alloy X-750 has a forging range of 980 to 1205 °C (1800 to 2200 °F). Below 980 °C (1800 °F), the metal is stiff and hard to move, and attempts to work it may cause splitting. All heavy forging should be done at about 1040 °C (1900 °F), and the metal should be reheated whenever it cools to below that temperature. Forgings can be finished with some light reduction between 980 and 1040 °C (1800 and 1900 °F).

As a general rule, alloy X-750 should be air cooled rather than liquid quenched from the forging temperature. Liquid quenching can cause high residual stresses that may result in cracking during subsequent heating for further hot work or for heat treatment. Parts with large cross sections and pieces with variable cross sections are especially susceptible to thermal cracking during cooling. In very large cross sections, furnace cooling may be necessary to prevent thermal cracking.

Alloy 800. Hot working of alloy 800 starts at 1205 °C (2200 °F), and heavy forging is done at temperatures down to 1010 °C (1850 °F). Light working can be accomplished down to 870 °C (1600 °F). No working should be done between 870 and 650 °C (1600 and 1200 °F). As with alloy 600, thermal cracking is not a problem, and workpieces should be cooled rapidly through the range between 540 and 760 °C (1000 and 1400 °F) to ensure freedom from sensitization.

Alloy 825 has a forging range of 870 to 1175 °C (1600 to 2150 °F). It is imperative that some reduction be accomplished in the range between 870 and 980 °C (1600 and 1800 °F) during final forging to ensure maximum corrosion resistance.

Cooling after forging should be done at a rate equal to or faster than air cooling. Heavy sections may become sensitized during cooling from the forging temperature and therefore be subject to intergranular corrosion in certain media. A stabilizing anneal of 1 h at 940 °C (1725 °F) restores resistance to corrosion. If the forged piece is to be welded and used in an environment that could cause intergranular corrosion, the piece should be given a stabilizing anneal to prevent sensitization from the heat of welding, regardless of the cooling rate after forging.

Alloy 925. The hot-working characteristics of alloy 925 (UNS N09925) are similar to those of alloy 825 at temperatures to 1095 °C (2000 °F). At higher temperatures, alloy 925 has lower ductility and higher strength. The forging range is 870 to 1175 °C (1600 to 2150 °F). For maximum corrosion resistance and highest mechanical properties after direct aging, final hot working should be done in the range of 870 to 980 °C (1600 to 1800 °F).

Alloys 722 and 751 (UNS N07722 and N07751, respectively) are forged using the same procedures and temperatures as those for alloy X-750.

Alloys 903, 907, and 909 (UNS N19903, N19907, and N19909, respectively) are best forged in three stages in order to obtain the desired properties after aging. The initial breakdown of 40% minimum reduction should be performed at 1060 to 1120 °C (1940 to 2050 °F). For intermediate forging at a minimum of 25% reduction, these alloys should be heated between 995 and 1050 °C (1825 and 1925 °F). The final heating for alloys 907 and 909 should be 980 to 1025 °C (1800 to 1875 °F) for a minimum reduction of 20% over a falling temperature range (finishing at ≤925 °C, or 1700 °F). The final heating for alloy 903 should be 870 °C (1600 °F) with a final forging reduction of 40% minimum.

Thermomechanical processing (TMP) refers to the control of temperature and deformation during processing to enhance specific properties. Special TMP sequences have been developed for a number of nickel-base alloys, and the design of TMP sequences relies on a knowledge of the melting and precipitation temperatures for the precipitates in the alloy of interest. Precipitates in nickel-base (as well as iron- and cobalt-base) alloys include various types of metal carbides such as *M*C carbides (*M* = titanium, niobium, etc.), M_6C carbides (*M* = molybdenum and/or tungsten), or $M_{23}C_6$ carbides (*M* = chromium). However, the primary precipitate of concern in the processing of precipitation-strengthened nickel-base alloys is the γ′ strengthening precipitate. γ′ is an ordered face-centered cubic (fcc) compound in which aluminum and titanium combine with nickel to form $Ni_3(Al,Ti)$. In nickel-iron alloys such as alloy 718, titanium, niobium, and, to a lesser extent, aluminum combine with nickel to form ordered fcc γ′ or ordered body-centered tetragonal γ″. Nickel-iron-base alloys are also prone to the formation of other phases, such as hexagonal close-packed Ni_3Ti (η), as in titanium-rich alloy 901, or orthorhombic Ni_3Nb (δ) in niobium-rich alloy 718.

Table 24 lists first melting and precipitation temperatures for several nickel-base alloys. Early forging practice of nickel- and nickel-iron-base alloys consisted of forging from and solution heat treating at temperatures well in excess of the γ′ solvus temperature. High-temperature solution treatment dissolved all of the γ′, annealed the matrix, and promoted grain growth (typical grain size ≈ ASTM 3 or coarser). This was followed by one or more aging treatments that promoted controlled precipitation of γ′ and carbide phases. Optimal creep and stress-rupture properties above 760 °C (1400 °F) were thus achieved. Later in the development of forging practice, it was found that using preheat furnace temperatures slightly above the recrystallization temperature led to the development of finer grain sizes (ASTM 5 to 6). Coupling this with modified heat treating practices resulted in excellent combinations of tensile, fatigue, and creep properties.

Forging temperature is carefully controlled during the thermomechanical processing of nickel- and nickel-iron-base alloys to make use of the structure control effects of second phases such as γ′. Above the optimal forging temperature range (Table 25), the structure-control phase goes into solution and loses its effect. Below this range, extensive fine precipitates are formed, and the alloy becomes too stiff to process.

Some examples of specific TMP sequences are given below. State-of-the-art TMP practices for nickel-base alloys rely on the following microstructural effects (Ref 2):

- Dynamic recrystallization is the most important softening mechanism during hot working.
- Grain boundaries are preferred nucleation sites for recrystallization.
- The rate of recrystallization decreases with the temperature and/or the extent of deformation.
- Precipitation that may occur during the recrystallization can inhibit the softening process.
- Recrystallization cannot be completed until the precipitate coarsens to a relatively ineffective morphology.

Waspaloy. A typical TMP treatment of nickel-base alloys is that used for Waspaloy (UNS N07001) to obtain good tensile and creep properties. This consists of initial forging at 1120 °C (2050 °F) and finish forging below approximately 1010 °C (1850 °F) to produce a fine, equiaxed grain size of ASTM 5 to 6. Solution treatment is then done at 1010 °C (1850 °F), and aging is conducted at 845 °C (1550 °F) for 4 h, followed by air cooling plus 760 °C (1400 °F) for 16 h and then air cooling.

René 95. Initial forging of René 95 is done between 1095 and 1140 °C (2000 and 2080 °F). Following an in-process recrystallization anneal at 1175 °C (2150 °F), finish forging (reduction ≈ 40 to 50%) is then imposed below the γ′

Table 24 Critical melting and precipitation temperatures for several nickel-base alloys

Alloy	UNS No.	First melting temperature, °C	First melting temperature, °F	Precipitation temperature, °C	Precipitation temperature, °F
Alloy X	N06002	1260	2300	760	1400
Alloy 718	N07718	1260	2300	845	1550
Waspaloy	N07001	1230	2250	980	1800
Alloy 901	N09901	1200	2200	980	1800
Alloy X-750	N07750	1290	2350	955	1750
M-252	N07252	1200	2200	1010	1850
Alloy R-235	. . .	1260	2300	1040	1900
René 41	N07041	1230	2250	1065	1950
U500	N07500	1230	2250	1095	2000
U700	. . .	1230	2250	1120	2050
Astroloy	N13017	1230	2250	1120	2050

Source: Ref 35

Table 25 Structure control phases and working temperature ranges for various heat-resistant alloys

Alloy	UNS No.	Phases for structure control	Working temperature range °C	°F
Nickel-base alloys				
Waspaloy	N07001	γ' (Ni_3(Al,Ti)	955–1025	1750–1875
Astroloy	N13017	γ' (Ni_3(Al,Ti)	1010–1120	1850–2050
IN-100	. . .	γ' (Ni_3(Al,Ti)	1040–1175	1900–2150
René 95	. . .	γ' (Ni_3(Al,Ti)	1025–1135	1875–2075
Nickel-iron-base alloys				
901	N09901	η (Ni_3Ti)	940–995	1725–1825
718	N07718	δ (Ni_3Nb)	915–995	1675–1825
Pyromet CTX-1	. . .	η (Ni_3Ti), δ (Ni_3Nb), or both	855–915	1575–1675

Source: Ref 36

solvus, typically between 1080 and 1105 °C (1975 and 2025 °F). The large grains formed during high-temperature recrystallization are elongated and surrounded by small recrystallized grains that form during finish forging.

Alloy 901. The thermomechanical processing of alloy 901 is often done to produce a fine-grain structure that enhances fatigue strength. This is accomplished by using the η (Ni_3Ti) phase, which is introduced in a Widmanstätten form at the beginning of processing by a heat treatment at 900 °C (1650 °F) for 8 h. Forging is then conducted at 955 °C (1750 °F), which is below the η solvus; the forging deformation is completed below the recrystallization temperature. A fine-grain structure is generated by a subsequent recrystallization treatment below the η solvus. The needlelike η phase will become spherical during forging and will restrict grain growth. Aging is then conducted according to standard procedures.

Isothermal Forging. Nickel-base alloys that are hard to work or are typically used in the cast condition can be readily forged when in a powder-consolidated form. The most common forging technique using powder preforms is isothermal forging. In this process, powder is produced by inert gas atomization and is compacted into billet form by extrusion. The billets are fabricated below the γ' solvus temperature for alloys such as IN-100 in order to maintain a fine grain size and a fine distribution of precipitates. In this condition, the material exhibits superplastic properties that are characterized by large tensile elongations (during sheet forming) and good die-filling capacity (during forging). Multiples of the extruded bar are then isothermally forged into a variety of complex turbine engine and other high-temperature parts.

The key to successful isothermal forging of nickel-base alloys is the ability to develop a fine grain size before forging and to maintain it during forging. With regard to the latter, a high volume percentage of second phase is useful in preventing grain growth. Therefore, alloys such as IN-100, René 95, and Astroloy, which contain large amounts of γ', are readily capable of developing the superplastic properties necessary in isothermal forging. In contrast, Waspaloy, which contains less than 25 vol% γ' at isothermal forging temperatures, is only marginally superplastic. Nickel-iron-base alloys such as alloys 718 and 901 have even lower volume fractions of precipitate and are therefore even less frequently used in isothermal forging.

As the term implies, isothermal forging consists of forging with the workpiece and the dies at the same temperature. Because this temperature is often of the order of 980 to 1095 °C (1800 to 2000 °F), the dies are usually made of molybdenum for elevated-temperature strength. The isothermal forging system must be operated in a vacuum or inert atmosphere to protect such die materials from oxidation.

Compared to conventional forging, isothermal forging deformation rates are slow; hydraulic press speeds of approximately 2.5 mm/min (0.1 in./min) are typical. However, the slower production rate is largely offset by the ability to forge complex shapes to closer tolerances, which leads to less machining and substantial material savings. In addition, a large amount of deformation is accomplished in one operation, pressures are low, and uniform microstructures are achieved. For example, the as-forged weight of a finish-machined 68 kg (150 lb) Astroloy disk is about 110 kg (245 lb) for a conventional forging versus 72 kg (160 lb) for the corresponding isothermal forging.

Titanium Alloys (Ref 37)

As a class of materials, titanium alloys are among the most difficult metal alloys to forge, ranking behind only refractory metals and nickel-cobalt-base superalloys. Therefore, titanium alloy forgings, particularly closed-die forgings, are typically produced to less highly refined final forging configurations than are typical of aluminum alloys. However, precision titanium alloy forgings can be produced to the same design and tolerance criteria as aluminum alloy precision forgings. Titanium alloy forgings also may be equivalent to, or more refined than, carbon or low-alloy steel forgings, because of reduced oxidation or scaling tendencies in heating.

Thermomechanical processing is another important aspect in forging of titanium alloys. Working history and forging parameters used in titanium alloy forging have a significant impact on the final microstructure (and therefore the resultant mechanical properties) of the forged titanium alloys—perhaps to a greater extent than in any other commonly forged material. Therefore, the forging process in titanium alloys is used not only to create cost-effective forging shapes, but also, in combination with thermal treatments, to create unique and/or tailored microstructures to achieve the desired final mechanical properties. By the design of the working process history from ingot to billet to forging, and particularly the selection of metal temperatures and deformation conditions during the forging process, significant changes in the morphology of the allotropic (α, β) phases of titanium alloys are achieved that in turn dictate the final mechanical properties and characteristics of the alloy. The classes of titanium alloys (α, $\alpha + \beta$, β) exert a strong influence on forging part design and forging-process selection.

Classes of Titanium Alloys. Titanium and its alloys exist in two allotropic forms: the hexagonal close-packed (hcp) α phase and the body-centered cubic (bcc) β phase. The more difficult to deform α phase is usually present at low temperatures, while the more easily deformed β phase is present at high temperatures. However, the addition of various alloying elements (including other metals and such gases as oxygen, nitrogen, and hydrogen) stabilizes either the α or β phase. The temperature at which a given titanium alloy transforms completely from α to β is termed the beta transus (β_t). The β_t is a critical temperature in titanium alloy forging-process criteria.

Based on the predominant allotropic form(s) present at room temperature, titanium alloys are divided into three major classes:

- α/near-α alloys
- $\alpha + \beta$ alloys
- β/metastable β alloys

Table 26 lists most of the commonly forged titanium alloys by alloy class, along with the major alloying elements constituting each alloy. Each of these types of titanium alloys has unique forging-process criteria and deformation behavior. Further, the forging-process parameters, often in combination with subsequent thermal treatments, are manipulated for each alloy type to achieve the desired final forging microstructure and mechanical properties (heat treatment serves a different purpose in titanium alloys from that in aluminum alloys or alloy steels, as discussed in the paragraphs that follow).

Alpha/near-α titanium alloys contain elements that stabilize the hcp α phase at higher temperatures. These alloys (with the exception of commercially pure titanium, which is also an α alloy) are among the most difficult titanium alloys to forge. Typically, α/near-α titanium alloys have modest strength but excellent elevated-temperature properties. Forging and TMP processes for α alloys are typically designed to develop optimal elevated-temperature properties, such as strength and creep resistance. The β_t of α/near-α alloys typically ranges from 900 to 1065 °C (1650 to 1950 °F).

Alpha-beta titanium alloys represent the most widely used class of titanium alloys (Ti-6Al-4V

Table 26 Recommended forging temperature ranges for commonly forged titanium alloys

Alloy	β_t °C	β_t °F	Process(a)	Forging temperature(b) °C	Forging temperature(b) °F
α/near-α alloys					
Ti-C.P.(c)	915	1675	C	815–900	1500–1650
Ti-5Al-2.5Sn(c)	1050	1925	C	900–1010	1650–1850
Ti-5Al-6Sn-2Zr-1Mo-0.1Si	1010	1850	C	900–995	1650–1925
Ti-6Al-2Nb-1Ta-0.8Mo	1015	1860	C	940–1050	1725–1825
			B	1040–1120	1900–2050
Ti-6Al-2Sn-4Zr-2Mo(+0.2Si)(d)	990	1815	C	900–975	1650–1790
			B	1010–1065	1850–1950
Ti-8Al-1Mo-1V	1040	1900	C	900–1020	1650–1870
IMI 685 (Ti-6Al-5Zr-0.5Mo-0.25Si)(e)	1030	1885	C/B	980–1050	1795–1925
IMI 829 (Ti-5.5Al-3.5Sn-3Zr-1Nb-0.25Mo-0.3Si)(e)	1015	1860	C/B	980–1050	1795–1925
IMI 834 (Ti-5.5Al-4.5Sn-4Zr-0.7Nb-0.5Mo-0.4Si-0.06C)(e)	1010	1850	C/B	980–1050	1795–1925
α-β alloys					
Ti-6Al-4V(c)	995	1825	C	900–980	1650–1800
			B	1010–1065	1850–1950
Ti-6Al-4V ELI	975	1790	C	870–950	1600–1740
			B	990–1045	1815–1915
Ti-6Al-6V-2Sn	945	1735	C	845–915	1550–1675
Ti-6Al-2Sn-4Zr-6Mo	940	1720	C	845–915	1550–1675
			B	955–1010	1750–1850
Ti-6Al-2Sn-2Zr-2Mo-2Cr	980	1795	C	870–955	1600–1750
Ti-17 (Ti-5Al-2Sn-2Zr-4Cr-4Mo(f)	885	1625	C	805–865	1480–1590
			B	900–970	1650–1775
Corona 5 (Ti-4.5Al-5Mo-1.5Cr)	925	1700	C	845–915	1550–1675
			B	955–1010	1750–1850
IMI 550 (Ti-4Al-4Mo-2Sn)	990	1810	C	900–970	1650–1775
IMI 679 (Ti-2Al-11Sn-4Zr-1Mo-0.25Si)	945	1730	C	870–925	1600–1700
IMI 700 (Ti-6Al-5Zr-4Mo-1Cu-0.2Si)	1015	1860	C	800–900	1470–1650
β/near-β/metastable β alloys					
Ti-8Al-8V-2Fe-3Al	775	1425	C/B	705–980	1300–1800
Ti-10V-2Fe-3Al	805	1480	C	705–785	1300–1450
			B	815–870	1500–1600
Ti-13V-11Cr-3Al	675	1250	C/B	650–955	1200–1750
Ti-15V-3Cr-3Al-3Sn	770	1415	C/B	705–925	1300–1700
Beta C (Ti-3Al-8V-6Cr-4Mo-4Zr)	795	1460	C/B	705–980	1300–1800
Beta III (Ti-4.5Sn-6Zr-11.5Mo)	745	1375	C/B	705–955	1300–1750
Transage 129 (Ti-2Al-11.5V-2Sn-11Zr)	720	1325	C/B	650–870	1200–1600
Transage 175 (Ti-2.7Al-13V-7Sn-2Zr)	760	1410	C/B	705–925	1300–1700

(a) C, conventional forging processes in which most or all of the forging work is accomplished below the β_t of the alloy for the purposes of desired mechanical property development. This forging method is also referred to as α-β forging. B, β forging processes in which some or all of the forging is conducted above the β_t of the alloy to improve hot workability or to obtain desired mechanical property combinations. C/B, either forging methodology (conventional or β) is employed in the fabrication of forgings or for alloys, such as β alloys, that are predominately forged above their β_t but may be finish forged at subtransus temperatures. (b) These are recommended metal temperature ranges for conventional α-β, or β forging processes for alloys for which the latter techniques are reported to have been employed. The lower limit of the forging temperature range is established for open-die forging operations in which reheating is recommended. (c) Alloys for which there are several compositional variations (primarily oxygen or other interstitial element contents) that may affect both β_t and forging temperature ranges. (d) This alloy is forged and used both with and without the silicon addition; however, the β_t and recommended forging temperatures are essentially the same. (e) Alloys designed to be predominately β forged. (f) Ti-17 has been classified as an α-β and as a near-β titanium alloy. For purposes of this article, it is classified as an α-β alloy

is the most widely used of all titanium alloys) and contain sufficient β stabilizers to stabilize some of the β phase at room temperature. Alpha-beta titanium alloys are generally more readily forged than α alloys and are more difficult to forge than some β alloys. Typically, α-β alloys have intermediate-to-high strength with excellent fracture toughness and other fracture-related properties. Forging and TMP processes for α-β alloys are designed to develop optimal combinations of strength, fracture toughness, and fatigue characteristics. The β_t of α-β alloys typically ranges from 870 to 1010 °C (1600 to 1850 °F).

Beta/Metastable β *Alloys* are those alloys with sufficient β stabilizers that the bcc β-phase is the predominant allotropic form present at room temperature. Beta titanium alloys are usually easier to fabricate than other classes of titanium alloys, although β alloys may be equivalent to, or more difficult to forge than, α-β alloys under certain forging conditions. Beta titanium alloys are characterized by very high strength with good fracture toughness and excellent fatigue characteristics; therefore, forging and TMP processes are designed to optimize these property combinations. The β_t of β titanium alloys ranges from 650 to 870 °C (1200 to 1600 °F).

Forging Processes. Titanium alloy forgings are produced by all of the typical mechanical methods of forging, including open-die (or hand) forging, closed-die forging, upsetting, roll forging, orbital forging, spin forging, mandrel forging, ring rolling, and forward and backward extrusion. Selection of the optimal forging method for a given forging shape is based on the desired forging shape, the sophistication of the design of the forged shape, the cost, and the desired mechanical properties and microstructure. In many cases, two or more forging methods are combined to achieve the desired forging shape, to obtain the desired final part microstructure, and/or to minimize cost.

From a metallurgical standpoint, the methods of forging titanium also are classified according to whether the forging temperature is below or above the β_t. The two basic categories are (a) conventional (α-β) forging, where the alloy temperature during forging is predominantly below the β_t and (b) β forging, where the alloy temperature during forging is predominantly above the β_t. These are the two principal metallurgical approaches to the forging of titanium alloys. However, within these fundamental approaches, there are several possible variations that blend these two techniques into processes that are used commercially to achieve controlled microstructures that tailor the final properties of the forging to specification requirements and/or intended service applications. In particular, manipulation of the α phase is important during the conventional (α-β) forging of the α and α-β alloys. By the selection of metal temperatures and deformation conditions, the forging process can have a significant influence on the α-phase morphology, which in turn dictates the mechanical properties and characteristics of the alloy. In fully β-stabilized alloys, manipulation of the α phase through forging-process techniques is less prevalent; therefore, fully β-stabilized alloys are typically forged above the β_t of the alloy.

Table 26 lists recommended metal temperatures for 27 commonly forged α, α-β, and β titanium alloys. With some exceptions, these alloys can be forged to the same degree of severity; however, the power and/or pressure requirements needed to achieve a given forging shape may vary with each individual alloy and particularly with alloy class. As a general guide, metal temperatures of β_t – 28 °C (50 °F) for α-β forging and β_t + 42 °C (75 °F) for beta forging, are recommended.

The upper limits of the temperature ranges in Table 26 are based on (a) prudent proximity (from furnace temperature variations and minor composition variations) to the nominal β_t of the alloy in the case of conventional (α-β) forging and (b) forging without undue metallurgical risks (such as excessive grain growth) in the case of β forging. The lower limit of the specified ranges is the temperature at which forging should be discontinued in the case of open-die forging to avoid excessive cracking and/or other surface-quality problems.

Conventional (α-β*) forging of titanium alloys,* in addition to implying the use of die temperatures of 540 °C (1000 °F) or less, is the term used to describe a forging process in which most or all of the forging deformation is conducted at temperatures below the β_t of the alloy. For α, α-β, and metastable β alloys, this forging technique involves working the material at temperatures where both the α and β phases are present, with the relative amounts of each phase being dictated by the composition of the alloy and the actual temperature used. With this forging technique, the resultant as-forged microstructure is characterized by deformed or equiaxed primary α in a transformed-β matrix. The volume fraction of primary α is dictated by the alloy composition and the actual working history and temperature. Alpha-beta forging is typically used to develop optimal strength/ductility combinations and optimal high/low-cycle fatigue properties. With conventional (α-β) forging, the effects of working on microstructure, particularly α-phase morphology changes, are cumulative; therefore, each successive α-β working operation adds to the structural changes achieved in earlier operations.

Beta forging, as the term implies, is a forging technique for α, α-β, and metastable β alloys in which most or all of the forging work is done at temperatures above the β_t of the alloy. In commercial practice, β forging techniques typically involve supertransus forging in the early and/or intermediate stages with controlled amounts of final deformation below the β_t of the alloy. Actual final subtransus working criteria are dependent on the alloy, the forging design, and the mechanical property combinations sought.

In β forging, the working influences on microstructure are not fully cumulative; with each working-cooling-reheating sequence above β_t, the effects of the prior working operations are at least partially lost because of recrystallization from the transformation upon heating above the β_t of the alloy. Beta forging techniques are used to develop microstructures characterized by Widmanstätten or acicular primary α morphology in a transformed β matrix. This forging process is typically used to enhance fracture-related properties, such as fracture toughness and fatigue crack propagation resistance, and to enhance the creep resistance of α and α-β alloys. In fact, several α alloys are designed to be β forged to develop the desired final mechanical properties. There is often a loss in strength and ductility with β forging as compared to α-β forging. Beta forging, particularly of α and α-β alloys, has the advantages of significant reduction in forging unit pressures and reduced cracking tendency, but it must be done under carefully controlled forging-process conditions to avoid nonuniform working, excessive grain growth, and/or poorly worked structures, all of which can result in final forgings with unacceptable or widely variant mechanical properties within a given forging or from lot to lot of the same forging.

The effect of temperature on the flow stresses of titanium alloys varies with alloy class. This effect is illustrated in Fig. 18 for three alloys, each representative of one class of titanium alloy. For example, it is evident that the more difficult-to-forge α alloy (Ti-8Al-1Mo-1V in Fig. 18a) displays the greatest sensitivity to metal temperature. The flow stress at 10/s and 900 °C (1650 °F) is two to three times that of the alloy at 1010 °C (1850 °F) (the latter temperature is below the β_t of the alloy). In comparison, the α-β alloy (Ti-6Al-4V in Fig. 18b) also displays sensitivity to metal temperature, but to a lesser extent than the Ti-8Al-1Mo-1V alpha alloy, especially at higher levels of total strain. In Fig. 18(b), at 1000 °C (1830 °F), Ti-6Al-4V is being deformed at or above the nominal β_t of the alloy, in which the structure is entirely bcc and considerably easier to deform. Finally, the Ti-10V-2Fe-3Al β alloy displays even less metal temperature sensitivity, also at higher levels of total strain. At 815 °C (1500 °F), Ti-10V-2Fe-3Al is being deformed above its β_t with an attendant reduction in flow stresses in comparison to subtransus deformation at 760 °C (1400 °F). However, at this high strain rate, the flow-stress reduction achieved by deforming the β alloys above its β_t is less than the flow-stress reduction achieved by deforming the α-β alloys above its β_t.

As with other forged materials, many titanium alloys display a strain-softening behavior at the strain rates typically used in conventional forging techniques. As shown in Fig. 18(a) to (c), strain softening is typically observed when such alloys are forged below their β_t. It is observed to a much lesser extent when these alloys are β forged (for example, 1000 °C in Fig. 18b for Ti-6Al-4V and 815 °C in Fig. 18c for Ti-10V-2Fe-3Al). The differences in strain-softening behavior are a function of the differences in microstructure present during the deformation above or below the transus of the alloy. The equiaxed α in a β-matrix structure, typical of subtransus forging, has been found to redistribute strain and to promote dislocation movement more effectively than acicular α in a transformed β structure, leading to increased strain softening in the former.

Flow stresses describe the lower limit of the deformation resistance of titanium alloys as represented by ideal deformation conditions and are therefore rarely present during actual forging

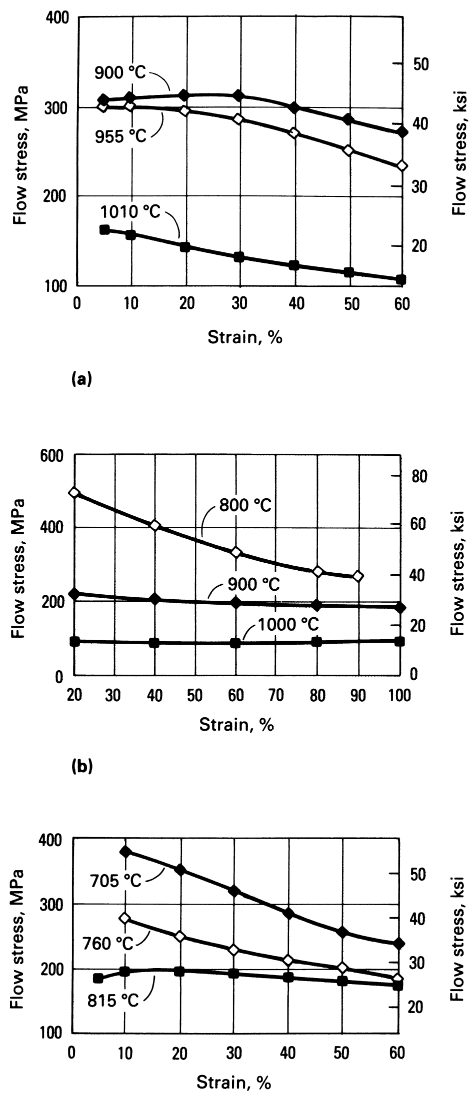


Fig. 18 Effect of forging temperature on flow stress of titanium alloys at 10/s strain rate. (a) α alloy Ti-8Al-1Mo-1V. (b) α-β alloy Ti-6Al-4V. (c) Metastable β alloy Ti-10V-2Fe-3Al

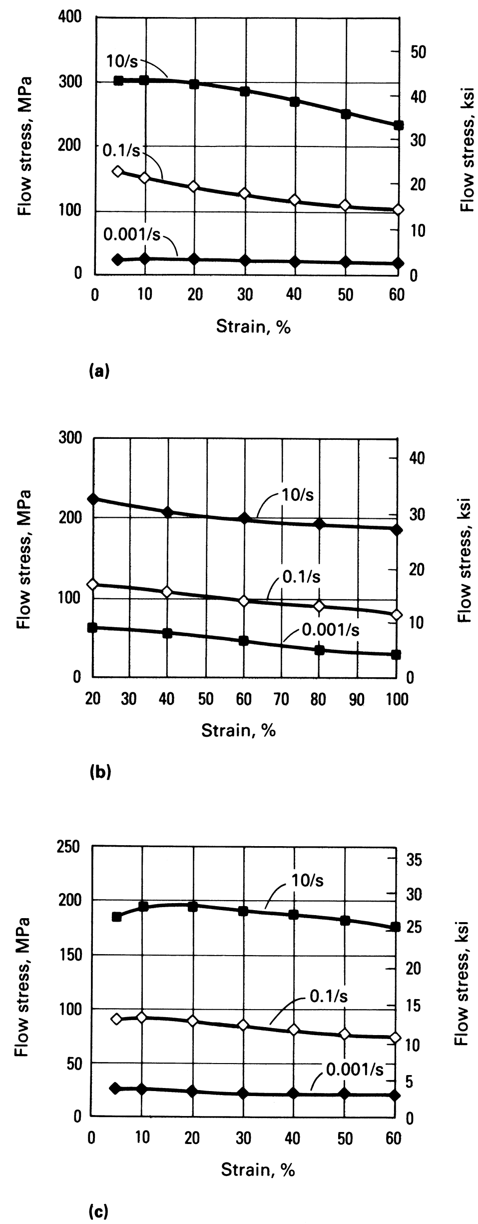


Fig. 19 Effect of three strain rates (0.001, 0.1, and 10/s) on flow stress of three titanium alloys forged at different temperatures. (a) α alloy Ti-8Al-1Mo-1V at 955 °C (1750 °F). (b) α-β alloy Ti-6Al-4V at 900 °C (1650 °F). (c) Metastable β alloy Ti-10V-2Fe-3Al at 815 °C (1500 °F)

processes. However, flow-stress information, as a function of such forging-process variables as temperature and strain rate, is useful in designing titanium alloy forging processes. Because of other forging-process variables, such as die temperature, lubrication, prior working history, and total strain, actual forging pressures or unit pressure requirements may significantly exceed the pure flow stress of any given alloy under similar deformation conditions.

Effect of Deformation Rate. Deformation resistance of titanium alloys is very sensitive to strain rate—considerably more so than aluminum alloys or alloy steels. The strain-rate sensitivity for representative alloys from each of the three classes is shown in Fig. 19 for an α alloy (Ti-8Al-1Mo-1V), an α-β alloy (Ti-6Al-4V), and a β alloy (Ti-10V-2Fe-3Al). For each of these alloys, as the deformation rate is reduced from 10/s to 0.001/s, the flow stress can be reduced by up to ten times. For example, the flow stress for Ti-6Al-4V at 900 °C (1650 °F), 50% strain, and 10/s is 205 MPa (30 ksi); at 0.001/s, the flow stress is 50 MPa (7 ksi), a fourfold reduction.

From the known strain-rate sensitivity of titanium alloys, it appears to be advantageous to deform these alloys at relatively slow strain rates in order to reduce the resistance to deformation in forging; however, under the nonisothermal conditions present in the conventional forging of titanium alloys, the temperature losses encountered by such techniques far outweigh the benefits of forging at slow strain rates. Therefore, in the conventional forging of titanium alloys with relatively cool dies, intermediate strain rates are typically employed as a compromise between strain-rate sensitivity and metal temperature losses to obtain the optimal deformation possible with a given alloy. In isothermal forging, major reduction in resistance to deformation of titanium alloys can be achieved by slow-strain-rate forging techniques under conditions where metal temperature losses are minimized through dies heated to temperatures at or close to the metal temperature.

With rapid-deformation-rate forging techniques, such as the use of hammers and/or mechanical presses, deformation heating during the forging process becomes important. Because titanium alloys have relatively poor coefficients of thermal conductivity, temperature nonuniformity may result, giving rise to nonuniform deformation behavior and/or excursions to temperatures that are undesirable for the alloy and/or final forging mechanical properties. As a result, in the rapid-strain-rate forging of titanium alloys, metal temperatures are often adjusted to account for in-process heat-up, or the forging process (sequence of blows, and so on) is controlled to minimize undesirable temperature increases, or both. Therefore, within the forging temperature ranges outlined in Table 26, metal temperatures for optimal titanium alloy forging conditions are based on the type of forging equipment to be used, the strain rate to be employed, and the design of the forging part.

Temperature Control and Heating. As noted, titanium alloys have a relatively narrow temperature range for conventional forging. Titanium alloy heating equipment should be equipped with pyrometric controls that can maintain ±14 °C (±25 °F) or better. Titanium alloy stock heating equipment is often temperature uniformity surveyed in much the same manner as with heat treating furnaces. Continuous rotary furnaces used for titanium alloys typically have three zones: preheat, high heat, and discharge. Most furnaces are equipped with recording/controlling instruments, and in some batch furnace operations separate load thermocouples are used to monitor furnace temperature during preheating operations.

In addition to highly controlled heating equipment and heating practices, the temperature of heated titanium alloy billets can be verified with contact pyrometry or noncontact optical pyrometers. The latter equipment must be used with care because it is emissivity sensitive and may provide different temperature indications when the metal is observed inside the hot furnace versus when the metal has been removed from the furnace. In most closed-die and open-die forging operations, it is desirable to have titanium alloy metal temperatures near the upper limit of the recommended temperature ranges. In open-die forging, the lower limit of the recommended ranges is usually the point at which forging must be discontinued to prevent excessive surface cracking.

Heating Time. It is good practice to limit the exposure of titanium alloys in preheating to times just adequate to ensure that the center of the forging stock has reached the desired temperature in order to prevent excessive formation of scale and a case. Actual heating times will vary with the section thickness of the metal being heated and with furnace capabilities. Because of the relatively low thermal conductivity of titanium alloys, necessary heating times are extended in comparison to aluminum and alloy steels of equivalent thickness. Generally, 1.2 min/mm (30 min/in.) of ruling section is sufficient to ensure that titanium alloys have reached the desired temperature.

Heating time at a specific temperature is critical in titanium alloys for the reasons outlined previously. Long soaking times are not necessary and introduce the probability of excessive scale or a case. Generally, soaking times should be restricted to 1 to 2 h, and if unavoidable delays are encountered, where soaking time may exceed 2 to 4 h, removal of the metal from the furnace is recommended.

Table 27 Die temperature ranges for the conventional forging of titanium alloys

Forging process/equipment	Die temperature °C	Die temperature °F
Open-die forging		
Ring rolling	150–260	300–500
	95–260	200–500
Closed-die forging		
Hammers	95–260	200–500
Upsetters	150–260	300–500
Mechanical presses	150–315	300–600
Screw presses	150–315	300–600
Orbital forging	150–315	300–600
Spin forging	95–315	200–600
Roll forging	95–260	200–500
Hydraulic presses	315–480	600–900

Heating of Dies. Dies are always preheated in the closed-die conventional forging of titanium alloys, as noted in Table 27, with die temperature varying with the type of forging equipment used. Dies for titanium alloy forging are usually preheated in remote die-heating systems, although on-press equipment is sometimes used. Remote die-heating systems are usually gas-fired die heaters, which can slowly heat the die-blocks to the desired temperature range before assembly into the forging equipment.

With some conventional forging processes, particularly the hydraulic press forging of titanium alloys, the temperature of the dies may increase during forging. Die damage may occur without appropriate cooling. Therefore, titanium alloy dies are often cooled during forging using wet steam, air, or occasionally water.

For those conventional forging processes in which die temperatures tend to decrease, on-press heating systems ranging from rudimentary to highly sophisticated are used. The techniques used include gas-fired equipment, induction-heating equipment, resistance-heating equipment, or combinations of these methods.

Lubrication is also a critical element in the conventional forging of titanium alloys and is the subject of engineering and process-development emphasis in terms of the lubricants used and the methods of application. With titanium alloy conventional forging, a lubrication system is used that includes ceramic precoats of forging stock and forgings, die lubrication, and, for certain forging processes, insulation.

Ceramic Glass Precoats. Most titanium alloy forging stock and forgings are precoated with ceramic precoats prior to heating for forging. These ceramic precoats, which are formulated from metallic and transition-element oxides and other additives, provide several functions, such as:

- Protection of the reactive titanium metal from excessive contact with gaseous elements present during heating
- Insulation or retardation of heat losses during transfer from heating to forging equipment
- Lubrication during the forging process

The formulation of the ceramic precoat is varied with the demands of the forging process being used, the alloy, and the forging type. Modification of the ceramic precoat formulation usually affects the melting or softening temperature, which ranges from 595 to 980 °C (1100 to 1800 °F) for most commercially available precoats for titanium alloys. Experience has shown that ceramic precoats with a viscosity of 20 to 100 Pa · s (200 to 1000 P) at operating temperature provide optimal lubricity and desired continuous film characteristics for protecting the metal during heating and for preventing galling and metal pickup during forging. The actual formulations of ceramic precoats are often proprietary to the forger or the precoat manufacturer. Ceramic precoats are usually colloidal suspensions of the ceramics in mineral spirits or water, with the latter being the most common. Finally, most conventional titanium forging die design techniques include allowances for ceramic precoat thickness in sinking the die cavity to ensure the dimensional integrity of the final forging.

Ceramic precoats are applied using painting, dipping, or spraying techniques; state-of-the-art dipping and/or spraying processes are fully automated. Necessary ceramic precoat thicknesses vary with the precoat and the specific forging process, but generally fall in the range of 0.01 to 0.1 mm (0.0005 to 0.005 in.). Most ceramic precoats require a curing process following application to provide sufficient green strength for handling. Curing procedures range from drying at room temperature to automated furnace curing at temperatures to approximately 150 °C (300 °F).

Die lubricants are also used in the conventional closed-die forging of titanium alloys. Such die lubricants are subject to severe demands and are formulated to modify the surface of the dies to achieve the desired reduction in friction under conditions of very high metal temperatures and die pressures and yet leave the forging surfaces and forging geometry unaffected. Die-lubricant formulations for titanium alloys are usually proprietary, developed either by the forger or the lubricant manufacturer. Die-lubricant composition is varied with the demands of the specific forging process; however, the major active element in titanium alloy die-lubricants is graphite. In addition, other organic and inorganic compounds are added to achieve the desired results because of the very high temperatures present. Carriers for titanium alloy die lubricants vary from mineral spirits to mineral oils to water.

Titanium alloy die lubricants are typically applied by spraying the lubricant onto the dies. Several pressurized-air or airless systems are employed, and with high-volume, highly automated titanium alloy forging processes, die-lubricant application is also automated by single or multiaxis robots. Some state-of-the-art application systems can apply very precise patterns or amounts of lubricant under fully automated conditions.

Insulation. In the conventional forging of titanium alloys in relatively slow-strain-rate processes such as hydraulic-press forging, insulative materials in the form of blankets are often used to reduce metal-temperature losses to the much cooler dies during the initial deformation stages. The insulative blankets are usually fabricated from fiberglass that is formulated to provide the necessary insulative properties. Blanket thickness varies with specific materials of fabrication and desired insulative properties, but generally ranges from 0.25 to 1.3 mm (0.010 to 0.050 in.). If insulative blankets are used, allowance is made in die-sinking tolerances for modification of die-cavity dimensions to ensure the dimensional integrity of the finished forging. Insulative blankets are usually applied to the dies immediately before insertion of the hot metal for forging.

Summary

Engineering materials represent the functional connection between design and manufacturing, and the selection of a given manufacturing process is very much dependent on the specific product design and the selected material. Selection and application of a manufacturing method and material depend on many factors during the various stages of conceptual, configuration, and detailed design. This is discussed in *ASM Handbook,* Volume 20, *Material Selection and Design.* This chapter, which is adapted from content of that book, briefly summarizes some of the general design factors that influence the selection and application of metal bulk forming operations, as discussed in more detail in subsequent chapters of this book. This chapter also summarizes the forgeability characteristics of major alloy systems and some key factors that may influence forging practices, based on content from *ASM Handbook,* Volume 14, *Forming and Forging.*

ACKNOWLEDGMENT

Portions of this article were adapted from content in *Forming and Forging,* Vol 14, *ASM Handbook* (formerly 9th ed., *Metals Handbook,* 1988) and the following articles in *Materials Selection and Design,* Vol 20, *ASM Handbook,* 1997:

- G.E. Dieter, "Relationship between Materials Selection and Processing"
- B.L. Ferguson, "Design for Deformation Processes"
- J.A. Schey, "Manufacturing Processes and Their Selection"
- H.W. Stoll, "Introduction to Manufacturing and Design"

REFERENCES

1. *Forming and Forging,* Vol 14, *ASM Handbook* (formerly 9th ed. *Metals Handbook*), ASM International, 1988
2. T.G. Byrer, Ed., *Forging Handbook,* Forging Industry Association and American Society for Metals, 1985
3. K. Lange, *Handbook of Metal Forming,* McGraw-Hill, 1985
4. S. Kalpakjian, *Manufacturing Processes for Engineering Materials,* 3rd ed., Addison-Wesley, 1996
5. G.E. Dieter, Overview of Materials Selection Process, *Materials Selection and Design,* Vol 20, *ASM Handbook,* 1997, p 246–248
6. M.F. Ashby, *Materials Selection in Mechanical Design*, 2nd ed., Butterworth-Heineman, 1999
7. J. Schey, Manufacturing Processes and Their Selection, *Materials Selection and Design,* Vol 20, *ASM Handbook,* 1997, p 687–704
8. J.G. Bralla, *Design for Excellence,* McGraw-Hill, 1996
9. G. Boothroyd, P. Dewhurst, and W. Knight, *Product Design for Manufacture and Assembly,* 2nd ed., Marcel Dekker, 2002
10. H.E. Trucks, *Designing for Economical Production,* 2nd ed., Society of Manufacturing Engineers, 1987
11. R. Bakerjian, Ed., *Design for Manufacturability,* Vol 6, 4th ed., *Tool and Manufacturing Engineers Handbook*, Society of Manufacturing Engineers, 1992
12. J.A. Schey, *Introduction to Manufacturing Processes,* 3rd ed., McGraw-Hill, 2000
13. C.C. Gallagher and W.A. Knight, *Group Technology Production Methods in Manufacture,* John Wiley & Sons, 1986
14. I. Ham, Group Technology, Chapter 7.8, *Handbook of Industrial Engineering,* G. Salvendy, Ed., John Wiley & Sons, 1982
15. J.-Y. Jung and R.S. Ahluwalia , FORCOD: A Coding and Classification System for Formed Parts, *J. Manuf. Syst.,* Vol 10 (No. 3), 1991, p 223–232
16. M. Tisza, Expert Systems for Metal Forming, *J. Mater. Process. Technol.*, Vol 53 (No. 1–2), Aug 1995, p 423–432
17. G. Henzold, *Handbook of Geometric Toleranc-ing,* John Wiley, 1995
18. L. Edwards and M. Endean, Ed., *Manufacturing with Materials,* Butterworths, 1990
19. Data Card Index, *Manufacturing with Materials,* The Open University, Milton Keynes, 1990
20. E.B. Magrab, *Integrated Product and Process Design and Development,* CRC Press, 1997
21. G.E. Dieter, *Mechanical Metallurgy,* 3rd ed., McGraw-Hill, 1986
22. M. Gensamer, Strength and Ductility, *Trans. ASM,* Vol 36, 1945, p 30–60
23. P.W. Lee and H.A. Kuhn, Fracture in Cold Upset Forging—A Criterion and Model, *Metall. Trans.,* Vol 4, 1973, p 969–974
24. J.T. Winship, Fundamentals of Forging, *Am. Mach.*, July 1978, p 99–122
25. D.W. Hutchinson, "The Function and Proper Selection of Forging Lubricants," Acheson Colloids Co., 1984
26. *Open Die Forging Manual,* 3rd ed., Forging Industry Association, 1982, p 106–107
27. M.J. Donachie and S.J. Donachie, *Superalloys: A Technical Guide,* ASM International, 2002, p 92–95
28. H.J. Henning, A.M. Sabroff, and F.W. Boulger, "A Study of Forging Variables," Report ML-TDR-64-95, U.S. Air Force, 1964
29. W.D. Klopp, A-286 Datasheet (March 1987), *Aerospace Structural Metals Handbook,* CINDAS/USAF Handbooks Operation and

Purdue Research Foundation, Code 1601, 1995, p 6
30. W.D. Klopp, N-155 Datasheet (June 1989), *Aerospace Structural Metals Handbook,* CINDAS/USAF Handbooks Operation and Purdue Research Foundation, Code 1602, 1995, p 2
31. Forging of Refractory Metals, *Forming and Forging,* Vol 14, *ASM Handbook* (formerly 9th ed. *Metals Handbook*), ASM International, 1988, p 237–238
32. G.W. Kuhlman, Forging of Aluminum Alloys, Forming and Forging, Vol 14, *ASM Handbook* (formerly 9th ed. *Metals Handbook*), ASM International, 1988, p 241–254
33. H.H. Ruble and S.L. Semiatin, Forging of Nickel-Base Alloys, Forming and Forging, Vol 14, *ASM Handbook* (formerly 9th ed. *Metals Handbook*), ASM International, 1988, p 261–266
34. A.J. DeRidder and R. Koch, *MiCon 78: Optimization of Processing, Properties, and Service Performance Through Microstructural Control,* H. Abrams et al., Ed., American Society for Testing and Materials, 1979, p 547
35. T. Altan, F.W. Boulger, J.R. Becker, N. Ackgerman, and H.J. Henning, "Forging Equipment Materials, and Practices," MCIC-HB-03, Metals and Ceramics Information Center, 1973
36. D.R. Muzyka, *MiCon 78: Optimization of Processing, Properties, and Service Perfor-mance Through Microstructural Control,* H. Abrams et al., Ed., American Society for Testing and Materials, 1979, p 526
37. G.W. Kuhlman, Forging of Titanium Alloys, *Forming and Forging,* Vol 14, *ASM Handbook* (formerly 9th ed. *Metals Handbook*), ASM International, 1988, p 267–287

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 172-187
DOI:10.1361/hwpd2003p172

Chapter 12

Workability Theory and Application in Bulk Forming Processes*

Howard A. Kuhn, Consultant

WORKABILITY, as described in previous chapters, is not merely a property of a material but a characteristic of the material/process system. The role of the material is measured by a simple test or tests and should be expressed in a quantitative form that is applicable universally. This measure is taken to be a basic property of the material composition and structure, and it reflects the macroscopic outcome of the micromechanisms of plastic flow perturbed by such inhomogeneities as voids, inclusions, and grain boundaries. Such phenomena are dependent not only on the material structure but also on the process parameters (strain rate and temperature), which define the role of the process in determining workability.

The micromechanisms of ductile fracture in bulk forming processes are strongly influenced by the stress and strain environment imposed by the process. As shown in Fig. 10 of Chapter 2, overall measure of strain to fracture can be related to an overall value of hydrostatic stress in the material during processing. In the spectrum of processes (extrusion, rolling, forging, and wire drawing), the average hydrostatic stress becomes increasingly tensile, and the strain to fracture progressively decreases. Within each of these processes, however, the stress and strain states can be considered on a localized basis in the specific regions in which fractures initiate. These localized conditions are controlled by the geometry of the workpiece, die design, and friction at the die/workpiece interface. These three factors, in addition to the strain rate and temperature parameters already mentioned, embody the role of the process in determining workability.

This article focuses on the effects of mechanical plasticity on workability; that is, process control of localized stress and strain conditions to enhance workability. First, the nature of local stress and strain states in bulk forming processes is described, leading to a classification scheme that facilitates the application of workability concepts. This defines testing procedures and specific process measurements for an experimental approach to workability evaluation. Theoretical models and fracture criteria are then described and compared with experimental results. Finally, the application of workability concepts to forging, rolling, and extrusion processes is discussed.

Stress and Strain States

Forging, extrusion, and rolling processes generally are considered to involve the application of compressive force to material to impart a change in shape and dimensions. On close examination, however, it is clear that deformation resulting from the applied load causes secondary stress and strain states that vary from point to point throughout the deforming workpiece. These stress and strain states may include tension, so fracture can occur at certain locations in a material even though the primary (applied) load is compressive.

To explore this possibility further, it is useful to review briefly the von Mises yield criterion, which is the foundation of plasticity theory for isotropic materials. It gives the relationships between normal and shear stress components at yielding. From this, relationships between stress and strain components are derived:

$$\varepsilon_1 = \lambda\left[\sigma_1 - \frac{1}{2}\sigma_2 - \frac{1}{2}\sigma_3\right]$$
$$\varepsilon_2 = \lambda\left[\sigma_2 - \frac{1}{2}\sigma_3 - \frac{1}{2}\sigma_1\right]$$
$$\varepsilon_3 = \lambda\left[\sigma_3 - \frac{1}{2}\sigma_1 - \frac{1}{2}\sigma_2\right] \quad \text{(Eq 1)}$$

where ε denotes strain, σ denotes stress, and the subscripts 1, 2, and 3 designate the three directions in an orthogonal coordinate system. Here, λ is a proportionality factor dependent on the deformation history and flow stress curve of the material. The resulting strain in a given direction is affected by the stress in all three coordinate directions. In addition, these relationships satisfy the volume constancy condition, $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0$.

To illustrate these relationships, it is useful to consider a two-dimensional case of plane stress, $\sigma_3 = 0$:

$$\varepsilon_1 = \lambda\left[\sigma_1 - \frac{1}{2}\sigma_2\right]$$
$$\varepsilon_2 = \lambda\left[\sigma_2 - \frac{1}{2}\sigma_1\right] \quad \text{(Eq 2)}$$

Referring to Fig. 1 and Eq 2, a reduction in thickness (compressive strain in the 2-direction)

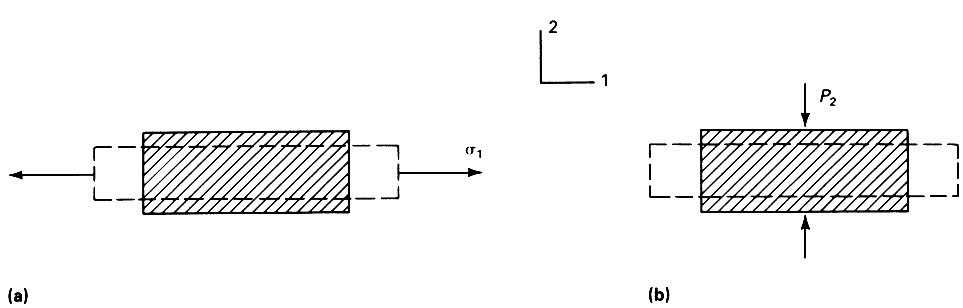


Fig. 1 Similarity of deformation under (a) horizontal tension and (b) vertical compression

*This article is revised from Howard A. Kuhn, Workability Theory and Application in Bulk Forming Processes, *Forming and Forging*, Vol 14, *ASM Handbook,* ASM International, 1988, p 388–404.

can be accomplished either by a compressive stress, P_2 (or $-\sigma_2$), or by a tensile stress, σ_1. Although the nature of the deformation is the same in both cases, accomplishing that deformation in the latter case could risk fracture because it involves a tensile stress.

For more general cases, Eq 2 can be rearranged to determine the following stresses:

$$\sigma_1 = \frac{4}{3\lambda}\left[\varepsilon_1 + \frac{1}{2}\varepsilon_2\right] = \frac{4}{3\lambda}\varepsilon_1\left[1 + \frac{1}{2}\frac{\varepsilon_2}{\varepsilon_1}\right]$$

$$\sigma_2 = \frac{4}{3\lambda}\left[\varepsilon_2 + \frac{1}{2}\varepsilon_1\right] = \frac{4}{3\lambda}\varepsilon_1\left[\frac{1}{2} + \frac{\varepsilon_2}{\varepsilon_1}\right]$$

$$\sigma_3 = 0 \qquad \text{(Eq 3)}$$

Therefore, the stresses depend on the localized strains ε_1 and ε_2 that result from metal flow. Equation 3 is a more convenient representation of plasticity relationships for workability study.

For example, the cylindrical surface of a compression test undergoes various combinations of axial and circumferential strains, depending on the aspect ratio and the friction at the die contact surfaces (Fig. 2). When no friction exists, the ratio of circumferential strain to axial strain is $\varepsilon_1/\varepsilon_2 = -\frac{1}{2}$. According to the first part of Eq 3, $\sigma_1 = 0$ for this case. The deformation in this case is referred to as homogeneous compression, because the only stress acting is σ_2 and it is uniform throughout the specimen. Therefore, the homogeneous compression test is suitable for measuring flow stress.

When friction exists at the die contact surfaces, material at these surfaces is constrained from moving outward, while material at the midplane is not constrained. As a result, bulging occurs, as shown in Fig. 2(a). Under these conditions, the circumferential strain ε_1 (>0) increases, and the localized axial compressive strain ε_2 (<0) decreases. From Eq 3 as $\varepsilon_1/\varepsilon_2$ becomes more negative ($<-\frac{1}{2}$), σ_1 becomes more positive. Therefore, increasing bulging due to friction during the compression of a cylinder increases the secondary tensile stress σ_1 and enhances the likelihood of fracture.

Similarly, at the edges of bars during rolling (Fig. 3), the elongation strain ε_1 is determined by the overall reduction in area. The localized compressive vertical strain ε_2, however, depends on the shape of the edge. Greater convexity and sharpness of the edge decrease the compressive vertical strain for a given reduction, which, from the first part of Eq 3, increases the secondary tensile stress σ_1 at the edge. Therefore, edge cracking during rolling is also due to secondary stress states.

More complex cases of the same type of secondary tensile stress states occur in forging (Fig. 4). During the forging of a hub shape, for example, the top surface of the hub is subjected to biaxial tension because of friction during flow around the die radius (Fig. 4a). This is identical to the conditions present at the nose of a billet that is being extruded or rolled. Similarly, during forging, the top surface of a rib undergoes tensile strain in the direction of curvature, and essentially no strain occurs along the length direction of the rib (Fig. 4b). In both cases, localized secondary tensile stresses are generated that may cause fracture. These stresses can be calculated from measured strain values using Eq 3.

Figures 2 to 4 show examples of plane stress, with the stress normal to the free surface being zero. Other regions of workpieces in bulk deformation processes, however, are subjected to three-dimensional stress states. For example, material at the die contact surfaces in forging, extrusion, and rolling (Fig. 5) is subjected to strains ε_1 and ε_2 in the plane of the surface, as in Fig. 2 to 4. In Fig. 5, however, this surface is also acted upon by pressure P_3 normal to the plane. Similarly, at internal locations of the workpieces in such processes as forging or wire drawing (Fig. 6), a material element of the central longitudinal plane is subjected to strains ε_1 and ε_2 and stress normal to the plane.

In Fig. 5 and 6, the material element can be thought of as the plane-stress elements in Fig. 2 to 4 with σ_3 acting normal to the plane. If ε_1 and ε_2 are taken to be the strains in this plane, then Eq 3 becomes:

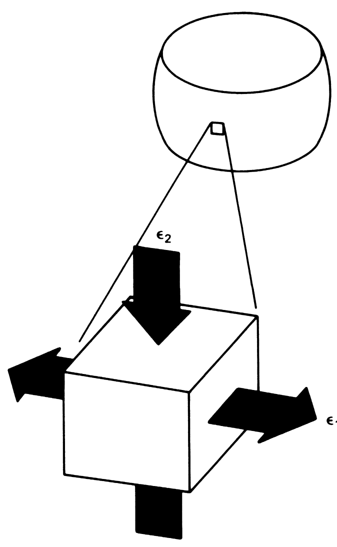
$$\sigma_1 = \frac{4}{3\lambda}\left[\varepsilon_1 + \frac{1}{2}\varepsilon_2\right] + \sigma_3$$

$$\sigma_2 = \frac{4}{3\lambda}\left[\varepsilon_2 + \frac{1}{2}\varepsilon_1\right] + \sigma_3$$

$$\sigma_3 = \sigma_3$$

In other words, for the same deformation (that is, the same values of ε_1 and ε_2 as in Fig. 2 to 4), the stresses σ_1 and σ_2 in Eq 3 are biased by σ_3. The stress normal to the surface, then, increases the hydrostatic stress component of the stress state by σ_3. This reflects the basic concept that hydrostatic stress does not affect yielding or plastic deformation. As shown in Fig. 10 of Chapter 2, "Bulk Working Behavior of Metals," however, the hydrostatic stress has a significant effect on fracture. In Fig. 5, the compressive stress (die pressure) P_3 would increase the strains at fracture, but in Fig. 6 the internal stress σ_3 may be tensile and decrease the strains at fracture.

The preceding discussion of stress and strain states in various regions of workpieces suggests a convenient method of classifying cracking defects. Figures 2 to 4 in this article illustrate the locations at which free surface cracks occur in forging and rolling, as well as at the nose ends of billets in extrusion or rolling. Figure 5 in this article shows examples of circumstances in which contact surface cracking occurs—for example, the fir tree defect in extrusion or the longitudinal surface cracks in rolled plates.



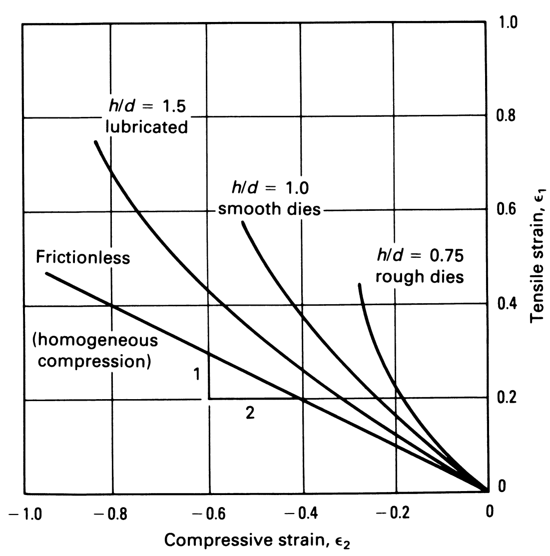


Fig. 2 Localized strains on (a) the bulging cylindrical surface of an upset test and (b) their variation with aspect ratio and friction conditions. Source: Ref 1

Empirical Criterion of Fracture

The stress and strain environments described in the previous section in this article suggest that a workability test should be capable of subjecting the material to a variety of surface strain combinations. A capability for testing under superimposed normal stress would also be desirable.

When considering workability tests, it is important to recognize that fractures initiate in localized regions where interaction between the stress and strain states and the material structure reaches a critical level. Orientation, shape, and volume fraction of inclusions and other inhomogeneities have a dominant effect on the fracture process. Therefore, it is critically important that workability test specimens contain material hav-

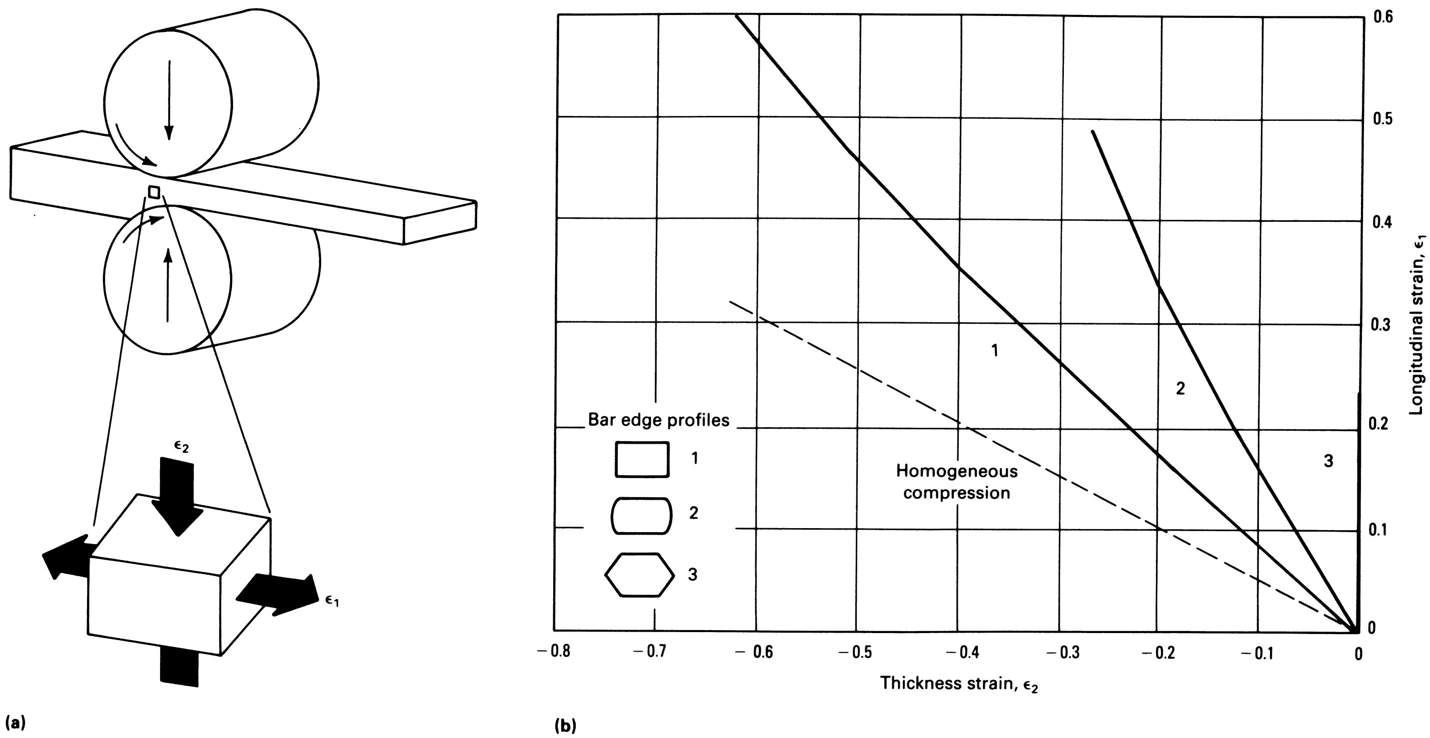


Fig. 3 Localized strains at the edges of bars during (a) rolling and (b) their variation with edge profile. Source: Ref 2

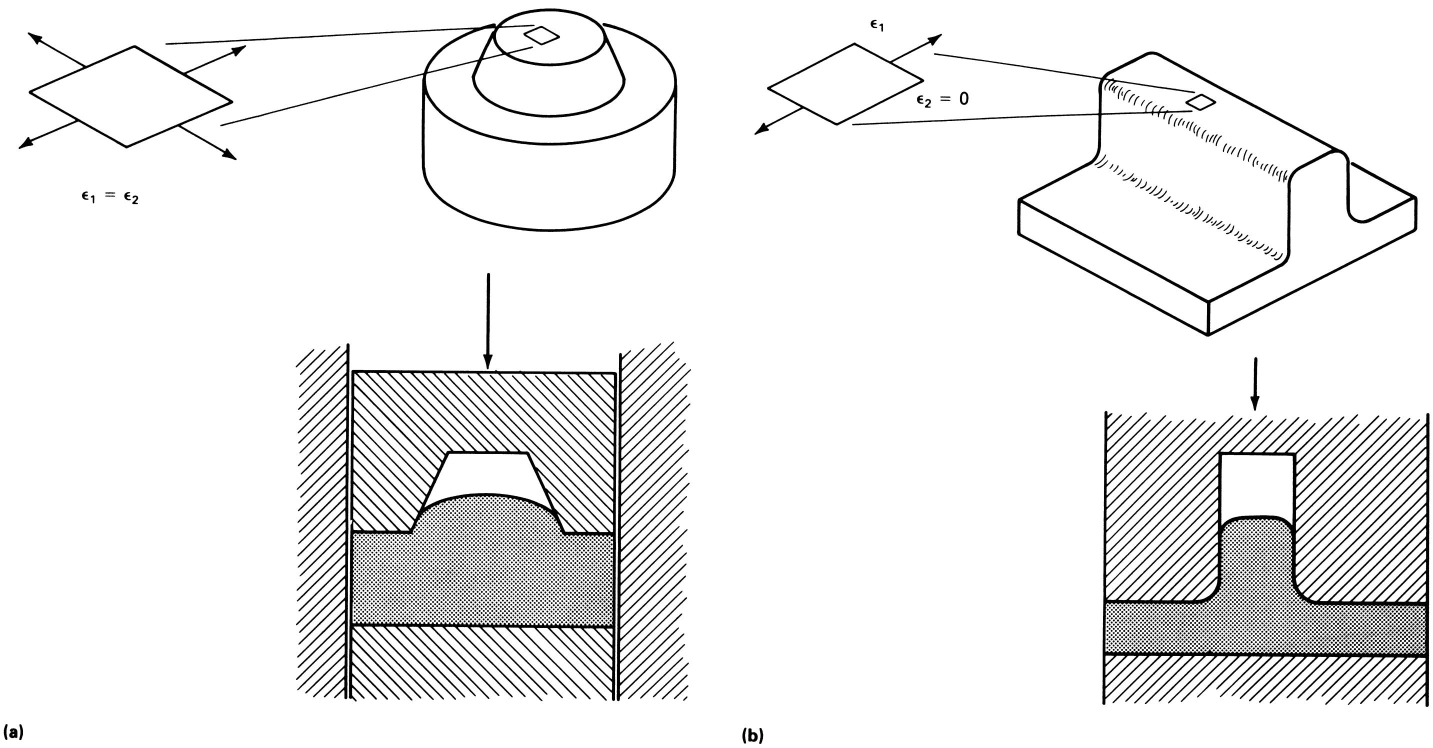


Fig. 4 Strains at the free surfaces of forgings. (a) Axisymmetric hub. (b) Rib-web forging

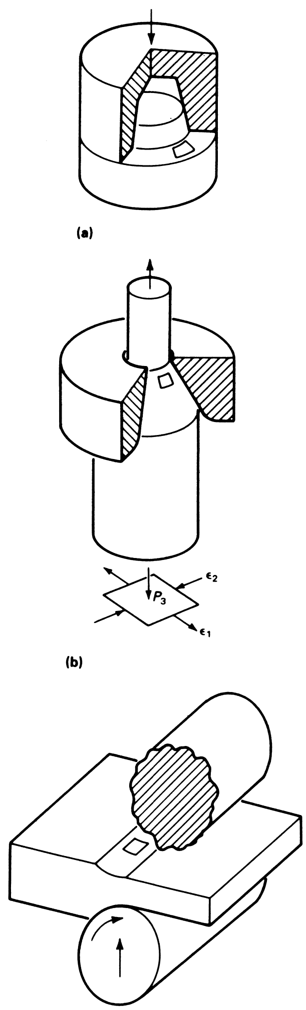


Fig. 5 An element of material at the die contact surfaces during (a) forging, (b) drawing or extrusion, and (c) rolling

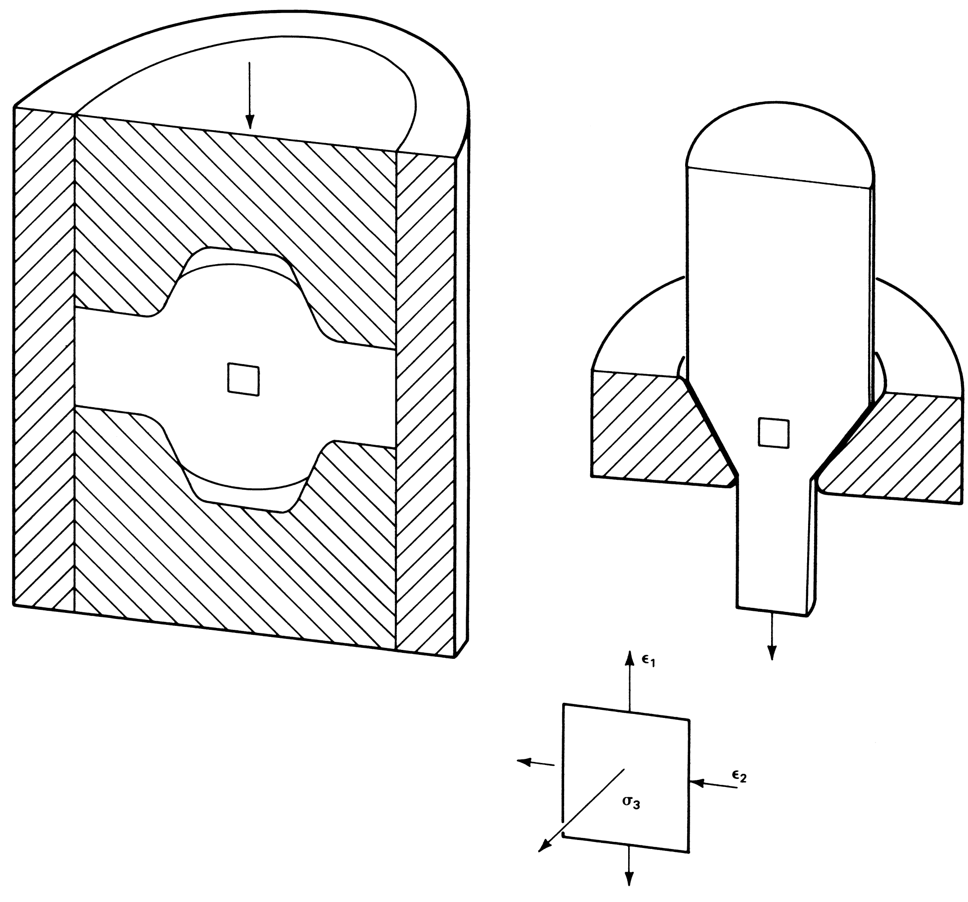


Fig. 6 An element of material at the center of a forging undergoing (a) double extrusion and (b) at the center of wire being drawn. The element undergoes strains ε_1, and ε_2, as in Fig. 2 to 4, with stress σ_3 normal to the 1–2 plane.

ing the same microstructural features as the material in the localized, potential fracture regions of the actual process.

Specifically, when evaluating a workpiece for surface fractures, specimen surfaces must contain the as-received surface of the workpiece under consideration because it may contain laps, seams, a decarburized layer, and so on, which affect fracture initiation. By the same argument, evaluation of material for internal fractures such as central burst must involve test specimens taken from the middle of the workpiece, where, for example, segregation of second phases may have occurred. Because of possible anisotropy effects, orientation of the critical stresses with respect to any inclusion alignment must be the same in the test specimens as it is in the actual process and material of interest.

The compression test has become a standard for workability evaluation. As shown in Fig. 2, a range of strain combinations can be developed at the cylindrical free surface simply by altering friction and geometry conditions. The influence of friction and consequent bulging on circumferential tensile stress development is clearly shown in Fig. 7. Compression with friction produces circumferential tension that leads to fracture, while frictionless compression prevents barreling, tension, and cracking as described in Fig. 2 and Eq 3.

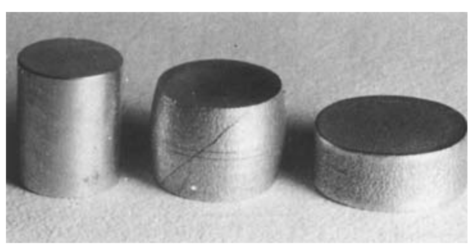


Fig. 7 Compression tests on 2024-T35 aluminum alloy. Left to right: undeformed specimen, compression with friction (cracked), compression without friction (no cracks)

Alterations of the compression test geometry have been devised to extend the range of surface strains available toward the vertical, tensile strain axis, ε_1 (Ref 3). Test specimens are artificially prebulged by machining a taper or a flange on the cylinders (Fig. 8). Compression then causes lateral spread of the interior material, which expands the rim circumferentially while applying little axial compression to the rim. Therefore, the tapered and flanged upset test specimens provide strain states consisting of

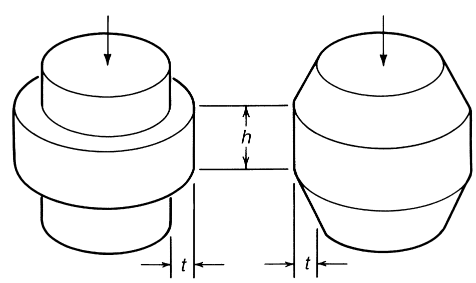


Fig. 8 (a) Flanged and (b) tapered prebulged compression test specimens. Lateral spread of interior material under compression expands the rim circumferentially while little axial compression is applied (see Fig. 9).

small compressive strain components. Each combination of height, h, and thickness, t, gives a different ratio of tensile to compressive strain. The strain states developed at the surfaces of straight, tapered, and flanged compression test specimens are summarized in Fig. 9.

The variety of strain combinations available in compression tests offers the possibility for material testing over most of the strain combinations that occur in actual metalworking processes. A number of samples of the same material and condition are tested, each one under different friction and geometry parameters. Tests are carried out until fracture is observed, and the local axial and circumferential strains are measured at fracture. Figures 10 to 12 give some examples of results for AISI 1045 carbon steel, 2024-T351 aluminum alloy at room temperature, and 2024-T4 alloy at a hot-working temperature. In some cases, the fracture strains fit a straight line of slope $-\frac{1}{2}$; in others, the data fit a dual-slope line with slope $-\frac{1}{2}$ over most of the range and slope -1 near the tensile strain axis. Similar data have been obtained for a wide variety of materials. In each case, the straight line behavior (single or dual slope) appears to be characteristic of all materials, but the height of the line varies with the material, its microstructure, test temperature, and strain rate.

The nature of the fracture loci shown in Fig. 10 to 12 suggests an empirical fracture criterion representing the material aspect of workability. The strain paths at potential fracture sites in material undergoing deformation processing (determined by measurement or mathematical analysis) can then be compared to the fracture strain loci. Such strains can be altered by process parameter adjustment, and they represent the process input to workability. If the process strains exceed the fracture limit lines of the material of interest, fracture is likely. Other approaches to establishing fracture criteria, as well as applications of the criteria, are given in the following two sections of this article, "Theoretical Fracture Models and Criteria" and "Applications."

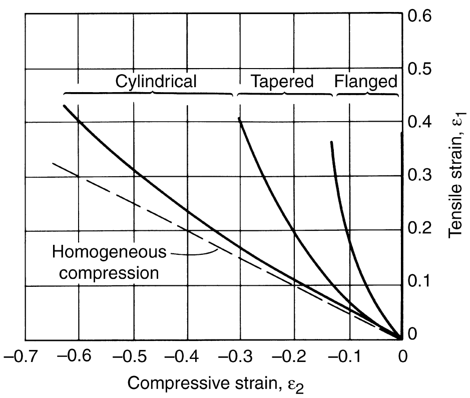


Fig. 9 Range of free surface strain combinations for compression tests having cylindrical (Fig. 2), tapered, and flanged (Fig. 8) edge profiles. The ranges shown are approximate and they may overlap a small amount.

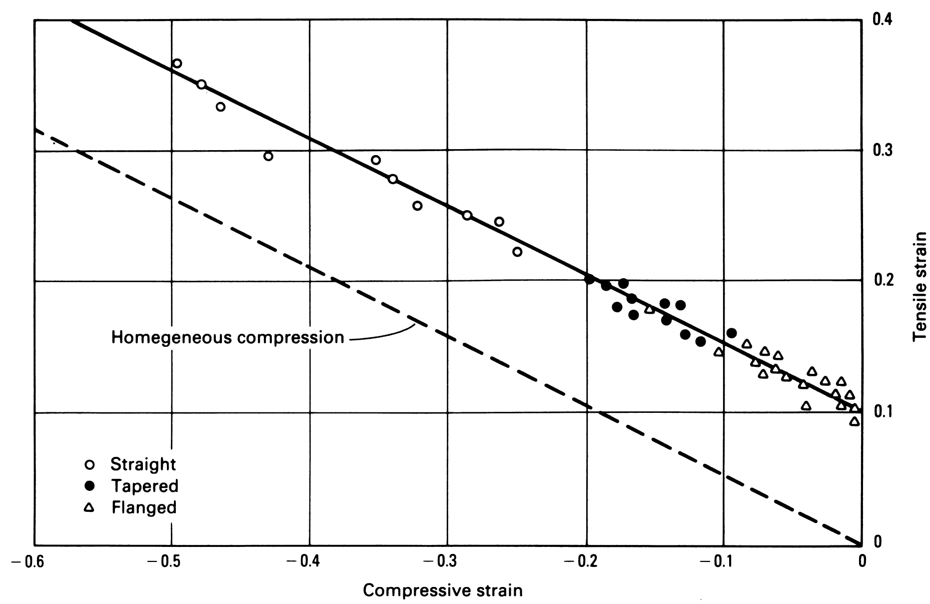


Fig. 11 Fracture locus for aluminum alloy 2024-T351 at room temperature

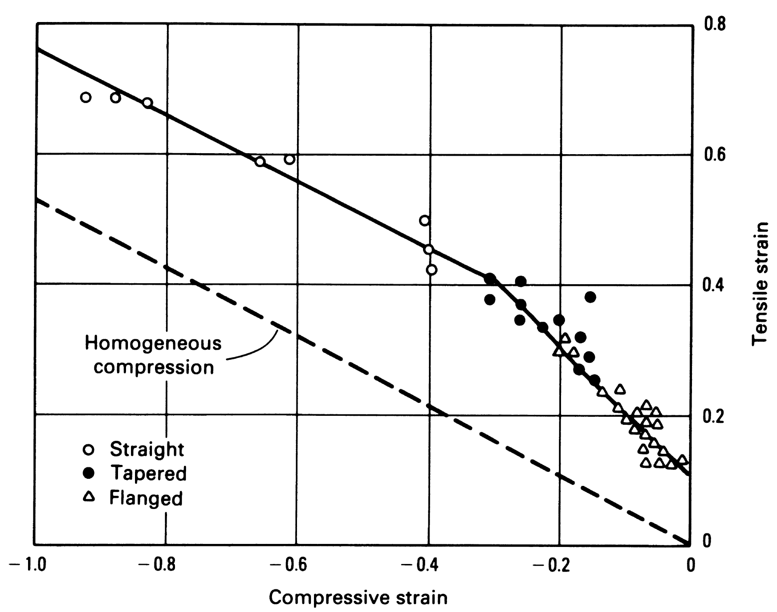


Fig. 10 Fracture locus for AISI 1045 cold-drawn steel

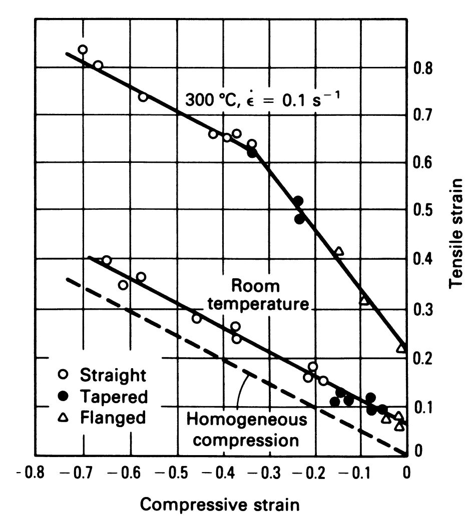


Fig. 12 Fracture locus for aluminum alloy 2024-T4 at room temperature and at 300 °C (570 °F). $\dot{\varepsilon} = 0.1\ s^{-1}$

Theoretical Fracture Models and Criteria

Fracture criteria for metalworking processes have been developed from a number of viewpoints, as described in Chapters 1 and 2. The most obvious approach involves modeling of the void coalescence phenomenon normally associated with ductile fracture. Another approach involves a model of localized thinning of sheet metal that has been adapted to bulk forming processes. In addition to models of fracture, criteria have been developed from macroscopic concepts of fracture. The Cockcroft criterion is based on the observation that both tensile stress and plastic deformation are necessary ingredients, which lead to a tensile deformation energy condition for fracture. The upper bound method has been used to predict fracture in extrusion and drawing. Other approaches are based on the calculation of tensile stress by slip-line fields.

Void Growth Model. Microscopic observations of void growth and coalescence along planes of maximum shear leading to fracture have led to the development of a model of hole growth (Ref 4). Plasticity mechanics is applied to the analysis of deformation of holes within a shear band. When the elongated holes come into contact, fracture is considered to have occurred (Fig. 13). When the McClintock model is evaluated for a range of applied stress combinations, a fracture strain line can be constructed.

Figure 14 shows the calculated results from the McClintock model in comparison with the experimental fracture line. The predicted fracture strain line has a slope of $-\frac{1}{2}$ over most of its length, matching that of the experimental fracture line. Near the tensile strain axis, the slope of the predicted line is -1, matching that of actual material results shown in Fig. 10 and 12.

Localized Thinning Model. In sheet forming, the observation that a neck forms before fracture, even under biaxial stress conditions in which localized instability cannot occur, has prompted consideration of the effects of inhomogeneities in the material. For example, a model of localized thinning due to a small inhomogeneity has been devised (Ref 5). Beginning with the model depicted in Fig. 15, plasticity mechanics is applied to determine the rate of thinning of the constricted region t_B in relation to that of the thicker surrounding material t_A. When the rate of thinning reaches a critical value, the limiting strains are considered to have been reached, and a forming limit diagram can be constructed. The analysis includes the effects of crystallographic anisotropy, work-hardening rate, and inhomogeneity size.

This model was applied to free surface fracture in bulk forming processes because of evidence that localized instability and thinning also precede this type of ductile fracture (Ref 6). Two model geometries were considered, one having a groove in the axial direction (Z-model) and the other having a groove in the radial direction (R-model) as shown in Fig. 15. Applying plasticity mechanics to each model, fracture is considered to have occurred when the thin region, t_B, reduces to zero thickness. When these fracture strains are plotted for different applied stress ratios, a fracture strain line can be constructed. As shown in Fig. 16, the predicted fracture line matches the essential features of the experimental fracture lines. The slope is $-\frac{1}{2}$ over most of the strain range and approximately -1 near the tensile strain axis. Again, this model is in general agreement with the dual-slope fracture loci shown in Fig. 10 and 12.

Cockcroft Model. The Cockcroft criterion of fracture is not based on a micromechanical model of fracture but simply recognizes the macroscopic roles of tensile stress and plastic deformation (Ref 7). It is suggested that fracture occurs when the tensile strain energy reaches a critical value:

$$\int_0^{\bar{\varepsilon}_f} \sigma^* \, d\bar{\varepsilon} = C$$

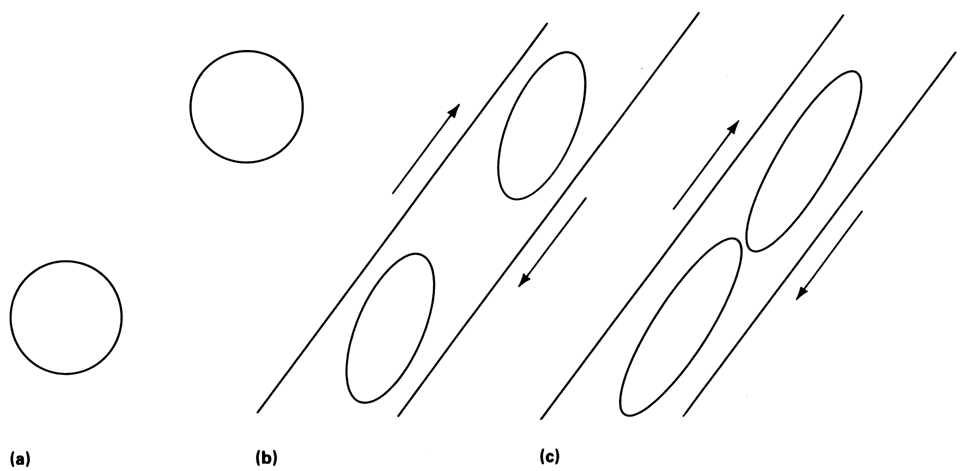


Fig. 13 McClintock model of void coalescence by shear from (a) initial circular voids, through (b) growth, and (c) void contact

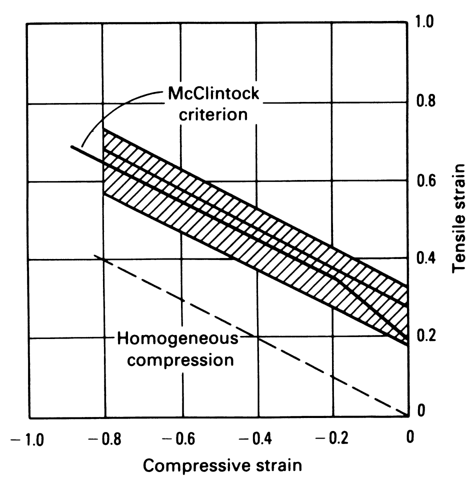


Fig. 14 Fracture strain locus predicted by the McClintock model of void growth. The shaded area represents typical experimental fracture loci such as Fig. 10 to 12.

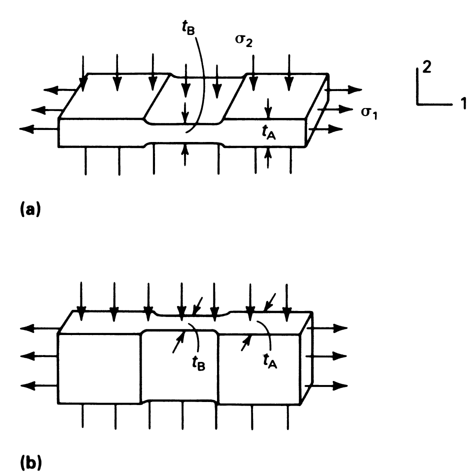


Fig. 15 Models for the analysis of localized thinning and fracture at a free surface. (a) R-model. (b) Z-model. Source: Ref 6

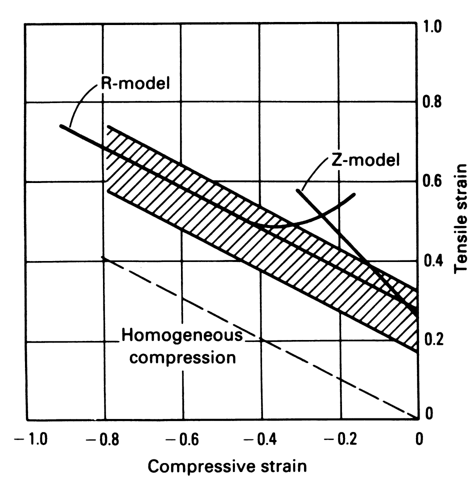


Fig. 16 Fracture strain locus predicted by the model of localized thinning. The shaded area represents typical experimental fracture loci.

where σ* is the maximum tensile stress; $\bar{\varepsilon}$ is the equivalent strain; and *C* is a constant determined experimentally for a given material, temperature, and strain rate. The criterion has been successfully applied to cold-working processes. It has also been reformulated to provide a predicted fracture line for comparison with the experimental fracture strain line. Figure 17 shows that the fracture strain line predicted by the Cockcroft criterion is also in reasonable agreement with experimental results. This criterion does not show the dual-slope behavior of the previous models and some actual materials. The question of why some materials exhibit the dual-slope fracture locus and others only a single slope remains the subject of speculation, discussion, and further study.

The upper bound method of plasticity analysis requires the input of a flow field in mathematical function form. The external work required to produce this flow field is determined through extensive calculation. This value for external work is an upper bound on the actual work required. Through optimization procedures, the flow field can be found that minimizes the calculated external work done, and this flow field is closest to the actual metal flow in the process under analysis. The upper bound method has been applied to a number of metalworking processes (Ref 8).

In consideration of metal flow fields, perturbations can be incorporated that simulate defect formation. In some situations, the external work required to create the flow field with defects is less than the work required to create the flow field without defects (sound flow). According to the upper bound concept, defects would occur in these situations.

The upper bound method has been applied to the prediction of conditions for central burst formation in extrusion and wire drawing. As shown in Fig. 18, for die angles less than α_1, sound flow requires less drawing force than flow with central burst formation. Above α_2, extrusion with a dead metal zone near the die requires lower drawing force. In the range of die angles between α_1 and α_2, central burst is energetically more favorable (Ref 9).

Repeated calculations using the upper bound method provide the combinations of die angle and reduction that cause central burst (Fig. 19). Friction at the die contact surface affects the results. If operating conditions are in the central burst region, defects can be avoided by decreasing the die angle and/or increasing the reduction so that the operating conditions are in the safe zone. This example is a clear illustration of the role of process parameters (in this case, geometric conditions) in workability. An application of this method is given in the section "Applications" in this article.

It should be pointed out that the upper bound method for defect prediction gives only a necessary condition. The strain-hardening and strain rate hardening characteristics of the material have been included in the analysis (Ref 10), but the microstructural characteristics have been omitted. Therefore, when operating in the central burst range illustrated in Fig. 19, fracture can occur; whether or not it will depends on the material structure (voids, inclusions, segregation, and so on). When operating in the safe area shown in Fig. 19, central burst will not occur, regardless of the material structure.

Tensile Stress Criterion. The role of tensile stress in fracture is implicit yet overwhelmingly clear throughout the discussion of fracture and fracture criteria. The calculation of tensile stresses in localized regions, however, requires the use of advanced plasticity analysis methods such as slip-line fields or finite-element analysis. One result of slip-line field analysis that has wide application in workability studies is discussed here.

Double indentation by flat punches is a classical problem in slip-line field analysis (Fig. 20).

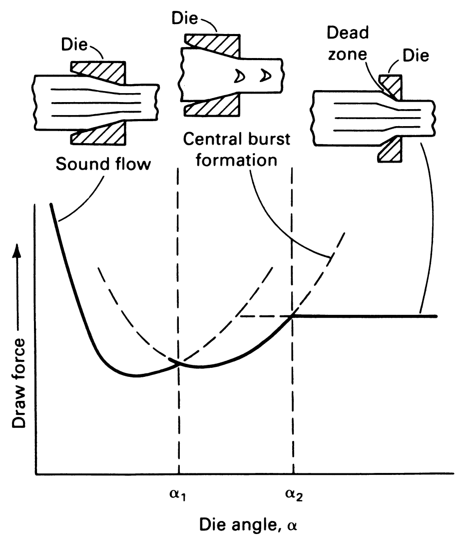


Fig. 18 Variation in mode of flow with die angle in wire drawing. The mode requiring the smallest force at any die angle is the active mode. This is a schematic for one value of reduction. Source: Ref 9

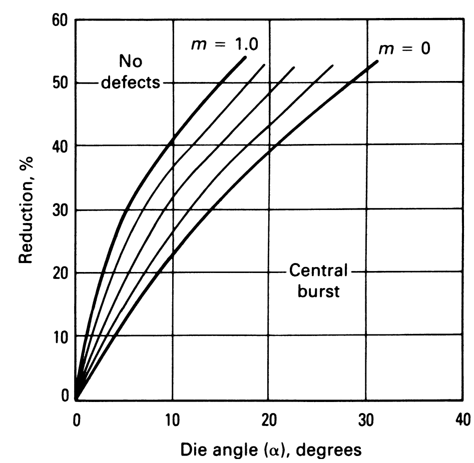


Fig. 19 Upper bound prediction of central burst in wire drawing. Increasing friction, expressed by the friction factor, *m*, increases the defect region of the map. Source: Ref 10

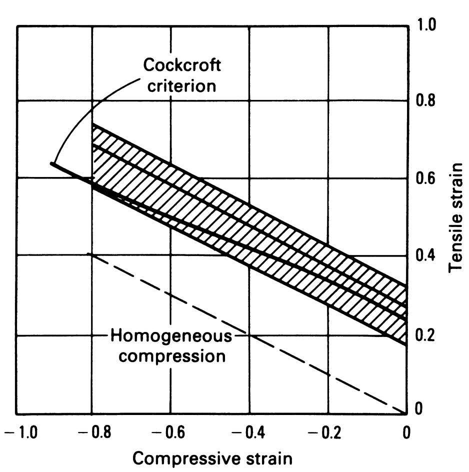


Fig. 17 Fracture strain locus predicted by the Cockcroft criterion

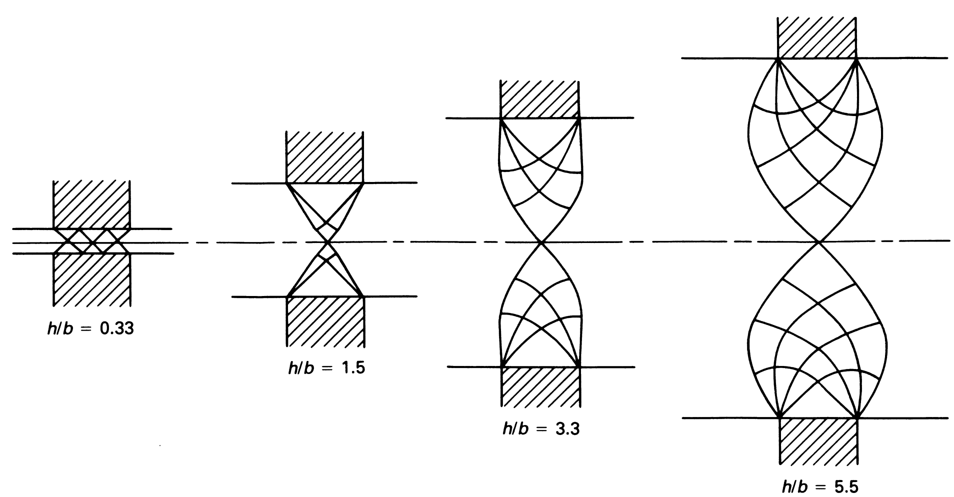


Fig. 20 Slip-line fields for double indentation at different *h*/*b* ratios. Source: Ref 11

The boundaries of the deformation zone change as the aspect ratio h/b (workpiece thickness-to-punch width) increases. For $h/b > 1$, the slip-line field meets the centerline at a point, and for $h/b < 1$, the field is spread over an area nearly as large as the punch width.

This tooling arrangement and deformation geometry approximates several other metalworking processes, as shown in Fig. 21. For similar h/b ratios in these processes then, the stresses throughout the deformation zone of the process can be approximated by those calculated from slip-line field analysis for double indentation.

Most of the work on double indentation has focused on the calculation of punch pressure and the extrapolation of these results to other processes—for example, those in Fig. 21. In a very detailed study of double indentation, remarkable agreement was found between experimental results and slip-line field results (Ref 12).

For workability studies, however, it is necessary to locate and to calculate the critical tensile stresses. It can be shown that the hydrostatic stress is always greatest algebraically at the centerline of the material and that this stress is tensile for $h/b > 1.8$. The calculated results for punch pressure and centerline hydrostatic stress are given in Fig. 22. Therefore, it is necessary to specify die and workpiece geometric parameters such that $h/b < 1.8$ in order to avoid tensile stress and potential fracture at the centerline of the processes shown in Fig. 21.

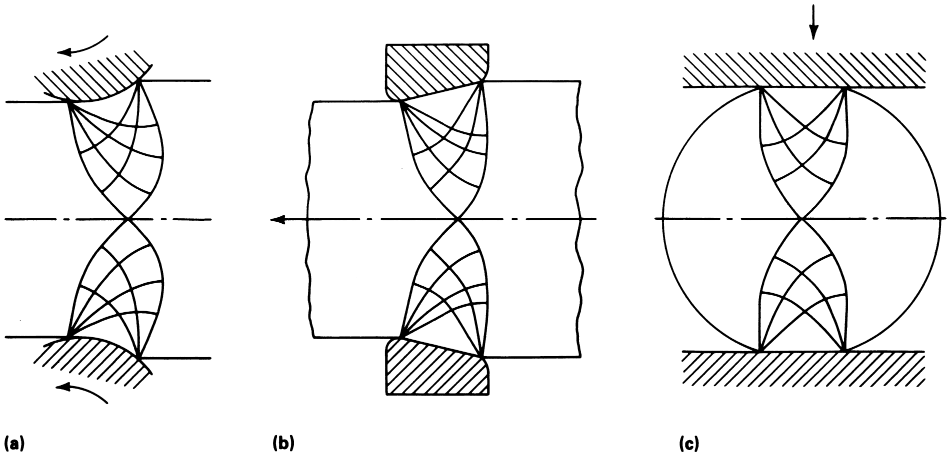


Fig. 21 Slip-line fields for (a) rolling, (b) drawing, and (c) side pressing. These fields are similar to those for double indentation shown in Fig. 20.

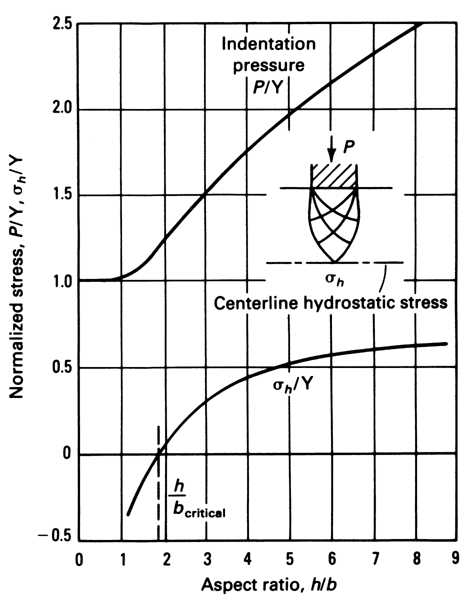


Fig. 22 Variation of the normalized indentation pressure (P/Y where Y is the yield strength) and the normalized centerline hydrostatic stress (σ_h/Y) with h/b ratio as calculated from slip-line field analysis

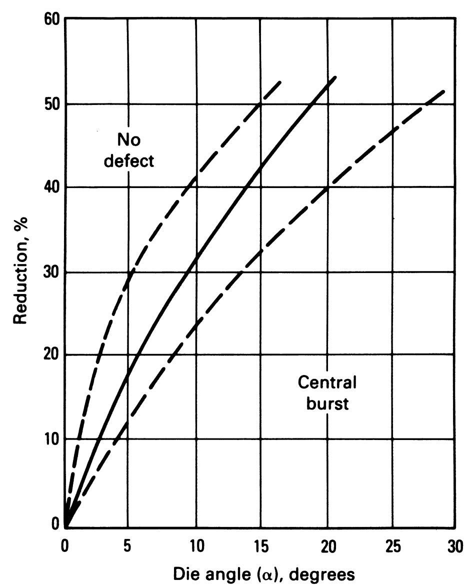


Fig. 23 Prediction of central burst in wire drawing by the tensile stress criterion and slip-line field analysis of double indentation. The range of predictions by upper bound analysis (Fig. 19) is shown by dashed lines.

For example, in extrusion, h/b is approximated by:

$$h/b = \frac{\alpha\left[1+(1-R)^{1/2}\right]^2}{R}$$

where α is the die half-angle and R is area reduction. Taking $h/b = 1.8$, the relationship between α and R that produces tensile hydrostatic stress at the centerline can be calculated. The result is given in Fig. 23, which is shown to be similar to the relationship predicted by upper bound analysis (Fig. 19). The correlation is remarkable in view of the dissimilarity in die shape between extrusion and double indentation. Furthermore, the similarity emphasizes that the flow mode for defect formation in the upper bound method is physically equivalent to the development of tensile hydrostatic stress.

As in the case of the upper bound method, the existence of a tensile hydrostatic stress does not ensure fracture, but it is a necessary ingredient. The material structure must be considered in conjunction with the tensile stress. In other words, the upper bound and tensile stress criteria are useful for defining approximate deformation limits and successful process parameters, but more detailed criteria or models are required to provide more exact values. An experimental-analytical approach in which an experimental value reflecting the inherent material ductility is determined first would be most useful (this value is used to define a point on the fracture strain locus), followed by development of the rest of the fracture-limit line, as in Fig. 10 to 12, 14, 16, and 17.

Applications

The fracture criteria discussed previously in this article can be used as tools for troubleshooting fracture problems in existing processes or for designing/modifying processes for new products. In either case, graphical representation of the criteria permits independent consideration of the process and material parameters in quantitative or qualitative form.

An example is the bolt-heading process shown in Fig. 24(a). If it is required to form a bolt head diameter D from the rod of diameter d, the required circumferential strain is ln (D/d), indicated by the horizontal dashed line in Fig. 24(b). The strain paths that reach this level, however, depend on process parameters, as shown previously in Fig. 2(b), and the fracture strain loci vary with material, as shown in Fig. 10 to 12. Referring to Fig. 24(b), if the strain path labeled a describes the strain state at the expanding free surface for one set of processing conditions and the material used has a forming limit line labeled A, then, in order to reach the required circumferential strain, the strain path must cross the fracture line, and fracture is likely to occur. As shown, one option for avoiding defects is to use material B, which has a higher fracture limit. Another option is to alter the process so that strain path b is followed by the material. The latter option represents a process change, which in this case involves improved lubrication, as shown in Fig. 2. This procedure has been quantified and implemented in a computerized tool for upsetting process design (Ref 13).

In more complex cases, other means are available for altering the strain path, such as modification of die design, workpiece (preform) design, and redistribution of lubricant. Examples of application to powder forging preform design and other metalworking processes are given in Ref 14 and Ref 15.

The workability concept presented earlier in this article provides a useful supplement to the experience and intuition of the die designer, because it presents a graphical and quantitative description of the relationship between material and process parameters. Some examples of the application of the workability analysis procedures already described follow.

Bar Rolling. As shown in Fig. 3, the strains at the edges of bars during rolling are similar to those at the bulging free surface of a cylinder during compression. It should be possible, then, to predict fracture in bar rolling from compression tests on the alloy of interest. This is pertinent in current attempts to roll ingots of high alloy content into bar form. The complete workability study of bar rolling includes physical modeling of bar rolling to obtain the strain states at the edges of the bar, compression tests to obtain the material fracture limits, and comparison of the two sets of results to establish roll pass reduction limits.

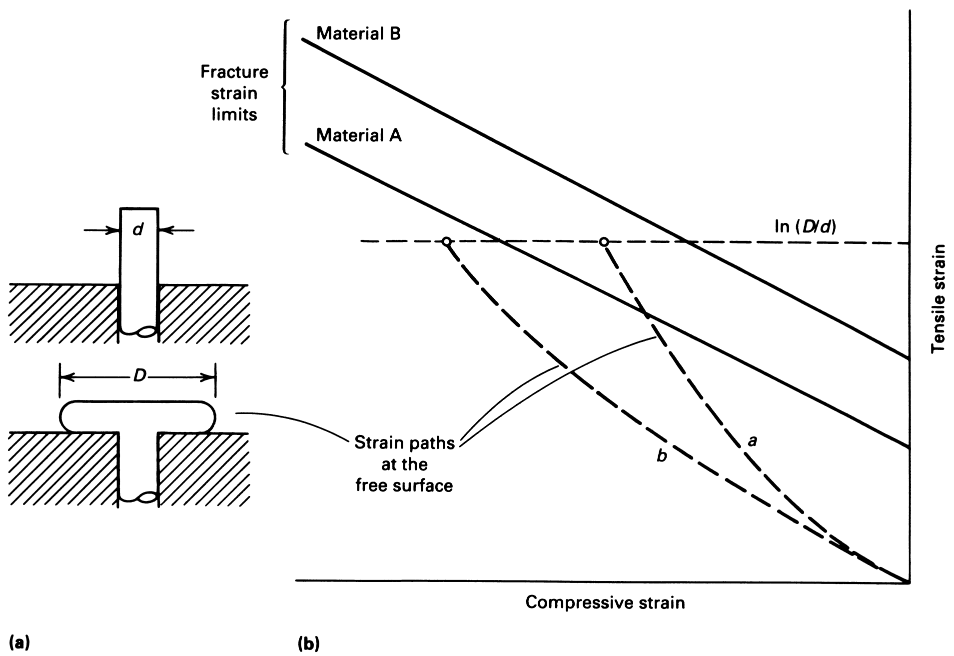


Fig. 24 Upsetting (a) of bar diameter *d* to head diameter *D*. (b) Material fracture strain limits are superimposed on strain paths reaching the final required strain. Strain path b (low friction) prevents fracture for both materials. Material B avoids fracture for either strain path.

Such a study is illustrated by the analysis of cracking during the rolling of 2024-T351 aluminum alloy bars. The intent was to roll square bars into round wire without resolutioning. Rolling was done on a two-high reversible bar mill with 230 mm (9 in.) diam rolls at 30 rpm (approximate strain rate: $4s^{-1}$). The roll groove geometry is shown in Fig. 25. Defects occurred primarily in the square-to-diamond passes (1–2 and 3–4), but the two diamond-to-square passes (2–3 and 4–5), the square-to-oval pass (5–6), and the oval-to-round pass (6–7) also were examined for completeness.

Lead was used as the simulation material for the physical modeling of bar rolling. Pure (99.99%) lead was cast and extruded into 25 mm (1 in.) round bars and then squared in the box pass (step 1, Fig. 25). Grids were placed on the lateral edges of the bars by an impression tool, and the grid spacing was measured before and after each pass for calculation of the longitudinal, ε_1, and vertical, ε_2, strains. Different reductions in area were achieved by feeding various bar sizes and by changing roll separation distances. A transverse slice was cut from the bars after each pass for measurement of the cross-sectional area and calculation of the reduction.

Results of the strain measurements are summarized in Fig. 26, in which tensile strain is plotted simultaneously with the compressive strain and reduction. As expected, the square-to-diamond passes involve the least compressive vertical strain, and the square-to-oval pass has the greatest compressive strain. The tensile strain versus reduction plot is the same for all cases, reflecting volume constancy.

Compression tests were performed on the 2024 aluminum alloy at room temperature and at 250 °C (480 °F) at a strain rate of $4s^{-1}$ to determine fracture limit lines. Straight, tapered, and flanged specimen profiles were used. Results are given in Fig. 27. Superposition of Fig. 26 onto Fig. 27 gives the rolling deformation limits.

To test the workability predictions, aluminum alloy bars were rolled at room temperature and at 250 °C (480 °F). Grid and area reduction measurements were made for the square-to-diamond passes. Figure 28 shows the measured strains at room temperature, which agree with

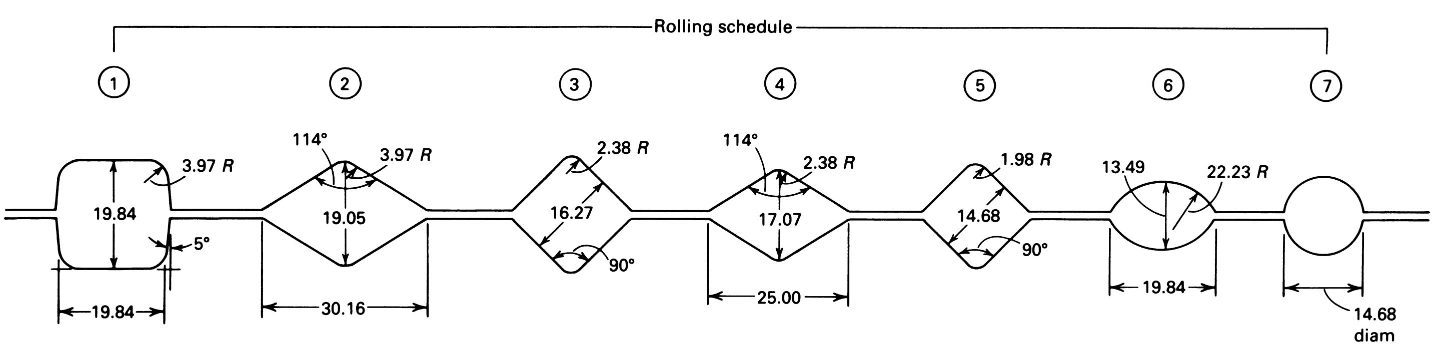


Fig. 25 Roll groove geometry for rolling square bars into round wire. Dimensions given in millimeters

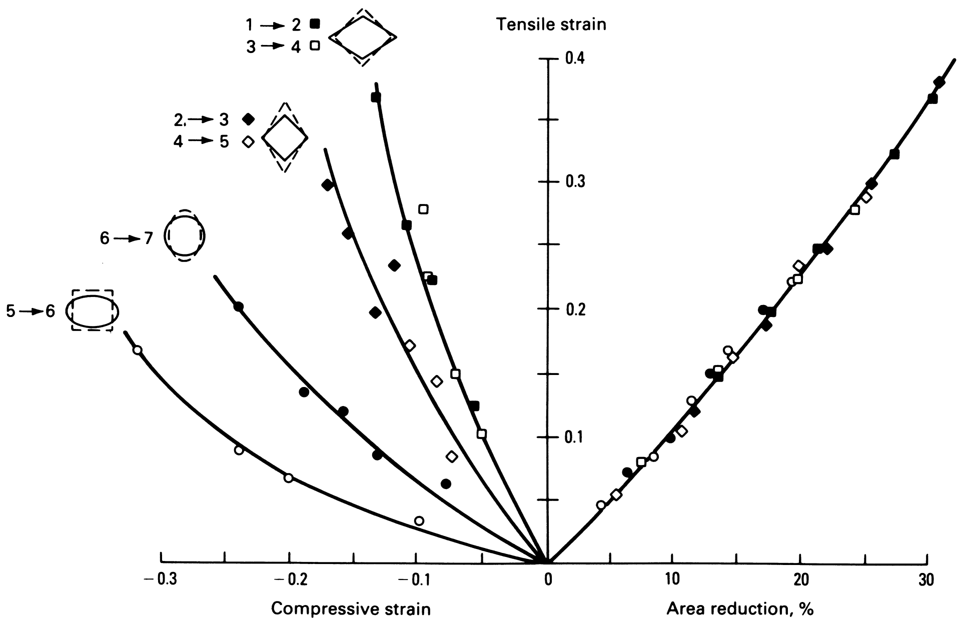


Fig. 26 Measured localized strains during the rolling of lead bars. Left side shows longitudinal tensile strain versus vertical compressive strain. Right side shows longitudinal strain versus cross-sectional area reduction at room temperature.

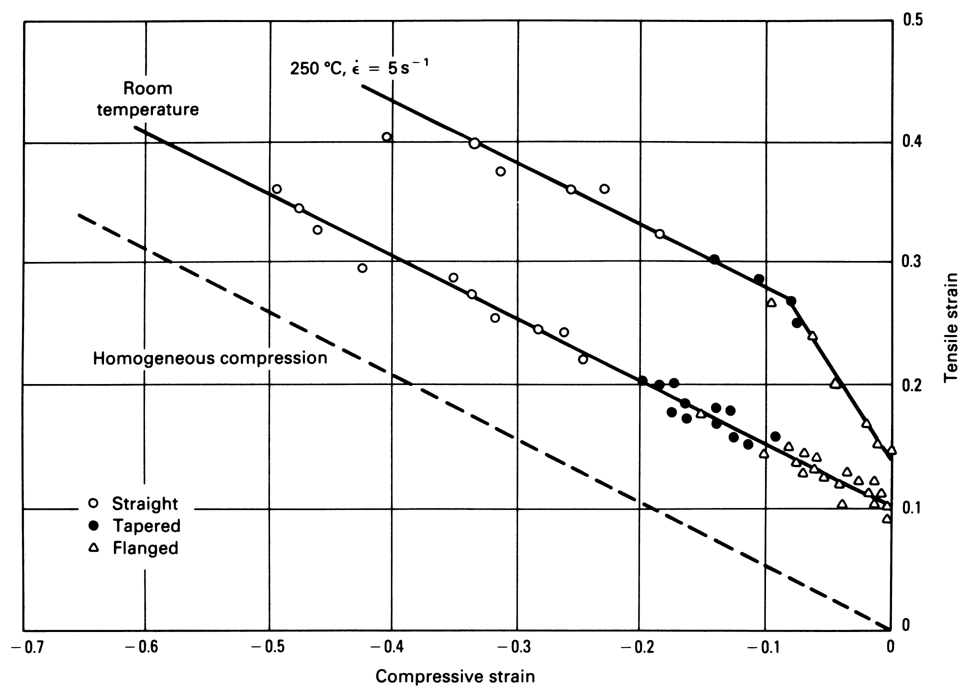


Fig. 27 Fracture strain lines for 2024 aluminum alloy in the T351 temper, measured by compression tests at room temperature and at 250 °C (480 °F)

those measured in lead bars for the same pass (Fig. 26). Open circles indicate fracture, and closed circles indicate no fracture. The fracture line for the aluminum alloy at room temperature is superimposed as the dashed line. It is clear that edge cracking in bar rolling conformed with the material fracture line, and the limiting reduction is approximately 13% for this combination of material and pass geometry. Similarly, at 250 °C (480 °F), there was conformance between fracture in bar rolling (Fig. 29) and the fracture line of the alloy (Fig. 28). In this case, the limiting reduction is approximately 25%.

Extrapolating the preceding results for cold rolling, the limiting reduction for diamond-to-square passes would be approximately 15%; for the oval-to-round pass, approximately 20%; and for the square-to-oval pass, approximately 25%. Similarly, in the hot rolling of this aluminum alloy, the reduction limit for diamond-to-square passes would be approximately 27%; for oval-to-round passes, approximately 30%; and for square-to-oval passes, approximately 40%. The latter two are beyond the reduction normally used because of fin formation, so cracking occurs rarely in such passes.

Example 1: Preform Design for a Ball Bearing Race. A low-load high-torque ball bearing outer race was cold forged from a low-alloy steel powder preform (Fig. 30). The preforms were compacted from 4600 grade powder with carbon added to give 0.20% C in the sintered material. The sintered preforms were 80% of theoretical density.

Initial efforts led to cracking through the preform along a diagonal beginning at the point of contact between the punch and preform (Fig. 31). The large shear stress developed by the contact was beyond the fracture limit of the porous preform, and two solutions were considered that would avoid such stresses (Fig. 32). The first solution was the use of a flat preform that involves back extrusion flow into the outer rim, and the second was the use of a tall, thin-wall preform that involves radial inward flow into the inner flange. The first option was rejected because it would generate circumferential tension that would most likely cause fracture. The second option is desirable because compression is applied at the top face; this option was pursued through physical modeling.

The primary concern with the second option (Fig. 32b) was the large amount of radially inward deformation required to form the inner flange. As a result, this option was examined by physical modeling. Model preforms were produced from sintered 601AB aluminum alloy powder and gridded on the inside surface (Fig. 33a). Grid displacements were measured after each of several increments of deformation, and the calculated strains were plotted along with the fracture line of the material (Fig. 33b). It is clear that both the axial and circumferential strains are compressive throughout the process and do not exceed the fracture line. Some wrinkling of the inside surface occurred, but this was smoothed out when the surface contacted the mandrel under pressure.

Actual production of straight-wall preforms as in Fig. 33(a) was not feasible, because the height-to-thickness ratio is too large for compaction. A compromise was developed in which the preform angle was 17° instead of 30°, as used in the original preform, or 0°, as used in the physical model (Fig. 33b). This ensured initial punch contact at the top face of the preform and generated compressive strains on the inside surface, as in Fig. 33(b). Cold-forging trials on these preforms produced no cracks, developed the desired full density in the ball path region, and showed the added benefit of a smooth ball path surface that did not require grinding.

Example 2: Back Extrusion of Copper Alloy. A low-ductility dispersion-strengthened copper alloy was back extruded into a cup shape, as shown in Fig. 34. The deformation was carried out at room temperature on a mechanical press. Crack formation on the rim caused high rejection rates. The original slugs for this process were smaller in diameter (16 mm, or 0.625 in.) than the die inside diameter (19 mm, or 0.75 in.), as shown in Fig. 35(a). Deformation of such preforms involves circumferential expansion strain (equal to ln (D/d), where D is the die bore diameter and d is the slug diameter) along with very little compressive strain at the rim (Fig. 35b). For this case, the circumferential strain is ln (0.75/0.625) = 0.18. Workability analysis would then require only measurement of the material fracture line and comparison with the required circumferential strain 0.18.

Fracture strains were measured in flange compression tests as shown in Fig. 36, giving a minimum circumferential strain of 0.2, which is sufficiently above the required strain for avoidance of fracture. A hydraulic press was used (giving a strain rate of approximately 0.5 s^{-1}), on the assumption that there is no strain rate effect at room temperature. Because the workability analysis showed that fracture should not be a

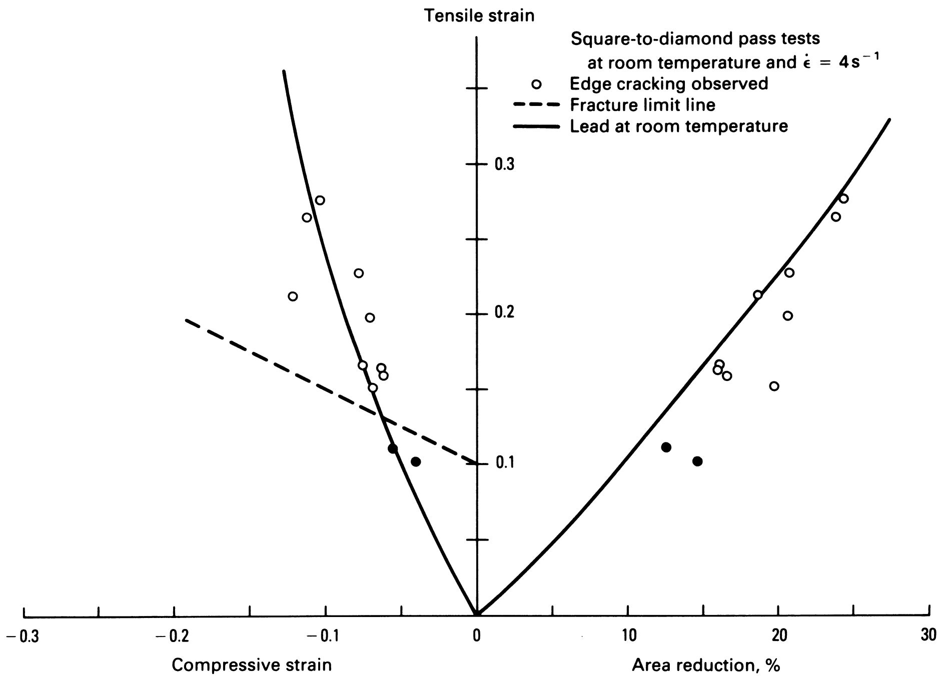


Fig. 28 Superposition of fracture line (dashed) on measured strains during rolling of 2024-T351 aluminum alloy bars at room temperature. Solid line represents the strain path measured during rolling of the lead model material shown in Fig. 26.

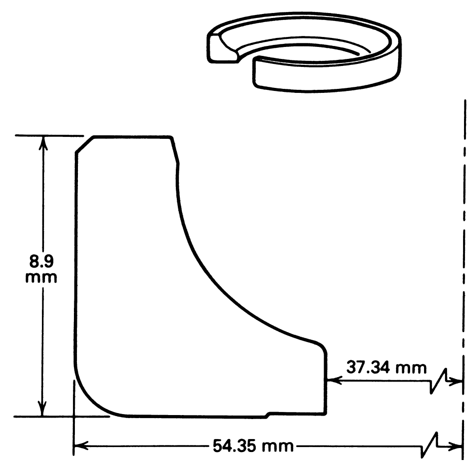


Fig. 30 Ball bearing outer race that was cold forged from sintered powder preform of 4620 low-alloy steel

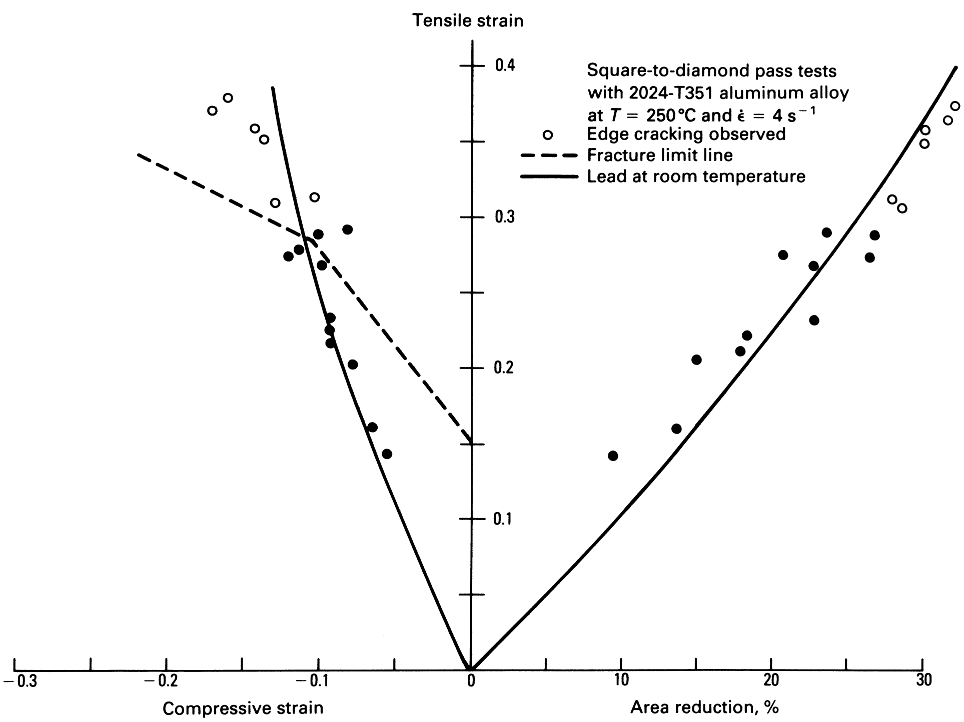


Fig. 29 Superposition of fracture line (dashed) on measured strains during the rolling of 2024-T351 aluminum alloy bars at 250 °C (480 °F). Solid line represents the strain path measured during rolling of the lead model material shown in Fig. 26.

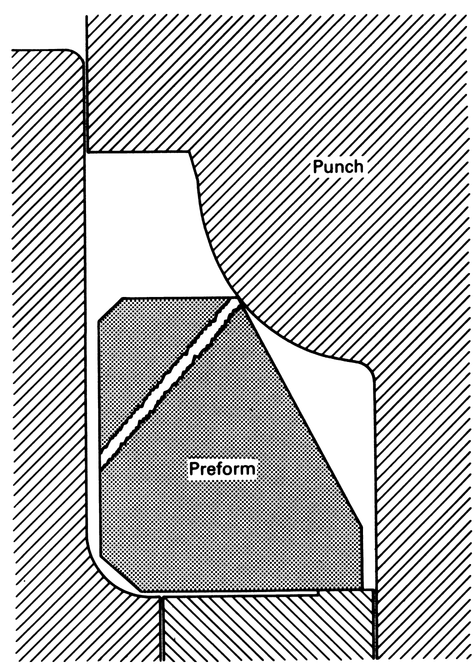


Fig. 31 Cracks initiated at the point of contact between the punch and preform in the original preform design. The preform had a taper of 30° on the inside diameter.

problem, the effect of strain rate was explored further.

Tests were performed on the same alloy using controlled strain rate servohydraulic test equipment at strain rates of 5, 10, and 15 s^{-1}; the third strain rate given is close to that in the production mechanical press. Figure 37 shows the surprising result that the fracture limit line decreases with increasing strain rate. In particular, the minimum circumferential strain falls below the required value of 0.18 for successful forming of the rim; this explains the occurrence of fracture on the production press. The problem was corrected by using slugs of larger diameter to decrease the circumferential tension and by preforming a taper on the top face (Fig. 38), which produced some axial compression in the material at the rim. The strain path then avoided crossing the fracture line (Fig. 39), and the rejection rate during production on the mechanical press was nil.

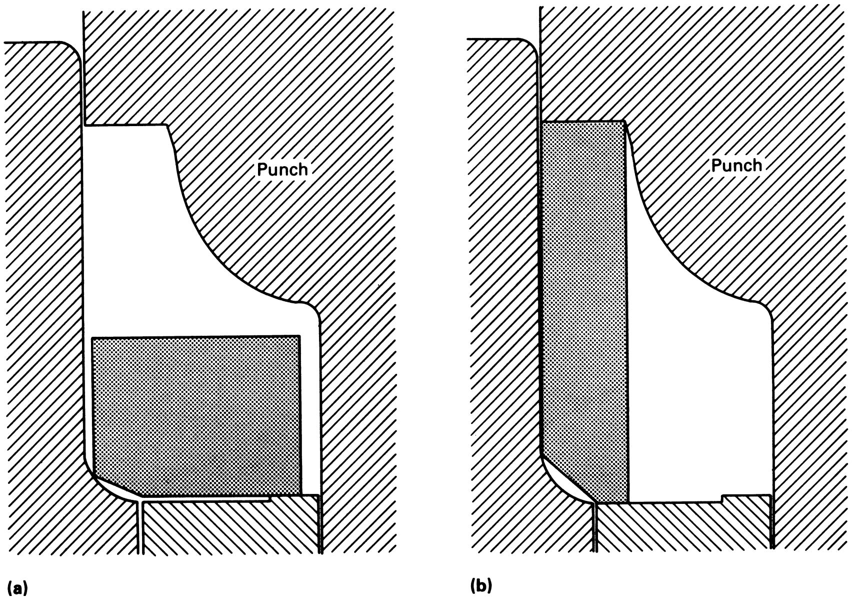


Fig. 32 Preform alternatives for forging the ball bearing outer race shown in Fig. 30. (a) Back extrusion. (b) Compression and radial inward flow

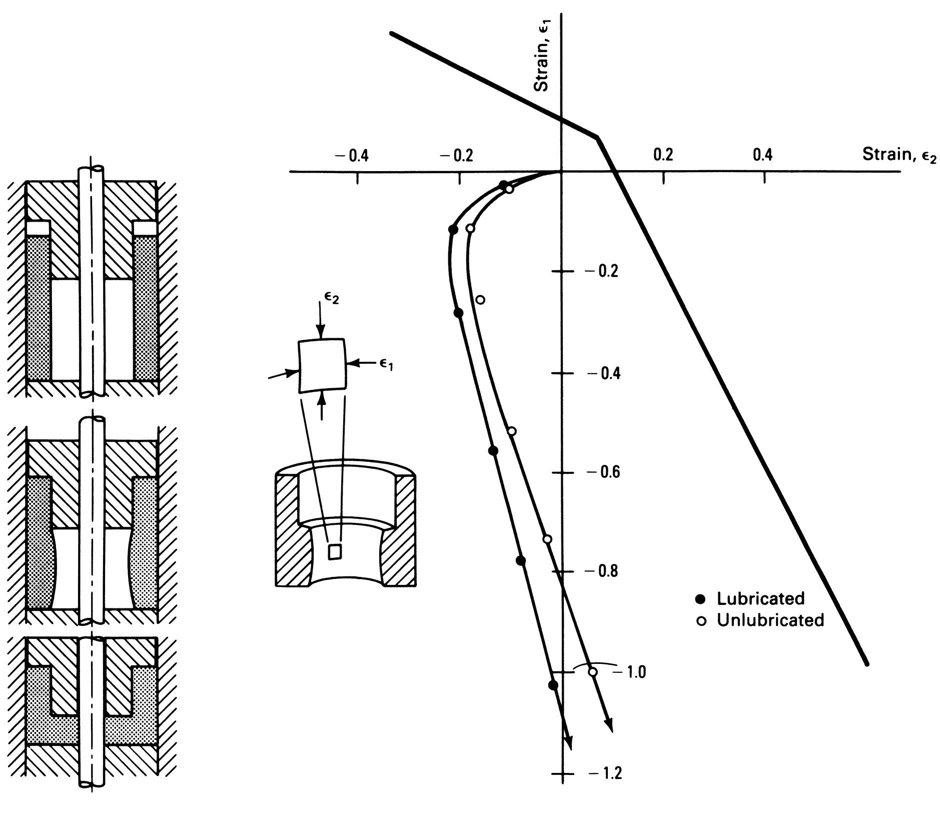


Fig. 33 (a) Physical model for the second option (Fig. 32b) and (b) the measured strains during forging of the preform. The heavy line is the material fracture line. It is clear that the strain path never crosses the fracture line and that defects are prevented.

Contact Surface Fracture and Internal Fracture. All of the previous applications and examples involved free surface fractures and could be treated directly by the fracture line. Consideration of contact surface fractures (Fig. 5) and internal fracture (Fig. 6), however, requires modification of this approach or use of a new approach. In the following, an example is given of the application of the upper bound and tensile stress criteria to central burst in extrusion. The empirical workability concept described previously is then modified for application to contact surface fracture as well as central burst.

Example 3: Central Burst During Extrusion. Central burst can occur in extrusion when light reductions and large die angles are used (Fig. 19), and it is encountered in the production of shafts for transmissions and suspension systems. A test of the central burst criterion was carried out by processing shafts from hot-rolled 1024 steel bars 22 mm ($\frac{7}{8}$ in.) in diameter (Ref 16). The processing sequence consisted of initial drawing followed by three extrusion steps in a boltmaker:

Process	Reduction, %	Die half-angle, degrees
Drawing	8	9
Extrusion	22	22.5
Extrusion	23	22.5
Extrusion	16	22.5 or 5

All passes are in the central burst area of Fig. 40, except for the last pass with a 5° die angle.

A total of 1000 shafts were processed with the 22.5° die, and 500 shafts were processed with the 5° die. All shafts were tested ultrasonically for internal defects. Central bursting was de-

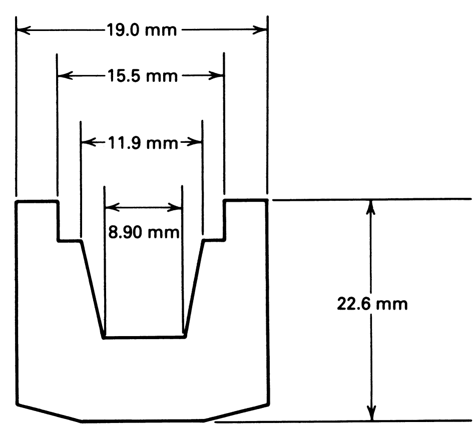


Fig. 34 Part that was back extruded from copper

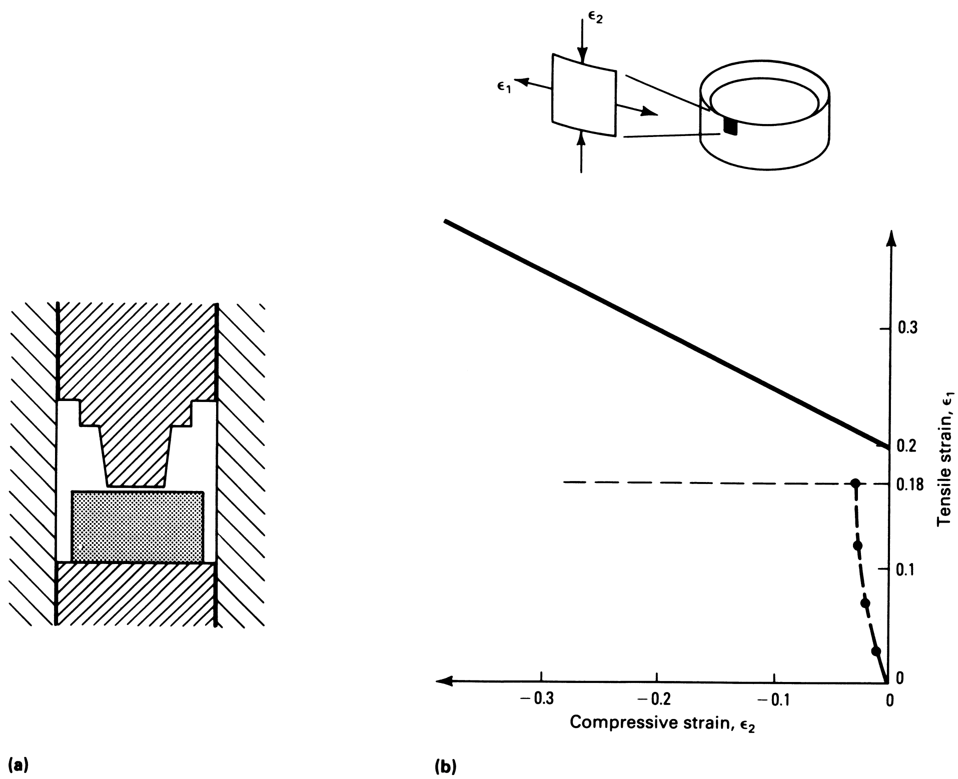


Fig. 35 Back extrusion of the cup shape shown in Fig. 34. (a) The preform slug was 16 mm (0.625 in.) in diameter, the die was 19 mm (0.75 in.) in diameter. (b) Strains at the cup rim where fracture occurred consist of circumferential tension to a value of 0.18 and very little compressive strain. The heavy line is the material fracture strain line.

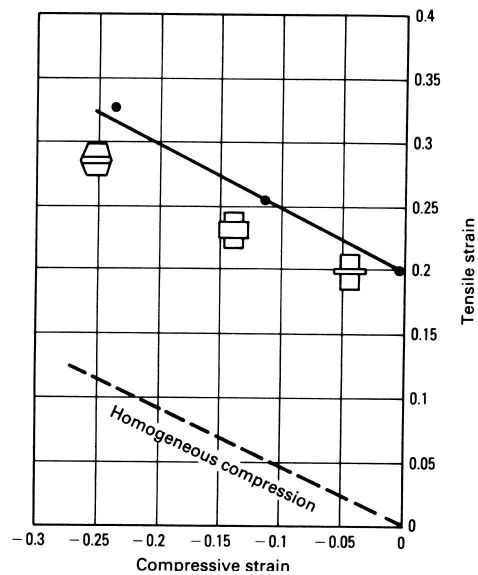


Fig. 36 Fracture strain line for copper alloy, determined by flanged and tapered compression tests. Specimen geometries used for each test are shown also.

tected in 4.5% of the shafts extruded with the 22.5° die, and no defects were detected in the shafts extruded with the 5° die. These results show that the upper bound central burst criterion is a necessary condition. It was further shown in Ref 16 that central burst was avoided in other heats with slightly different compositions because their strain-hardening coefficients were larger than the original heat. This confirmed the predicted results in Ref 10.

Modified Empirical Criterion. It was shown previously in this article that measured free surface strains at fracture fit a linear or bilinear line that constitutes a fracture locus for the material tested (Fig. 10 to 12). This is a convenient representation of the complexities of ductile fracture that are controlled by stress and deformation. The experimental fracture locus is also reproduced by several theoretical fracture criteria (Fig. 14, 16, and 17).

For contact surface and internal fractures, however, the surface on which the strains can be monitored is subjected to stress normal to that surface. It was shown in Eq 3 that stress states leading to a given set of surface strains differ only by a hydrostatic stress component, and this component is equal to the applied stress normal to the surface on which the strains are monitored. Experience shows that this hydrostatic stress affects fracture, and it should also affect the fracture strain locus. It should be possible, then, to use the theoretical fracture criteria to predict the effects of hydrostatic stress on the fracture strain locus.

The simplest criterion described previously in this article is that due to Cockcroft; therefore, it was modified to predict the effects of stress normal to the plane (Fig. 5 and 6) on the fracture strains ε_1 and ε_2. The result (Fig. 41) shows that superimposed pressure ($P > 0$) increases the

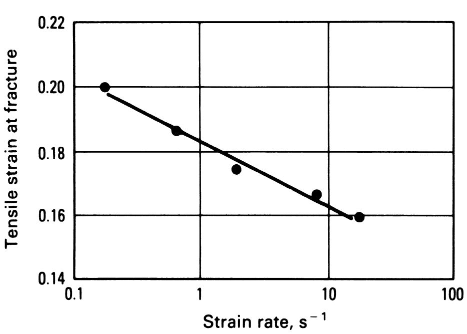


Fig. 37 Decrease in circumferential tensile strain at fracture with increasing strain rate for the copper alloy tested in Fig. 36. Results are for thin-flanged compression specimens, which have the lowest fracture strain.

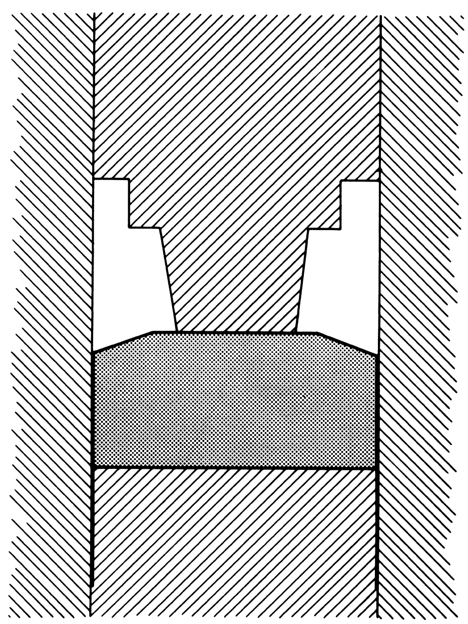


Fig. 38 Modified preform slug for the back extrusion of the cup shape shown in Fig. 34. The slug diameter is 18.8 mm (0.74 in.) and has a 5° taper on the top surface.

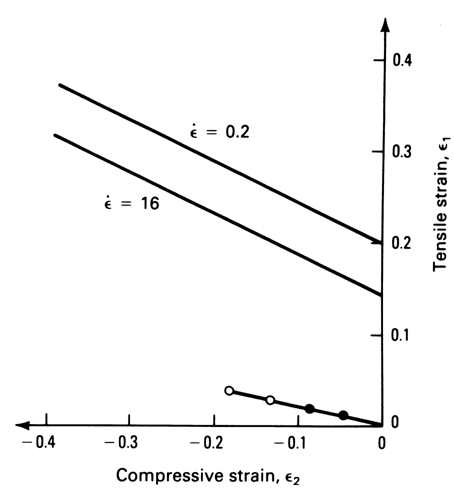


Fig. 39 Comparison of measured strains at the cup rim during back extrusion of the modified preform slug shown in Fig. 38. The strains do not exceed the material fracture strain line for low or high strain rate forming.

height of the fracture strain line and also increases its slope slightly. Superimposed tension ($P < 0$) decreases the height of the fracture line, decreases its average slope, and gives it a slight downward curvature. It is clear that the increase in strains to fracture due to additional pressure is unlimited as pressure increases, but the strains to fracture due to additional tension are limited by zero as tension increases. This result is discussed with regard to internal fracture and die contact surface fracture in the following paragraphs.

Central Burst in Forgings. Internal fractures along the centerline of extruded or drawn bars were discussed earlier (Fig. 18, 19, 23, and 40). Similar fractures are observed in forged shapes such as that shown in Fig. 42 for heat-treated 6061 aluminum alloy. Here, as the outer region is compressed between dies, material flows radially inward and then vertically into the opposed hubs. This develops a hydrostatic tensile stress state at the center (Fig. 6a), and fracture is a strong possibility.

Through visioplasticity analysis on split, gridded specimens (Fig. 42), the strain and stresses at the center of the workpiece were calculated for several increments of deformation. The hydrostatic stress state at the center of the specimen is not always tensile; initially, it is compressive, and then it reverses, becoming tensile as the flange thickness is reduced and flow into the hub occurs. Meanwhile, the strains at the center are increasing monotonically as deformation progresses. This is illustrated in Fig. 43 by the steps 0-1-2-3. As deformation proceeds, the strains at the center increase, but the hydrostatic pressure is also increasing, so the fracture line moves upward. Then, as the flange thickness approaches one-half of the hub base diameter (die orifice diameter), the hydrostatic stress becomes tensile, so the fracture line decreases in height. The strains at the center continue to rise, however, and cross the fracture line, leading to the central burst. The calculated hydrostatic tension at fracture was $0.3Y$. This approach could be used for predicting central burst in drawing and extrusion to provide a material-dependent criterion, as opposed to the more simplistic upper bound and tensile stress criteria described previously.

Die Contact Surface Fracture. Frequently, cracks occur during forging on surfaces that are in contact with the dies (Fig. 5). One common location of such defects is the vicinity of a die or punch corner. From the observation of a variety of such defects, it appears that a common characteristic is an abrupt change in frictional shear traction distribution in the region of the crack. High friction to retard metal flow in advance of the crack location is one method for preventing such defects.

A technique for studying die contact surface cracks was developed by means of a disk compression test and dies having a rough surface in the central region and a smooth surface in the outer region. Figure 44 shows the top view of a 6061 aluminum alloy disk compressed between such dies. In the transition region between the rough central die surface and the smooth outer region, radial cracks initiate and propagate outward. Such cracks occurred at approximately 30% reduction when the smooth outer region was lubricated with Teflon. The cracks occurred at approximately 45% reduction when grease lubrication was used in the outer smooth region. No cracks occurred even for very large reductions when the smooth outer region was not lubricated.

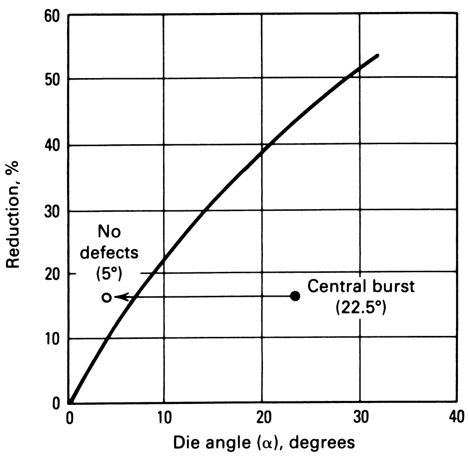


Fig. 40 Location of process conditions on a theoretical central burst map. For an angle of 22.5°, central burst occurred in 4.5% of the extruded shafts. For a die angle of 5°, no central burst occurred.

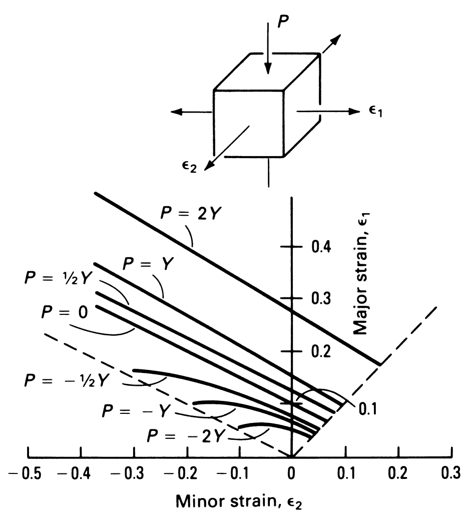


Fig. 41 Movement of the fracture strain line due to superimposed hydrostatic stress. Applied stress is represented in terms of multiples of the yield strength, Y. Negative values of P indicate hydrostatic tension. Calculations are based on a modification of the Cockcroft criterion.

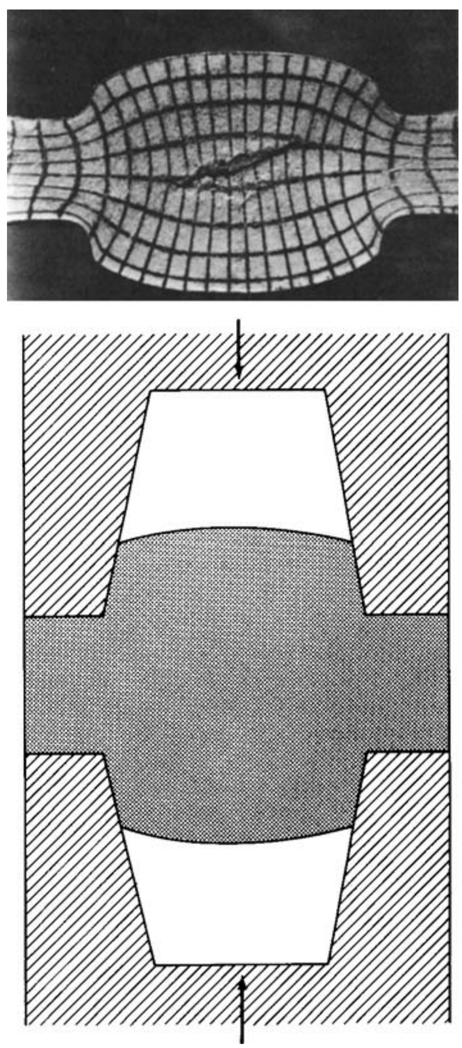


Fig. 42 Internal fracture during the double-extrusion forging of aluminum alloy 6061. Grid deformations on the middle longitudinal plane are shown. Stress and strain states are defined by Fig. 6(a).

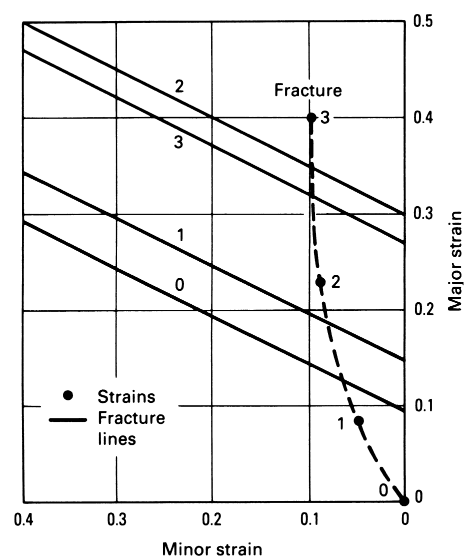


Fig. 43 Progression of surface strains and fracture line at the central internal location of the double-extrusion forging shown in Fig. 42. The fracture line rises from 0 to 1 to 2 as internal pressure increases and then falls to point 3 as the internal stress becomes tensile.

Grid marks placed on the die contact surface of the disks were used to measure the distribution of surface strains in the radial direction. Figure 45 gives an example of such measurements. In the rough central region the strains are zero, while in the smooth outer region the strains are equal and constant. In the transition, however, the circumferential strain, ε_θ, jumps abruptly from zero to its constant value in the outer region, and the radial strain, ε_r, overshoots to a very high value before returning to its constant value in the smooth outer region.

The strains shown in Fig. 45 were the same regardless of the friction condition in the smooth outer region. Therefore, fractures in the transition region occur because of the combination of large tensile surface strains and low hydrostatic stress state. This explains the occurrence of cracks at low reduction when Teflon is used, and no occurrence of fracture when no lubricant is used. The Teflon, having a near-zero friction coefficient, results in very low radial back pressure on the transition region, while grease and no lubricant provide progressively larger back pressures.

By means of visioplasticity analysis, the stresses were determined at the contact surface in the vicinity of the transition region. The resulting normal die pressure plus the surface radial and circumferential strains define the stress and strain states in the transition region and can be illustrated on a forming limit diagram. Figure 46 shows the change in surface strains and the increase in the fracture line due to increasing normal pressure during compression of a disk with grease lubricant in the outer region (indicated by the increments of reduction to 45%). The fracture line increases at a slower rate than the strains increase with increasing pressure, and at 45% reduction the strain path exceeds the fracture line and cracks are observed. For Teflon lubricant, the crossover occurs at about 30% reduction, while in the case of no lubricant the fracture line moves progressively away from the strain path.

Example 4: Fir Tree Defect. The criterion for contact surface fracture can be applied qualitatively for interpretation of the fir tree defect in extrusion. Such defects occur on the surfaces of extruded bars as well as in localized areas of forgings containing ribs. In a section of a rib-web forging from a preform of sintered aluminum alloy powder, small cracks formed with a regular spacing on the rib surfaces (Fig. 47). Such defects occurred only when the thickness of the rib was greater than approximately one-half the web length; that is, extrusion reduction was less than one-half. Each crack formed by shear at the corner as material flowed from the web into the vertical rib.

The stress and strain state on a surface material element is shown in Fig. 48. As material in the web is compressed, the surface element experiences tensile strain in the direction of flow, and the strain increases as the element approaches the die corner. At the same time, there is a compressive stress from the die onto the surface element. This pressure diminishes, however, as the element nears the die corner and almost disappears as the element moves around the corner.

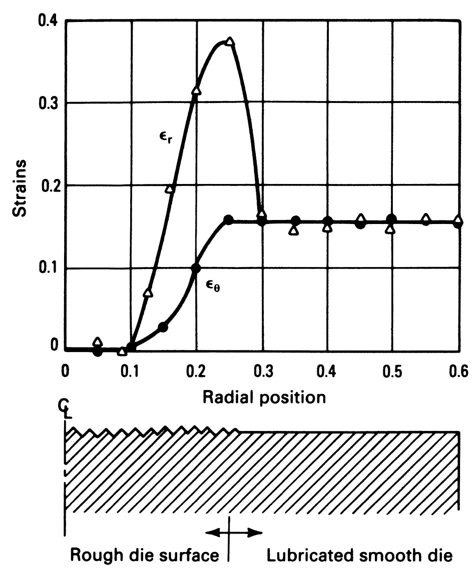


Fig. 45 Radial variation of contact surface strains after 30% compression of the disk shown in Fig. 44

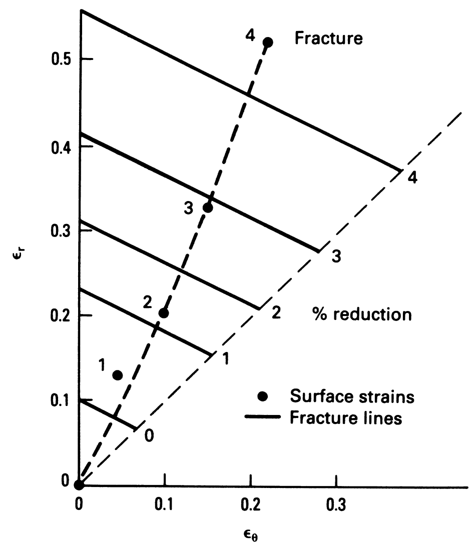


Fig. 46 Progression of surface strains and fracture line at the transition region between rough and smooth zones of the compressed disk shown in Fig. 45. Points 1, 2, 3, and 4 represent 10, 20, 30, and 45% reduction, respectively.

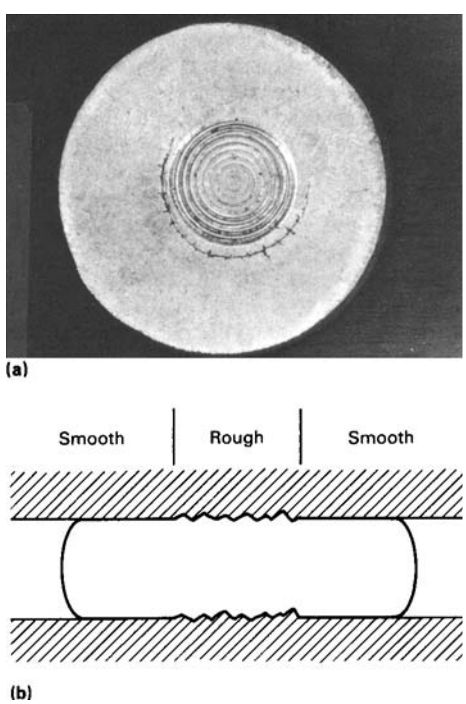


Fig. 44 (a) Top view of aluminum alloy 6061 disk compressed between dies. (b) Cracks form at the transition region between rough and smooth areas of the die.

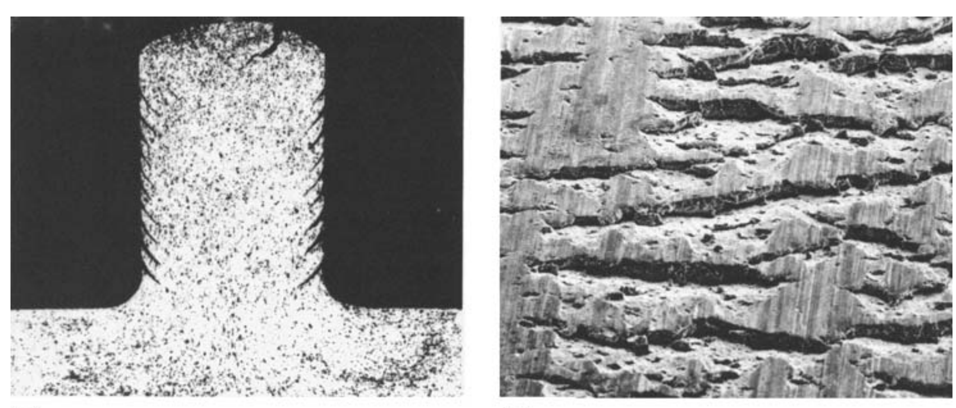


Fig. 47 Die contact surface cracking during forging extrusion of aluminum alloy powder compact. (a) Cross section. (b) Normal to vertical rib surface. Note also the cracks at the top free surface. Stress and strain states are defined in Fig. 4(b).

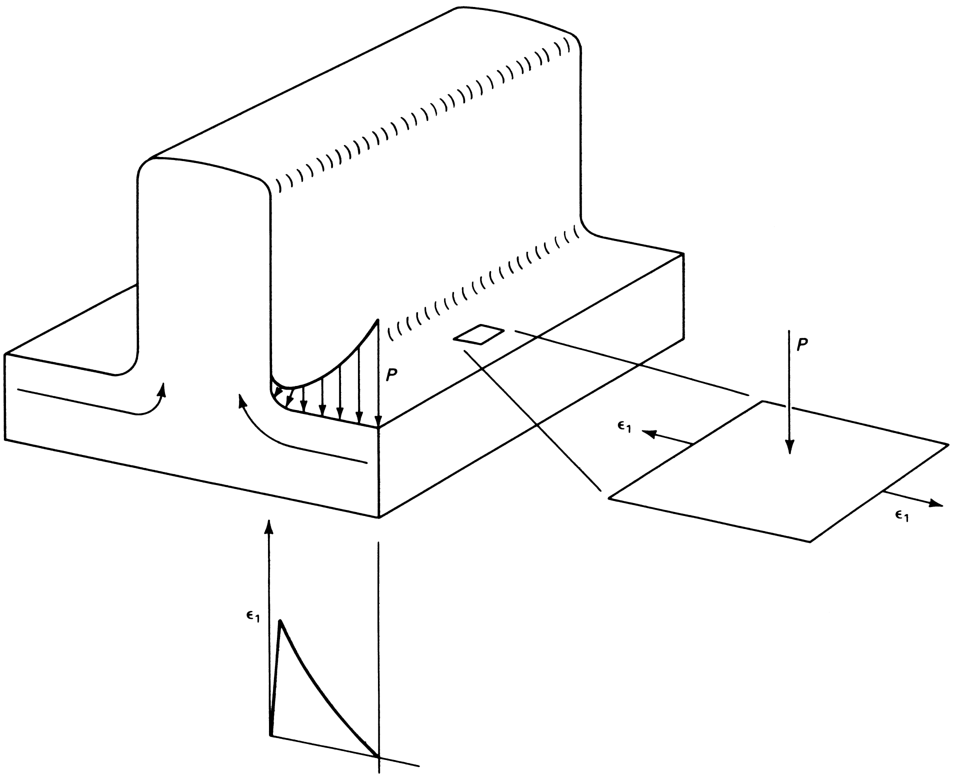


Fig. 48 Increase in strain ε_1, and decrease in die contact pressure P on a surface element as it moves from the web into the rib section of the extrusion forging shown in Fig. 47.

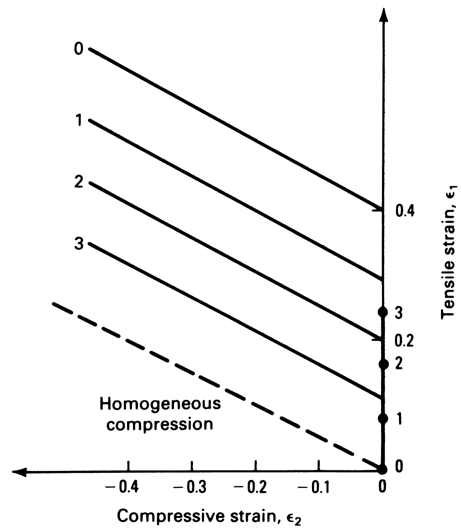


Fig. 49 Progression of strain ε_1, and decline of fracture line due to decrease in pressure P for an element moving from the web to rib section during the forging extrusion shown in Fig. 48. At stage 3, the strain exceeds the fracture line.

This can be illustrated schematically on the fracture strain diagram shown in Fig. 49. Because the deformation is in a state of plane strain, the strain path is represented as a vector of increasing length along the vertical axis. Meanwhile, the fracture line decreases in height because the pressure acting normal to the element is progressively decreasing. When the strains cross the fracture line, fracture occurs.

This phenomenon does not take place when extruding ribs of small thickness, because the extrusion reduction, and therefore the pressure, is larger, which maintains the fracture line at a high level. For thick ribs, two solutions were considered. One approach is to increase friction along the rib walls by roughening the die surface or avoiding lubrication of the die rib. This produces greater back pressure at the die corner, elevating the fracture line and preventing cracking. Such an approach is difficult to implement and can be used only with segmented dies because the formed rib cannot be removed from the die. The second approach is to use a draft angle on the rib, which has the same effect as increased friction. An angle of 10° prevented fracture in the current case, but other alloys may require a smaller or larger angle. A quantitative analysis combining the pressure effect on the fracture line and plasticity analysis would provide a method of predicting the draft angle in order to prevent fracture.

REFERENCES

1. H.A. Kuhn, P.W. Lee, and T. Erturk, A Fracture Criterion for Cold Forging, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 95, 1973, p 213–218
2. H. Ortiz and H.A. Kuhn, Physical Modeling of Bar Rolling for Workability Study, *Physical Modeling of Bulk Metalworking Processes, Conf. Proc.,* ASM International, 1987
3. E. Erman and H.A. Kuhn, Novel Test Specimens for Workability Measurements, *Compression Testing of Homogeneous Materials and Composites,* STP 808, ASTM International, 1983, p 279–290
4. F.A. McClintock, *J. Appl. Mech. (Trans. ASME),* Vol 90, 1968, p 363
5. Z. Marciniak and K. Kuczynski, A Model of Localized Thinning in Sheet Metalforming, *Int. J. Mech. Sci.,* Vol 9, 1967, p 609
6. P.W. Lee and H.A. Kuhn, Fracture in Cold Upset Forging—A Criterion and Model, *Metall. Trans. A,* Vol 4, 1973, p 969–974
7. M.G. Cockcroft and D.J. Latham, Ductility and the Workability of Metals, *J. Inst. Met.,* Vol 96, 1968, p 33–39
8. B. Avitzur, *Metal Forming—Processes and Analysis,* McGraw-Hill, 1968
9. W.M. Evans and B. Avitzur, "Die Design for Drawing Extrusion," Paper MF67–582, Society of Manufacturing Engineers, 1967
10. B. Avitzur, "Strain Hardening and Strain Rate Effects in Plastic Flow through Conical Converging Dies," Paper 66-Prod-17, American Society of Mechanical Engineers, 1966
11. R. Hill, On the Inhomogeneous Deformation of a Plastic Lamina in a Compression Test, *Philos. Mag.,* Vol 41, 1950, p 733
12. J.F. Nye, Experiments on the Plastic Compression of a Block between Rough Plates, *J. Appl. Mech. (Trans. ASME),* Vol 19, 1952, p 337
13. J.J. Shah and H.A. Kuhn, An Empirical Formula for Workability Limits in Cold Upsetting and Bolt Heading, *Proceedings of 13th NAMRC,* Society of Mechanical Engineers, 1985
14. C.L. Downey and H.A. Kuhn, Application of a Forming Limit Criterion to Design of Preforms for Powder Forging, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 97H, 1975, p 121
15. H.A. Kuhn, Deformation Processing of Sintered Powder Materials, *Powder Metallurgy Processing,* H.A. Kuhn and A. Lawley, Ed., Academic Press, 1978, p 99
16. Z. Zimerman, H. Darlington, and E.H. Kottcamp, Jr., Selection of Operating Parameters to Prevent Central Bursting during Cold Extrusion, *Mechanical Working and Steel Processing,* The Metallurgical Society, 1970, p 405

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 188-207
DOI:10.1361/hwpd2003p188

Chapter 13

Workability in Forging

S.L. Semiatin, Air Force Research Laboratory, Materials and Manufacturing Directorate

WORKABILITY IN FORGING depends on a variety of material, process-variable, and die-design features. In general, forging is a term that can be used to describe a wide variety of bulk metalworking processes, and forging engineers must consider a variety of workability tests to diagnose forging problems for a wide range of possible forging operations. In basic form, however, all forging processes consist of compressive deformation of a piece of metal, known as the workpiece, between a pair of dies. Depending on the geometry of the dies, varying amounts of lateral constraint may be imposed on the workpiece, a fact that enables forging operations to be classified very broadly into two categories: open-die forging and closed-die forging (Ref 1–3).

In open-die forging, the lateral constraint is minimal, and the amount and distribution of lateral metal flow are controlled by factors such as total reduction in the workpiece dimension parallel to the forging direction, frictional boundary conditions, and heat-transfer effects (when the dies and workpiece are at different initial temperatures). Simple two-dimensional examples of such operations include axial compression of right-circular cylinders between flat, parallel dies (a generic *upsetting* operation) and pressing of long, round bars along their lateral surfaces, also between flat, parallel dies—a process aptly known as *sidepressing*. The former is perhaps the simplest type of axisymmetric deformation, and the latter the simplest type of plane-strain deformation commonly used in forging practice. Nevertheless, such basic operations find wide commercial application in processing sequences for complex parts. For instance, upsetting is often used to "pancake" round billets prior to secondary processing via a closed-die forging operation. Furthermore, open-die forging is frequently used in conjunction with annealing treatments during the breakdown of cast ingots in order to obtain recrystallized, wrought microstructures.

As its name implies, closed-die forging is done in closed or impression dies that impart a well-defined shape to the workpiece. As might be expected, the degree of lateral constraint varies with the shape of the dies and the design of the peripheral areas where flash is formed, as well as with the same factors that influence metal flow in open-die forging (amount of reduction, frictional boundary conditions, and heat transfer between the dies and the workpiece). Simple closed-die forgings may also be largely axisymmetric or plane strain in nature. Gears for automotive and industrial uses and rotor blades for steam and gas turbine engines are among the most common parts in these categories that are made using closed-die techniques. In addition, closed-die forging is often employed for manufacture of more complex parts in which the metal flow is three-dimensional in nature. These include structural shapes consisting of thin webs and long, thin protrusions known as ribs, a class of parts with wide application in the aircraft industry.

The complete description of any forging process, be it basically an open-die or a closed-die process, requires specification of various process variables besides the die geometry and workpiece material. These include working speed, die temperature, workpiece preheat temperature, and lubrication. Working speed is determined by the kind of equipment used and varies depending on whether hydraulic presses (relatively low forging speeds), mechanical and screw presses (intermediate speeds), or hammers (high speeds) are selected (Ref 1).

Because forging is a relatively complex process, fracture or generation of undesirable defects in the workpiece during deformation may take a wide variety of forms, several of which are summarized in Table 1. For this reason, it is not surprising that no single workability test may be relied on for determination of forgeability. However, a number of test techniques have been developed for gaging forgeability depending on alloy type and microstructure, die geometry, and process variables. This chapter summarizes some of the more common tests and illustrates their application in practical forging situations.

The term *workability* is also used in reference to the determination of flow stresses and total working loads in metal-forming practice (which thus determines equipment utilization factors such as required press capacity, die materials and die wear, and certain other die-design features). Such topics are summarized in other chapters of this Handbook and in other sources (Ref 1, 4) and thus are not discussed here. In addition, workability can also refer to the design of thermomechanical processes to control and optimize microstructure for specific service applications. The development of desirable microstructure during forging plays a critical role in the design of forging processes. The general aspects of microstructure evolution during thermomechanical processing are briefly reviewed in Chapter 3, "Evolution of Microstructure during Hot Working," in terms of both phenomenological as well as mechanism-based approaches.

Table 1 Common metallurgical defects in forging

Temperature regime	Metallurgical defects in: Cast grain structure	Metallurgical defects in: Wrought (recrystallized) grain structure
Cold working	(a)	Free-surface fracture Dead-metal zones (shear bands, shear cracks) Centerbursts Galling
Warm working	(b)	Triple-point cracks/fractures Grain-boundary cavitation/fracture
Hot Working	Hot shortness Centerbursts Triple-point cracks/fractures Grain-boundary cavitation/fracture Shear bands/fractures	Shear bands/fractures Triple-point cracks/fractures Grain-boundary cavitation/fracture Hot shortness

(a) Cold working of cast structures is typically performed only for very ductile metals (e.g., dental alloys) and usually involves many stages of working with intermediate recrystallization anneals. (b) Warm working of cast structures is rare.

Workability Tests for Open-Die Forging of Cast Structures

In open-die forging, metal-flow patterns and stress states are not highly complex, and forgeability is determined primarily by material structure and properties and process conditions and only secondarily by die geometry. Material-structure variables include grain structure and texture, crystal structure, and the presence of second phases or solid-solution elements. Material properties include temperatures at which melting, recrystallization, and phase changes occur; the variation of flow stress with strain, strain rate, and temperature; and physical properties such as density, specific heat, and thermal conductivity. Of the process variables mentioned previously, workpiece preheat temperature is the most important.

The variation of workability with workpiece preheat temperature depends a great deal on grain size and grain structure, factors that influence the choice of workability tests (Fig. 1). When the grain size is large, as in conventionally cast ingot structures, cracks may initiate easily (and then propagate along the grain boundaries) due to the large stress concentrations associated with slip within the grains or unrelaxed sliding at *triple points* (where three grain boundaries intersect). Also, as is often the case in cast structures, impurities are segregated to the center and top of the ingot (producing areas of low workability there), and chemical elements are not distributed uniformly on either a microscopic or a macroscopic scale (Ref 5). Thus, the temperature range over which ingot structures can be forged is very limited. Typically, they must be hot worked, that is, worked at temperatures greater than three-fifths of the absolute melting point of the alloy. The upper temperature limit in this range is usually slightly below the melting point of the metal. The melting point of an alloy in the as-cast condition is usually lower than that of the same alloy in the finer-grain, recrystallized state because of chemical inhomogeneities and the presence of low-melting-point compounds, which are often found at the grain boundaries. Forging at temperatures too close to the melting points of these compounds may lead to grain-boundary cracking when heat developed during deformation raises the workpiece temperature and causes melting of the compounds. This fracture mechanism is called hot shortness and is sometimes avoided by forging at deformation rates sufficiently low to allow the heat developed by deformation to be dissipated into the dies, by use of lower working temperatures, or by use of homogenization heat treatments prior to working. The last of these methods is especially useful for the breakdown of highly alloyed metals such as nickel-base superalloys.

The lower temperature limit for hot forging of cast structures is the temperature at which there is some, but not a great deal of, dislocation multiplication. Dislocations are needed to build a reservoir of stored energy that, in combination with thermal energy (supplied during hot deformation or during postdeformation heat treatment), is useful in breaking down the coarse grains and generating a much finer recrystallized grain structure. However, the benefit of dislocation generation must be balanced against the reduction with temperature of grain-boundary mobility, a phenomenon that leads to dynamic recrystallization and offsets the tendency for intergranular fracture due to grain-boundary sliding (Ref 6). Low forging temperatures may also lead to precipitation of alloying elements that remain in solution at higher temperatures, an occurrence that tends to pin matrix dislocations and also to cause grain-boundary failure. Precipitation at grain boundaries may also be harmful when it provides sites for grain-boundary cavitation. From a process-variable standpoint, the lower temperature limit for working of cast (and wrought) structures can be further restricted by the effects of die chilling. Because hot working is conventionally carried out with a hot workpiece and usually much cooler tool steel dies, heat transfer brought about by die chilling may lead to workpiece temperatures during deformation that are much lower than the preheat temperature. In terms of die design, secondary tensile stresses are often set up in open-die forging (Ref 7).These stresses may assist the grain-boundary fracture process, giving rise to gross centerbursts, and should be evaluated in workability tests as well.

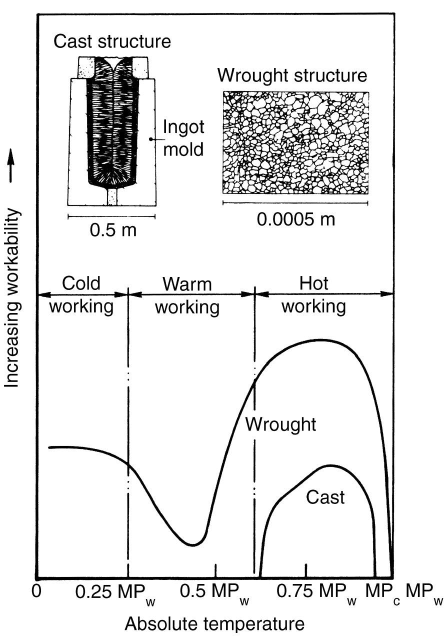


Fig. 1 Schematic illustration of relative workability of cast metals and wrought and recrystallized metals at cold, warm, and hot working temperatures. The melting point (or solidus temperature) is denoted as MP_c (cast metals) or MP_w (wrought and recrystallized metals).

From the previous discussion, it is clear that workability tests for open-die forging of cast structures should be designed in order to determine realistic temperature regimes in which a recrystallized structure may be obtained and fracture avoided. Two such tests are the wedge-forging test and the sidepressing test.

The wedge-forging test (Ref 2, 8) is ideal for determining breakdown temperature ranges for ingot structures. In this test, a wedge-shaped piece of metal is machined from the ingot (Fig. 2) and forged between flat, parallel dies. The overall dimensions of the wedge are selected on the basis of initial grain size. Large-grain materials require large specimens, and fine-grain materials require small specimens. In either case, a specimen large enough to be representative of the bulk of the ingot should be chosen. In many cases, a number of specimens are required because of the variety of grain shapes (columnar versus equiaxed), grain sizes, and grain orientations relative to the working direction that pertain to ingot breakdown. Specimen design and subsequent interpretation of workability data may require that the crystallographic textures that are developed during ingot solidification also be taken into account. For example, columnar grains typically grow via solidification along a preferred (rapid-growth) direction and thus possess noticeable fiberlike textures.

The wedge test is different from other workability tests (e.g., tension test, uniaxial compression test) designed for small specimens of material with relatively fine wrought grain sizes that are subjected to nominally uniform deformation throughout the test section. In contrast, wedge specimens forged between flat, parallel dies suffer strains that vary with position. The variation of strain can be estimated to a first order from measurements of local thickness reduction and width increase. If higher accuracy is desired, finite-element-method (FEM) simulations can be used to predict local strains and strain rates and thus to correlate process variables to failure or microstructure evolution (Ref 9, 10). Typical results for the effective strain developed in wedge forging are shown in Fig. 3.

Machined wedge specimens are forged (or, equivalently, flat rolled) in the equipment in which the actual forging is to take place. This permits evaluation of the effects of working speed, and hence of die chilling due to heat transfer, on workability. Wedge preheat temperatures are selected on the basis of previous ex-

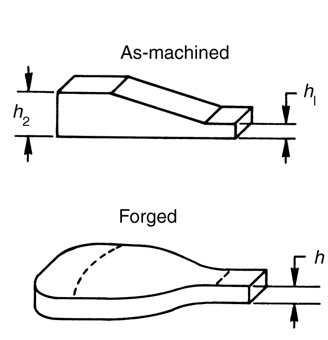


Fig. 2 Specimens for the wedge test. (Top) As-machined specimen. (Bottom) Specimen after forging

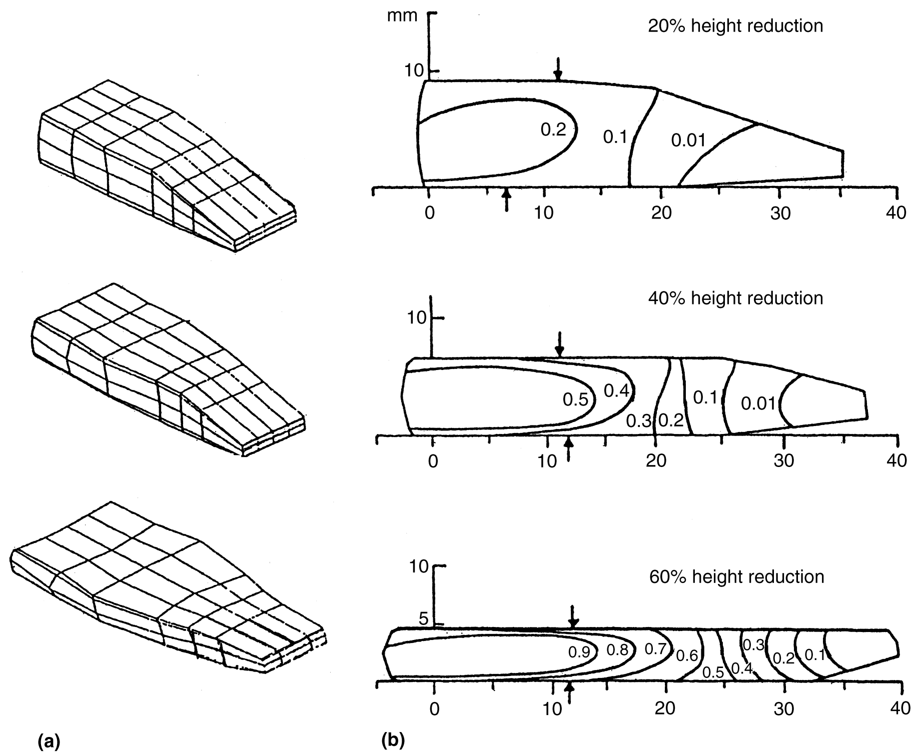


Fig. 3 Finite-element-model predictions for the wedge-forging test, assuming a friction shear factor of 0.4. (a) Grid distortions. (b) Strain contours along the symmetry plane. Source: Ref 9

perience with the alloy or, for new alloys, previous experience with similar alloys. If no such information is available, a temperature of approximately 0.8 to 0.9 times the measured solidus temperature is a good starting point. This choice can be modified based on examination of data for possible binary and multicomponent phases that may be formed from the alloy constituents and that are known to possess low melting points or that precipitate from solution at high temperatures. Even if no such phases are expected to be present, very high temperatures often are not advisable, because recrystallization may be accompanied by grain growth and a less workable microstructure.

Following deformation, wedge specimens are quenched or furnace annealed to determine dynamically and statically recrystallized grain sizes as functions of working temperature, amount of deformation, and annealing temperature and time. Figures 4 and 5 illustrate the kinds of results that can be obtained from a wedge test in which deformation is imposed by rolling rather than by forging. The material was an experimental high-nitrogen stainless steel (nominal composition, in weight percent: 24 Cr, 16 Ni, 5 Mn, 0.1 C, 0.65 N) produced by plasma arc remelting. The von Mises effective strain, $\bar{\varepsilon}$, (Ref 7, 12) at various locations was estimated solely from measurements of the final thickness of the wedge specimen in order to obtain the compressive thickness strain profile. Because the wedge was rolled, rather than forged between flat, parallel dies, the width strain was small; therefore, the longitudinal tensile strain was taken to be equal in magnitude to the thickness strain because of the constant-volume assumption.

Estimates of the percentage of recrystallized structure at the wedge specimen midwidth for the high-nitrogen stainless steel for specimens oil quenched immediately after deformation (Fig. 4) illustrate the important influence of preheat temperature and strain on dynamic recrystallization (Ref 11, 13). The data were taken from test specimens deformed at temperatures between 1090 and 1230 °C (1995 and 2245 °F), with each specimen providing a data point for each constant-$\bar{\varepsilon}$ curve. Also illustrated in Fig. 4 is the fact that cracking of wedge specimens occurred for $T \leq$ 1090 °C (1995 °F) and $T \geq$ 1200 °C (2190 °F). In both regimes, failure was intergranular. It was found from metallography and isothermal-transformation results that the alloy formed chromium nitride precipitates, the kinetics of formation of which were greatest at approximately 1050 °C (1920 °F) (Ref 14). Because the rolls were operated at a temperature much lower than the wedge preheat temperatures, a large amount of chilling,

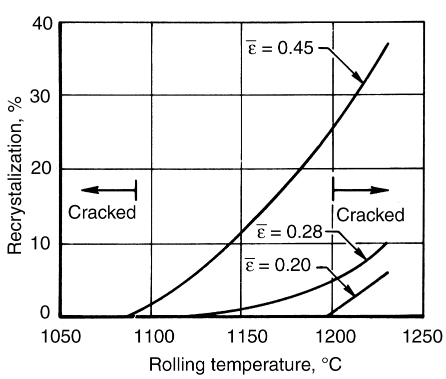


Fig. 4 Dependence of recrystallization on rolling temperature and deformation for high-nitrogen stainless steel wedge-test specimens cut from cast ingot. Source: Ref 11

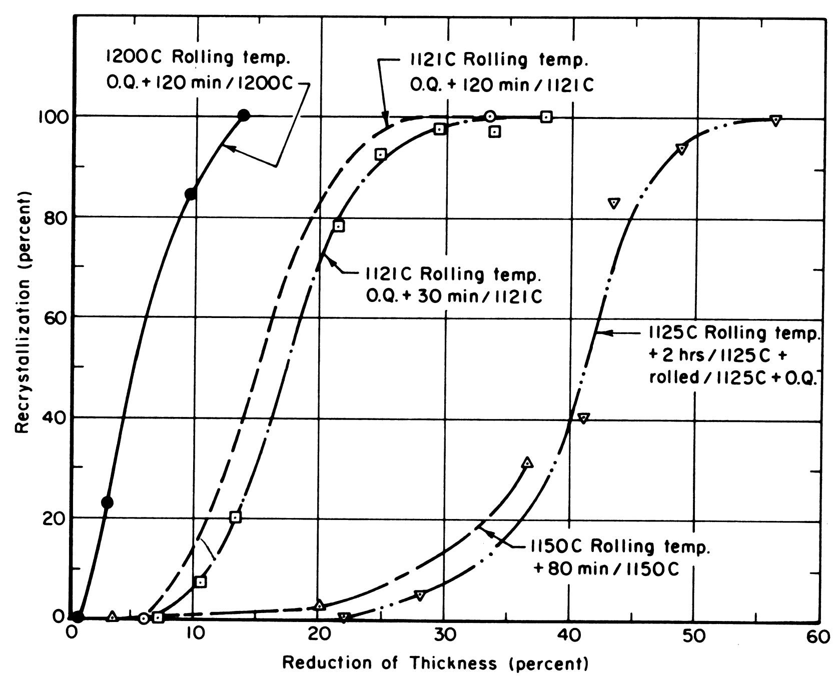


Fig. 5 Percentage of recrystallized structure in wedge-test specimens after various thermomechanical treatments performed on a high-nitrogen stainless steel. OQ, oil quench. Source: Ref 11

or cooling of the workpiece surface in contact with the rolls during deformation, must have occurred. Because of this effect, specimens preheated to 1090 °C (1995 °F) actually experienced much lower temperatures, at which precipitation of any dissolved chromium and nitrogen was likely. This precipitation would most likely occur on dislocations, impeding their motion and leading to intergranular failure. Precipitation in this alloy was also found to occur along the grain boundaries, preventing recrystallization and healing of intergranular microcracks.

Data in Fig. 5 for rolled and subsequently annealed wedge specimens of the high-nitrogen stainless steel demonstrate the usefulness of the wedge test not only as a workability technique but also as a means of devising detailed thermomechanical processing sequences for obtaining recrystallized microstructures in cast ingots. These static recrystallization trends are similar to those for annealing of cold-worked and hot-worked fine-grain metals (Ref 5, 15) and establish the ability of residual hot work to promote recrystallization of a cast structure.

A workability method that is very similar to the wedge test involves hot forging of so-called double-cone samples whose size is established based on the grain size of the cast microstructure (Fig. 6a). Finite-element modeling is often used to establish local deformations for such tests in order to quantify fracture and microstructure observations, much as in the wedge test. Typical results from such simulations are shown in Fig. 6(b).

Sidepressing Test. Because specimen size can be varied, the sidepressing test is another method well suited for establishing the workability of large-grain cast structures. It is particularly useful for estimating the interaction of incipient grain-boundary cracks, on the one hand, and secondary tensile and hydrostatic stresses, often present in open-die forging, on the other. In this test, round bars are laterally pressed between flat, parallel dies. As discussed in Ref 17 and 18, the maximum tensile stresses in this instance occur at the center of the bar at the beginning of the test (Fig. 7). As deformation proceeds, the bar assumes more of a rectangular cross section. As this occurs, the magnitudes of the tensile stresses decrease, and the location of the maximum stress is shifted away from the center of the bar. Similar changes in magnitude and location of the maximum tensile stress can be made to occur at early stages of deformation by changing from flat dies to dies that encompass more of the workpiece (Fig. 7); that is, secondary tensile stresses are minimized as less of an open-die geometry is employed and thus as more constraint is imposed. Moreover, friction and chilling may act in concert to promote bulging and thus to increase (or maintain) high levels of tensile stress during sidepressing above that level resulting from geometry alone.

The application of sidepressing as a workability test for cast materials can also be illustrated using breakdown of ingots of high-nitrogen stainless steel as an example (Fig. 8). When round ingots were forged round-to-round via sidepressing at 1150 °C (2100 °F), at which temperature dynamic recrystallization was found to occur, centerbursting was observed (Fig. 8a). This must be attributed to the inability to recrystallize weak interfaces early enough in the deformation so as to avoid defect initiation and growth. Such an occurrence could be avoided (Fig. 8b) by forging at 1125 °C (2060 °F), at which dynamic recrystallization was also favorable, and by forging/sidepressing in a sequence of round-to-square operations. (Round-to-square forging at 1150 °C (2100 °F) also eliminated centerbursts, but minor surface cracks were still evident, presumably because of chilling.) In the latter instance, the decrease in tensile stresses can be assumed to have been beneficial in eliminating centerbursting and is in line with the typical press-forging practice for ingot breakdown.

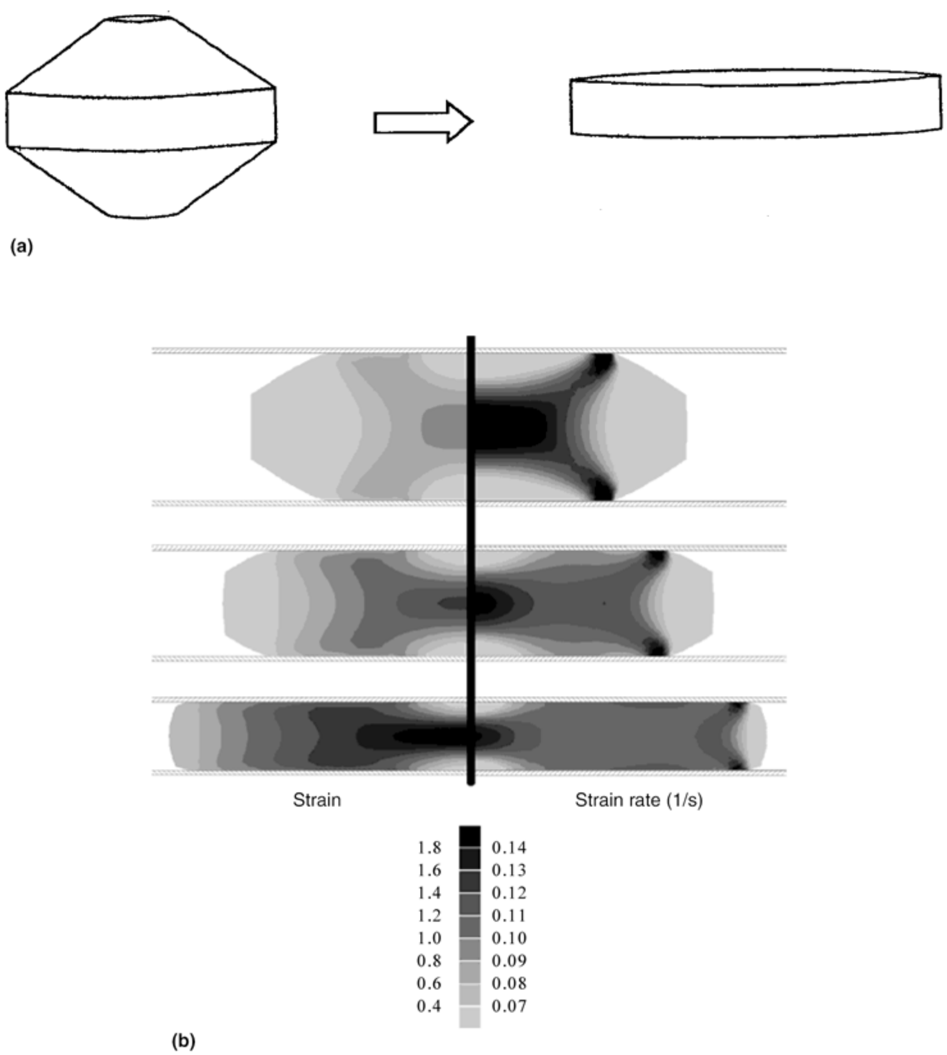


Fig. 6 The double-cone test. (a) Typical sample geometry. (b) FEM-predicted strain and strain-rate contours, assuming a friction shear factor of 0.5. Source: Ref 16

Workability Tests for Hot/Warm Open-Die Forging of Recrystallized Structures

Breakdown of coarse-grain cast structures to produce fine-grain wrought-and-recrystallized structures generally leads to improved workability in forging, as shown schematically in Fig. 1. Even though most wrought metals can be worked over wide ranges of temperature and deformation rate, care must be exercised in selection of forging temperature. Hot and warm forging require lower working loads than cold forging (performed at temperatures less than one-fourth of the melting or solidus temperature), but several types of defects can still arise in hot and warm open-die (and closed-die) forging. These defects include many of those previously discussed for cast structures (Table 1).

Examination of the common defects in hot and warm working of wrought-and-recrystallized structures (Table 1) reveals a convenient grouping that is helpful in devising different types of workability tests for open-die forging. This consists of two classes of workability problems: those related to fracture-controlled failure and those related to flow-localization-controlled failure. The first group comprises hot shortness, triple-point cracking, and grain-boundary cavitation (Fig. 9). The second group consists of shear banding (which can lead to fracture) that may be initiated by heat-transfer effects (due to chilling) as well as by flow softening (in the absence of chilling).

Workability Tests for Fracture-Controlled Defect Formation

Workability tests for fracture-related problems in forging must be able to determine the interactions among material aspects, process variables, and secondary tensile stresses. Material aspects include those mentioned before—such as crystal structure; phase changes; solidus, solutioning, and recrystallization temperatures; and mechanical and physical properties. Often, simple tests such as the hot compression, tension, and torsion tests are useful in the initial selection of forging temperature and strain rate and in gross estimation of forgeability. Usually, the estimate of forgeability is based on a parameter such as upset-test reduction at fracture, reduction of area in tension test, or shear strain to failure in torsion. A useful summary of such measurements has been made for different types of alloy systems (Fig. 10, 11) (Ref 2). Generally, pure metals and single-phase alloys exhibit the best ductility/forgeability except when grain growth occurs at high temperatures (Ref 21–23) (Fig. 10). On the other hand, alloys that contain low-melting-point phases (e.g., iron alloys with sulfur) or intermetallic second phases (e.g., gamma-prime-strengthened nickel-base superalloys) tend to be difficult to forge and have limited forging temperature ranges. Such behavior for alloys that contain strengthening phases is further illustrated in Fig. 12. This figure shows the relative forgeability of such an alloy system as a function of temperature and of percentage of alloy content that forms second phases. The trends shown are very representative of nickel-base superalloys, among which a lean alloy might be 718, and an alloy that is highly strengthened by precipitates might be alloy 700. As the amount of alloying increases, the probability of a low-melting-point phase being formed increases in much the same way as the liquidus temperature in a binary eutectic system decreases as the eutectic composition is approached. Further, the temperature for precipitation increases as the amount of second-phase elements, and hence the possibility for supersaturation, increases.

Workability Tests for Establishing Effects of Process Variables

Tests for establishing the effects of process variables on workability in forging include the high-speed isothermal tension test (or compression test) and the on-cooling tension test. The former test gives some estimate of the effects of strain rate and deformation heating on ductility. Done under nominally isothermal conditions,

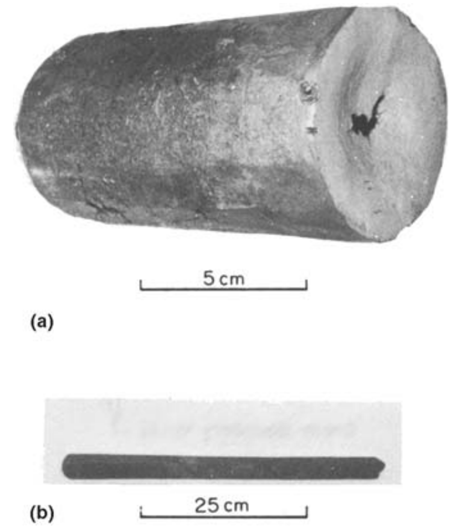


Fig. 8 Sidepressed bars of a high-nitrogen stainless steel cut from a cast ingot. (a) Bar forged round-to-round at 1150 °C (2100 °F), which developed center-burst. (b) Bar successfully forged round-to-square at 1125 °C (2060 °F). Source: Ref 11

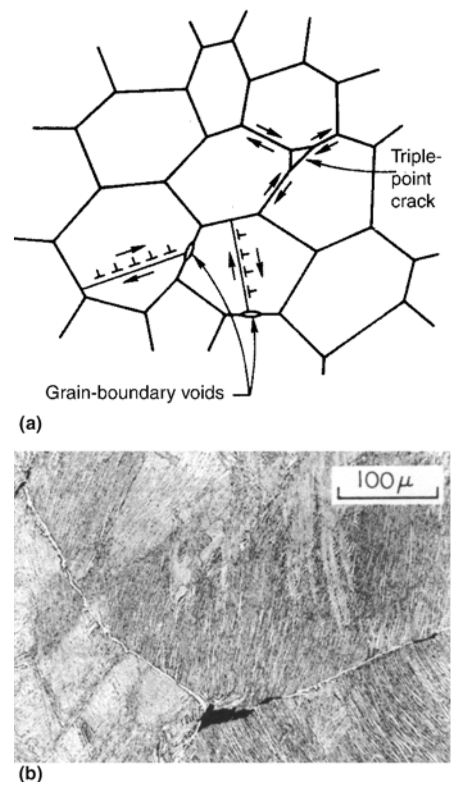


Fig. 9 Nucleation of grain-boundary voids (cavitation) and triple-point cracks at warm and hot working temperatures. (a) Schematic illustration showing how grain-boundary voids are formed under the action of matrix deformation and how grain-boundary sliding in the absence of grain-boundary migration and recrystallization may cause cracks to open at triple points. (b) Examples of grain-boundary voids and triple-point cracking at the prior beta grain boundaries in hot-forged Ti-6Al-2Sn-4Zr-2Mo-0.1Si with a colony-alpha starting microstructure. Source: Ref 19, 20

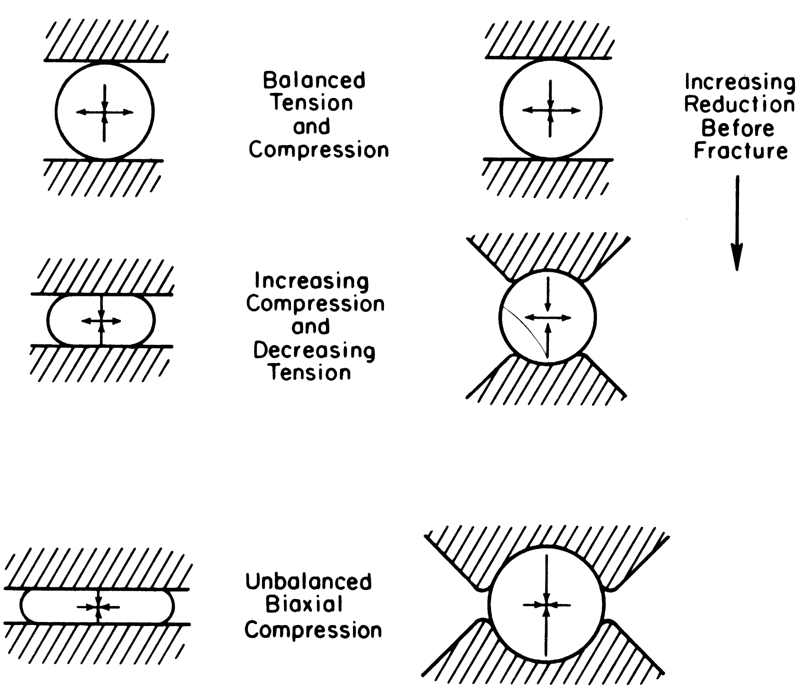


Fig. 7 Schematic illustration of the effects of billet shape and degree of enclosure on stress state in forging with good lubrication and no chilling. Source: Ref 17

the high-speed tension test (see Chapter 7, "Hot Tension Testing") may be used to gage when deformation-induced heating may lead to temperature increases sufficient to cause incipient melting (Ref 22). The temperature rise is easily estimated if the specimen gage length is great enough and the test fast enough to give rise to adiabatic conditions at the center of the gage section. Under these circumstances, the temperature rise is derivable from the measured flow curve and the material density and specific heat by assuming that a certain fraction (usually approximately 0.95) of the deformation work is transformed into heat. When the gage section of the tension specimen is short or the deformation slow, such that a substantial amount of heat transfer occurs between the gage section and the undeformed shoulders, temperature changes and workability observations must be interpreted with the aid of a numerical thermal simulation.

The on-cooling tension test (Ref 22) is useful in establishing the effects of die chilling in forging on workability. In such a technique, a series of nominally isothermal tension tests is first performed over a range of temperatures by simply heating to the test temperature and running the tests. Then, a second set of tests is run by heating to a variety of temperatures, holding at those temperatures for fixed or variable times, and then running the test after cooling to a lower temperature. It is obvious that such a test is useful for determining the effects of chilling as a function of working speed and thermal properties on the forgeability of alloys that form second phases on cooling or that undergo grain growth during high-temperature preheating. Further information on such testing is given in Chapter 7, "Hot Tension Testing," of this Handbook.

The standard tension test and on-cooling tension test may also be useful for evaluation of heat treatments prior to deformation which may improve forgeability. For example, a pretreatment to precipitate and overage (i.e., coarsen) second phases in various iron- and nickel-base alloys has been found to be very beneficial in subsequent working (Ref 24), an effect that can be evaluated most readily by means of simple workability tests.

Hot forging is often a multistep process, with intermediate dwells or reheats between each increment of deformation and changes in loading direction. Because of this, microstructure evolution and the propensity for fracture may be dependent on the precise thermomechanical history to which the workpiece is subjected. Hence, a number of studies have been undertaken to establish the effect of single- versus multi-blow deformation and strain-path changes on microstructure development (Ref 25–29). It is likely that similar studies will be undertaken in the future to establish the path-dependence of fracture as well.

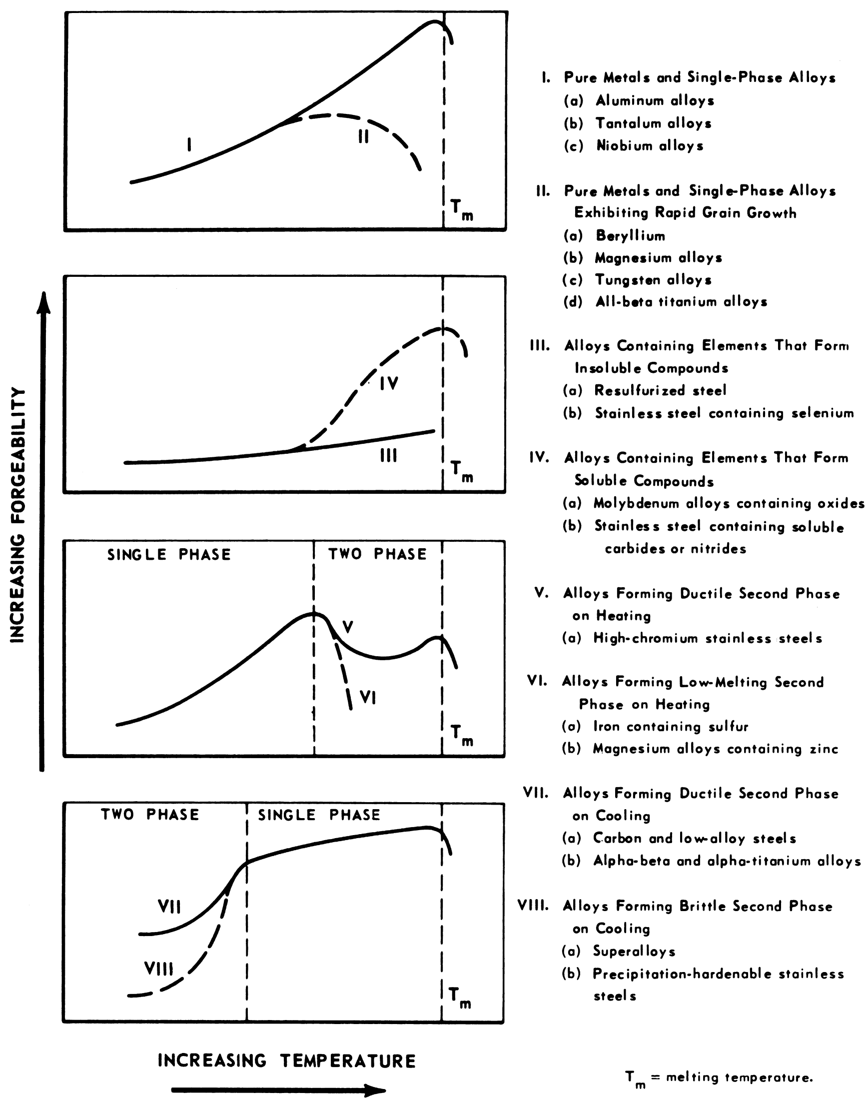


Fig. 10 Typical forgeability behaviors exhibited by different alloy systems. Source: Ref 2

Tests for Establishing Effects of Secondary Tensile Stresses on Forgeability

Sometimes the uniaxial tension and compression tests are insufficient for determining forgeability in actual open-die and closed-die forging operations. This may arise because of secondary and hydrostatic tensile stresses that can exacerbate a marginal workability problem. Several techniques have been developed specifically to establish forging guidelines in these instances. These include the notched-bar or U-notch upset test, the double-cone test, and pancake forging itself.

The notched-bar upset test (Ref 2, 8, 30), developed by the Ladish Company, is similar to the conventional upset test, except that axial notches are machined into the test specimens prior to compression. The notch magnifies the effects of secondary tensile stresses that may arise during a conventional upset test due to bulging caused by chilling and/or poor lubrication. The higher levels of tensile stress in the test are supposedly more typical of those in actual forging operations. In preparation of specimens for the test, a forging billet is quartered longitudinally, exposing center material along one corner of each test specimen. After sectioning, notches with either 1.0 or 0.25 mm (0.04 or 0.01 in.) radii are machined into the faces (Fig. 13); a weld button is often placed on one corner to identify center and surface material in metals that exhibit forgeability problems because of segregation. Assuming a uniform bulge, the faces with the deeper (1.0 mm, or 0.04 in.) notches are subjected to the highest tensile stresses and thus are more likely to indicate a workability problem.

The notched-bar specimens are heated to one or a number of possible forging temperatures and upset forged to a reduction in height of approximately 75%. Because of the stress-concentration effect, fractures in the form of

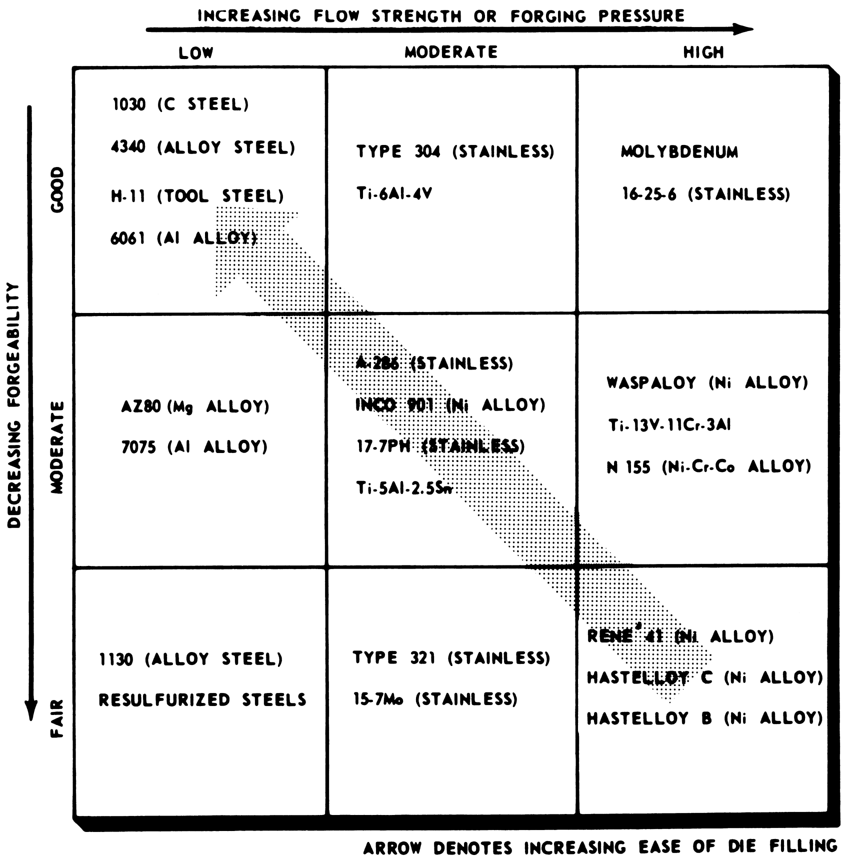


Fig. 11 Chart illustrating interaction of workability, flow strength, and die-filling capacity in forging. Large shaded arrow indicates increasing ease of die filling. Source: Ref 2

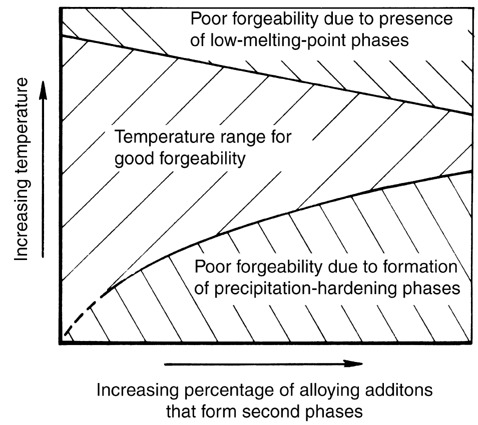


Fig. 12 Influence of solute content that forms second phases on melting and solution temperatures and thus on forgeability

ruptures are most likely to occur in the notched areas and may be assumed to have been initiated by hot shortness, triple-point cracking, grain-boundary cavitation, and so on. These ruptures may be classified according to the rating system shown in Fig. 14. A rating of 0 is applied if no ruptures are observed. If they are small, discontinuous, and scattered, the rating is 1. Higher rating numbers indicate increasing incidence and depth of rupture.

Examples of the application of the notched-bar test are those reported in Ref 2 and 8. In the former example, rolled rings of type 403 stainless steel with a notched-bar rating of 0 were sound, whereas those with a rating of 4 ruptured extensively. In a similar vein, the rupture sensi-

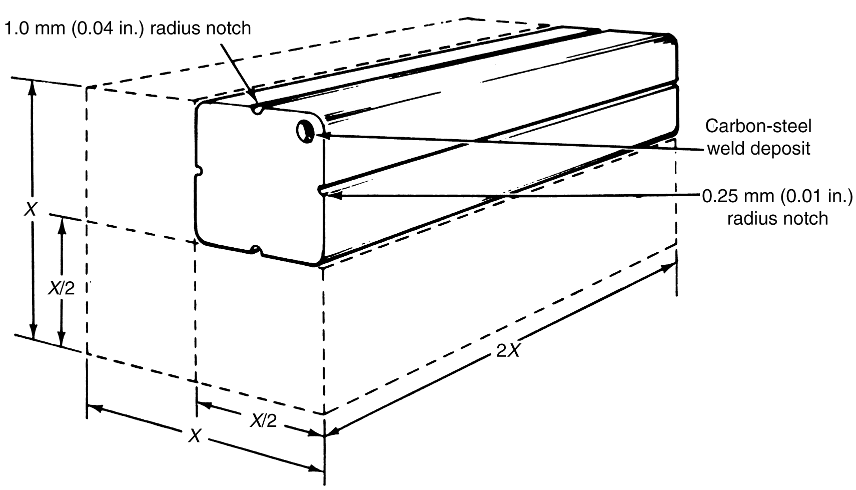


Fig. 13 Method of preparing specimens for notched-bar upset forgeability testing. Source: Ref 2, 30

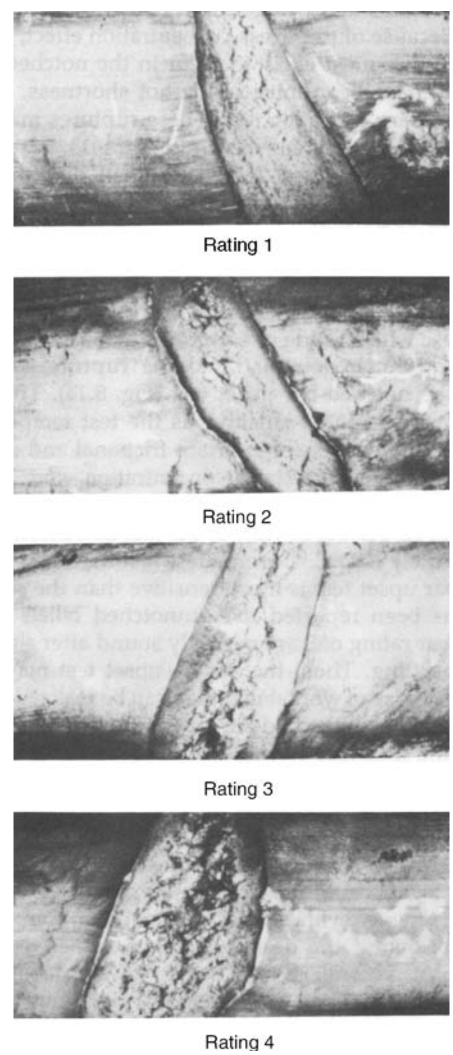


Fig. 14 Suggested rating system for notched-bar upset test specimens that exhibit progressively poorer forgeability. A rating of 0 would indicate freedom from ruptures in the notched area. Source: Ref 2, 30

tivity of Inconel 718 was demonstrated with the notched-bar upset test (Fig. 15). The alloy was shown to be much more notch-sensitive as the test temperature was increased. The combined effects of interface frictional and deformation heating in conjunction with the stress concentration were able to produce very large ruptures when the alloy was forged at and above 1150 °C (2100 °F). Such an effect is not surprising in view of the fact that Inconel 718 starts to melt at approximately 1200 °C (2190 °F) (Ref 31). From these results, one may conclude that the notched-bar upset test is more sensitive than the simple upset test. In fact, it has been reported that unnotched billets from heats having a notched-bar rating of 3 are perfectly sound after similar reductions in simple upsetting. Thus, the simple upset test may indicate a deceptively higher degree of workability than can be realized in an actual forging operation, and the notched-bar upset test may be particularly useful for identifying materials having marginal forgeability.

The double-cone test is another method employed to study the effects of secondary tensile stresses on defect formation. Similar to the notched-bar approach, this test involves a simple upsetting operation between flat, parallel dies. However, the tensile stresses are developed by a much different specimen geometry (Fig. 6). As the sample is upset, secondary tensile stresses are generated in the outside portion of the specimen. The generation of this stress state is analogous to that in sidepressing of round bars. Because of secondary tensile stresses, voids initiated by grain-boundary cavities and/or triple-point cracks and so on may be opened.

The comparison of FEM solutions for the stress field in the double-cone test with actual observations has confirmed the usefulness of this technique as a workability test. For example, cavitation observations in double-cone samples of Ti-6Al-2Sn-4Zr-2Mo-0.lSi with a colony-alpha microstructure forged isothermally (dies and workpiece at the same initial temperature) at 940 °C (1725 °F) and a strain rate of approximately 2.5 s^{-1} were found to correlate well with the (peripheral) regions in which circumferential tensile stresses were predicted to be developed (Ref 32, 33). Furthermore, the analysis suggested that the defects originated at low strains, at which the tensile stresses were relatively large, and that the tensile stresses diminished as the amount of compression increased, giving rise to less taper on the forging. In effect, the simulation suggested that the defects move toward the outside of the disk as the reduction increased, leaving behind a sound structure, a trend verified by comparison with experimental observations. Therefore, the test may be used to determine when voids may be formed and when they may be healed by further deformation.

The double-cone test is also finding increased use as a method to establish microstructure evolution during hot working of wrought alloys. For example, it has been used to determine the effect of process parameters on recrystallization and abnormal grain growth during the thermomechanical processing of nickel-base superalloys and microstructure evolution during hot working of beta-titanium alloys (Ref 16, 34).

Pancake forging comprises the upsetting of a round cylinder to large strains, thus producing a "pancake" whose diameter is much greater than its thickness. Because of the large reduction, friction, and chilling (during conventional hot forging), substantial barreling of the free surface occurs, giving rise to substantial tensile stresses. These stresses may cause free-surface fracture as well as subsurface cavitation without gross fracture. The analysis of free-surface fracture is described in Chapter 4, "Bulk Workability Testing." Because cavities can detract from subsequent load-carrying capability and fatigue resistance during service, it is important to quantify the occurrence of such defects, particularly when their severity is just below the level that causes surface fracture.

A method to quantify and interpret cavitation based on pancake forging is described in Ref 35. In this work, Ti-6Al-4V samples with a colony-alpha microstructure were hot forged to various reductions. Following deformation, the depth below the surface at which cavities were found (at 500×) was determined in each sample. The strain and stress state that corresponded to the regions of cavitation were determined using a FEM technique (Fig. 16). In a similar manner, cavitation was quantified in hot tension samples deformed at comparable strain rates and temperatures. A comparison of the results from the tension and forging experiments demonstrated that cavity initiation (as well as gross surface fracture) in the two different deformation modes could be correlated based on local values of the Cockroft and Latham (Ref 36) maximum tensile work, $C = \int(\sigma_t/\bar{\sigma})d\bar{\varepsilon}$, in which σ_t and $\bar{\sigma}$ denote the maximum tensile stress and the effective stress, respectively, and $\bar{\varepsilon}$ is the effective strain. (In uniaxial tension involving minimal necking, $\sigma_t \approx \bar{\sigma}$, the critical values of C for cavity initiation (C_i^*) and fracture (C_f^*) are simply the corresponding axial strains.) An example of this correlation is shown in Fig. 17. Hence, a critical damage parameter can be determined from pancake forging trials and accompanying FEM analysis. This critical damage parameter can then be used to predict the possible occurrence of cavitation in other forgings for which FEM simulations are also conducted to determine local values of the maximum tensile work.

Mechanistic Modeling of Cavitation. Unlike the well-developed phenomenological approaches for ductile fracture, such as that based on the Cockcroft and Latham criterion (Ref 36), the development of physics-based (mechanistic) models for the prediction of cavitation during hot forging is still in a relatively early stage of development. To a certain extent, this is a result of the complexity of the nucleation-and-growth processes themselves as well as the influence of complex stress states that characterize forging. For example, the classical surface-energy criterion for cavity nucleation at grain-boundary particles (Eq 17 in Chapter 7, "Hot Tension Testing") has a number of drawbacks. To replicate experimental observations of continuous nucleation, the criterion requires flow hardening in order to nucleate cavities at less favorable sites compared to those at which nucleation occurs initially. In addition, the

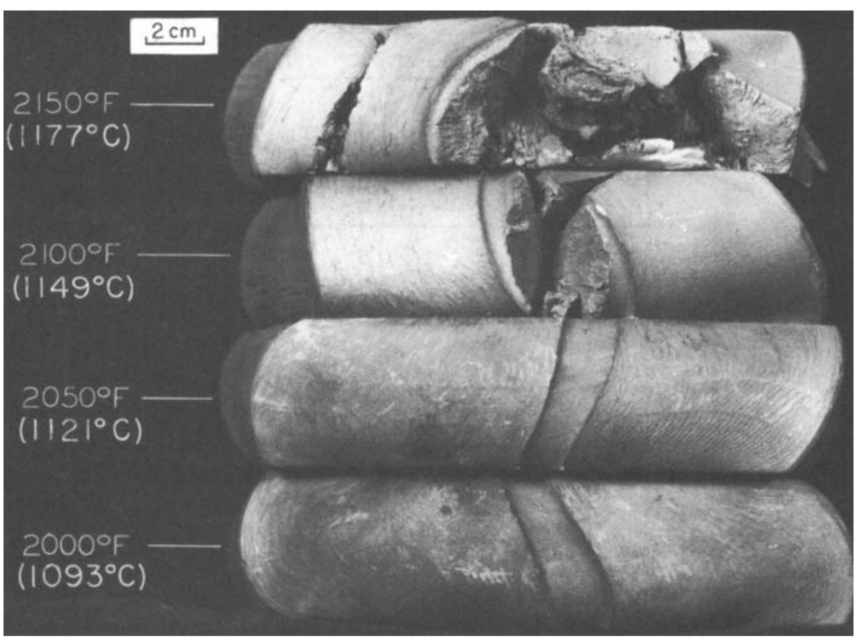


Fig. 15 Nickel alloy 718 notched-bar upset-test specimens showing rupture sensitivity associated with forging temperature. Source: Ref 8. Courtesy of Ladish Company

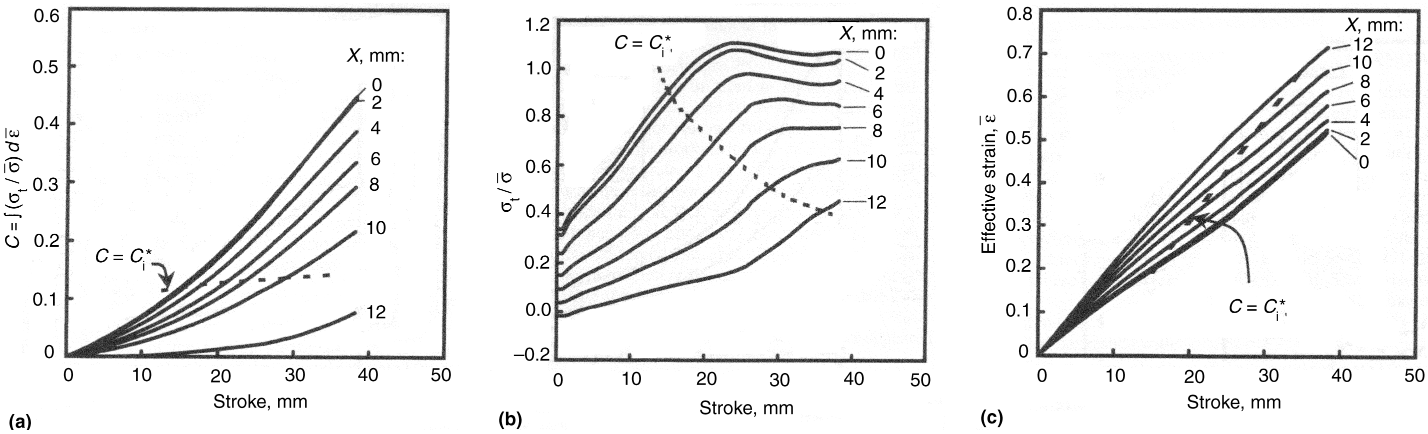


Fig. 16 Finite-element-modeling calculations of (a) Cockcroft and Latham tensile work ($\int(\sigma_t/\bar{\sigma})d\bar{\varepsilon}$), (b) $\sigma_t/\bar{\sigma}$, and (c) effective strain, $\bar{\varepsilon}$, as a function of stroke for a Ti-6Al-4V sample with a colony-alpha microstructure forged at 815 °C (1500 °F), 0.1 s^{-1} to a final height reduction of 50%. The results are plotted for material tracking points located at various distances from the free surface on the equatorial plane. The stroke at which cavitation was predicted to initiate (based on uniaxial tension measurements) is indicated by the intersections with the $C = C_i^*$ curves in the various figures. Source: Ref 35

stresses required for initiation and early growth are unrealistically high. Therefore, the formulation of other (constrained-plasticity) approaches based on cavity nucleation and early growth from preexisting inhomogeneities/nanovoids (at which locally higher stress triaxiality is developed) has been undertaken (Ref 37). In such approaches, the degree of constraint decays as the cavity grows from the nanometer to the micrometer scale, at which classical general plasticity-controlled analyses, such as are discussed in Chapter 7, "Hot Tension Testing," apply.

For cavities whose size is of the order of several micrometers or larger, a simple method was developed (Ref 38) for assessing the effect of stress state on the kinetics of cavity growth. The approach was based on the plasticity-controlled growth of an isolated spherical cavity. Under purely uniaxial tension conditions (i.e., $\sigma_m/\bar{\sigma} = ⅓$, in which σ_m is the hydrostatic stress, and $\bar{\sigma}$ denotes the effective stress), such growth is described by the following relation:

$$R = R_o \exp\left(\frac{\eta_o}{3}\varepsilon\right) \qquad \text{(Eq 1)}$$

In Eq 1, R and R_o are the instantaneous and initial cavity radii, respectively, ε is the axial strain, and η_o is the volumetric cavity growth rate parameter under uniaxial stress conditions ($\eta_o = d\ln V/d\varepsilon$). The effect of stress triaxiality on cavity growth as described in the preceding expression was incorporated by using the Rice and Tracey analysis (Ref 39). This prior analysis led to the determination of a measure of the cavity growth rate known as the void growth factor, D:

$$D = \frac{\dot{R}}{\dot{\varepsilon}R} = \frac{d\ln R}{d\varepsilon} \qquad \text{(Eq 2)}$$

or

$$R = R_o \exp(D\,\varepsilon) \qquad \text{(Eq 3)}$$

in which ε and $\dot{\varepsilon}$ denote the far-field strain and strain rate, respectively. The dependence of D on stress was shown to be as follows:

$$D = 0.558 \sinh(3\sigma_m/2\bar{\sigma}) + 0.008\,\nu \cosh(3\sigma_m/2\bar{\sigma}) \qquad \text{(Eq 4)}$$

in which ν denotes a function of the strain rates defined as $\nu = -3\dot{\varepsilon}_2/(\dot{\varepsilon} - \dot{\varepsilon}_3)$, with $\dot{\varepsilon}_1 \geq \dot{\varepsilon}_2 \geq \dot{\varepsilon}_3$ representing the principal strain rates. The dependence of D on the ratio of the mean to the effective stress, $\sigma_m/\bar{\sigma}$, is shown in Fig. 18(a). For example, for uniaxial stress conditions, $\sigma_m/\bar{\sigma} = ⅓$, $\nu = +1.0$, and $D = D_U = 0.30$. Equation 3 thus incorporates the effect of the local stress state (via D) on the growth of an individual cavity.

It was postulated (Ref 38) that the increase of the cavity growth rate parameter, η_o, (Eq 1) due to stress triaxiality is proportional to D/D_U, that is:

$$\eta = \eta_o \frac{D}{D_U} \quad \text{or} \quad \eta = \eta_o F(\sigma) \qquad \text{(Eq 5)}$$

in which η denotes the cavity growth rate parameter under the complex stress state, and $F(\sigma) = D/D_U$. $F(\sigma)$, and hence the average cavity growth rate, increases as $\sigma_m/\bar{\sigma}$ increases; for example, at the same strain level, the cavity size and volume fraction would be higher in the case of biaxial tension compared to uniaxial tension. The suitability of this approach was validated to a first order by a comparison to the cavitation observed in Ti-6Al-4V tension samples with sharp necks and thus high levels of hydrostatic stress (Ref 40) (Fig. 18b). In contrast, by reducing the level of hydrostatic stress, cavity growth can be suppressed. An example of such an effect is the elimination of cavitation in a gamma-titanium-aluminide alloy by using superimposed pressure during hot forging (Ref 41).

As described in Chapter 7, "Hot Tension Testing," expressions are readily determined to relate η to the overall evolution of cavitation when continuous nucleation and coalescence occur in addition to the growth of single cavities.

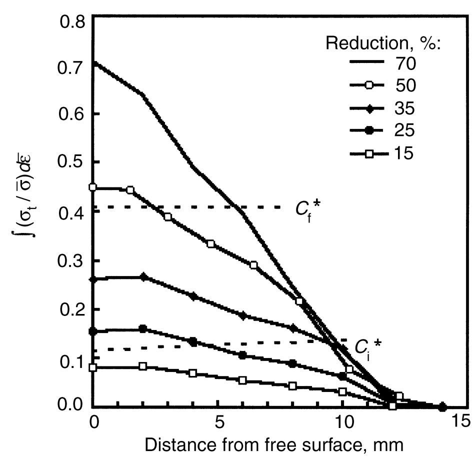
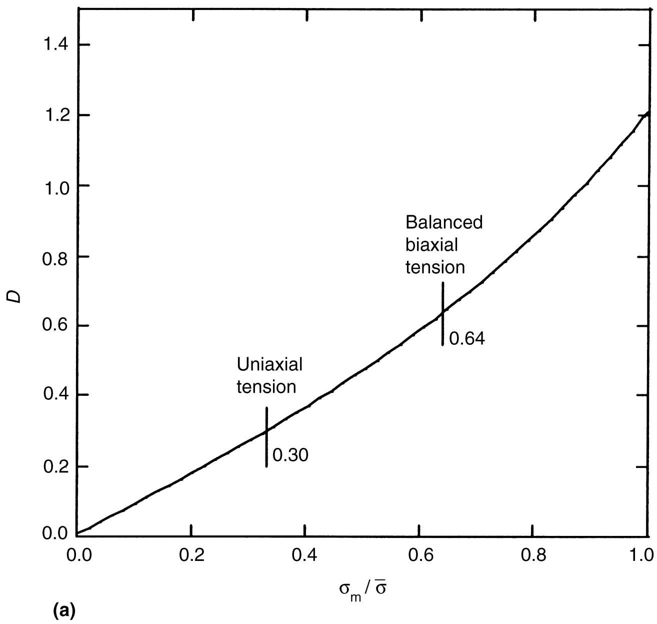


Fig. 17 Comparison of FEM calculations of the Cockcroft and Latham tensile work ($\int(\sigma_t/\bar{\sigma})d\bar{\varepsilon}$) at the end of the stroke to the critical values established from tension tests for cavity initiation (C_i^*) and surface fracture (C_f^*) in pancake forgings of Ti-6Al-4V with a colony-alpha microstructure conducted to various height reductions at 815 °C (1500 °F), 0.1 s^{-1}. The predicted depths of cavitation and the critical reduction for surface fracture were well predicted by such comparisons. Source: Ref 35

Workability Tests for Flow-Localization-Controlled Failure

Nonisothermal Upset Test. Workability problems in hot and warm open-die forging that are related to flow localization are very common, and various terms, such as *chill zone, dead-metal zone,* and *locked metal,* have become synonymous with defects that result from this phenomenon. In open-die forging, where lateral constraint to metal flow is minimal, the occurrence of flow localization is a result of heat



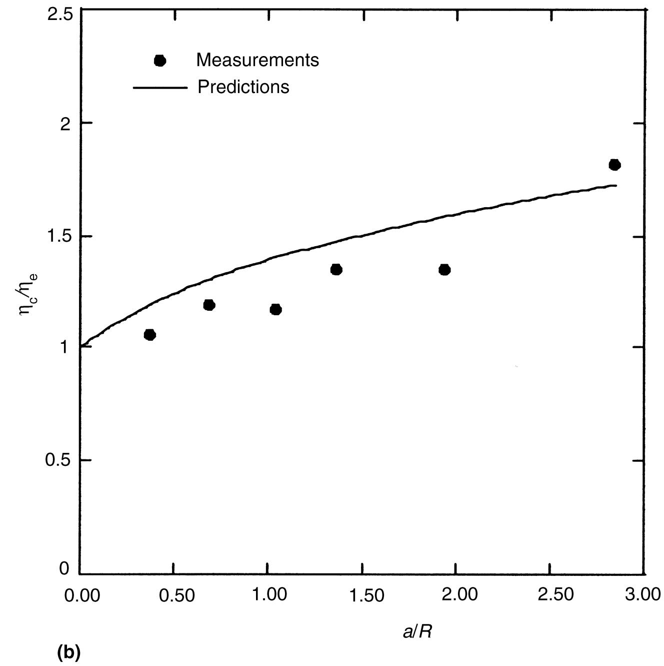


Fig. 18 Effect of stress state on cavitation. (a) Dependence of the void growth factor, *D*, on the ratio of the mean to effective stress ($\sigma_m/\bar{\sigma}$). Source: Ref 38, 39. (b) Comparison of measured and predicted values of the ratio of the cavity growth parameter at the center and edge (η_c/η_e) of Ti-6Al-4V hot tension specimens as a function of the local value of the ratio of sample radius to profile radius (*a*/*R*). Source: Ref 40

transfer between the hot workpiece and cold dies or of excessive friction. As such, it is relatively easy to devise workability tests for investigating and resolving these problems when they occur.

Probably the simplest workability test that can be employed to study the effects of heat transfer on flow localization is the nonisothermal upset test. As in the uniform, isothermal hot compression test, a cylindrical workpiece is upset between flat, parallel dies. Workpiece and die temperatures, as well as working speed, lubrication, and dwell time on the dies prior to forging, can be varied. Following deformation, specimens are sectioned and metallographically prepared in order to determine the extent of chilling and formation of shear bands (regions of intense localized deformation) between the chill zones and the remainder of the deforming bulk. As shown in Fig. 19 for several Ti-6Al-2Sn-4Zr-2Mo-0.1Si (Ti-6242Si) specimens that had a starting equiaxed-alpha microstructure, the chill zones are usually revealed as areas that etch differently than the rest of the specimen.

By varying the process variables in a series of nonisothermal upset tests, the extent of chill zones and the severity of shear banding can be estimated. These results may be interpreted using a variety of mathematical techniques, such as the upper-bound technique (Ref 4) and a simplified analytical treatment based on a one-dimensional heat-transfer analysis and the stress-equilibrium equation (Ref 42). Although very approximate in nature, the latter technique is capable of quantifying the three most important material/processing components that determine the magnitude and extent of strain-rate (and thus strain) gradients: the material flow-stress dependence on temperature, the material strain-rate sensitivity, and the magnitude of the temperature gradient. The first two are properties that are measured in isothermal hot compression tests, whereas the third can be predicted from heat-transfer analysis. For Ti-6242Si, the dependence of flow stress on temperature at typical forging temperatures is very strong; also, the rate sensitivity is modest (Ref 42–44). The heat-transfer analysis incorporates the densities and thermal properties (specific heat, thermal conductivity) of the dies and the workpiece, the initial die and workpiece temperatures, the deformation rate, and the heat-transfer coefficient characterizing the interface. Simplifying assumptions can be made regarding heat generation and heat transport that make a closed-form solution possible. With these data and assumptions, the extent of chill zones in simple tests such as those done on Ti-6242Si (Fig. 19) can be predicted with fairly good accuracy (Ref 42).

The nonisothermal sidepressing test is another workability technique that can be used to gage the interactions of material properties and process variables during flow localization in hot forging. As in the nonisothermal upset test, several test specimens are sidepressed between flat, parallel dies at a variety of workpiece temperatures, die temperatures, working speeds, and so on, and the formation of defects is determined by metallography. Unlike the upset test, with its axisymmetric chill zones, however, flow localization is manifested primarily by shear-band formation and propagation. The absence of well-defined chill zones in the nonisothermal sidepressing test may be attributed to two main factors. First, the amount of contact area in sidepressing of round bars starts at zero (i.e., line contact) and builds up rather slowly at the beginning of deformation. Second, because the operation is basically a plane-strain operation, surfaces of zero extension, along which block shearing can initiate and propagate, are present in sidepressing. These surfaces are the natural ones along which shear strains can concentrate into shear bands. Formation of well-defined chill zones prior to shear-band formation would prob-

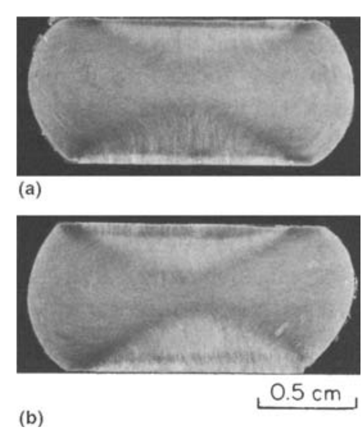


Fig. 19 Axial cross sections of specimens of Ti-6Al-2Sn-4Zr-2Mo-0.1Si with an equiaxed-alpha starting microstructure that were nonisothermally upset at 954 °C (1750 °F) to 50% reduction in a mechanical press ($\dot{\bar{\varepsilon}} \approx 30\ s^{-1}$) between dies at 191 °C (375 °F). Dwell times on the dies prior to deformation were (a) 0 s and (b) 5 s. Source: Ref 42

ably involve more deformation work and hinder the propagation of shear bands. This is not to say, however, that temperature gradients are not important in supplying a driving force for flow localization.

Sidepressing results for the Ti-6242Si alloy illustrate the kinds of localization behavior that can be observed (Ref 42). In Fig. 20, cross sections of bars sidepressed in a mechanical press (strain rate $\approx$ 30 s^{-1}), using dies preheated to 191 °C (375 °F), are shown. When the bars were preheated to a temperature much below the transus temperature (913 °C, or 1675 °F), at which the flow stress is very sensitive to temperature, shear bands whose intensity increases with reduction were developed (Fig. 20a–c). It has been found that, in sidepressing, the propagation stages of shear banding appear to follow an identical pattern, regardless of die temperature, working speed, and so on, which points out the importance of geometry in determining the shapes of the shear bands. These so-called propagation stages are summarized somewhat schematically in Fig. 21, which is discussed more thoroughly in Ref 42. In contrast to the previously mentioned results, no shear banding was observed (Fig. 20d–f) when the workpiece temperature is approximately 982 °C (1800 °F), which is close to the beta transus temperature for this alloy. At this temperature, the dependence of flow stress on temperature is minimal. Hence, the high strain rates employed in this instance, in conjunction with the low-temperature sensitivity of the flow stress, may be surmised to have minimized stress gradients, and hence strain gradients, during sidepressing.

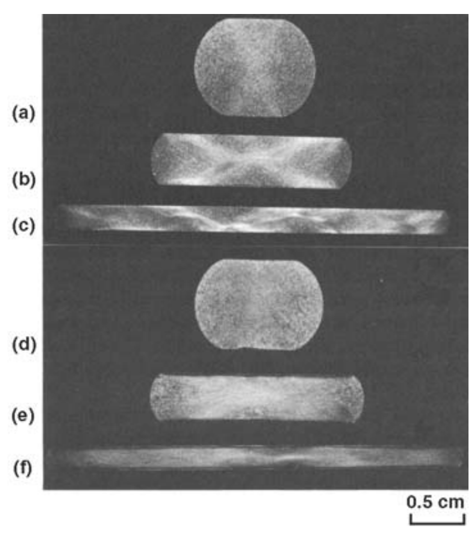


Fig. 20 Transverse metallographic sections of specimens of Ti-6Al-2Sn-4Zr-2Mo-0.1Si with an equiaxed-alpha starting microstructure that were nonisothermally sidepressed with zero dwell time in a mechanical press ($\dot{\varepsilon} \approx 30\ s^{-1}$) between dies heated to 191 °C (375 °F). Specimen preheat temperatures, T_s, and percent reductions, *R* (relative to the initial specimen diameter), were (a) T_s = 913 °C (1675 °F); *R* = 14, (b) T_s = 913 °C (1675 °F); *R* = 54, (c) T_s = 913 °C (1675 °F); *R* = 77, (d) T_s = 982 °C (1800 °F); *R* = 21, (e) T_s = 982 °C (1800 °F); *R* = 57, (f) T_s = 982 °C (1800 °F); *R* = 79. Source: Ref 42

The effect of lower working speed, and therefore greater heat transfer, on shear-band formation in sidepressing was also determined in sidepressing tests on Ti-6242Si and serves to illustrate the extremes in behavior that can be investigated by means of this test. These experiments were conducted in a hydraulic press (strain rate $\approx$1 s^{-1}); for a given reduction, deformation thus lasted approximately 30 times as long as in the mechanical-press trials. At low reductions, shear bands similar in appearance to those observed previously were formed in the hydraulic-press specimens (Fig. 22). On the other hand, at high reductions, shear cracks were found to develop along these shear bands. The same effect was found regardless of whether the workpiece preheat temperature was much below the transus temperature or rather close to it (Ref 42), a trend that, as has been mentioned, was significant for the mechanical-press tests. From these results, it may be concluded that long deformation times may drop the average workpiece temperature from one at which the flow stress is not too temperature-sensitive to one at which it is and at which thermal gradients may lead to shear bands, and long deformation times many drop the average workpiece temperature into a regime of low ductility. The latter conclusion explains why cracking was observed in hydraulic-press trials, inasmuch as process-simulation results revealed that the average workpiece temperature had indeed fallen to a level at which ductility is low.

Isothermal Compression and Sidepressing Tests. It has also been found that flow localization may occur during hot forging in the absence of chilling or frictional effects (Ref 43, 45, 46). In these cases, localization is a result of flow softening, or negative strain hardening. Such an effect may occur during hot working of metals because of dynamic recrystallization, an unstable microstructure, changes in texture, redistribution of solid-solution elements, changes in the morphology of second phases, and so on (Ref 45).

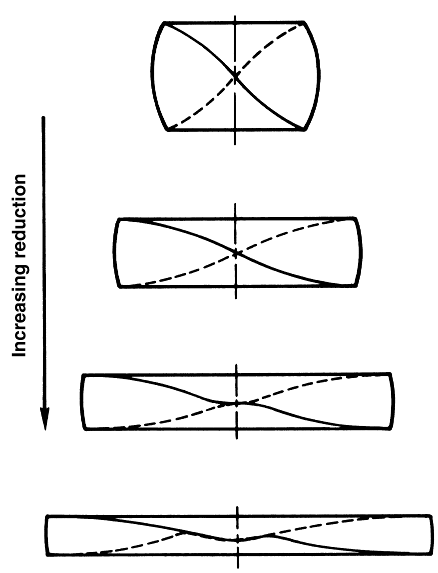


Fig. 21 Schematic representation of the mechanism of shear-band formation in sidepressing. Source: Ref 42

The constant-strain-rate, isothermal hot compression test (see Chapter 6, "Hot Compression Testing") is the most useful workability test that can be used to detect or predict flow localization during forging in which the dies are preheated to the same temperature as the workpiece, that is, during isothermal forging. Flow-stress data and observations are useful in establishing guidelines for flow localization for various alloy systems and various deformation states. To this end, the flow-stress data are analyzed to determine the normalized flow-softening rate, $\gamma' = (1/\sigma)\, d\sigma/d\varepsilon$, where $\sigma(\varepsilon)$ is the flow curve determined under constant-strain-rate conditions, and the strain-rate-sensitivity exponent is $m = \partial \log\sigma/\partial \log\dot{\varepsilon}$ at fixed strain and temperature, where $\dot{\varepsilon}$ denotes strain rate. This determination is made as a function of strain for flow curves measured at a variety of strain rates and temperatures. It has been found that for materials that exhibit $\alpha_c \equiv (\gamma' - 1)/m \geq 5$ at a given strain, strain rate, and temperature, nonuniform flow in uniaxial compression is likely (Ref 47, 48).

The previously mentioned conclusion is supported by data for a number of alloys, such as Ti-6242Si, Ti-l0V-2Fe-3Al, and nickel alloy 700. Observations for Ti-6242Si deformed at a temperature of 913 °C (1675 °F) and a strain rate of 2 s^{-1} are shown in Fig. 23 (Ref 47). When the

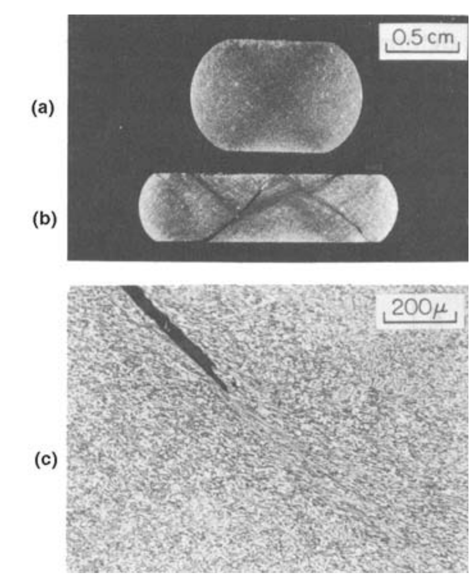


Fig. 22 (a, b) Transverse metallographic sections and (c) micrograph of region with shear band and crack from section shown in (b) of Ti-6Al-2Sn-4Zr-2Mo-0.1Si specimens with an equiaxed-alpha starting microstructure that were nonisothermally sidepressed in a hydraulic press ($\dot{\varepsilon} \approx 30\ s^{-1}$). Specimen preheat temperature, die temperature, and dwell time were 913 °C (1675 °F), 191 °C (375 °F), and 14 s, respectively. Reductions were (a) 25% and (b) 53%. Source: Ref 42

alloy had an equiaxed-alpha microstructure, α_c values less than zero were estimated from the flow-stress data, and upset testing gave rise to uniform deformation at this temperature and strain rate (Fig. 23a). In contrast, when the alloy had a nonequilibrium colony-alpha microstructure, the flow curves exhibited large amounts of flow softening. At 913 °C (1675 °F) and $\dot{\bar{\varepsilon}} = 2$ s^{-1}, α_c values of approximately 5 were estimated using the flow curves; the very nonuniform flow observed in compression (Fig. 23b) was ascribed to the high levels of α_c. It is also interesting to note that the nonuniform flow observed in isothermal compression is analogous to the formation of chill zones in nonisothermal upsetting, inasmuch as it is evidenced on a bulk scale rather than as an internal defect similar to a shear band.

The occurrence of nonuniform deformation during hot compression testing of beta-annealed Ti-l0V-2Fe-3Al further demonstrates the applicability of the α_c parameter in predicting nonuniform flow. When this alloy was deformed at 704 °C (1300 °F), the maximum values of α_c over the strain interval from 0 to 0.7 were approximately 2.5 ($\dot{\bar{\varepsilon}} = 10^{-3}$ s^{-1}), 5 ($\dot{\bar{\varepsilon}} = 10^{-1}$ s^{-1}), and 5 ($\dot{\bar{\varepsilon}} = 10$ s^{-1}). At 816 °C (1500 °F), α_c was less than zero at strain rates of 10^{-3}, 10^{-1}, and 10 s^{-1}. Observation of nonuniform flow and flow localization at 704 °C (1300 °F) and $\dot{\bar{\varepsilon}} = 10^{-1}$ and 10 s^{-1} and the absence of such localization at 704 °C (1300 °F) and $\dot{\bar{\varepsilon}} = 10^{-3}$ s^{-1} and at 816 °C (1500 °F) and $\dot{\bar{\varepsilon}} = 10^{-3}, 10^{-1}$, and 10 s^{-1} (Fig. 24) are in line with the previous results for Ti-6242Si.

A related parameter, $\alpha_p \equiv \gamma'/m$, where γ' and m have the same definitions as before, may be used to estimate the tendency for flow localization in plane-strain forging operations such as sidepressing. As in nonisothermal sidepressing, localization under isothermal conditions due to flow softening is in the form of shear bands that initiate and propagate in a manner similar to that previously discussed. In addition, in analogy to the $\alpha_c \geq 5$ criterion for nonuniform flow in isothermal compression, it has been found that shear bands in isothermal plane-strain operations such as sidepressing typically occur when processing conditions are such that $\alpha_p \geq 5$ (Ref 46). Figures 25 and 26 illustrate this behavior for alloys Ti-6242Si and Ti-l0V-2Fe-3Al, respectively. Figure 25 also shows the results of a FEM simulation of sidepressing of Ti-6242Si that confirms the usefulness of the α_p-parameter correlation.

Information on the effects of temperature and strain rate on flow localization, such as that obtained in isothermal compression and sidepressing tests, is often summarized in workability diagrams such as that shown in Fig. 27 for the occurrence of shear bands in plane-strain sidepressing of Ti-6242Si (Ref 46). Closed and open circles, denoting process conditions for which shear bands were and were not observed, are noted on the diagrams, one for an equiaxed-alpha microstructure and the other for a colony-alpha microstructure obtained via beta annealing. It can be seen that the two regions (called "Safe" and "Fail" in analogy to sheet metal forming-limit diagrams) are fairly well delineated by lines that correspond to temperature and strain-rate conditions for which $\alpha_p \geq 5$ pertains over a large range of strain. Workability diagrams for other hot working defects, such as cavitation and triple-point cracking, may also be obtained from experimental workability tests as well as predicted from mechanistic models (Ref 49).

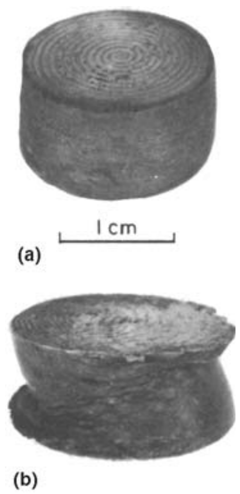


Fig. 23 Specimens of Ti-6Al-2Sn-4Zr-2Mo-0.1Si from isothermal hot compression tests at 913 °C (1675 °F), $\dot{\bar{\varepsilon}} \approx 2$ s^{-1}. Starting microstructures were (a) equiaxed alpha and (b) colony alpha. Source: Ref 47

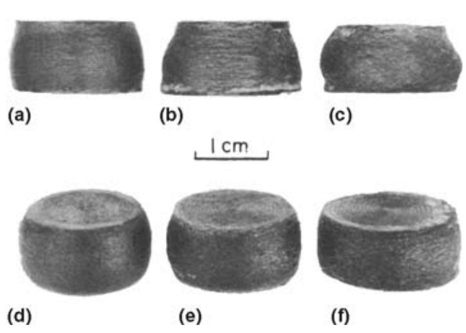


Fig. 24 Specimens of Ti-10V-2Fe-3Al from isothermal hot compression tests at (a, b, c) 704 °C (1300 °F) and (d, e, f) 816 °C (1500 °F). Strain rates were (a, d) 10^{-3} s^{-1}, (b, e) 10^{-1} s^{-1}, and (c, f) 10 s^{-1}. Prior to testing, the alloy had been beta annealed to yield an equiaxed-beta starting microstructure.

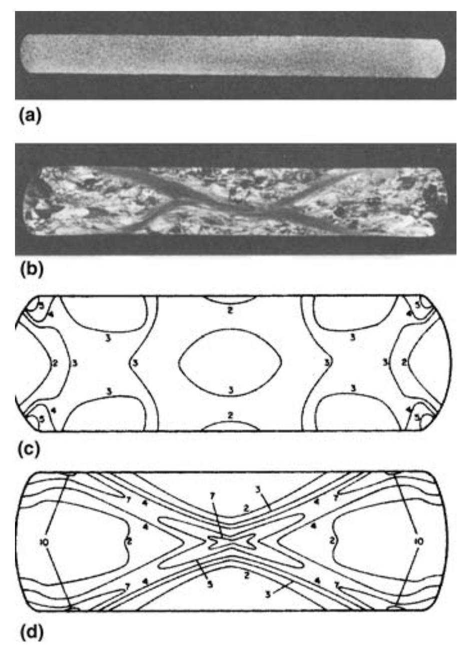


Fig. 25 (a,b) Transverse metallographic sections of bars of Ti-6Al-2Sn-4Zr-2Mo-0.1Si isothermally sidepressed at 913 °C (1675 °F), $\dot{\bar{\varepsilon}} \approx 2$ s^{-1}. (c, d) FEM-simulation predictions of contours of constant strain rate. Specimen in (a) and simulation in (c) are for the alloy with an equiaxed-alpha starting microstructure (no shear bands), and specimen in (b) and simulation in (d) are for the alloy with a colony-alpha starting microstructure (that exhibited shear bands). Source: Ref 46

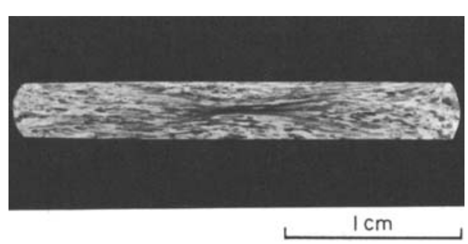


Fig. 26 Transverse metallographic section of specimen of Ti-10V-2Fe-3Al isothermally sidepressed at 704 °C (1300 °F), $\dot{\bar{\varepsilon}} \approx 10$ s^{-1}, which exhibited shear bands

Workability Tests for Cold Open-Die Forging of Recrystallized Structures

As the forging temperature is lowered into the cold working regime (i.e., to temperatures less than or equal to one-fourth of the melting or solidus temperature), workability problems change from failure related largely to intergranular cracking and cavitation to failure by a mode involving transgranular ductile fracture or shear banding. Workability tests used to diagnose these problems are discussed in turn.

Workability Tests for Ductile Fracture During Cold Forging. Ductile fracture consists of the nucleation of cavities at second-phase particles and inclusions and the growth and eventual coalescence of these cavities under the action of tensile stresses to cause fracture. In cold forging, ductile fracture occurs most readily at free surfaces where barreling has set up secondary tensile stresses. A simple workability test for analyzing these kinds of failures is the upset test in which various combinations of tensile and compressive stress fields are developed by use of various lubrication conditions and initial specimen height-to-diameter ratios. From tests of this sort, it has been found that surface fracture often follows a critical strain criterion, $\varepsilon_1 + a\varepsilon_2 =$ constant, where ε_1 and ε_2 are the tensile and com-

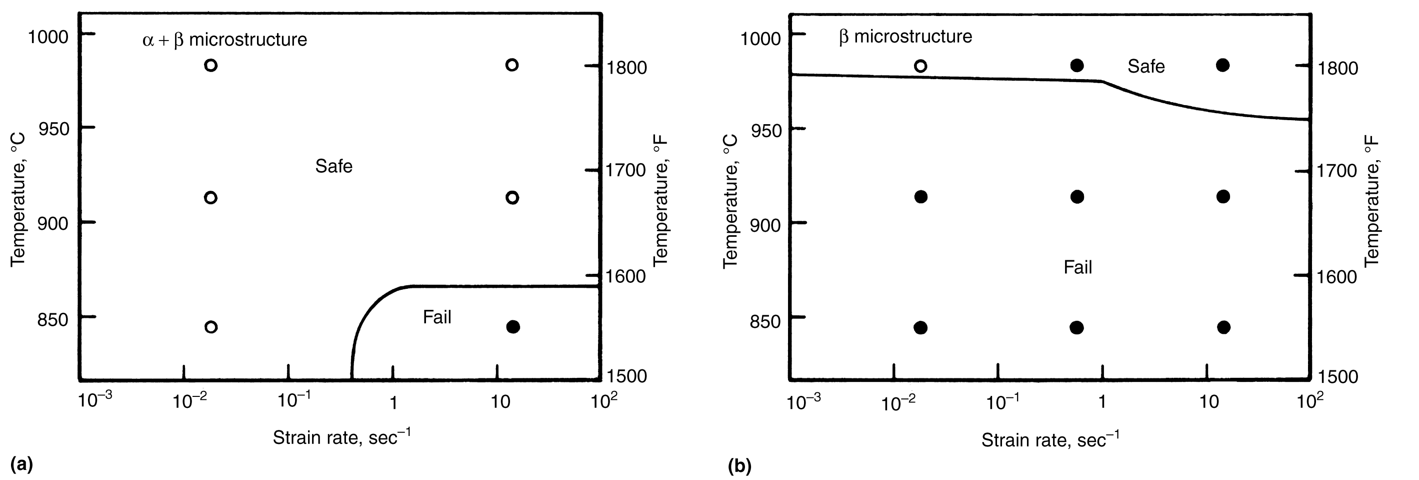


Fig. 27 Workability maps for occurence of shear bands in isothermal sidepressing of Ti-6Al-2Sn-4Zr-2Mo-0.1Si with (a) equiaxed-alpha starting microstructure and (b) colony-alpha starting microstructure. Workability predictions based on $\alpha_p = 5$ (—). Forging conditions in which shear bands were (•) and were not (o) observed are noted. Source: Ref 46

pressive principal surface strains at fracture, respectively, and *a* denotes the slope, usually equal to −½ (Ref 50, 51). This test technique and the resulting criterion are discussed in detail in Chapter 5, "Cold Upset Testing." Determination of values of the constant in this fracture criterion is very useful in ranking the resistance of various materials and different lots of a given material to ductile fracture. Because the determination of such fracture loci involves surface strains, a modified upset technique, known as the collar test, has been developed (Ref 52, 53). In this method, circular cylinders with a flange (collar) at midheight (Fig. 28) are upset to the point of surface fracture. The hoop (tensile) and axial (compressive) strains are determined simply by measuring the final diameter and thickness of the flange.

Several tests other than those employing simple upsetting of cylindrical specimens have been devised to obtain a measure of workability in cold forging. These include the grooved-compression test and an indentation-type method using a truncated cone. Both are useful for determining the deformation responses of metals, such as low-carbon steels, that do not readily fracture in ordinary upsetting. In the grooved-compression test (Ref 54–57), cylindrical specimens with various initial aspect ratios and with axial grooves are upset under various conditions of lubrication in much the same way that standard upset testing for workability analysis is conducted. Fracture is initiated in the groove at the specimen midheight (where tensile stresses are greatest), and the average circumferential fracture strain at that point is taken as a measure of workability. Available data also show that, for a given material, the value of this circumferential fracture strain is independent of initial groove depth; however, as groove depth is increased, the axial reduction at which this strain is reached is diminished. For this reason, the grooved-compression test appears to be useful for establishing the effects of surface flaws on forgeability and in ranking the quality of different lots of the same material. In this regard, it is similar to the previously discussed notched-upset test used for hot workability. It has been suggested (Ref 54) that this test be used for establishing an index of surface quality for steel wire used in cold heading.

The truncated-cone test (Ref 58) was developed in an attempt to minimize the effects of surface flaws and the random errors they produce on workability data. Unlike the simple upset and grooved-upset tests, the truncated-cone test involves indentation by a conical tool of a cylindrical specimen whose initial diameter is greater than that of the truncation (Fig. 29). By this means, cracking is made to occur beneath the surface of the workpiece at the tool-material interface, thereby eliminating the influences of free-surface flaws and surface finish on workability limits. The reduction (measured at the specimen axis) at which cracking occurs may be used to compare the workability of different materials. Alternatively, the reduction at which a fixed crack width is obtained or the width of the crack at a given reduction has been suggested as an index of workability.

As for hot working, FEM modeling provides a useful tool for the evaluation of the workability of complex cold forgings. Detailed stress, strain, and strain-rate histories are predicted and can be coupled with an appropriate fracture criterion, such as the ductile-fracture relation of Cockcroft and Latham (Ref 36) or that of Lee and Kuhn (Ref 51) (which is derivable from the Cockcroft and Latham criterion), to provide overall deformations at which failure should occur. The usefulness of such a method has been described (Ref 59), and researchers successfully predicted critical reductions for surface fracture during upsetting of aluminum alloy 7075-T6 using a FEM simulation and a measured fracture criterion of the Lee-Kuhn type.

Ductile fracture may also occur in the workpiece interior during cold forging (Ref 60). As in free-surface fracture and centerbursting at hot working temperatures, secondary tensile stresses play a large role in such failures. In addition, the voids generated by these stresses may form gross centerbursts, as in hot forging. Methods of establishing the sources of centerbursts are discussed in Chapter 21, "Workability and Process Design in Extrusion and Wire Drawing," inasmuch as these are the principal operations in which these defects occur.

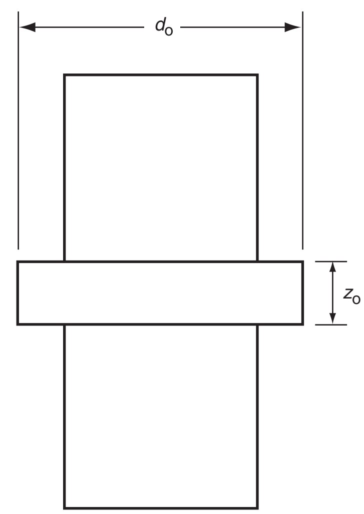


Fig. 28 Typical specimen geometry for the collar test. Source: Ref 52

Workability Tests for Shear Banding during Cold Forging. Shear bands are defects that may be found in cold forging as well as in hot forging. In open-die forging, shear-band formation may be attributed almost wholly to the formation of dead-metal zones as a result of high levels of friction at the die-workpiece interface. Therefore, workability techniques are relied on to evaluate the effects of lubricants on friction and metal flow.

The most common test for evaluation of friction is the ring compression test (Ref 61) (Chapter 6, "Hot Compression Testing"). In this

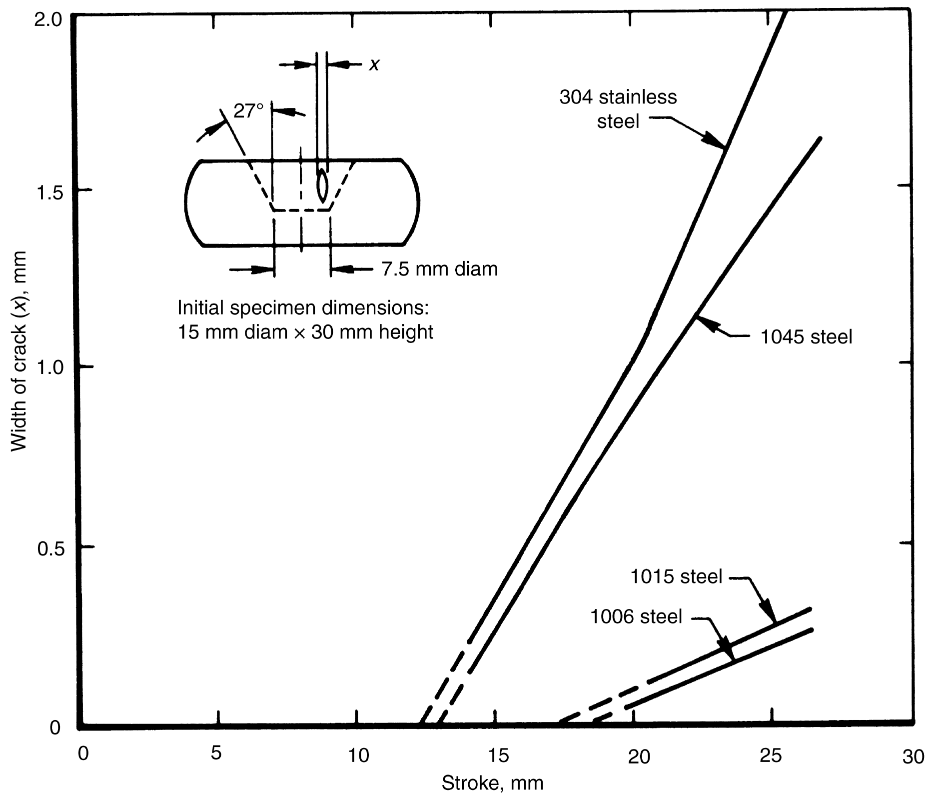


Fig. 29 Relationship between crack width and stroke in truncated-cone indentation test for workability of various steels at cold forging temperatures. Source: Ref 58

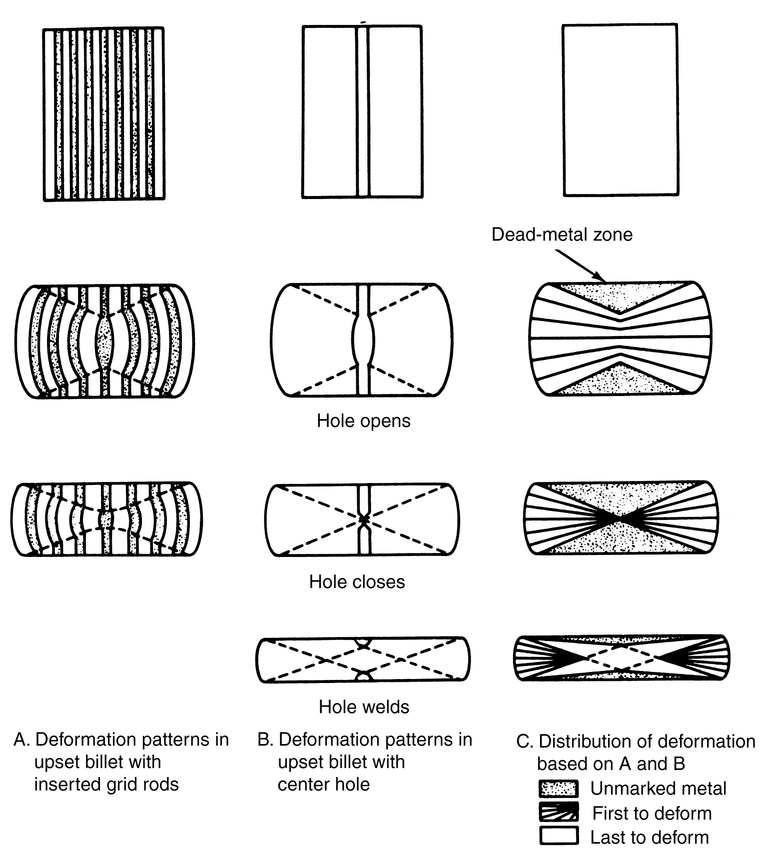


Fig. 30 Deformation patterns in cylindrical billets upset between flat, parallel dies. Source: Ref 2, 62, 63

test, rings of the workpiece material are upset to various reductions using different lubricants. With good lubrication, the inside diameter of the ring increases, whereas the reverse occurs under conditions of poor lubrication. Through measurements of reduction and percentage change in inside diameter, a quantitative measure of lubricity may be obtained. This measure may be in terms of a coefficient of friction, μ, or a friction factor, m, depending on the analysis used to reduce the test data. In addition, the flow curve, with corrections for friction, can also be determined through analysis of ring-test data.

Other techniques for evaluation of friction include those in which deformation patterns are visualized through the use of grid wires inserted, or small-diameter holes drilled, into the workpiece (Fig. 30). Either method may be used to study the formation of dead-metal zones. As shown in Fig. 30, the dead-metal zones in simple upsetting form early in deformation and tend not to deform at all until the reduction is high enough to bring them into contact. Prior to this, shear bands may develop between the dead-metal zones and the deforming bulk, moving radially outward. Once the dead-metal zones come into contact, however, deformation becomes more uniform, and high levels of hydrostatic compression are set up. This stress state tends to close holes in test specimens with such intentional initial defects (Fig. 30).

Workability in Closed-Die Forging

The defects mentioned previously may also occur in closed-die or impression-die forging of metals. Thus, the workability tests discussed previously may be used to evaluate the formation of defects under conditions in which more constraint is applied to metal flow. In addition, other defects occur in closed-die forging that may or may not involve formation of fractures or regions of localized flow. These defects often result from factors such as improper choice of the starting (or preform) shape of the workpiece that is used in the forging operation, poor die design, and poor choice of lubricant or process variables. All of these may contribute to formation of flaws such as laps, flow-through defects, extrusion defects, cold shuts, and so on.

Laps are defects that form when metal folds over itself during forging (Ref 2, 64). This may occur in rib-web forgings at a variety of locations. One such location is the web in a forging in which the preform web is too thin. During finish forging, such a web may buckle, causing a lap to form (Fig. 31). Another location is a rib in which metal is made to flow nonuniformly (Fig. 32, 33). Most often, a lap at this location is a result of an excessively sharp fillet radius in the forging die.

Flow-through defects are flaws that form when metal is forced to flow past a recess after it has filled or has ceased to deform because of chilling (Ref 2) (Fig. 34, 35). Similar to laps in appearance, flow-through defects may be shallow but are indicative of an undesirable grain-flow pattern or shear band that extends much deeper into the forging.

A special compression test using cylindrical samples with a reduced section (Fig. 36a) has been developed to quantify the effect of material properties on the severity of flow localization that may give rise to flow-through and other kinds of defects (Ref 67, 68). Following forging, a parameter know as the distributed gage volume (DGV) is calculated from measurements of the original gage volume, V_o, and final apparent gage volume, V_f, between the specimen ends, that is, DGV = $[(V_o - V_f)/V_o] \times 100\%$. The DGV provides a measure of the ability to distribute deformation. Greater ability to distribute deformation results in greater penetration of the reduced section into the specimen ends and thus higher values of DGV. Processing conditions that produce high values of DGV are attractive, in that they are least likely to give shear bands, flow-through defects, and so on. Typical results from the test are shown in Fig. 36(b).

Flow-through defects may also occur when trapped lubricant forces metal to flow past an impression.

Extrusion-type defects are formed when centrally located ribs formed by extrusion-type flow draw too much metal from the main body or web of the forging. A defect similar to a pipe cavity is thus formed (Fig. 37) (Ref 2). Means of minimizing the occurrence of these defects include increasing the thickness of the web or designing the forging with a small rib opposite the larger rib.

Most of the defects just summarized are found in hot forging, which is most common for impression-die forging. Because of this, defect formation may also involve entrapment of oxides (and lubricant). When this occurs, the metal is incapable of rewelding itself back together under the high forging pressures. The term *cold shut* is often applied in conjunction with laps, flow-through defects, and so on in describing the flaws that are generated.

As summarized in Ref 64 and 69, three primary techniques are available for investigation and correction of metal-flow defects such as those described previously. These techniques involve the application of empirical guidelines in designing preforms and finish-forging dies, physical modeling studies, and computer-aided design techniques. Computer-aided design techniques are useful in the calculation of required volume distributions in forging preforms and the determination of neutral planes and general metal-flow patterns. With modern FEM techniques, preform geometries and metal-flow patterns that avoid defect formation in complex forgings are readily established.

Empirical Guidelines for Die and Preform Design. Several simple guidelines for design of forging dies and preforms to avoid

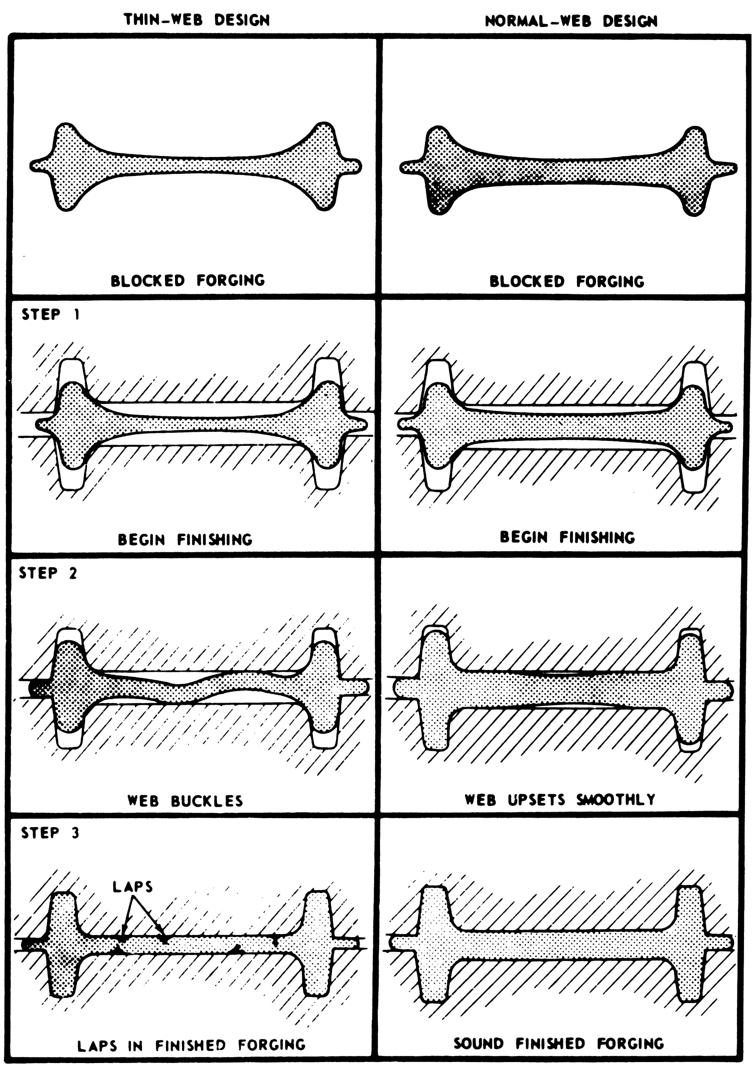


Fig. 31 Typical deformation sequence in closed-die forging of a rib-web part, showing how laps can be generated if preform geometry is selected improperly. Source: Ref 2

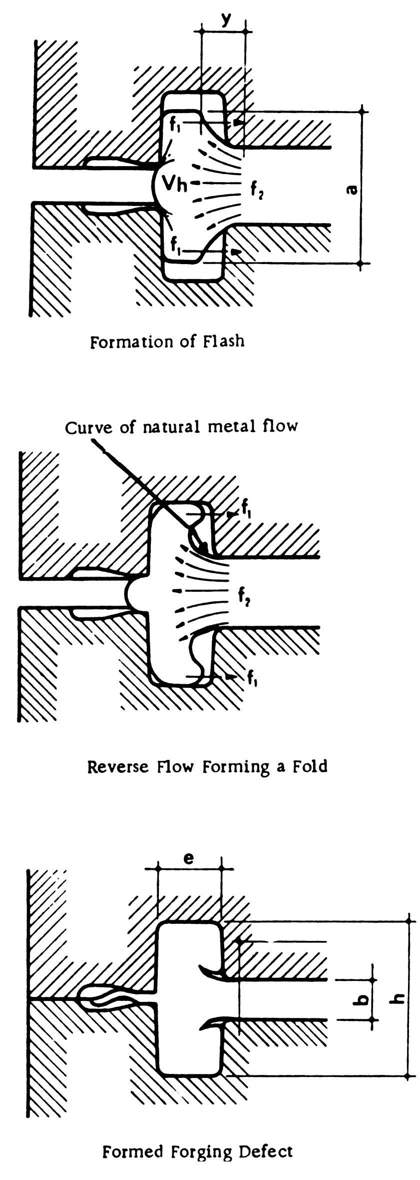


Fig. 32 Lap formation in the rib of a rib-web part due to improper preform geometry. Source: Ref 65

defect formation have been developed through years of experience. By and large, these guidelines are for conventional hot forging in which the dies are much cooler than the workpiece. From the viewpoint of dies, the most important design parameters are draft angles, corner and fillet radii, and rib and web dimensions (Ref 1–3). Some of these guidelines are summarized subsequently. More information on this topic can be found in Chapter 14, "Process Design in Impression Die Forging."

Draft angles are needed to allow removal of parts after forging; they are generally in the range of 1 to 7°, with lower values being used when part-knockout mechanisms are employed or when surface area must be minimized because of frictional or chilling effects. In addition, easily forged materials, such as aluminum alloys, tend to require small draft angles (~1°), carbon and alloy steels require moderate draft angles (3 to 5°), and hard-to-work alloys, such as nickel-base superalloys, need large draft angles (7° or greater) because of higher forging pressures and a greater tendency for sticking in the dies (Ref 2).

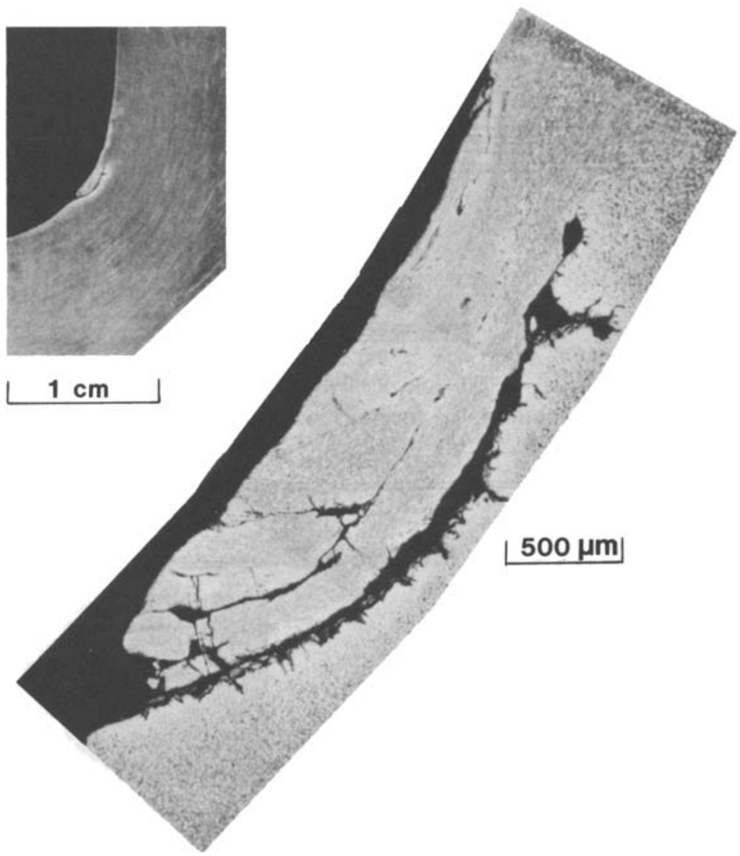


Fig. 33 Lap defect in Ti-6Al-4V bulkhead forging. Source: Ref 66

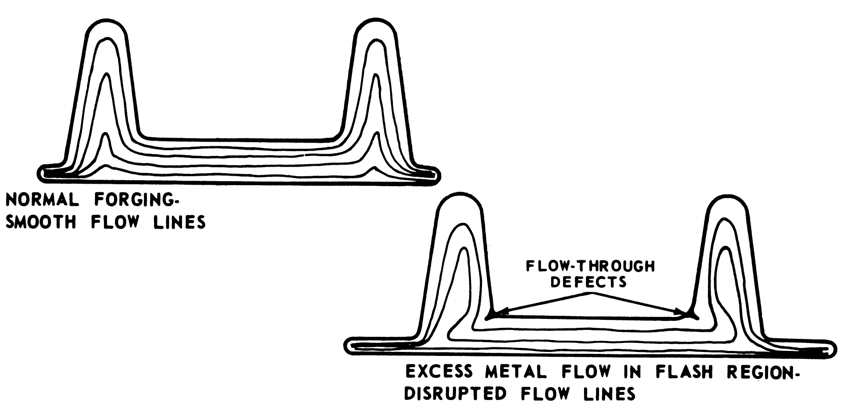


Fig. 34 Metal-flow patterns in forgings without and with flow-through defects. Source: Ref 2

Radii are also important in die design, because an insufficient radius may lead to lap formation or incomplete die fill. As with draft angles, required radii increase with increasing forging difficulty of the workpiece alloy. Typical values for making 2.5 cm (1 in.) high ribs are given in Table 2. Guidelines for minimum dimensions of ribs and webs in conventional hot forgings may also be suggested, based on experience (Table 3) (Ref 2). The important or deciding factors here are the cooling conditions of the forging, which depend on the workpiece alloy and deformation speed, among other things, and the minimum web dimension, which determines the amount of lateral constraint.

Because the formation of metal-flow defects also depends a great deal on the shape of preforms, it is not surprising that guidelines for preform design have been devised. Primarily intended for conventional hot forging, these guidelines provide a basic idea of how much metal flow can be imposed under typical forging conditions. A number of useful relations between preform and finish forging dimensions have been suggested for steel, aluminum-alloy, and titanium-alloy rib-web structural forgings (Ref 1). Those for aluminum and titanium alloys are summarized in Table 4.

Physical Modeling. Although FEM computer modeling has become the most common technique for the design of preform and finish forging dies, an alternative approach, physical modeling, is still finding use as a heuristic tool as well as a method to validate the detailed flow patterns predicted by FEM analyses. Based on the application of the principles of similarity, physical modeling makes use of small-scale dies and model materials.

For the model system to have applicability, it is best that the laws of similarity be adhered to. The laws of similarity consist of rules of geometric, physical, and boundary-condition similitude. From a geometric standpoint, model dies must be designed so as to approach, as nearly as possible, exact scale models of the envisioned forging dies—that is, all linear dimensions, radii, angles, and so forth of the model dies must be in fixed proportions to those of the real tooling. A similar relation between model and actual workpiece dimensions should be followed. Physical similarity pertains largely to selection of a model material whose deformation and thermal response are similar to those of the intended workpiece material. Boundary-condition similarity is related to interface friction and heat-transfer considerations.

In physical modeling of forging, model specimens/workpieces are usually sectioned and placed back together after a grid pattern has been inscribed on one of the flat faces. By this means, the flow pattern may be visualized, and regions where defects are likely to form in an actual forging may be determined. Such a technique is similar to the grid technique used on for diagnosing problems in sheet metal forming.

Physical modeling has been applied often to simulation of both conventional and isothermal hot forging (Ref 70–73). These studies are usually performed at room temperature with model materials that are rate-sensitive and non-strain-hardening, as are typical forging alloys at hot working temperatures. Lead, plasticine, and wax are the most common choices for materials. With these materials, model dies need not be of high strength and can be made from inexpensive steels, aluminum alloys, or even hard plastics. Because the tests are usually conducted at room temperature, boundary-condition similarity with regard to heat transfer is satisfied for isothermal but not for conventional hot forging. Therefore, information regarding the possible effects of chilling on metal flow may not be obtainable from simple modeling studies. On the other hand, it may be possible to adjust the frictional boundary conditions to compensate for this deficiency.

The work described in Ref 74 provides a good example of the usefulness of physical modeling. In this work, densification behavior during the forging of ferrous-powder compacts was rationalized based on the deformation of wedge samples made from plasticine (Fig. 38). Variations in both normal and shear strains, readily apparent on the surface and cross sections of the plasticine samples, were used to explain density gradients in actual forgings.

Figure 39 gives an example of how a lap defect is visualized in a physical model of a closed-die-forging process (Ref 75). Here, the grid lines give valuable information not only on the development of the lap in the forging but also on the occurrence of a dead-metal zone in contact with the bottom die. Similar work has been done in demonstrating the occurrence of laps in an engine impeller part (Ref 72). It was shown by the use of the model forgings that the laps could be avoided by changes in preform design.

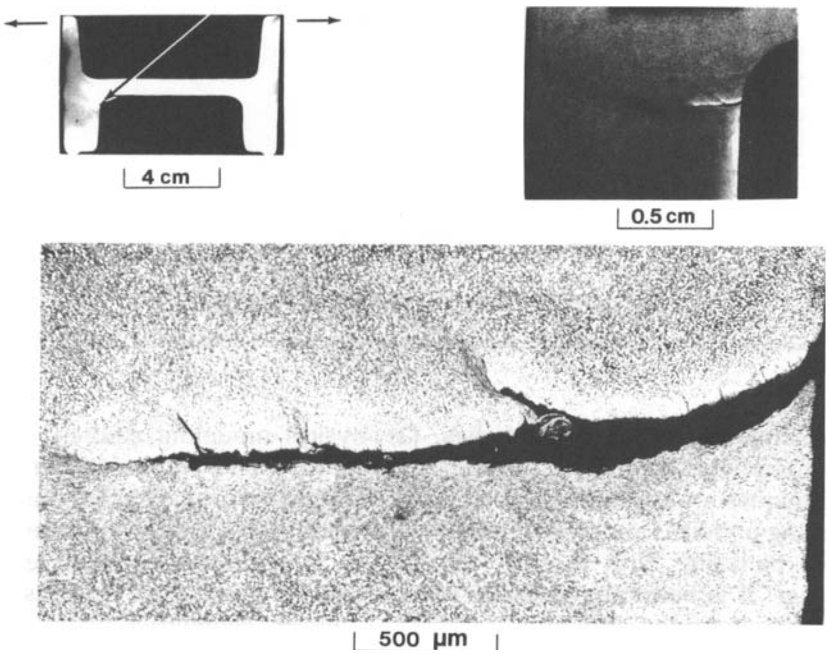


Fig. 35 Flow-through defect in Ti-6Al-4V rib-web structural part. Source: Ref 66

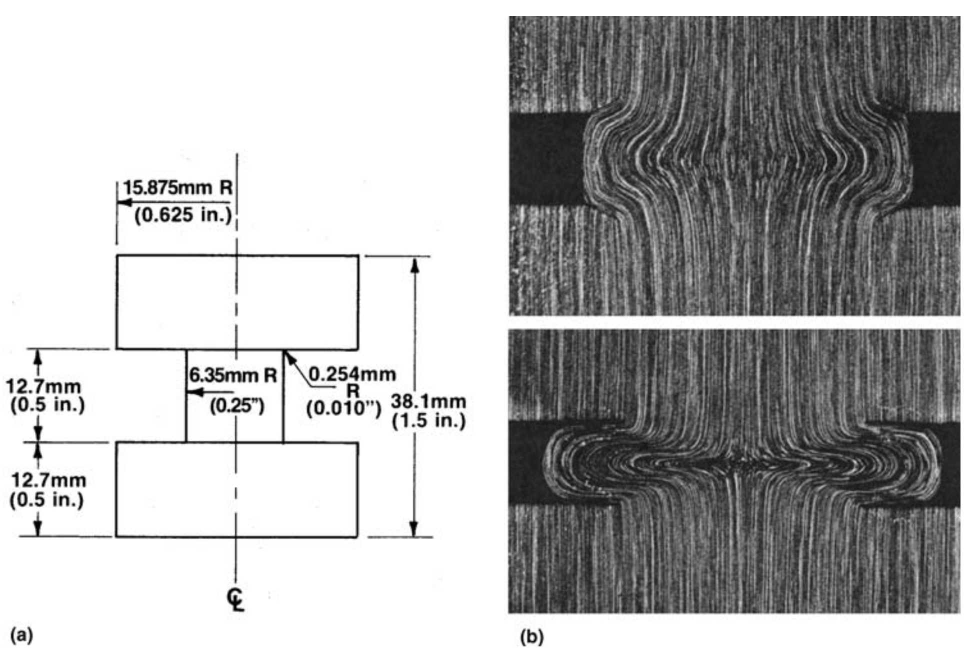


Fig. 36 The reduced-section compression test. (a) Typical specimen geometry. (b) Specimens of JBK-75 hot forged at 650 °C (1200 °F) (top) or 760 °C (1400 °F) (bottom). Source: Ref 67, 68

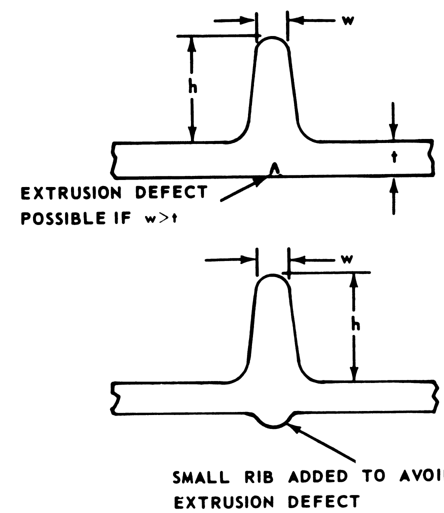


Fig. 37 Extrusion-type defect in centrally located rib (top), and die-design modification used to avoid defect (bottom). *w*, width; *h*, height; *t*, thickness. Source: Ref 2

Table 2 Fillet and corner radii suggested for forgings of several alloys with 2.5 cm (1 in.) high ribs

Alloy	Fillet radius, mm	Corner radius, mm
2014 (aluminum)	6	3
AISI 4340 (alloy steel)	10–13	3
H11 (hot work steel)	10–13	5
17-7PH (stainless steel)	6–13	5
A286 (superalloy)	13–19	6
Ti-6Al-4V	13–16	6
Unalloyed molybdenum	13	13

Note: From the viewpoint of the forging or workpiece. Source: Ref 2

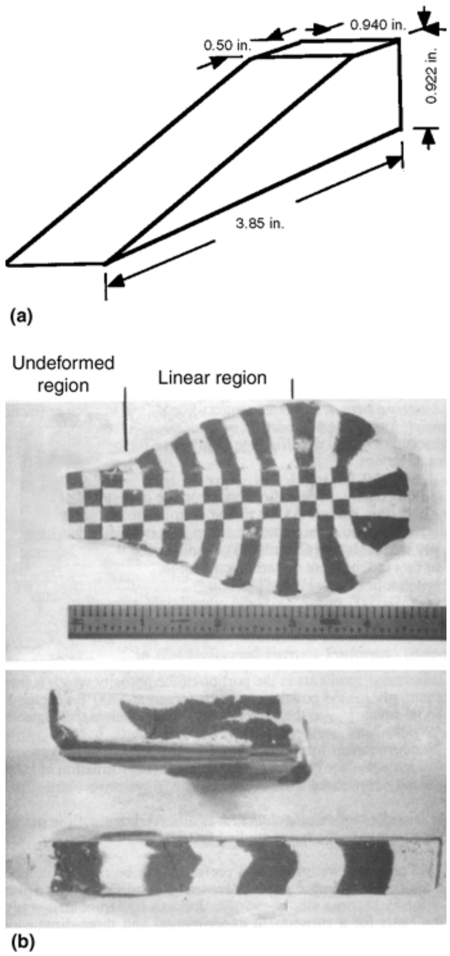


Fig. 38 Physical modeling of the wedge-forging test. (a) Preform geometry. (b) Plan view, longitudinal section, and transverse section of forged plasticine wedge sample. Source: Ref 74

Summary

Workability in forging depends on a variety of material, process-variable, and die-design features. Because of the almost unlimited range of possible forging operations, the forging engineer must use a variety of workability tests to diagnose forging problems. For example, open-die forging operations do not impose the same amount of lateral constraint on the workpiece as closed-die forging, and thus, different techniques are used in the evaluation of workability for open-die forging and closed-die forging. Workability may also be associated with equipment use (e.g., press capacity, die materials, die, etc.) and evolution of microstructure, as discussed in other chapters of this Handbook.

For open-die forging, the wedge-forging (or rolling), sidepressing, and double-cone tests are useful for determining optimal processing parameters for breaking down cast structures and for avoiding fracture. For forging of wrought and recrystallized structures, a number of tests are available for establishing the effects of process variables, such as temperature and deformation rate, and of secondary tensile stresses on working limits controlled by fracture or flow localization preceding fracture. These include standard tension, compression, and sidepressing tests as well as variations of these tests introduced to simulate chilling (e.g., nonisothermal compression test) and the development of complex stress states (notched-upset test, double-cone test).

Techniques for predicting the occurrence of forging defects in closed-die forging, which may not involve fracture or flow localization, are also available. These methods include physical modeling, in which easy-to-deform model materials, such as lead, plasticine, and wax, are deformed in inexpensive model dies. In such a way, features of preform or die design that may cause laps or other undesirable defects may be established economically. The use of computer simulation techniques to diagnose workability problems in both open- and closed-die forging has become widespread with the advent of inexpensive computers and powerful FEM software. With such tools, the forging engineer can readily determine stresses and metal-flow patterns and thus predict the occurrence of metal-flow defects, gross shear bands, voids, and so on.

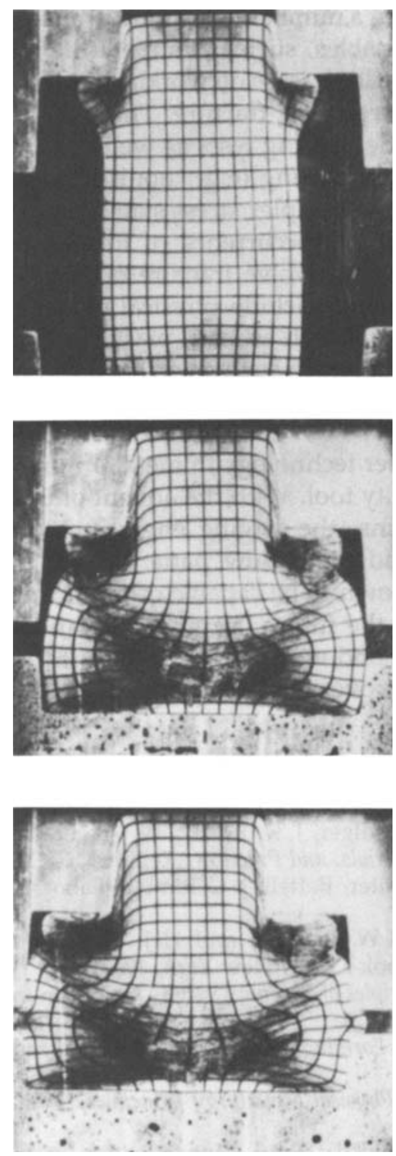


Fig. 39 Forging sequence of gridded lead billet used in physical modeling studies to establish the effect of die and preform design on the occurrence of metal-flow defects such as laps. Source: Ref 75. Courtesy of S. Kobayashi

Table 3 Suggested minimum section thicknesses for rib-web forgings of several alloys

	Minimum rib thickness(a), mm, for forgings of given plan area, cm^2, of:			Minimum web thickness(a), mm, for forgings of given plan area, cm^2, of:		
Alloy	Up to 25	25 to 250	250 to 750	Up to 25	25 to 250	250 to 750
2014 (aluminum)	3	5	. . .	3	6	8
AISI 4340 (alloy steel)	3	5	. . .	5	8	13
H11 (hot work steel)	3	5	. . .	. . .	10	. . .
17-7PH (stainless steel)	. . .	6	. . .	. . .	10	. . .
A286 (superalloy)	. . .	6	. . .	. . .	10	. . .
Ti-6Al-4V	. . .	6	. . .	. . .	8	. . .
Unalloyed molybdenum	. . .	10	. . .	. . .	10	. . .

(a) Values listed are for rib-web forgings with 7° draft and standard fillet and corner radii. Source: Ref 2

Table 4 Relations between preform and finished forging dimensions for rib-web structural parts

Dimensions in finished forgings(a)	Dimensions in preform(a): Aluminum alloys	Dimensions in preform(a): Titanium alloys
Web thickness, t_F	$t_p \approx (1–1.5)\ t_F$	$t_p \approx (1.5–2.2)\ t_F$
Fillet radii, R_{FF}	$R_{PF} \approx (1.2–2)\ R_{FF}$	$R_{PF} \approx (2–3)\ R_{FF}$
Corner radii, R_{FC}	$R_{PC} \approx (1.2–2)\ R_{FC}$	$R_{PC} \approx (2)\ R_{FC}$
Draft angle, δ_F	$\delta_p \approx \delta_F$ (2–5°)	$\delta_p \approx \delta_F$ (3–5°)
Width of rib, W_F	$W_p \approx W_F - 0.8$ mm	$W_p \approx W_F - (1.6–3.2$ mm)

(a) The first subscript letter for each dimension indicates finish forging (F) or preform (P). Source: Ref 1

REFERENCES

1. T. Altan, F.W. Boulger, J.R. Becker, N. Akgerman, and H.J. Henning, "Forging Equipment, Materials, and Practices," Report MCIC-HB-03, Metals and Ceramics Information Center, Battelle's Columbus Laboratories, Columbus, OH, Oct 1973
2. A.M. Sabroff, F.W. Boulger, and H.J. Henning, *Forging Materials and Practices*, Rheinhold Book Corp., 1968
3. J.A. Schey, *Principles of Forging Design*, American Iron and Steel Institute, 1965
4. B. Avitzur, *Metal Forming: Processes and Analysis,* McGraw-Hill, 1968
5. R.E. Reed-Hill, *Physical Metallurgy Principles,* Van Nostrand, 1973
6. F.N. Rhines and P.J. Wray, Investigation of the Intermediate Temperature Ductility Minimum in Metals, *Trans. ASM*, Vol 54, 1961, p 117
7. W.A. Backofen, *Deformation Processing*, Addison-Wesley, 1972
8. A.J. DeRidder and R. Koch, "Forging and Processing of High-Temperature Alloys," *MiCon 78: Optimization of Processing, Properties, and Service Performance Through Microstructural Control,* ASTM STP 672, H. Abrams, G.N. Maniar, D.A. Nail, and H.D. Solomon, Ed., American Society for Testing and Materials, 1979, p 547
9. J.J. Park and S. Kobayashi, Three-Dimensional Finite-Element Analysis of Block Compression, *Int. J. Mech. Sci.,* Vol 26, 1984, p 165
10. V.K. Jain, L.E. Matson, H.L. Gegel, and R. Srinivasan, Physical Modeling of Metalworking Processes II: Comparison of Viscoplastic Modeling and Computer Simulation, *Physical Modeling of Metalworking Processes*, E. Erman and S.L. Semiatin, Ed., TMS, 1987, p 127
11. D.D. Bhatt, G.E. Meyer, and A.L. Hoffmanner, "Optimum Parameters for Deformation Processing of a Plasma-Arc Melted High-Nitrogen Stainless Steel," paper presented at the Soviet-American Symposium on Problems of Electrometallurgy and Welding, Sept 1977 (Kiev, Ukraine)
12. G.E. Dieter, *Mechanical Metallurgy*, McGraw-Hill, 1976
13. G.F. Torkhov, Y.V. Latash, R.R. Fessler, A.H. Clauer, E.E. Fletcher, and A.L. Hoffmanner, Development of Melting and Thermomechanical-Processing Parameters for a High- Nitrogen Stainless Steel Prepared by Plasma-Arc Remelting, *J. Met.,* Vol 30 (No. 12), Dec 1978, p 20
14. E.E. Fletcher and R.R. Fessler, "Characterization of an Ingot of High-Nitrogen Stainless Steel Produced by Plasma-Arc Remelting," paper presented at the Soviet-American Symposium on Problems of Electrometallurgy and Welding, 5–7 Sept 1977 (Kiev, Ukraine)
15. J.J. Jonas, Recovery, Recrystallization, and Precipitation Under Hot-Working Conditions, *Proc. Fourth Int. Conf. on the Strength of Metals and Alloys*, 30 Aug–3 Sept 1976 (Nancy, France), p 976
16. M. Jackson, R.J. Dashwood, L. Christodoulou, and H.M. Flower, Application of Novel Technique to Examine Thermomechanical Processing of Near β Alloy Ti-10V-2Fe-3Al, *Mater. Sci. Technol.,* Vol 16, 2000, p 1437
17. A.L. Hoffmanner, "Plasticity Theory as Applied to Forging of Titanium Alloys," paper presented at the Symposium on the Thermal-Mechanical Treatment of Metals, 1 May 1970 (London), Institute of Metals
18. E.P. Unksov, *An Engineering Theory of Plasticity,* Butterworth and Co., 1961
19. S.L. Semiatin, Battelle's Columbus Laboratories, unpublished research, March 1979
20. H. Gegel, S. Nadiv, and R. Raj, Dynamic Effects on Flow and Fracture During Isothermal Forging of a Titanium Alloy, *Scr. Metall.,* Vol 14, 1980, p 241
21. K.A. Bywater and T. Gladman, Influence of Composition and Microstructure on Hot Workability of Austenitic Stainless Steels, *Met. Technol.,* Vol 3, 1976, p 358
22. R.E. Bailey, Predicting the Hot-Working Characteristics of Superalloys by Hot Tensile Testing, *Met. Eng. Q.,* Vol 15, May 1975, p 43
23. N. Cederblad and N.J. Grant, Hot Workability of Nimonic 115 as a Function of Strain Rate, *Metall. Trans. A*, Vol 6, 1975, p 1547
24. R.J. Quigg, Effect of Phase Changes on the Workability and Mechanical Properties of Udimet 700, *High Temperature Materials II*, G.M. Ault, W.F. Barclay, and H.P. Munger, Ed., Interscience Publishers, 1963, p 245
25. A. Laasraoui and J.J. Jonas, Recrystallization of Austenite After Deformation at High Temperatures and Strain Rates—Analysis and Modeling, *Metall. Trans. A,* Vol 22, 1991, p 151
26. M.C. Mataya, Simulating Microstructural Evolution During the Hot Working of Alloy 718, *JOM*, Vol 51 (No. 1), Jan 1999, p 18
27. A.A. Korshunov, F.U. Enikeev, M.I. Mazurskii, G.A. Salishchev, A.V. Muravlev, P.V. Chistyakov, and O.V. Dmitriev, Effect of Method of High-Temperature Loading on Transformation of Lamellar Structure in VT9 Titanium Alloy, *Russ. Metall.,* Vol 3, 1994, p 103
28. C.M. Sellars and Q. Zhu, The Effect of Strain Path on Mechanical Behaviour and Microstructure, *20th Risø Int. Symp. on Materials Science,* J.B. Bilde-Sørenson et al., Ed., Risø, 1999, p 167
29. S.B.Davenport, R.L. Higginson, and C.M. Sellars, The Effect of Strain Path on Material Behaviour during Hot Rolling of FCC Metals, *Philos. Trans. R. Soc. (London) A,* Vol 357, 1999, p 1645
30. R.P. Daykin, Ladish Co., unpublished research, 22 June 1951
31. Anonymous, "High-Temperature, High-Strength Nickel-Base Alloys," The International Nickel Company, Inc., 1977
32. J.C. Malas, H.L. Gegel, S.I. Oh, and G.D. Lahoti, Metallurgical Validation of a Finite-Element Program for the Modeling of Isothermal Forging Processes, *Experimental Verification of Process Models*, American Society for Metals, 1983, p 358
33. S.I. Oh, Finite Element Analysis of Metal Forming Problems with Arbitrarily Shaped Dies, *Int. J. Mech. Sci.,* Vol 24, 1982, p 479
34. B. Antolovich and M. Evans, Predicting Grain Size Evolution of Udimet Alloy 718 During the Cogging Process Through the Use of Numerical Analysis, *Superalloys 2000*, T.M. Pollock, R.D. Kissinger, R.R. Bowman, K.A. Green, M. McLean, S. Olson, and J.J. Schirra, Ed., TMS, 2000, p 39
35. S.L. Semiatin, R.L. Goetz, E.B. Shell, V. Seetharaman, and A.K. Ghosh, Cavitation and Failure During Hot Forging of Ti-6Al-4V, *Metall. Mater. Trans. A,* Vol 30, 1999, p 1411
36. M.G. Cockcroft and D.J. Latham, Ductility and Workability of Metals, *J. Inst. Met.,* Vol 96, 1968, p 33
37. A.K. Ghosh, D.H. Bae, and S.L. Semiatin, Initiation and Early Stages of Cavity Growth During Superplastic and Hot Deformation, *Mater. Sci. Forum,* Vol 304/306, 1999, p 609
38. P.D. Nicolaou and S.L. Semiatin, The Effect of Stress Triaxiality on Tensile Behavior of Cavitating Specimens, *J. Mater. Sci.,* Vol 36, 2001, p 5155
39. J.R. Rice and D.M. Tracey, On Ductile Enlargement of Voids in Triaxial Stress Fields, *J. Mech. Phys. Solids,* Vol 17, 1969, p 201
40. P.D. Nicolaou and S.L. Semiatin, An Experimental and Theoretical Investigation of the Influence of Stress State on Cavitation During Hot Working, *Acta Mater.,* Vol 51, 2003, p 613
41. D. Watkins, H.R. Piehler, V. Seetharaman, C.M. Lombard, and S.L. Semiatin, Effect of Hydrostatic Pressure on the Hot Working Behavior of a Gamma Titanium Aluminide, *Metall. Trans. A,* Vol 23, 1992, p 2669
42. S.L. Semiatin and G.D. Lahoti, The Occurrence of Shear Bands in Nonisothermal, Hot Forging of Ti-6Al-2Sn-4Zr-2Mo-0.lSi, *Metall. Trans. A,* Vol 14, 1983, p 105
43. S.L. Semiatin, G.D. Lahoti, and S.I. Oh, The Occurrence of Shear Bands in Metalworking, *Material Behavior under High Stress and Ultrahigh Loading Rates,* Plenum Press, 1983, p 119
44. S.L. Semiatin and G.D. Lahoti, Deformation and Unstable Flow in Hot Forging of Ti-6Al-2Sn-4Zr-2Mo-0.lSi, *Metall. Trans. A,* Vol 12, 1981, p 1705
45. J.J. Jonas and M.J. Luton, Flow Softening at

Elevated Temperatures, *Advances in Deformation Processing*, J.J. Burke and V. Weiss, Ed., Plenum Press, 1978, p 215
46. S.L. Semiatin and G.D. Lahoti, The Occurrence of Shear Bands in Isothermal, Hot Forging, *Metall. Trans. A*, Vol 13, 1982, p 275
47. S.L. Semiatin, G.D. Lahoti, and T. Altan, Determination and Analysis of Flow Stress Data for Ti-6242 at Hot-Working Temperatures, *Process Modeling—Fundamentals and Applications to Metals,* T. Altan, H. Burte, H. Gegel, and A. Male, Ed., American Society for Metals, 1980, p 387
48. J.J. Jonas, R.A. Holt, and C.E. Coleman, Plastic Stability in Tension and Compression, *Acta Metall.,* Vol 24, 1976, p 911
49. R.C. Koeller and R. Raj, Diffusional Relaxation of Stress Concentration at Second Phase Particles, *Acta Metall.,* Vol 26, 1978, p 1551
50. P.W. Lee and H.A. Kuhn, Fracture in Cold Upset Forging—A Criterion and Model, *Metall. Trans.,* Vol 4, 1973, p 969
51. H.A. Kuhn, P.W. Lee, and T. Erturk, A Fracture Criterion for Cold Forming, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 95H, 1973, p 213
52. R. Sowerby, N. Chandrasekaran, N.L. Dung, and O. Mahrenholtz, An Analysis of Some Upsetting Tests for Assessing the Cold Workability of Steels, *Fracture 84*, Vol 5, Pergamon Press, 1984, p 3239
53. M. Kaiso and M. Katsumata, Cold Forgeability in Medium Carbon Steel with Insufficiently Spheroidized Microstructure, *Tetsu-to-Hagane (J. Iron Steel Inst. Jpn.),* Vol 84, 1998, p 721
54. P.F. Thomason, The Free Surface Ductility of Uniaxial Compression Specimens with Longitudinal Surface Defects, *Int. J. Mech. Sci.,* Vol 11, 1969, p 65
55. L. Janicek and B. Maros, The Determination of the Cold Forgeability for Specimens with Axial Notches of Heat-Resisting and Corrosion-Resisting Chromium Steels, *J. Mater. Process. Technol.,* Vol 60, 1996, p 269
56. H. Sueyoshi, K. Suenaga, and R. Tanaka, Cold Forgeability and Machinability after Cold Forging of Hypo-Eutectoid Graphitic Steels, *J. Jpn. Inst. Met.,* Vol 53, 1989, p 206
57. K. Olsson, S. Karlsson, and A. Melander, The Influence of Notches, Testing Geometry, Friction Conditions, and Microstructure on the Cold Forgeability of Low Carbon Steels, *Scand. J. Metall.,* Vol 15, 1986, p 238
58. T. Okamoto, T. Fukuda, and H. Hagita, Material Fracture in Cold Forging—Systematic Classification of Working Methods and Types of Cracking in Cold Forging, *Sumitomo Search*, No. 9, May 1973, p 46
59. S.I. Oh and S. Kobayashi, Workability of Aluminum Alloy 7075-T6 in Upsetting and Rolling, *J. Eng. Ind. (Trans. ASME)*, Vol 98B, 1976, p 800
60. S.K. Suh and H.A. Kuhn, Three Fracture Modes and Their Prevention in P/M Forging, *Modern Developments in Powder Metallurgy*, Vol 9, Met. Powder Ind. Tech., 1977, p 407
61. V. DePierre, G. Saul, and A.T. Male, "The Relative Validity of Coefficient of Friction and Interface Friction Shear Factor Under Conditions of Bulk Plastic Deformation," Report AFML-TR-70–243, Air Force Materials Laboratory, Wright-Patterson Air Force Base, OH, Oct 1970
62. H. Wagner and F.W. Boulger, "A Study of Possible Methods for Improving Forging and Extruding Processes for Ferrous and Non-Ferrous Materials," final report on Air Force Contract 33(600)-2627Z, Battelle Memorial Institute, Columbus, OH, June 1957
63. A. Tomlinson and J.D. Stringer, The Closing of Internal Cavities in Forgings by Upsetting, *J. Iron Steel Inst.,* Vol 188, 1958, p 209
64. T.L. Subramanian, N. Akgerman, and T. Altan, "Application of CAD/CAM to Precision Isothermal Forging of Titanium Alloys," Report AFML-TR-77-108, Battelle's Columbus Laboratories, Columbus, OH, July 1977
65. A. Chamouard, *Closed-Die Forging*, Part I, Dunod, Paris, 1964 (in French)
66. F.N. Lake and D.J. Moracz, "Comparison of Major Forging Systems," Report AFML-TR-71-112, TRW, Inc., Cleveland, May 1971
67. M.C. Mataya and G. Krauss, A Test to Evaluate Flow Localization during Forging, *J. Appl. Metalwork.*, Vol 2, 1981, p 28
68. M.C. Mataya, M.J. Carr, and G. Krauss, Flow Localization and Shear Band Formation in a Precipitation-Strengthened Austenitic Stainless Steel, *Metall. Trans. A*, Vol 13, 1982, p 1263
69. N. Akgerman, J.R. Becker, and T. Altan, Preform Design in Closed-Die Forging, *Metall. Met. Form.*, Vol 40 (No. 5), May 1973, p 135
70. T. Altan, H.J. Henning, and A.M. Sabroff, The Use of Model Materials in Predicting Forming Loads in Metalworking, *J. Eng. Ind. (Trans. ASME)*, Vol 92B, 1970, p 444
71. K. Yagishita, H. Tsukamoto, T. Egawa, S. Oomori, and J. Ibushi, "A Study of Simulative Model Test for Metal Forming Using Plasticine," Mitsubishi Technical Bulletin 91, Mitsubishi Heavy Industries, Ltd., July 1974
72. C.W. Corti, G.H. Gessinger, and A.H. Shabaik, Superplastic Isothermal Forging: A Model Metal Flow Study, *J. Mech. Work. Technol.,* Vol 1, 1977, p 35
73. E. Erman and S.L. Semiatin, *Physical Modeling of Metalworking Processes*, TMS, 1987
74. T.L. Hartman, D.F. Bickford, and H.R. Piehler, Densification and the Nature of Residual Porosity in Hot-Deformed Ferrous Preforms, *Physical Modeling of Metalworking Processes*, E. Erman and S.L. Semiatin, Ed., TMS, 1987, p 9
75. S. Kobayashi, University of California, unpublished research, 1981

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 220-231
DOI:10.1361/hwpd2003p220

Chapter 15

Modeling Techniques in Forming Processes

W.T. Wu, J.T. Jinn, and C.E. Fischer, Scientific Forming Technologies Corp.

THE OBJECTIVE OF MANUFACTURING is the production of a consistent quality product at a minimal cost. Generally effective goals include shortening the lead time in the design cycle, reducing tooling cost and machine downtime at the production stage, and developing a stable process with a minimal reject rate. The ability to predict the performance of a particular manufacturing process and to compare it with alternative manufacturing processes at an early stage in the process design cycle furthers these goals by reducing costly trial-and-error design iterations using production equipment.

Numerous analytical techniques have been developed to improve the process designer's ability to evaluate a process and to predict various aspects of the metal forming process. Early methods relied on simple analytical techniques, such as the slab method, the slip-line method, the upper bound method, and heuristics to predict forming load, critical ratios, workability limits, and die design features. Complex analytical equations were converted to charts, which could be applied by the designer.

As computer technology became prevalent in the engineering and manufacturing industry, analytical techniques, such as the upper bound method, were used to develop specialized computer programs, which could be used to analyze a particular process, such as tube sinking, strip rolling, or extrusion. The finite element method (FEM) for metal forming applications was first introduced in early 1970 (Ref 1). The continuous improvements in computer technology and FEM finally made an important impact in the metal forming industry in the mid-1980s. Due to its unique capability in describing complex shapes, boundary conditions, and realistic material thermomechanical response, the development of a general-purpose metal forming analysis software has been realized. The method has been used as an essential tool for product and process design engineers to reduce development time and cost. Due to the demand from the industry to produce more accurate simulation models, the FEM has continuously evolved from two-dimensional analysis into the true three-dimensional models since the late 1980s and early 1990s (Ref 2–4).

This chapter reviews the overall development of modeling techniques for forming processes, including:

- Slab method
- Slip-line method
- Upper bound method
- Finite element method

Modeling Techniques

The most fundamental calculations used in metal forming analysis involve a forming load estimate, which is useful in selecting the size of equipment required to form the product. The simplest formula takes the form:

$$P = K\,\bar{\sigma}\,A \qquad \text{(Eq 1)}$$

where P is the forming load, $\bar{\sigma}$ is the mean flow stress of a workpiece material under an idealized state of deformation, A is the planar area of the workpiece, and K is an empirically determined correction factor for a particular forming process. The correction factor reflects on the effects of nonuniformity of deformation and friction between the workpiece and tool.

This estimation has been improved upon by more elaborated approximate solution techniques such as the slab method, the slip-line method, the upper bound and lower bound methods, Hill's general method, and, finally, the finite element methods. These methods have been employed to allow for the estimation of not only the forming load, but also the material flow pattern and the stress distribution. The first two methods, i.e., the slab method and the slip-line method, solve the field equations in a differential form directly with certain degrees of simplification. On the other hand, the other methods solve the field equations in an integral form derived from the associated energy theorems. Except for the FEMs, application is generally limited to metal forming processes with simple geometry and idealized materials laws.

The following sections briefly review some classical solution techniques for the slab method, the slip-line method, and the upper bound method. For detailed discussion, refer to textbooks, e.g., Ref 5 to 9. The FEM is discussed in detail in a separate section.

Slab Method. As implied in its name, the slab method is based on the assumptions such that the deformation of a workpiece can be approximated with the deformation of a series of slabs, and the shape of slab (either flat or cylindrical shape) does not change in the course of deformation. Each slab having an infinitesimal thickness is sliced so that it contacts with the tool. These methods assume that the shear stress due to friction between the workpiece and tool does not affect the stress distribution within the slab, but it is considered in the force equilibrium. The stress distribution within each slab is simplified further such that the stress component normal to the thickness direction can be varied as a function of the thickness coordinate. The other nontrivial stress components are then related to the stress component normal to the thickness direction according to the yield criterion and the flow rule for a rigid-perfectly-plastic material. Then, the differential equilibrium equation expressed in terms of stresses and the thickness coordinate is solved with the associated boundary conditions. The slab method has been applied widely to various plane strain or axisymmetric forming problems such as upsetting, extrusion, drawing, and rolling.

Figure 1 illustrates a simple example of the slab method applied to a ring compression problem. In fact, the ring compression test has been used widely as a convenient tool to evaluate the friction factor between the tool and workpiece. In bulk metal forming industries, the constant shear friction model (based on shear strength of the workpiece) has been preferred to the well-known Coulomb friction model (based on contact pressure). In Fig. 1(a), the inner radius R_i, the outer radius R_o, and the height H characterize the geometry of the ring. Now consider an infinitesimal element depicted in Fig. 1(b). The equilibrium equation in the r direction can be written as follows:

$$\sigma_r\,Hrd\theta - (\sigma_r + d\sigma_r)H(r+dr)d\theta + 2\sigma_\theta \sin\left(\frac{d\theta}{2}\right)H\,dr \mp 2mkr\,dr\,d\theta = 0 \qquad \text{(Eq 2)}$$

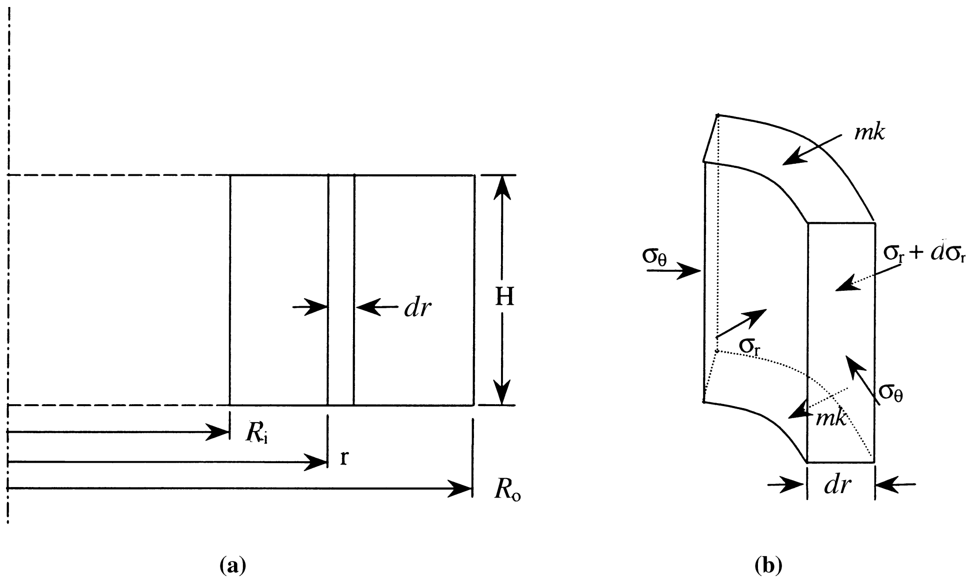


Fig. 1 A ring compression problem. (a) Geometry of ring. (b) Slab element

where m and k are the shear friction factor and the equivalent shear stress, respectively, and the sign $\mp$ is used to express that the direction of frictional stress changes from outward to inward at the neutral point as r increases. The neutral point is denoted by the radius r_n. Assuming that $\sigma_r = \sigma_\theta$ and neglecting higher order terms, Eq 2 can be simplified as:

$$d\sigma_r = \mp 2\alpha k\, dr \quad \text{(Eq 3)}$$

where $\alpha = \dfrac{m}{H}$ is a constant representing the coining effect, that is, a portion of the total forming load required to overcome friction increases as the height of ring decreases. It is noted that the assumption $\sigma_r = \sigma_\theta$ holds when the radial component of velocity u is a function of homogeneous degree one in r so that

$$\frac{du}{dr} = \frac{u}{r}$$

or

$$\dot{\varepsilon}_r = \dot{\varepsilon}_\theta \quad \text{(Eq 4)}$$

By applying the boundary condition $\sigma_r = 0$ at $r = R_i$ and $r = R_o$, Eq 4 can be integrated in the form:

$$\sigma_r = \begin{cases} 2\alpha k\,(r - R_i) & \text{when} \quad r \le r_n \\ 2\alpha k (R_o - r) & \text{when} \quad r \ge r_n \end{cases} \quad \text{(Eq 5)}$$

Since the stress must be continuous at r_n, the location of neutral point is determined as:

$$r_n = \frac{1}{2}(R_o + R_i) \quad \text{(Eq 6)}$$

By applying von Mises yield criterion:

$$\bar{\sigma} = |\sigma_z - \sigma_r| = Y \quad \text{(Eq 7)}$$

the axial component of stress can be obtained as follows:

$$\sigma_z = Y + \sigma_r \quad \text{(Eq 8)}$$

where Y is the yield stress of a material. It is noted that $Y = \sqrt{3}\ k$ for a material complying with von Mises yield criterion. The forming load can then be evaluated as:

$$P = \int_{R_i}^{r_n} \sigma_z 2\pi r\, dr + \int_{r_n}^{R_o} \sigma_z 2\pi r\, dr = K_{Slab}\, \bar{\sigma}\, A \quad \text{(Eq 9)}$$

where K_{Slab} is the correction factor or normalized forming load defined by:

$$K_{Slab} = 1 + \frac{1}{2\sqrt{3}} \frac{m}{H}(R_o - R_i) \quad \text{(Eq 10)}$$

The slip-line method is another simple and powerful classical solution method, although its application is limited to plane-strain problems for a rigid-perfectly plastic material (Ref 5, 6). In this method, the equilibrium equations for a plane-strain state are first transformed into the hyperbolic differential equations expressed in terms of the mean stress, the maximum/minimum shear stress, and the direction of maximum/minimum shear stress. The characteristics of the hyperbolic differential equations are known as the *slip-lines.* The slip-line field then can be constructed by networking two kinds of slip-lines representing the maximum and minimum constant shear lines that are orthogonal to each other. Several useful techniques have been proposed to construct the slip-line field graphically depending on the configuration of a problem and the associated boundary conditions. The forming load can then be obtained by determining integral constants for particular slip lines from the known state of stresses at some points. The slip-line method has also been applied successfully to various plane-strain forming problems such as indentation, extrusion, drawing, and rolling.

Upper-Bound Method. Unlike the two previously discussed methods, the upper-bound method (UBM) is based on the energy principle, known as the *upper-bound theorem.* The upper-bound theorem states that the rate of total energy associated with any kinematically admissible velocity field defines an upper bound to the actual rate of total energy required for the deformation. Hence, for a given class of kinematically admissible velocity fields, the velocity field that minimizes the rate of total energy is the lowest upper bound, and therefore is nearest the actual solution. Here, the kinematically admissible velocity field is used to denote a velocity field that satisfies the incompressibility requirement for a rigid-plastic material and the prescribed velocity boundary conditions. However, the velocity field may be discontinuous on a finite number of imaginary internal surfaces. The rate of total energy generally consists of three terms such that:

$$\dot{E}_T = \int_V \bar{\sigma}\,\dot{\bar{\varepsilon}}\, dV + \int_{S_D} k|\Delta v_t|\, dS + \int_{S_F} mk|\Delta v_s| dS \quad \text{(Eq 11)}$$

where $\bar{\sigma}$ and $\dot{\bar{\varepsilon}}$ are the equivalent stress and the equivalent strain rate, respectively; Δv_t is the magnitude of velocity discontinuity tangent to the velocity discontinuity surfaces S_D; and Δv_s is the magnitude of sliding velocity on the contact surface S_F.

Each term in the right-hand side of Eq 11 represents the rate of plastic deformation energy, the rate of energy dissipation associated with internal velocity discontinuity, and the rate of energy dissipation due to friction between the tool and workpiece, respectively. The second term, also known as the *jump condition,* can be omitted when a class of continuous velocity fields is considered. Among various classical solution methods, the upper-bound method has been applied most extensively to various two-dimensional or three-dimensional forming problems because it delivers a fast and accurate solution as long as the trial velocity field can be provided closer to the actual velocity field. However, it is not easy to choose a good trial velocity field using a combination of analytic functions for geometrically complicated problems. In order to relax in such a difficulty, the upper-bound elemental technique (UBET), based on the concept of a "unit rectangular deforming region" (Ref 10) has been developed and applied to rather complex forming problems and preform design applications.

In order to compare the characteristics of solutions with different solution methods, i.e., the upper-bound and slab methods, the ring compression problem illustrated in Fig. 1 is employed as an example here. Also, a simple form of the trial velocity field is chosen so that an explicit form of solution can be obtained. The trial velocity field used in the present example has the form (Ref 11):

$$u(r) = \frac{D}{2}\left(r - \frac{r_n^2}{r}\right)$$
$$v = 0$$
$$w(z) = -Dz \tag{Eq 12}$$

where $D = \frac{V_D}{H}$, V_D is the die velocity, H is the height of ring, and r_n is the radius at a neutral point. It is noted that this velocity field contains only one unknown, i.e., r_n, and it is adequate for investigating the effect of the location of neutral point on the forming load. The neutral radius r_n is defined by:

$$u = 0 \quad \text{at} \quad r = r_n$$
$$u < 0 \quad \text{for} \quad r < r_n$$
$$u > 0 \quad \text{for} \quad r > r_n \tag{Eq 13}$$

Because the radial component of velocity, u, is a function of r only, the present trial velocity field cannot reproduce the so-called barreling or bulging phenomenon due to friction.

All nonzero strain-rate components are then written as:

$$\dot{\varepsilon}_r = \frac{D}{2}\left(1 + \frac{r_n^2}{r^2}\right)$$
$$\dot{\varepsilon}_\theta = \frac{D}{2}\left(1 - \frac{r_n^2}{r^2}\right)$$
$$\dot{\varepsilon}_z = -D \tag{Eq 14}$$

It is noted that the velocity field satisfies the incompressibility condition such that:

$$\dot{\varepsilon}_r + \dot{\varepsilon}_\theta + \dot{\varepsilon}_z = 0 \tag{Eq 15}$$

Since the present velocity field satisfies the continuity requirement, the incompressibility condition, and the prescribed velocity boundary condition, it is proved to belong to a class of kinematically admissible velocity fields. It is also noted that the second term in Eq 11 can be omitted because there is no internal velocity discontinuity.

The rate of total energy in Eq 11 can be rewritten as:

$$\dot{E}_T = \int_V \bar{\sigma}\dot{\bar{\varepsilon}}\, dV + \int_{S_F} mk|\Delta v_s|\, dS$$
$$= \int_{R_i}^{R_o} \bar{\sigma}\dot{\bar{\varepsilon}}(2\pi r H)\, dr + 2\int_{R_i}^{R_o} mk|u|(2\pi r)\, dr \tag{Eq 16}$$

After substituting the following expression of the equivalent strain-rate:

$$\dot{\bar{\varepsilon}} = D\sqrt{1 + \frac{1}{3}\left(\frac{r_n}{r}\right)^4} \tag{Eq 17}$$

and the radial velocity u defined in Eq 12 into Eq 16, the rate of total energy $\dot{E}_T$ can be integrated explicitly as:

$$\frac{\dot{E}_T}{2\pi k V_D} = \frac{1}{2}\left(\sqrt{r_n^4 + 3R_o^4} - \sqrt{r_n^4 + 3R_i^4}\right)$$
$$- \frac{r_n^2}{2}\left\{\log\left(\frac{r_n^2 + \sqrt{r_n^4 + 3R_o^4}}{r_n^2 + \sqrt{r_n^4 + 3R_i^4}}\right) - \log\left(\frac{R_o}{R_i}\right)^2\right\}$$
$$+ \frac{m}{H}\left\{\frac{4}{3}r_n^3 - (R_i + R_o)r_n^2 + \frac{1}{3}\left(R_i^3 + R_o^3\right)\right\} \tag{Eq 18}$$

The unknown coefficient, i.e., the neutral radius r_n, is then determined so that the rate of total energy attains to its minimum value such that:

$$\frac{d\dot{E}_T}{dr_n} = 0 \tag{Eq 19}$$

After some manipulation, the neutral radius can be written as:

$$r_n = \frac{1}{2}(R_i + R_o) + \frac{1}{4}\frac{H}{m}\log\left(\frac{\beta_o^2 + \sqrt{3 + \beta_o^4}}{\beta_i^2 + \sqrt{3 + \beta_i^4}}\right) \tag{Eq 20}$$

where $\beta_i = \frac{r_n}{R_i}$ and $\beta_o = \frac{r_n}{R_o}$. Equation 20 can then be solved for r_n numerically with specific values of the ring geometry, i.e., R_i, R_o, and H, and the shear friction factor m.

The forming load can be evaluated directly by substituting the value of neutral radius into Eq 18 and with use of the relation:

$$P = \frac{\dot{E}_T}{V_D} \tag{Eq 21}$$

The correction factor or normalized forming load K_{UBM} can be defined by:

$$K_{UBM} = \frac{P}{\bar{\sigma} A} \tag{Eq 22}$$

Finite Element Method

Metal forming simulation is often classified as a class of highly nonlinear continuum mechanics problems because it is accompanied by large deformation (geometric nonlinearity), nonlinear materials behavior (material nonlinearity in both deformation and temperature), and frictional contact (nonlinear boundary condition). Starting from the mid-1980s, the FEM has shown great success in the axisymmetric applications such as disk forging, cold forging of cylindrical fasteners, and so forth. The approximation of a two-dimensional cross section in a three-dimensional part, using the plane-strain assumption, is an alternative to achieving some understanding of the three-dimensional forming process. Although a considerable amount of research has been done in developing the FEM for metal forming simulation since the pioneering work (Ref 1) was presented in 1973, rigorous three-dimensional simulation of metal forming problems still remains a challenging task from the standpoint of computational efficiency, solution accuracy, graphics visualization, mesh generation and automatic remeshing, and so on. As computer technology and FEM advance, wider and more complicated metal forming processes are being investigated. It is believed that the further development of FEM will be continuously challenged by the need from the industry to make the modeling more accurate, more practical, and more affordable.

Since the two-dimensional FEM implementation has been discussed elsewhere in detail (Ref 12), this section focuses on the three-dimensional implementation.

Preliminary Assumptions

In order to narrow down the discussion to the most practical applications among a variety of metal forming simulations, the following preliminary assumptions are first introduced.

Quasi-Static Analysis. In most metal forming processes, dynamic effects can be neglected except for high strain-rate processes in which a realistic deformation mode cannot be obtained without considering the effect of stress wave propagation and in which the magnitude of kinetic energy is comparable to that of deformation energy. In cases where dynamic effects can be neglected, the progress of deformation is analyzed in a manner such that every instantaneous state of a body in the course of deformation is satisfied with the equilibrium conditions. Such a method of analysis is generally called the *quasi-static analysis.*

It is noteworthy that even for forming processes between medium and low strain-rate ranges, the explicit method of solution originally designed for dynamic analyses is sometimes used for the sake of computational efficiency. For the explicit method, however, the size of time step must be small enough, typically the order of microseconds, to satisfy the stability criterion of the explicit time marching scheme. It means that the computational efficiency may not be achieved because the number of solution steps increases so greatly compared with that of the implicit method. Thus, the so-called mass scaling technique (Ref 13), which may be interpreted as the introduction of artificial inertia term, is usually employed in order to increase the size of time step. In these cases, the mass scaling factor must be selected carefully to get the solution with reasonable accuracy and simulation time. In sheet forming applications, the explicit method has been very successful.

Rigid-plastic analysis is more advantageous for computational efficiency and robustness than elasto-plastic analysis is. This method has been used predominantly for the majority of bulk forming processes where the elastic deformation is negligible compared with the plastic deformation and the distribution of residual stresses is not of major concern. By neglecting the elastic portion of deformation, the rigid-plastic formulation (Ref 14) turns out to be very similar to that of fluid flow problems except for the presence of yielding and so it is sometimes called the *flow formulation.* The velocity field satisfying the equilibrium equations, constitutive equations, and boundary conditions instantaneously is obtained at each state in the course of deformation. Therefore, it is necessary to adopt an appropriate scheme for updating deformed configurations from the velocity field obtained at each state.

Unlike the elasto-plastic formulation, the rigid-plastic formulation does not have any ambiguity related to the choice of the objective stress rates and the decomposition of deformation gradient into the elastic and plastic parts (Ref 15, 16).

Updated Lagrangian (UL) Formulation. There are several different formulations (Ref 17, 18) for continuum mechanics problems with large deformation and/or large rotation, e.g., total Lagrangian (TL), updated Lagrangian (UL), Eulerian, and arbitrary Lagrangian Eulerian (ALE) methods. In the first two methods, the motion of a continuum is described in terms of the coordinates of a material particle in an arbitrarily chosen reference configuration. The TL method uses the initial undeformed configuration as a fixed reference configuration, but the UL method uses the most updated configuration during the progress of deformation as the reference configuration. Both methods are amenable to solid mechanics problems in which the configuration of a boundary is not fixed in space and changes in the course of deformation. Compared with the TL method, the UL method can use the simplified kinematics of a continuum assuming that the reference configuration is updated continuously within a small amount of increment of deformation according to the desired degree of solution accuracy. Thus, the UL method has been used widely in most metal forming simulation. It is worthwhile to mention that this method generally requires frequent remeshing (or rezoning) when mesh is severely distorted in the process of large deformation. Furthermore, it may not be an effective method for a certain class of forming problems, such as extrusion, rolling, and machining, in which the steady state solution is of major concern.

In the Eulerian method, the motion of a continuum is described in terms of the spatial coordinates of a material particle in the current configuration. Therefore, this method is particularly suitable for well-constrained fluid flow problems in which the domain of interest is fixed in the space. However, there are inherent difficulties in applying this method to the moving boundary or free surface problems. The ALE method has been developed to remove these difficulties by introducing the concept of mesh motion independent of the motion of material particle. The motion of a mesh can be chosen arbitrarily, but it must preserve the boundary of a continuum in the course of deformation. When the motion of mesh is set the same as the motion of material, it becomes the same as the Lagrangian methods. Several different schemes regarding the selection of mesh motion have been proposed according to the specific area of applications. The ALE method has been considered a prominent method, especially for the steady state metal forming applications. However, it has disadvantages, such as the increase of problem size due to additional variables for defining the mesh motion and the requirement of additional computation due to the evaluation of convection terms in updating the state variables.

Description of the Problem

The governing partial differential equations and the associated boundary conditions for the rigid-plastic and rigid-viscoplastic problems can be written as (Ref 12):

$$\sigma_{ij,j} = \sigma'_{ij,j} + \sigma_{m,i} = 0 \qquad \text{in } \Omega \tag{Eq 23}$$

$$u_{i,i} = 0 \qquad \text{in } \Omega \tag{Eq 24}$$

$$\sigma'_{ij} = \frac{2}{3}\frac{\bar{\sigma}}{\dot{\bar{\varepsilon}}}\dot{\varepsilon}_{ij} \qquad \text{in } \Omega \tag{Eq 25}$$

$$\dot{\varepsilon}_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i}) \qquad \text{in } \Omega \tag{Eq 26}$$

$$\bar{\sigma} = \bar{\sigma}(\bar{\varepsilon}, \dot{\bar{\varepsilon}}, T) \qquad \text{in } \Omega \tag{Eq 27}$$

$$u_i = \hat{u}_i \qquad \text{on } \Gamma_u \tag{Eq 28}$$

$$\sigma_{ij}\mathbf{n}_j = \hat{t}_i \qquad \text{on } \Gamma_t \tag{Eq 29}$$

$$\text{Friction \& contact conditions on } \Gamma_c \tag{Eq 30}$$

Here, the open domain Ω and its associated boundary Γ represent the current configuration of a body according to the UL formalism. The subscripts u, t, and c in Γ are used to denote three different types of the boundary associated with the boundary conditions: the prescribed velocity, the prescribed traction, and the frictional contact conditions, respectively. Indices i, j, and k are used to denote the components of a tensor and a comma denotes the spatial derivative with respect to the current configuration.

In Eq 23, i.e., the equilibrium equation, σ_{ij}, is the stress tensor, and σ'_{ij} and σ_m denote its deviatoric and volumetric components, respectively, such that:

$$\sigma'_{ij} = \sigma_{ij} - \sigma_m\delta_{ij} \tag{Eq 31}$$

and

$$\sigma_m = \tfrac{1}{3}\sigma_{kk} \tag{Eq 32}$$

where δ_{ij} is the Kronecker delta, and the repeated index denotes the summation. It is noted that the stress measure in Eq 23 must be the first kind of Piola-Kirchhoff stress tensor according to the framework of the referential description of continuum mechanics. However, it can be assumed that the reference configuration is updated as frequently as required, so it is therefore not distinguished from the Cauchy stress tensor.

Equation 24 represents the incompressibility condition where u_i is the velocity vector. Equation 25 represents the constitutive equation based on the so-called J_2 flow rule. Here, $\dot{\varepsilon}_{ij}$ is the strain-rate tensor or the symmetric part of the velocity gradient tensor as defined in Equation 26. $\bar{\sigma}$ and $\dot{\bar{\varepsilon}}$ are the effective stress and the effective strain-rate, respectively, defined by:

$$\bar{\sigma} = \left[\tfrac{3}{2}\sigma'_{ij}\sigma'_{ij}\right]^{\frac{1}{2}} \quad \text{and} \quad \dot{\bar{\varepsilon}} = \left[\tfrac{2}{3}\dot{\varepsilon}_{ij}\right]^{\frac{1}{2}} \tag{Eq 33}$$

Equation 27 represents an implicit form of the yield criterion as a function of the effective strain $\bar{\varepsilon}$, the effective strain rate $\dot{\bar{\varepsilon}}$, and the temperature T. Provided the increment of time between any two adjacent referential configurations is small enough, the effective strain can be evaluated approximately such that:

$$\bar{\varepsilon}_{new} = \bar{\varepsilon}_{old} + \dot{\bar{\varepsilon}}\Delta t \tag{Eq 34}$$

where Δt denotes the increment of time. Equation 27 represents the most popular form of the rigid-plastic constitutive equations, but more complicated forms of constitutive equations can, of course, be used within this structure of formulation such as the constitutive equations having internal state variables with the associated evolution equations.

Equations 28 and 29 are the prescribed velocity and traction boundary conditions, respectively. The hat symbol on u_i and t_i is used to denote prescribed values, and the vector $\mathbf{n}_i$ denotes the outward normal vector to the body at a point on the boundary. The friction and contact conditions are discussed in detail in the subsequent section.

Friction and Contact Conditions

The friction and contact conditions in this section are described between any two boundaries whether they belong to the same body (i.e., self-contact), rigid and deformable bodies, or two or more deformable bodies. Although several different methods (Ref 19–21) have been presented in describing the macroscopic contact and friction phenomena within the framework of continuum mechanics, the method of pointwise description of friction and contact conditions has been used in most practical applications. The contact condition and friction can be summarized as:

- *Contact condition (non-penetration condition):*
  a. Any material particle of a given body cannot penetrate into another
  b. The normal component of contact traction must be compressive for each body

c. A pair of contact points can separate only when the contact traction vanishes or becomes tensile

- *Friction condition:*
  a. The magnitude of the tangential component of contact traction must be less than or equal to that of the normal component multiplied by a coefficient of friction.
  b. The instantaneous relative motion in the tangential direction for a pair of contact points can take place when the equality in (a) above holds.
  c. The tangential relative motion must be along the same line as the tangential component of contact traction but in the opposite direction.

In most metal forming applications, two different types of friction laws have been widely used: the Coulomb's law of friction and the shear friction law. Statement (a) under "Friction condition" represents the Coulomb's law of friction, but it also represents the shear friction law by replacing the normal component of contact traction by the shear yield stress of the weaker material. Also, the coefficient of friction is generally referred to as the constant factor in the shear friction law.

This statement of contact and friction conditions can be expressed in a mathematical form as:

$$\dot{z}_i = \dot{\lambda}_n I(\mathbf{p}_n) n_i + \lambda_t I(f) \frac{(\mathbf{p}_t)_i}{|(\mathbf{p}_t)_i|} \quad \text{(Eq 35)}$$

where $\dot{z}$ is the relative velocity between two points in contact at an instant as:

$$\dot{z}_i = u_i^a - u_i^b \quad \text{(Eq 36)}$$

where the superscripts a and b denote the corresponding contact surfaces chosen arbitrarily, $\mathbf{n}_i$ is the outward normal vector with respect to the contact surface "*a*," **p** is the contact traction vector on the contact surface "*a*" as:

$$\mathbf{p}_i = \mathbf{p}_i^a = -\mathbf{p}_i^b \quad \text{(Eq 37)}$$

where the subscripts *n* and *t* denote the normal and tangential components, respectively, as:

$$\dot{z}_i = \dot{z}_n n_i + (\dot{z}_t)_i \quad \text{and} \quad \mathbf{p}_i = \mathbf{p}_n n_i + (\mathbf{p}_t)_i \quad \text{(Eq 38)}$$

Continuing with the description of terms in Eq 35, the symbol $|\cdot|$ denotes the absolute magnitude of a vector, and $\dot{\lambda}_n$ and $\dot{\lambda}_t$ are non-positive constants. $I(g)$ is an indicator defined for any scalar-valued function g as:

$$I(g) = \begin{cases} 1, & \text{if } g = 0 \\ 0, & \text{if } g \neq 0 \end{cases} \quad \text{(Eq 39)}$$

In order to describe the friction laws effectively, a slip function *f* is introduced for each pair of contact points similar to the yield function in the classical theory of plasticity. The slip function is defined by:

$$f(\mathbf{p}_i) \equiv |(\mathbf{p}_t)_i| + \mu \mathbf{p}_n \quad \text{(Eq 40)}$$

and

$$f(\mathbf{p}_i) \equiv |(\mathbf{p}_t)_i| - mk \quad \text{(Eq 41)}$$

corresponding to the Coulomb's law of friction and the shear friction law, respectively, where μ is the coefficient of friction, *k* is the shear yield stress of the weaker material, and *m* is a constant friction factor.

It is noted that the frictional dissipation functional associated with the friction law in Eq 38 or 39 has a non-differentiable form with respect to its primary variables, i.e., the velocities of each body. In the area of metal forming simulation, the following form of regularized friction laws has been widely used. For any $(\dot{z}_t)_i$:

$$(\mathbf{p}_t)_i = \mu\, \mathbf{p}_n \left\{ \frac{2}{\pi} \tan^{-1}\left(\frac{|(\dot{z}_t)_i|}{\dot{z}_o} \right) \right\} \frac{(\dot{z}_t)_i}{|(\dot{z}_t)_i|} \quad \text{(Eq 42)}$$

and

$$(\mathbf{p}_t)_i = -m\, k \left\{ \frac{2}{\pi} \tan^{-1}\left(\frac{|(\dot{z}_t)_i|}{\dot{z}_o} \right) \right\} \frac{(\dot{z}_t)_i}{|(\dot{z}_t)_i|} \quad \text{(Eq 43)}$$

where $\dot{z}_o$ is a positive constant. These regularized friction models approach the original friction laws asymptotically as $\dot{z}_o$ approaches zero. However, a very small value of $\dot{z}_o$ can make a difficulty in convergence, and a larger value may give a solution deviated from the original friction laws.

Mixed Variational Formulation

Boundary value problems (BVP) in continuum mechanics can be expressed in two different ways: the partial differential equations (PDE) with the associated boundary conditions and the variational equations with the appropriate function space. The solutions of these two different forms of BVP are referred to as the *strong solution* and the *weak solution,* respectively. The term *weak* is used in the sense that the requirement of continuity (differentiability) of solution is weakened in the variational form of BVP. If a strong solution requires the existence of a second derivative, the corresponding weak solution requires only the existence of a first derivative in the sense of distribution; that is, the first derivative needs to be continuous within each of a finite number of subdomains but not necessarily across the interboundary between subdomains.

The variational form of BVP can be obtained only when the quadratic form of functional exists so that the set of Euler equations, obtained by the vanishing of the first variation of the functional, is identical to the original PDE. However, the same form can be obtained by using the so-called weak formulation or the principle of virtual work, although certain mathematical features of variational problems with the quadratic functional such as the existence and uniqueness of solutions and the stability and accuracy of finite element solutions cannot be stated. Here, the weak formulation is used to accommodate a broad class of plastic constitutive models to the same framework of formulation.

The constraint conditions such as the incompressibility condition and the contact condition can generally be incorporated into the variational formulation by using one of two techniques: the penalty method or the Lagrange multiplier method. The penalty method has the advantage of simple implementation, but it has a drawback such that it can result in an overconstrained problem or an underconstrained problem depending on the choice of the penalty parameter. The overconstraint means the volumetric locking or the locking of contact surfaces, and the underconstraint means the inaccuracy of solution in the sense of incompressibility or nonpenetration. On the other hand, the Lagrange multiplier method can avoid the drawback of the penalty method, but it has a disadvantage concerning the increase of problem size because the Lagrange multipliers are treated as additional solution variables such as the velocity of material particles. The Lagrange multiplier can be interpreted as the hydrostatic stress for the incompressibility constraint and the normal contact traction for the nonpenetration condition.

In areas of metal forming simulation, it is popular to use the Lagrange multiplier method for the incompressibility condition and the penalty method for the contact condition. However, the penalty method has also been used successfully with a certain class of finite elements with the selective reduced integration scheme.

Finite Element Formulation

The finite element method can be distinguished from other approximate methods by the way it constructs the trial solution (e.g., kinematically admissible velocity field) with a finite number of piecewise continuous trial functions (polynomial functions in most applications) (Ref 22–24).

For a particular class of metal forming processes, it is always important to select an appropriate type of element (i.e., the order of polynomials, the geometric shape, the rule of numerical integration, and so on). Some issues related to the selection of element are discussed subsequently. From the standpoint of polynomial order, linear elements are generally preferred to quadratic or higher-order elements for most metal forming applications in which the friction and contact conditions are always present. With use of the friction and contact conditions described previously, the so-called node-to-segment contact situation cannot be avoided because these constraints need to be imposed point-wise. The node-to-segment contact can be treated in a simpler manner with linear elements rather than higher-order elements. Also, linear elements are generally easier to use without a-priori knowledge of solution than higher-order elements with the same degrees of freedom for a given problem.

Next, from the standpoint of geometric shape, there are two different kinds of linear elements: triangular or quadrilateral for two-dimensional elements and tetrahedral or hexahedral for three-dimensional elements. Shape functions for triangular and tetrahedral elements contain polynomial terms such as $(1, \xi, \eta)$ and $(1, \xi, \eta, \zeta)$, respectively. On the other hand, those for quadrilateral and hexahedral elements contain polynomial terms such as $(1, \xi, \eta, \xi\eta)$ and $(1, \xi, \eta, \zeta, \xi\eta, \eta\zeta, \zeta\xi, \xi\eta\zeta)$, respectively. Triangular and tetrahedral elements are known as *constant stress/strain elements* (CST) because all derivatives with respect to any component of coordinates vanish. In the rigid-plastic formulation, vanishing derivatives imply constant strain rates and, therefore, constant stresses, within an element. On the other hand, the quadrilateral and hexahedral elements have polynomial terms in the product form of coordinates, and thus, the velocity gradient in one component of coordinates is linear with respect to the other components of coordinates. Strain rates and stresses are linear within an element accordingly. Both kinds of linear elements have been used widely by weighing the pros and cons for a particular application.

Generally speaking, triangular and tetrahedral elements have more flexibility in filling meshes into any complicated shape than quadrilateral and hexahedral elements. It is noted that triangular and tetrahedral elements in the group of CST elements must be distinguished from the so-called degenerated elements whose strain rates and stresses are not constant within an element.

Degenerated elements can be obtained from coalescing adjacent nodes into the same node, for instance, mapping from the four-node quadrilateral parent element into a three-node triangular shape or mapping from the eight-node hexahedral parent element into a four-node tetrahedral shape. The performance of degenerated elements is not as good as that of the CST elements or the original quadrilateral and hexahedral elements. Also, a numerical integration scheme for the CST elements in evaluating the stiffness matrix is unnecessary because they have constant values of strain rate and stresses within an element. In other words, the integration is always exact since the integrand is constant with respect to coordinates. On the other hand, for quadrilateral and hexahedral elements, a numerical integration rule based mostly on the Gaussian quadrature formula is required in evaluating the stiffness matrix. The numerical integration is exact when a real element keeps the same shape as the parent element, i.e., a rectangular shape or a brick shape, although it is not possible except for problems with a very simple geometry. In such cases, the determinant of the Jacobian matrix mapping between the domains of the parent element and a real element is constant. Errors originated from the numerical integration increase as the shape of element is apart from that of the parent element.

Some disadvantages of CST elements are listed as follows. The CST elements are apt to have directionality in the mesh topology and thus in the solution because the mesh topology directly reflects on the form of trial solutions. Here, the directionality implies that triangular and tetrahedral elements are globally biased in a certain direction. Also, the CST elements may show more tendency of volumetric locking due to the incompressibility constraint condition than quadrilateral and hexahedral elements. The term *volumetric locking* means that each element endowed with a linear velocity field cannot deform properly when the same degree of satisfaction is required elementwise for both the deformation and the constraint condition. In order to avoid such a locking phenomenon, the degree of satisfaction for the constraint condition needs to be relaxed in a certain manner such as with the selective reduced integration scheme or with the mixed formulation as described previously. Another important aspect to be considered in determining the element shape, especially for large deformation problems, may be whether an efficient and robust method for automatic remeshing is available with the shape of element. In areas of metal forming simulation, quadrilateral elements are preferred to triangular elements for the two-dimensional analysis, but tetrahedral elements have been used more than hexahedral elements for the three-dimensional analysis because of the versatility in automatic remeshing.

As mentioned previously, linear tetrahedral elements cannot be used for metal forming simulation without an appropriate resolution of the volumetric locking problem. The most common approach to enforce incompressibility includes the penalty method and the Lagrange multiplier as discussed previously. These methods, however, are limited to the quadrilateral element in two-dimensional and hexahedral element in three-dimensional. Furthermore, the resulting matrix is ill conditioned for a conjugate-gradient solver due to large penalty values for incompressibility. The mixed formulation uses the Lagrange multiplier method to achieve incompressibility condition. Generally speaking, the order of polynomials for the velocity shape function needs to be higher than that for the mean stress shape function because the mean stress is originally a duality variable paired with the volumetric strain rate that is the trace of the velocity gradient. For instance, when the linear shape function is used for both variables, the volumetric strain rate becomes constant within an element, but the mean stress is linear. It is obvious that a linear function cannot be matched with a constant value in a general case and a mesh system may tend to lock as a result. When the quadratic shape function is used for the velocity in order to make the well-posed problem, a considerable increase of problem size cannot be avoided.

The MINI element is effectively designed for the three-dimensional simulation of incompressible problems such as fluid flow, rubber elasticity, and rigid-plasticity problems under the framework of the mixed formulation. In the MINI element, the shape function for the velocity is enriched with bubble terms associated with an additional bubble node, although the shape function for the mean stress remains as a linear function interpolated with the values at four vertex nodes. The bubble node is located at the centroid of tetrahedron and has only the velocity degrees of freedom. Here, the word *bubble* means that its value always vanishes along the boundary of the element. This element does not have the volumetric locking problem. Moreover, the total number of equations can be maintained as the same as that with the tetrahedral elements because the velocity components at the bubble nodes can be eliminated at the element level by means of the static condensation.

Although the bubble node is introduced mainly for a systematic stabilization of the ill-posed finite element system, a considerable amount of computational effort is additionally required for the static condensation and recovery of the velocity of the bubble node. However, such additional efforts can be minimized with a few modifications of the standard MINI element. In fact, it turns out that the derivation of this modified MINI element can be conceived as a systematic stabilization of the standard mixed tetrahedral element because the effect of the bubble node appears on only the block diagonal terms associated with the mean stress.

Heat Transfer

In the metal forming industry, heating is frequently used to increase material workability and control forming loads. During the forming process, heat is generated from plastic and friction work. Heat is also lost through contact with colder dies and through convection and radiation with the environment. Since materials properties vary considerably with temperature, accurate temperature prediction is required. For convenience, the deformation analysis and thermal analysis can be loosely coupled in such a way that plastic work and friction are considered as heat source in the thermal analysis while the updated temperature field is used to determine the flow stress behavior during the deformation analysis.

The governing equation for heat transfer can be expressed as:

$$k_1 T_{ii} + \dot{q} - \rho C\dot{T} = 0 \qquad \text{(Eq 44)}$$

where ρ is the density, C is the specific heat capacity, T is the temperature, t is time, k is the thermal conductivity, and $\dot{q}$ is a heat generation term. Heat generation in metal forming is due to work of plastic deformation and friction. Heat generation due to plastic deformation is given by

$$\dot{q}_{pw} = \kappa \int \bar{\sigma}\,\dot{\bar{\varepsilon}}\,dV \qquad \text{(Eq 45)}$$

where κ is a deformation efficiency term, representing the fraction of the work of deformation converted to heat.

The boundary condition for the tool-workpiece contact surface includes friction heating and heat exchange via temperature difference of two objects.

$$\dot{q}_1 = \int_{S1} \eta f_{s1} |u_{s1}| dS_1 + \int_{S1} H \Delta T dS_1 \qquad \text{(Eq 46)}$$

where u_s is the sliding velocity, f_s is the friction stress, η is the percentage of friction energy absorbed by the object, H is the lubricant heat transfer coefficient, and ΔT is the temperature difference between two objects.

The boundary condition of the free surface includes convection heat and radiation heat from/to the environment:

$$\dot{q}_2 = \int_{S2} h_c (T - T_\infty) dS_2 + \int_{S2} \sigma\varepsilon (T^4 - T_\infty^4) dS_2 \qquad \text{(Eq 47)}$$

where h_c is the convection heat transfer coefficient, T_∞ is the environment temperature, σ is the Stefan-Boltzmann radiation constant, and ε is the emissivity of the surface. It is also noted that the inclusion of the view factor for radiation heat calculation (which is not addressed here) can improve the accuracy of the thermal model, especially for the hot forging condition.

By substituting Eq 45 into Eq 44 and introducing a small, arbitrary temperature variation δT, and applying the divergence theorem, Eq 44 can be written in the form:

$$\int_V k_1 \; T_i \; \delta T_j \; dV + \int_V \rho C \dot{T} \; \delta T \; dV + \int_V \kappa \sigma_{ij} \dot{\varepsilon}_{ij} \; \delta T \; dV - \int_{Sq} q_n \; \delta T \; dS = 0 \qquad \text{(Eq 48)}$$

where q_n is the heat flux across the boundary. It includes the convection heat and radiation heat to the environment for the free surface and friction heat and heat gain or loss to the contacting surface. The temperature distribution function can be expressed through nodal temperatures and shape functions (Ref 12).

After discretization, Eq 48 can be further expressed in matrix form (Ref 24) as:

$$\mathbf{C}\dot{\mathbf{T}} + \mathbf{K}_c\mathbf{T} = \mathbf{Q} \qquad \text{(Eq 49)}$$

where $\mathbf{C}$ is the heat capacity matrix, $\mathbf{K}_c$ is the heat conduction matrix, $\mathbf{T}$ is the node point temperature vector, and $\dot{\mathbf{T}}$ is a vector containing the time rate of change of temperature of node points. The heat flux vector, $\mathbf{Q}$, for metal forming simulations considers plastic work of deformation, heat generation due to sliding contact friction, and heat flux due to lubricant conduction, convection, and radiation.

Example Simulations

To demonstrate the capability of the modeling techniques, four examples are included in the following section. The first example is the ring compression, and three methods (slab, UBM, and FEM) are used to study the process. The second example is the cold forging of an electrode, the third example is the hot forging of a crankshaft, and the last example considers material cutting. Only FEM is used in the last three examples due to the complexity of the processes.

Ring Compression

The ring compression test is one of the methods used to determine the friction factor. The process is generally assumed to be two-dimensional or axisymmetric. Due to the simplicity of the workpiece and die geometry, ring compression is used as the first example to demonstrate and compare the capability of slab, UBM, and FEM.

Two-ring test FEM simulations with outer diameter (OD), inner diameter (ID) and height ratios of 6:3:2 and 6:3:0.5 were carried out with a constant friction factor of 0.4. A constant flow stress of 70 MPa (10 ksi) was used in the simulation. The predicted shapes using the FEM and the associated neutral line within the workpiece at different stages of deformation in both cases are shown in Fig. 2 and 3, respectively. In the case with the ratio 6:3:2, the neutral line within the workpiece is a function of height. In the case with 6:3:0.5, the neutral line remains to be a vertical line. Furthermore, the ID and OD surface bulged due to the friction force on the two contacting surfaces.

To simplify the mathematical derivation, the assumed velocity field (Eq 12, 13) in the UBM

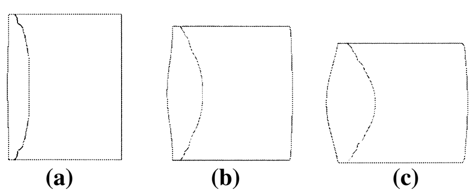


Fig. 2 The effect of friction on the location of the neutral point for a thick ring compression test. (a) 2%. (b) 10%. (c) 20%. The cross section of the ring is shown with the curved line denoting the neutral point over the height.

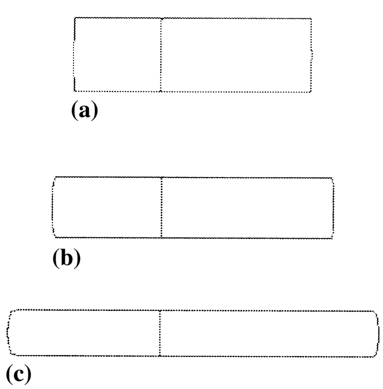


Fig. 3 The effect of friction on the location of the neutral point in a thin ring compression test. (a) 4%. (b) 20%. (c) 40%. The cross section of the ring is shown with the curved line denoting the neutral point over the height.

is based on a rectangular workpiece shape. The radial velocity is a function of the radius and the neutral radius, the axial velocity is a function of the height position, and the neutral radius is a function of friction factor and shear stress.

This simplification/assumption introduces error when the ID and OD surfaces of workpiece start to bulge during the compression. As the workpiece height increases, it is known that the unstable flow (or buckling mode) will gradually dominate the deformation. In such situations, the assumed UBM velocity field is moving away from the true velocity field. To demonstrate the buckling mode, another ring compression example with OD, ID, and height ratio of 6:3:6 was carried out. In this case, the flow stress is assumed to be in the form of $\sigma = 10\varepsilon^{0.1}$. A friction factor of 0.2 was used. The predicted FEM mesh and various deformation stages are shown in Fig. 4.

Figure 5 shows the effect of friction on the location of a neutral point with three different aspect ratios of a ring: i.e., 6:3:0.5, 6:3:1, and 6:3:2, in terms of the OD, the ID, and the H, in sequence. From this figure, it is noted that the upper-bound solutions approach to the mean radius of ring asymptotically as the magnitude of shear friction factor increases. It is because, except for point sticking, neither local nor global sticking phenomenon can be represented with the present velocity field. It is also shown that the upper-bound solutions correlate better with the FEM solutions as the height of ring reduces.

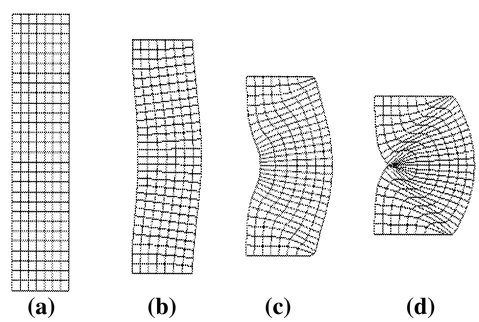


Fig. 4 A two-dimensional buckling simulation. (a) 1%. (b) 20%. (c) 35%. (d) 50%

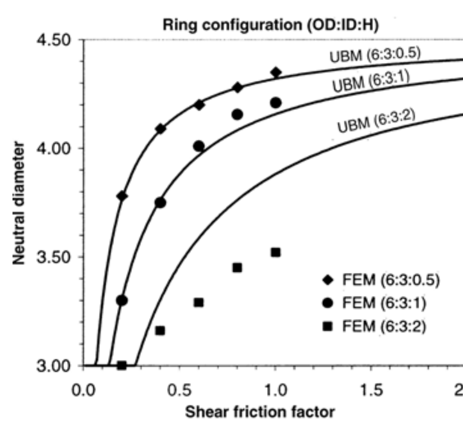


Fig. 5 The effect of friction on the location of the neutral point. UBM, upper-bound method; FEM, finite-element method

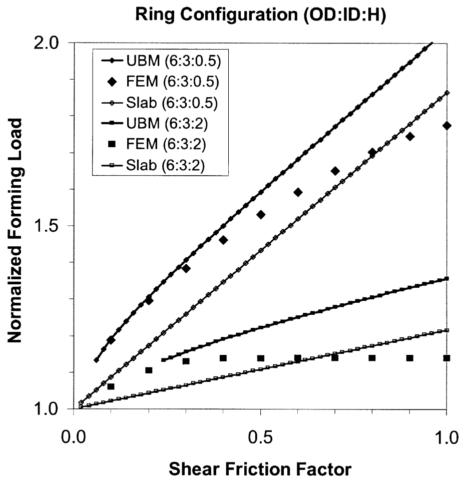


Fig. 6 The variation of forming load with the magnitude of friction factor. UBM, upper-bound method

The normalized forming loads obtained using the UBM are compared with those obtained using the slab method presented in Eq 10 as well as the corresponding FEM solutions. It is noteworthy that the forming load estimated using both the UBM and the FEM based on the velocity formulation can be used as a measure of the upper limit of the actual load. As shown in Fig. 6, the forming load estimated using the FEM is smaller than that using the UBM, and thus it can be said that the present UBM solution becomes less accurate than the FEM solution as the magnitude of friction factor increases. On the other hand, the load estimated using the slab method can be used as a measure of lower limit when the stress field used in the slab method satisfies all requirements of the kinetically admissible stress field: the equilibrium condition, the yield criterion, and the prescribed traction boundary condition. As shown in Fig. 6, however, the present slab method does not provide a lower-bound solution because the stress field does not take into account the shear stress due to friction, and hence, it violates the Cauchy stress law on the traction boundary.

Electrode Cold Forging Process

Folding, also known as *lapping,* is a severe forging defect that process designers would like to predict and avoid in the manufacturing process. There are many challenges to modeling a fold properly, especially in the three-dimensional UL method. The methodologies, to mesh a workpiece with very thin fold and to account for self-contact condition in the FEM formulation, are important for a robust and accurate FEM model.

This section demonstrates the modeling of the forming of a copper electrode (Ref 25). The actual progression is shown in Fig. 7. The workpiece undergoes a total of five operations: shearing, squaring, pancaking, backward extrusion, and final forming. The pancaking and backward extrusion operations were modeled in two dimensions, and the final forming was modeled in three dimensions. The shearing and squaring operations were not modeled in this study.

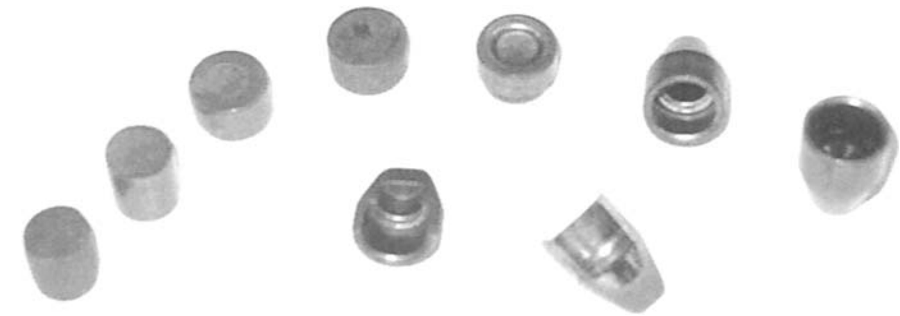


Fig. 7 Progression of electrode forging

Two laps were developed during the backward extrusion as shown in Fig. 8(a) and (b). One lap is seen on the outside wall (Fig. 8b) and extends the entire circumference. The other lap occurred only partially on the inside wall (Fig. 8a). It is believed that eccentricity of the forming process caused the internal lap. From visual inspection, the lap occurred in every tooth in the final part as shown in Fig. 8(c). Figure 8(c) also shows a change in the texture at the inner wall of the electrode, above the teeth. Figure 8(d) shows a lap that occurred at the bottom of the cavity, which is seen as a round pattern pointed to by the arrow. Due to the size of the part, it is very difficult to visualize the fold, other than the one on the tooth tip.

The strain distribution and predicted geometry at the end of each operation through extrusion are shown in Fig. 9. It is seen that the external fold is successfully predicted in the zoomed area of the figure. Due to the axisymmetric assumption, the internal fold is not seen in the simulation results. Further study, considering tooling eccentricity, will be carried out in the near future. Due to the symmetry condition, 1/12 of the workpiece was modeled in the simulation to reduce the simulation time.

The predicted geometry at early stage of the tooth forming is shown in Fig. 10. It is clearly seen that the material is pushed both upward and downward through the contact of the punch. Folds occurred on the wall above the teeth and below the teeth. It is also noted that the partially formed tooth is seen as a concave shape. As the punch moved farther down, the fold began to wrap around the top of the teeth (Fig. 11). Figure 12 shows that this material peeled down into the bottom of the cavity as shown in Fig. 8(d). Also, the fold above the teeth is shown to smear, which correlates well to Fig. 8(c) as the cause of the uneven texture. The final modeled part is seen in Fig. 13. As shown in the real part, the tooth tip is the last area to fill. A clear fold is predicted in the same area.

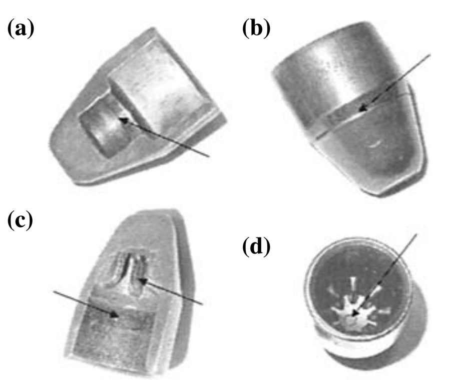


Fig. 8 Arrows highlight the defects seen in the forming of the electrode.

Fig. 9 Forming results after squaring and backward extrusion.

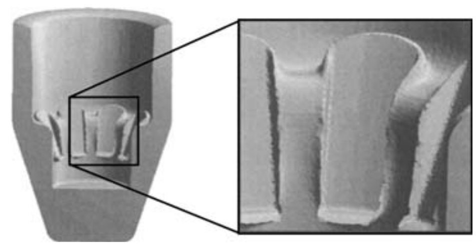


Fig. 10 Early stage of tooth forming. The material flow can be seen as driven up and down in forming the teeth.

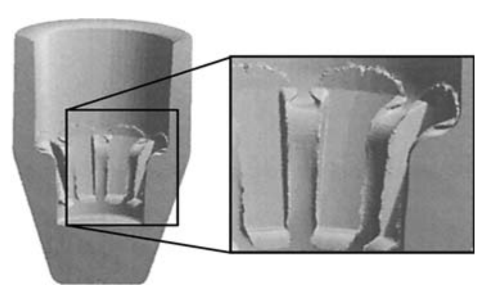


Fig. 11 As forming of the teeth continues, note the smearing of the material above the teeth.

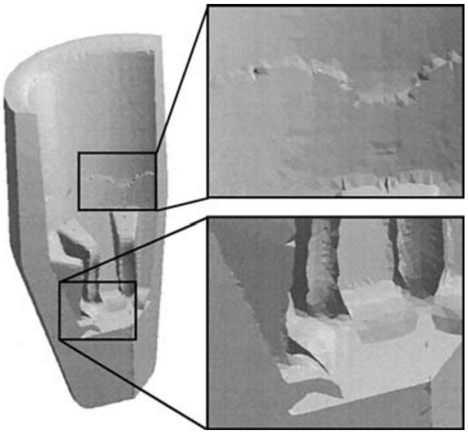


Fig. 12 The material peeling down as the teeth are formed

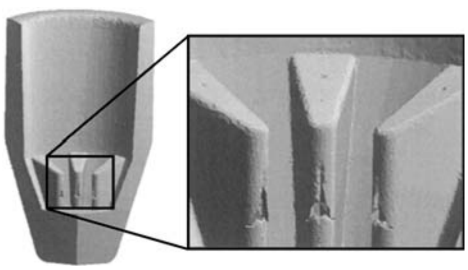


Fig. 13 The final part shape

Due to the small part size, it is very difficult to visualize the folds as well as understand the overall folding development. The simulation has, however, clearly revealed the complete history of the fold development.

Crankshaft Hot Forging Process

This forging process and the corresponding simulation (Ref 26) were carried out at the Hyundai Motor Company. The objective of the numerical simulation was to evaluate the formability and to predict the forging load for both a crankshaft and a connecting rod. In this section, only the crankshaft simulation is presented.

Specific interests included:

- The flow pattern and potential defects such as under fill or folding
- The forging load for tool stress analysis and press selection
- Stress and strain distribution for possible evaluation of microstructure
- Improvement and optimization of the existing forging processes

An accurate hot forging model involves many material variables and process variables. It requires proper characterization of the material such as flow stress behavior and thermal data as well as interface properties such as the friction factor and the lubricant heat transfer coefficient. In the case of a hot forging process, the thermal data will affect the accuracy of the temperature prediction. The deformation pattern is sensitive

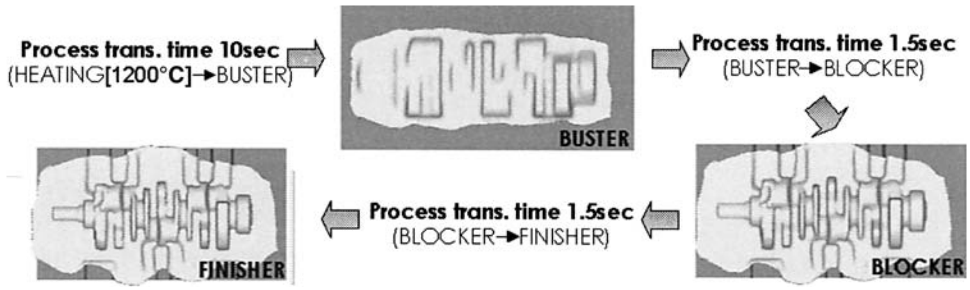


Fig. 14 Outline of forging process of crankshaft forging

to the temperature distribution inside the workpiece. Being able to describe the flow stress behavior within the process window is extremely important for an accurate simulation. Furthermore, the flow stress is generally sensitive to the strain rate at a hot forging condition; ram speed will also play an important role, not only to predict accurately in flow pattern, but also in the prediction of the load. Interface properties such as friction and heat transfer coefficient are also important variables that will influence the heat loss rate of the workpiece, material flow, and forging load. Therefore, an accurate model should take into account the coupling between the process variables, for example, ram speed, friction factor, heat transfer coefficient, and the material data (e.g., flow stress representation and thermal data).

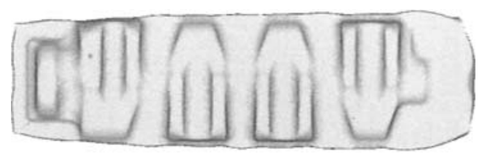
Toward this goal, a parametric study was first carried out to evaluate these critical process variables to assure the accuracy of the model. The variables under study included the flow stress, friction factor, heat transfer coefficient, and billet temperature.

The materials properties for AISI-1045 and AISI-1055 were selected for the simulation. There are three stages to forge the crankshaft: busting, blocking, and finishing.

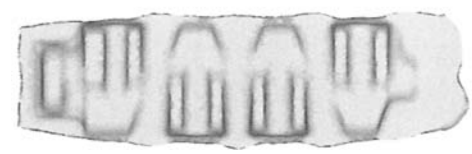
The workpiece is first heated to 1200°C (2192 °F). All dies are heated to 200°C (392 °F). As shown in Fig. 14, after 10 s of air transfer, the workpiece is placed in the buster dies for the first forging operation. After 1.5 s air transfer, the workpiece is removed from buster and placed in the blocker for the second forging process. After another 1.5 s air transfer, the workpiece is removed from blocker and placed in the finisher for the final forging process. A 6500 ton mechanical press is used in the forging and simulation. To accurately model the temperature evolution during the entire forming processes, the simulation was carried out in a nonisothermal manner, including heat transfer analysis to account for the air transfer time.

During the parametric study, it was found that an accurate lubricant heat transfer coefficient was very critical to obtain a similar flash profile as the actual part. The actual workpiece at the end of busting is shown in Fig. 15. The predicted shapes based on two lubricant heat transfer coefficients are given in Fig. 16. From these figures, it is noted that the waviness of the flash

Fig. 15 Crankshaft forging after buster stage



(a)



(b)

Fig. 16 Comparison of buster results for (a) high heat transfer and (b) low heat transfer

profile reduces as the heat transfer coefficient increases. This is due to the fact that heat loss from the die-workpiece interface is faster using a higher lubricant heat transfer coefficient. The difference in temperature distribution influences the material flow in the die cavity and flash.

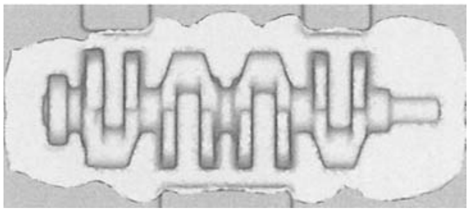
At the end of the blocker operation, the predicted shape and the actual shape are in excellent agreement as shown in Fig. 17. It is also noted that the folding defects (also known as *laps*) are successfully predicted by the simulation as shown in Fig. 18 and 19. The prediction of the lap shape and the propagation pattern are well correlated with the corresponding experimental results.

To validate the model further, the flash thickness and the corresponding forging load are measured and compared with the numerical prediction with excellent agreement as shown in Fig. 20.

Cutting Modeling

Machinability is of primary interest in the materials cutting process. It is affected by many

(a)



(b)

Fig. 17 Comparison of (a) actual to (b) simulated blocker result

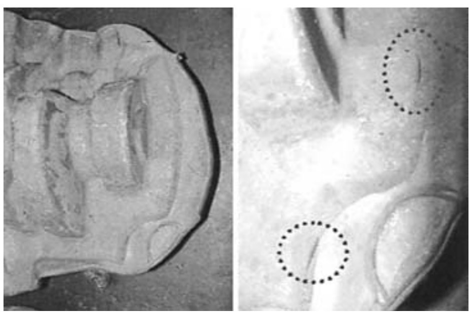


Fig. 18 Actual lap on crankshaft forging

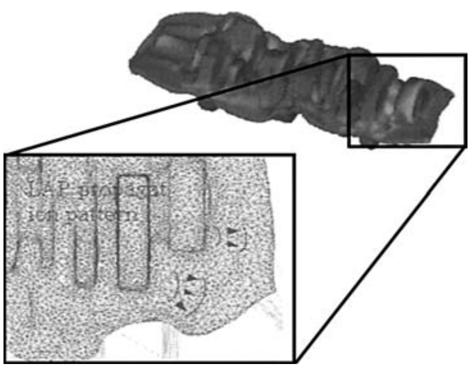


Fig. 19 Simulated lap on crankshaft forging

Stage			Simulation		Measured	
LOAD (TON)	BUSTER	FLASH thickness(mm)	2,200	10.0	2,100	9.9
	BLOCKER	FLASH (mm)	2,850	7.3	3,100	7.4
	FINISHER	FLASH (mm)	3,500	6.2	3,400	6.4

Fig. 20 Results of actual and simulated forming loads

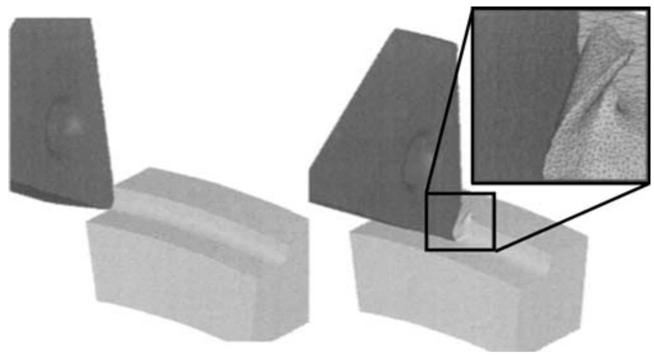


Fig. 21 Initial and final stages of Lagrangian analysis of chip formation

factors, among which are materials behaviors, insert shape, and cutting condition. A good understanding of the interactions among the chip flow, heat generation, residual stress, tool stress, and tool wear is crucial in order to optimize the design of the process and tooling (Ref 27).

To understand the thermal-mechanical response on an insert (tool wear) during a long period of cutting, the conventional updated Lagrangian transient approach, especially for three dimensional analysis, is not efficient since it requires enormous CPU time to reach a steady state. A hybrid procedure using both UL formulation for transient analysis and Eulerian formulation for steady state has been developed (Ref 27). In this hybrid procedure, the transient UL method is first used to predict the initial chip formation as shown in Fig. 21. The solution at the end of the transient run is then used as an initial guess for the steady state analysis.

In most solid mechanics problems, the domain, i.e., the steady state configuration of a deformable body, is unknown a priori and must be determined as a part of the solution. From the standpoint of the FEM, each node has degrees of freedom for both the velocity and the position. Thus, the velocity field needs to be satisfied for the equilibrium equations, the constitutive equations and the boundary conditions, and the configuration needs to be satisfied for the free surface condition and the contact condition. Here, the free surface condition implies that the velocity on the free surface must be tangent to the surface and there is no traction on the surface.

A great number of different solution methods have been presented, including ALE methods, pseudo-solid domain mapping methods, adaptive h-p FE methods, streamline tracking methods, and methods of splines. In this current approach, two sets of coupled governing equations are solved for both the velocity and the position iteratively. First, the Eulerian velocity solution is obtained based on a given configuration. Next, the new configuration, i.e., the nodal position, is updated so that it satisfies the free surface condition and contact condition with the given velocity field. The first part is based on the standard flow formulation of plastic materials. For the sake of computational effectiveness, the second part is divided into two levels—that is, the determination of the position of surface nodes and internal nodes.

Employing surface elements that have both membrane and edge bending stiffness then solves the position of surface nodes. The free surface condition is also embedded into the elemental equations so that each free surface element is positioned to be parallel to the velocity at the center of element. Since the resulting set of equations is nonlinear due to the geometric nonlinearity, the Newton-Raphson iterative method is used. Once the position of surface nodes is obtained, employing incompressible elastic solid elements solves for the position of internal nodes.

State variables such as stresses and strains are updated in two different ways: simple interpolation scheme of existing solutions and streamline tracking scheme, depending on whether the updated Lagrangian solutions are available. A result of steady state thermomechanical machining simulation is presented in Fig. 22. The steady state insert temperature is shown in Fig. 23. The work to validate the cutting force, temperature, and wear using the model is being carried out (Ref 27 and 28).

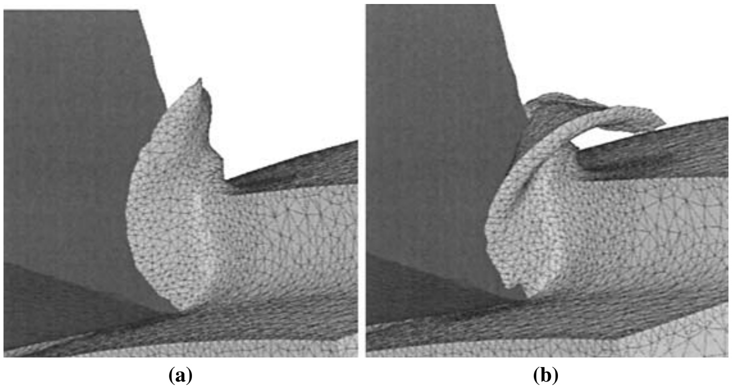


Fig. 22 Chip shape before and after Eulerian calculation

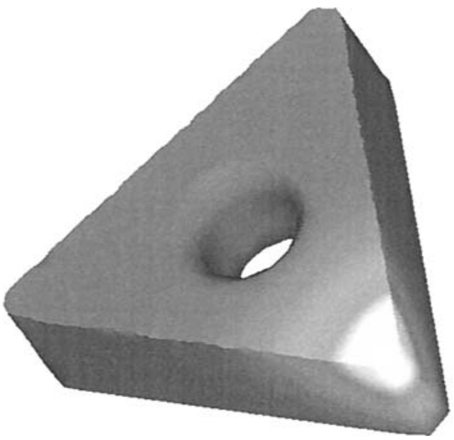


Fig. 23 Insert temperature after cutting

Current and Future Works

Modeling techniques for metal forming have continuously evolved throughout the years. Because of the unique capability of FEMs to describe the complex geometry and boundary condition of the forming process, the method has proven itself by numerous success stories experienced by the industry or reported by research organizations, such as reducing the trials and errors on the shop floor, and shortening the lead time to develop a new product, and so on.

Specifically, the FEM has been used by the forming industry primarily for the following purposes:

- The forming flow can be analyzed to evaluate the die/model design and to avoid flow-related defects such as folds, suck-in defects, flow-through defects, under fill, and ductile fracture.
- The tool stress can be analyzed to improve the tool life.
- The process window control (e.g. press speed, forging temperature, lubricant, and heat treatment procedure) can be modeled to ensure that the resulting microstructure properties (such as grain size) meet the requirement.
- The forging load can be predicted for equipment selection.

The solution accuracy depends largely on the validity of the assumptions that are made for the simulation. Minimizing the time and cost of production trials and design iterations continues to be the goal of FEM development.

Toward this goal, ongoing research is focused on the following areas.

Computational Efficiency. Three-dimensional FEM transient metal forming analysis is very computationally intensive. Although the computing power has been improving dramatically every year, the size of the FEM model for simulation has also been increasing quickly to improve the model accuracy. Methodologies to improve the computing speed, such as parallel computing and efficient contact searching algorithms, are under continuous exploration and development. Steady state or quasi-steady state methods are being investigated for certain classes of forming processes, such as extrusion, rolling, and ring rolling.

Robustness and Ease of Use. The required geometry data is generated by a computer-aided design (CAD) system. Generally speaking, the data may not meet the requirement for computer-aided engineering (CAE) analysis. An error-tolerant system is necessary for a robust CAE tool. In addition, realistic graphics presentation is also important for post-processing the simulation result. To facilitate the use of the FEM model, special purpose preprocessors aimed for specific manufacturing process are being developed.

Accuracy. The FEM approach is an error optimization process. Generally speaking, the larger the model is, the better the solution accuracy is expected. Certain assumptions, due to the lack of understanding of many real phenomena, may introduce inevitable error. Specifically, they are as follows:

Material Characterization:

- Flow stress data should cover the process window of the actual forming process, including temperature, strain, and strain rate. The extrapolation of the data range usually results in unexpected error.
- Elastic response has been neglected for large plastic deformation using rigid-plastic and rigid-viscoplastic formulation. For processes in which spring back or residual stress is of interest, an elasto-plastic formulation should be used. At high temperatures, materials can be highly rate sensitive, in which case an elasto-viscoplastic formulation should be considered.
- Microstructure modeling. Currently, most modeling considers isotropic behavior. Recrystallization, texture, damage, phase transformation, grain size effect, precipitate size, and distribution could potentially affect the deformation behavior. Understanding the microstructure properties evolution during the thermal mechanical process and its associated effect with respect to the mechanical properties, such as flow stress, Young's modulus, and creep behavior, is very crucial to the model accuracy since these properties are generally the input data to the FEM analysis.

Process Conditions, Including the Friction and the Lubricant Heat Transfer Coefficient. During the forming process, the surface of the workpiece is severely deformed/stretched, wherein the lubricant film thickness and film temperature are continuously changing in time. Consequently, this will affect the material flow behavior. Proper characterization of the lubricant behavior is important for the accuracy of the model.

Integrated FEM for Complete Manufacturing Processes. Most discrete mechanical components go through a manufacturing cycle, which typically includes ingot breakdown, forging, trimming, heat treating, welding, machining, and assembling and installation. A complete material flow and thermal mechanical history is important to model the behavior properly under the service condition. It is also necessary to have a better understanding of the material evolution and modeling throughout the entire manufacturing process.

Optimization. The FEM provides a prediction of the result of a proposed manufacturing process but still relies on an experienced designer to interpret the results of the analysis and modify the process based on prior knowledge and experience. Some recent research efforts have sought to use computational resources to enhance and optimize process designs based on a starting design, and improvement of the design is based on the sensitivity analysis of the design variables. Because the amount of simulations are generally large for the sensitivity analysis during optimization, these techniques are still somewhat limited to two-dimensional analysis but are emerging and evolving quickly.

With advances in the FEM, more manufacturing processes/problems with greater complexity are being studied and investigated. In addition to the development of an integrated manufacturing software, utilizing optimization techniques to systematically achieve design objectives is also the focus of current and future work. Due to the required computing resource in optimization and to minimize the user's time, the speed and the accuracy, the robustness of the FEM procedure, and computer graphics in three dimensions will continuously be challenged.

ACKNOWLEDGMENTS

The authors would like to thank Professor Soo-Ik Oh, Seoul National Unversity, for providing the upper bound method of ring compression, and Mr. D.H. Seo, Hyundai Motor Company, for their crankshaft forging example. The authors would also like to thank Mr. Tony Perez and Dr. Daqing Jin, both of the Timken Company, for their collaborate effort on material cutting modeling under the DOE funded CTMP project.

REFERENCES

1. C.H. Lee and S. Kobayashi, New Solutions to Rigid-Plastic Deformation Problems Using a Matrix Method, *J. Eng. Ind. (Trans. ASME),* Vol 95, 1973, p 865
2. J. J. Park and S.I. Oh, Application of Three Dimensional Finite Analysis to Metal Forming Process, *Proc. NAMRC XV*, (Bethlehem, PA), 1987, p 296
3. T. Coupez, N. Soyris, and J.L. Chenot, 3-D Finite Element Modeling of the Forging Process with Automatic Remeshing, *J. Mater. Process. Technol.,* Vol 27, 1991, p 119–133
4. G. Li, W.T. Wu, and J.P. Tang, DEFORM-3D: A General Purpose 3-D Finite Element Code for the Analysis of Metal Forming Processes, *Metal Forming Process Simulation in Industry,* 28–30, Sept 1994, (Baden-Baden, Germany), Internationale Kongress und Tagungsorganisation
5. E.G. Thomsen, C.T. Yang, and S. Kobayashi, *Mechanics of Plastic Deformation in Metal Processing*, Macmillan, New York, 1965
6. W. Johnson and P.B. Mellor, *Engineering Plasticity,* Van Nostrand and Reinhold, London, 1973
7. B. Avitzur, *Metal Forming: Processes and Analysis,* McGraw-Hill, New York, 1968
8. R.A.C. Slater, *Engineering Plasticity: Theory and Application to Metal Forming Processes,* Macmillan, London and Basingstoke, 1977
9. T. Altan, S.I. Oh, and H.L. Gegel, *Metal Forming: Fundamentals and Applications*, American Society for Metals, Metals Park, Ohio, 1983
10. H. Kudo, Some Analytical and Experimental Studies of Axi-Symmetric Cold Forging and Extrusion, Parts I & II, *Int. J. Mech. Sci.,* Vol 2, 1960, p 102–127; Vol 3, 1961, p 91–117
11. S.I. Oh, "Ring Compression Modeling Using Upper Bound Method," Private Communication, 2002
12. S. Kobayashi, S.I. Oh, and T. Altan, *Metal Forming and The Finite Element Method*, Oxford University Press, New York, 1989
13. W.J. Chung, J.W. Cho, and T. Belytschko, On the Dynamic Effects of Explicit FEM in Sheet Metal Forming Analysis, *Eng. Comput.,* Vol 15, 1998, p 750–776
14. R. Hill, *The Mathematical Theory of Plasticity*, Oxford University Press, London, 1971
15. L.E. Malvern, *Introduction to The Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, New Jersey, 1969
16. A.E. Green and P.M. Naghdi, A General Theory of Elastic-Plastic Continuum, *Arch. Rational Mech. Anal.*, Vol 18, 1965, p 251–281
17. H.D. Hibbit, P.V. Marcal, and J.R. Rice, Finite Element Formulation for Problems of Large Strain and Large Displacements, *Int. J. Solids Struct.*, Vol 6, 1970, p. 1069–1086
18. J. Wang and M.S. Gadala, Formulation and Survey of ALE Method in Nonlinear Solid Mechanics, *Finite Elem. Anal. Des.*, Vol 24, 1997, p 253–269
19. J.T. Oden and J.A.C. Martins, Models and Computational Methods for Dynamic Friction Phenomena, *Comp. Meth. Appl. Mech. Eng.*, Vol 52, 1985, p 527–634
20. A. Curnier, A Theory of Friction, *Int. J. Solids Struct.*, Vol 20, 1984, p. 637–647
21. A. Klarbring, General Contact Boundary Conditions and The Analysis of Frictional Systems, *Int. J. Solids Struct.*, Vol 22, 1986, p 1377–1398
22. G. Strang and G.J. Fix, *An Analysis of The Finite Element Method,* Prentice-Hall, Englewood Cliffs, New Jersey, 1973
23. J.T. Oden and J.N. Reddy, *An Introduction to the Mathematical Theory of Finite Elements*, John Wiley & Sons, New York, 1976
24. O.C. Zienkiewicz and R.L. Taylor, Chapter 10, *The Finite Element Method,* 4th ed. Vol 2, McGraw-Hill Book Company
25. M. Foster, *Scientific Forming Technologies Corporation Simulation Database Archive*
26. D.H. Seo, Numerical Simulation of Crank Shaft Forging Process, *DEFORM Korean User's Group Meeting*, 29 Oct 2002, Changwon, Korea
27. Y.C. Yen, A. Jain, C. Avanachand, W.T. Wu, and T. Altan, "Computer Simulation of Orthogonal Cutting Using a Tool with Multiple Coatings," submitted to 2003 College International pour l'Etude Scientifique des Techniques de Production Mecanique (CIRP)
28. C. Pavanachand, J.T. Jinn, A. Perez, D.Q. Jin, et al., "Machining Modeling," *SFTC, Timken, DOE Project Final Report*, under preparation, 2003

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 232-247
DOI:10.1361/hwpd2003p232

Chapter 16

Rolling

ROLLING OF METALS is perhaps the most important metalworking process, because a greater volume of material is worked by rolling than by any other deformation process. A significant portion (on the order of ~90%) of steel, aluminum, and copper products may go through the rolling process at least one time during the production process. The principal advantage of rolling lies in its ability to manufacture products from relatively large pieces of metals at very high speeds in a somewhat continuous manner. Other methods of metalworking, such as forging, are slower. Most ingots are processed by hot rolling into blooms, slabs, and billets (Fig. 1), which are subsequently rolled into other products such as plate, sheet, bar, structural shapes, rod (for drawing into wire), and rounds for making seamless tubing.

The largest quantities of rolled forms are flat products (plate, sheet, strip, foil) from the processes of hot rolling followed by cold rolling. Smaller quantities are rolled into shapes (sections), usually hot. All flat products are made in a fairly standard sequence of operations:

- The starting material is an individually cast ingot or continuously cast slab that is hot reduced (broken down) on two- or four-high reversing mills. Unless the end product is wide plate, rolling is continued to form a long, coiled, wide band. If production volumes warrant it, rolling from approximately 100 mm (4 in.) to approximately 25 mm (1.0 in.) may take place on several in-line mills with the slab running out between the roll pairs (stands).
- Thick strip (band) is hot rolled to a thickness of approximately 1.5 to 5 mm (0.06 to 0.2 in.); if quantities are large, a continuous hot strip mill is employed in which the stands are placed close enough together for the strip to run through all stands simultaneously, with controlled tensions between stands. The hot strip may be the finished product in the form of a coil, cut-to-length plate, or sheet.
- Cold rolling on reversing or tandem mills is done if a thinner gage, a smoother surface finish, or a strain-hardened end product is required. The strip is cold rolled with accurate control of coiler and/or interstand tensions. If the material strain hardens excessively, intermediate process anneals are carried out. The resulting wide strip or sheet is typically 0.2 to 2 mm (0.008 to 0.08 in.) thick. For small quantities, individual sheets may be rolled by hand.
- The thinnest strip (less than, say, 0.05 mm, or 0.002 in.) is called foil and is produced in specialized cold rolling mills. The work rolls are invariably backed up; the number of backup rolls ranges from two (four-high mill) to eighteen (Sendzimir mill).

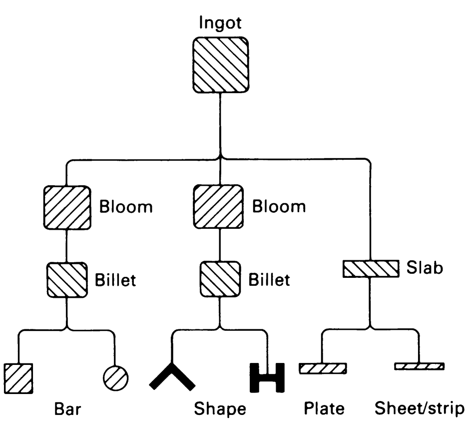


Fig. 1 Rolling sequence for fabrication of bars, shapes, and flat products from blooms, billets, and slabs

Narrow strip with rounded corners is obtained by flattening of wire.

The primary objectives of the rolling process are to reduce the cross section of the incoming material, to improve its properties, and to obtain the desired section at the exit from the rolls. In rolling, a squeezing type of deformation is accomplished by using two work rolls rotating in opposite directions. The process can be carried out hot, warm, or cold; thus, rolling processes are often classified as hot or cold. From a fundamental point of view, however, it is more appropriate to classify rolling processes on the bases of the complexity of metal flow during the process and the geometry of the rolled product. Thus, this chapter, based on adapted text from Ref 1 and 2, briefly introduces the rolling process for the basic product forms of strip, plate, and shapes.

In general, the rolling of solid sections can be divided into the subsequent categories:

- *Uniform reduction in thickness with no change in width:* This is the case with strip, sheet, or foil rolling where the deformation is in plane strain, that is, in the directions of rolling and sheet thickness. This type of metal flow exists when the width of the deformation zone is at least 20 times the length of that zone.
- *Uniform reduction in thickness with an increase in width:* This type of deformation occurs in the rolling of blooms, slabs, and thick plates. The material is elongated in the rolling (longitudinal) direction, is spread in the width (transverse) direction, and is compressed uniformly in the thickness direction.
- *Moderately nonuniform reduction in cross section:* In this case, the reduction in the thickness direction is not uniform. The metal is elongated in the rolling direction, is spread in the width direction, and is reduced nonuniformly in the thickness direction. Along the width, metal flow occurs only toward the edges of the section. The rolling of an oval section in rod rolling or of an airfoil section would be considered to be in this category.
- *Highly nonuniform reduction in cross section*: In this type of deformation, the reduction in the thickness direction is highly nonuniform. A portion of the rolled section is reduced in thickness, while other portions may be extruded or increased in thickness. As a result, in the width (lateral) direction, metal flow may be toward the center. Of course, in addition, the metal flows in the thickness direction as well as in the rolling (longitudinal) direction.

These points illustrate different deformation conditions of rolling. Except for strip rolling, metal flow in rolling is in three dimensions (in the thickness, width, and rolling directions). In addition, material flow during the rolling of shapes from cast billets of rectangular or round cross section is much more complicated than flat rolling. Determination of metal flow and rolling stresses in shape rolling is very important in the design and operation of rolling mills, and, so, numerical techniques are used to simulate metal flow in such complex rolling operations. However, this chapter only introduces general concepts on rolling of strip, plate, and shapes. Numerical techniques are not discussed in detail, but some references to numerical techniques are cited.

Flat Rolling

Flat rolling may appear to be a very simple process. The workpiece volume remains constant, and the reduction in thickness is first described under the idealized conditions of plane strain and homogeneous deformation. That is, the idealized (simplified) case of flat rolling assumes a rectangular body that is homogeneously deformed into another rectangular body of greater length. The reduction in thickness results mainly in a change of length of the material under the condition of plane strain. The deformation also is assumed to be homogeneous—that is, all vertical sections of the piece before rolling remain vertical during the rolling process. These two simplifying assumptions are useful as first approximation, although they are only valid under limited conditions. Plane-strain conditions are approximated when a flat is wide, and the assumption of homogeneous deformation restricts the validity of solutions to large L/h ratios, where L is the projected length for the arc of contact in the deformation zone, and h is the mean thickness of the section (as shown in Fig. 2). Friction and the pass geometry both influence the assumption of homogeneous deformation, and inhomogeneous deformation has several consequences (as noted in the section "Strain Distribution" in this chapter).

To initiate rolling, the workpiece first must be drawn into the roll gap by friction. This condition is reached when the horizontal component of the frictional force F is just equal to the horizontal component of the radial force P_r (Fig. 2). Thus:

$$F \cos \alpha > P_r \sin \alpha \quad \text{(Eq 1a)}$$

or

$$\frac{F}{P_r} > \tan \alpha \quad \text{(Eq 1b)}$$

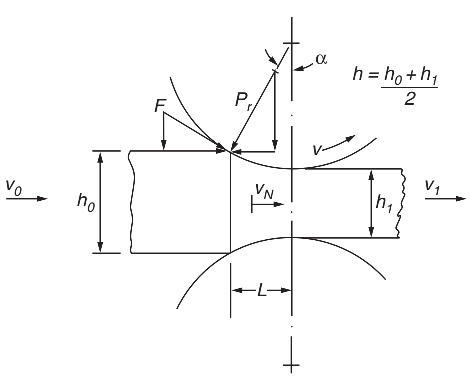


Fig. 2 Simplified schematic of rolling conditions

From the geometry of the pass, the maximum possible draft is:

$$\Delta h_{max} = \mu^2 R \quad \text{(Eq 2)}$$

where R is the radius of the roll, and μ is the coefficient of friction. By definition, $F/P = \mu$, which can be expressed also as tan f (where f is the friction angle). Therefore, it is usual to state that the section will enter the rolls unaided only if the friction angle f exceeds the contact angle α. This represents one of the instances where friction is desirable. This also would appear to offer a very simple way of determining the coefficient of friction: the roll gap is increased until the slab enters the rolls without being pushed. However, the value of this test is limited, because the frictional force is unidirectional at the point of acceptance, and lubricant films may be scraped off; neither of these conditions is representative of steady-state rolling. Nevertheless, the method is convenient under conditions of dry lubrication and for a very approximate ranking of lubricants.

Speed and Stress Distribution

Once the slab is accepted and drawn through the roll gap, the situation changes considerably. If idealized conditions of homogeneous deformation and plane strain are assumed, there is no spread, and so the reduction in thickness (Δh) causes an increase in length. Under these conditions, continuity can only be maintained if the products of thickness and velocity are constant at all points along the zone (arc) of contact. That is:

$$v_0 h_0 = v_n h_n = v_1 h_1$$

where v_0 and h_0 are the initial velocity and height before entering the deformation, and where v_n, h_n, and $v_1 h_1$ are equivalent products anywhere in the deformation zone and at the exit, respectively (Fig. 2). Accordingly, the exit speed of the slab, v_1, must increase in proportion to the elongation (which in turn is proportional to reduction). The rolls also move at some speed (v, Fig. 2) intermediate between the entry and exit speeds. Therefore, there is only one point in the arc of contact where the slab and roll move at the same speed ($v = v_n$). There is no relative slip at the point where $v = v_n$, and so, the point is described as the *no-slip point* or neutral point. In a three-dimensional presentation, it is referred to as the *neutral plane*.

With these variations in relative motion, slip between the slab and the roll occurs everywhere in the contact zone except at the neutral plane (or point). Between the entry point and neutral point, the work material moves more slowly than the rolls; this is referred to as *backward slip*. Between the neutral point and exit point, the strip moves more quickly than the roll; this is referred to as *forward slip*. The importance of relative slip between the workpiece and roll surfaces can hardly be overstated. In a typical rolling situation, forward slip seldom exceeds 15%; the rest of the speed differential is taken up in the backward slip zone. Therefore, the strip enters the roll gap at a high interface sliding velocity, which diminishes very rapidly, within a space of typically 20 mm (0.8 in.) or less, to zero in the neutral zone, only to pick up again as the strip leaves the roll gap. This relative sliding and its reversal, while helpful in many respects, also account for most of the problems encountered with lubrication in rolling, as discussed in more detail in the section "Relative Slip," in this chapter.

The Friction Hill. The amount of slip and the location of the neutral plane are influenced by the friction that opposes the relative slip. Frictional stresses oppose the slip, and the direction of the frictional stresses point (act) toward the neutral plane. These frictional stresses are described in terms of a so-called friction hill, which is related to the distribution of the pressure (p) between the rolls and the section in the contact arc (Fig. 3). The peak of the friction hill occurs at the neutral point, and the shape of it depends on the pass geometry and the magnitude of the friction. For example, as the amount of reduction is increased, the position of the neutral point approaches the exit. When the maximum possible reduction is attempted, the neutral point reaches the exit. This situation is very unstable, because any slight increase in reduction or drop in friction will cause the strip to stop moving and the rolls to start skidding over the strip.

Figure 4 illustrates some different situations and the shape of the friction hill. With very low friction (Fig. 4a), the inclination of the roll surfaces causes most of the material to flow backward, thus moving the neutral plane close to the exit. Deformation is almost homogeneous. The friction hill is low,

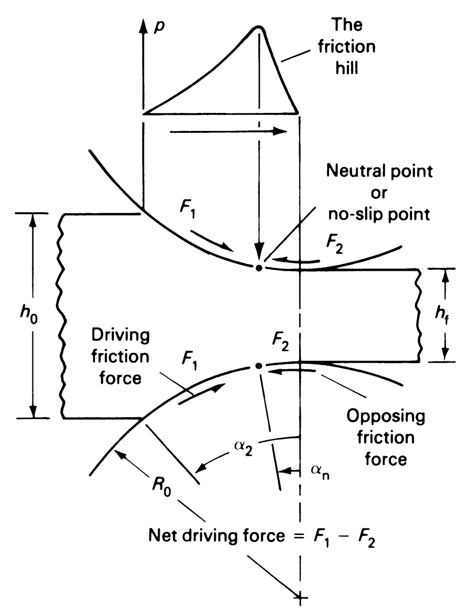


Fig. 3 Positive and negative friction directions during strip rolling

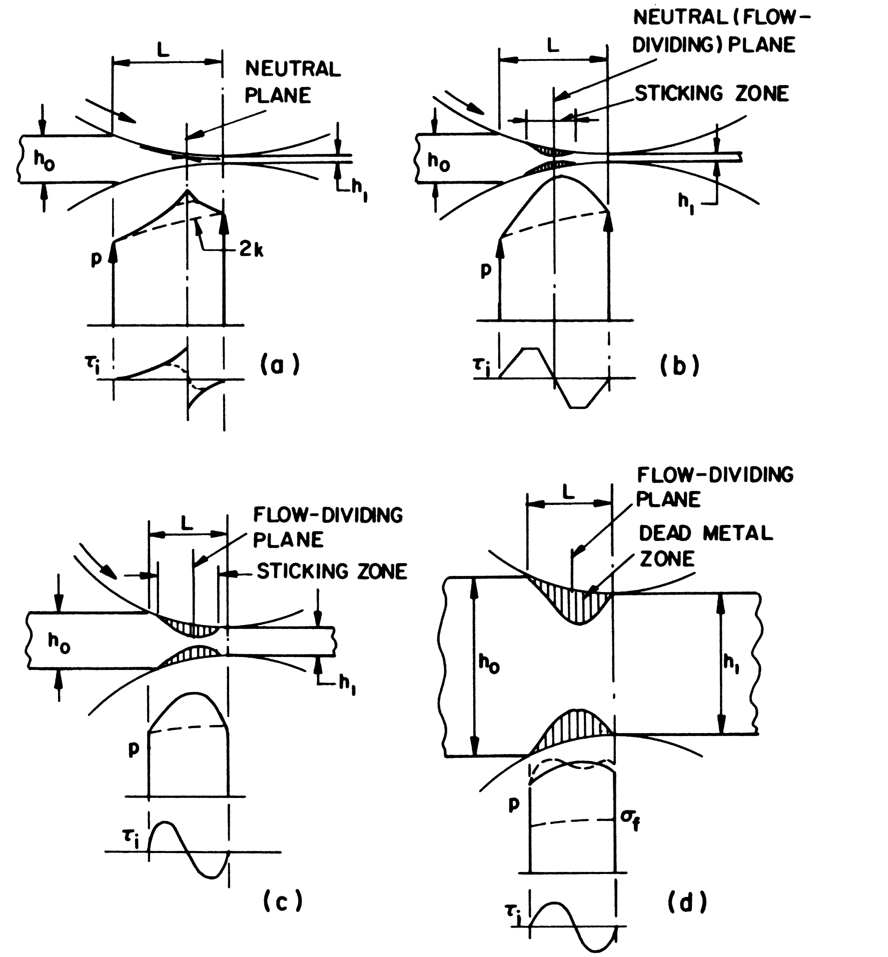


Fig. 4 Material flow, pressure, and shear stress distributions for various friction conditions and *L*/*h* ratios. (a) Very low friction, which moves the neutral plane close to the exit. (b) Intermediate friction, with a steeper friction hill and a neutral point further from the exit. If friction is high enough, sticking occurs, and the neutral plane broadens into a neutral zone. (c) Sticking within the entire contact zone. (d) Friction hill with double hump at low *L*/*h* ratios. See text for discussion.

and most of the pressure is attributable to the plane-strain flow stress (σ_0), equal to 1.15 σ_0 (or $2k$) if the material yields according to the von Mises criterion (where k is the shear flow stress according to von Mises). In cold working, the material strain hardens during its passage through the rolls, and the friction hill appears as in Fig. 4(a). Frictional stresses are low and frequently described by a constant μ, implying the τ_i distribution shown in Fig. 4(a) by broken lines. Direct measurements generally show a decay of the interface shear (friction) stress (τ_i) toward the neutral plane; even if there is some error in the measurements themselves, a variable μ (or friction factor m) within the arc of contact is more realistic.

With intermediate friction (Fig. 4b), the friction hill is steeper, the neutral point is farther from the exit, and, if friction is high enough, the following condition for sticking is satisfied near the neutral plane:

$$\tau_i = \mu p > k$$

where k is the shear flow stress according to von Mises. Because $k = 0.577\ \sigma_0$ with the von Mises criterion, the condition for sticking is sometimes considered to occur when $\mu_{max} = 0.577$, but this is really true only when full surface conformity is reached at $p = \sigma_0$ (Ref 3). With sticking, sliding is arrested, and the neutral plane broadens into a neutral zone. Shear stresses reach a limiting value, and correspondingly, the friction hill is rounded (Fig. 4b). Deformation is inhomogeneous, because a dead-metal zone forms adjacent to that part of the contact arc in which sticking friction prevails. Modeling of the shear stress distribution becomes more difficult. On rough rolls under unlubricated conditions, sticking may exist within the entire contact zone. Deformation becomes highly inhomogeneous, resembling backward and forward extrusion. Instead of a neutral zone, it is more realistic to speak of a flow-dividing zone. Because deformation is concentrated away from the roll surface, shear stresses on the roll decrease, and the friction hill is heavily rounded (Fig. 4c).

The previous examples are based on the situation of relatively thin strip—that is, $L/h > 2$ (where L is the projected length of the arc of contact, and h is the mean strip thickness). In these examples, inhomogeneous deformation results only from high friction. At lower L/h ratios, however, and especially when $L/h < 1$, inhomogeneity results from the process geometry itself. The friction hill may show a double hump (Fig. 4d). This situation is typical of early passes during the hot rolling of thick slabs, and separation of the effects of friction and process geometry becomes extremely difficult.

The prediction of the shape of the friction hill is one of the principal aims of rolling theory, but few theories predict the shape of the friction hill and the distribution of shear stresses in full agreement with actual measurements. To some extent, there is doubt also about the accuracy of the measurements themselves. In practice, the average pressure and roll force (discussed in the section "Modeling of Strip Rolling") are more practical and reliable indicators of the magnitude of friction; with higher friction, roll force is also higher.

Relative Slip

The Neutral Angle. As shown in Fig. 4, the neutral plane moves toward the middle of the arc of contact with increasing friction. Therefore, the angle of the neutral plane α_n (the angle between the neutral and exit planes in Fig. 3) is a sensitive indicator of friction balance and thus of the magnitude of friction. The neutral angle may be readily calculated if some simplifying assumptions are accepted. One of the most widely employed formulae (Ref 4) assumes that deformation is homogeneous, that slipping friction exists everywhere except in the neutral plane, and that radial roll pressure p and interface friction μ are constant along the arc of contact. These two latter assumptions are certainly unjustified, but it appears that many of the errors offset each other. Then, from the horizontal components of acting forces, the position of the neutral plane is:

$$\alpha_n = \frac{\alpha}{2} - \frac{1}{\mu}\left(\frac{\alpha}{2}\right)^2 \qquad \text{(Eq 3)}$$

where the α angle sustained at entry can be expressed from:

$$\sin\alpha = \frac{L}{R} = \left[\frac{h_0 - h_1}{R}\right]^{1/2} \qquad \text{(Eq 4)}$$

Substituting into Eq 3 obtains:

$$\alpha_n = \left[\frac{h_0 - h_1}{4R}\right] - \frac{1}{\mu}\left[\frac{h_0 - h_1}{4R}\right] \qquad \text{(Eq 5)}$$

and thus, α_n is completely determined by pass geometry and the coefficient of friction and, most conveniently, is independent of the flow stress of the material. This offers one of the most reliable methods of friction determination.

Rolling a strip into a coil permits the application of tensions. Back tension retards the movement of the strip and thus shifts the neutral plane toward the exit; front tension has the opposite effect. For rolling with tensions, the position of the neutral plane can be calculated only if the mean plane-strain flow stress ($2k$) is known. According to Ref 5:

$$\phi = \frac{\alpha}{2} - \frac{2k(h_0 - h_1) + h_0 t_0 - h_1 t_1}{4kR'\mu} \quad \text{(Eq 6)}$$

where t_0 is the back and t_1 the front tension, both expressed in units of stress; R' is the deformed roll radius (see Eq 28).

When friction is high enough for a sticking zone to develop, the surface of the strip exhibits forward slip only from the front end of the sticking zone. At a first glance, if sticking prevails over the whole arc of contact, forward slip should reduce to zero, at least on the strip surface. However, the center of the strip is now subjected to severe deformation, resembling forward and back extrusion, as noted in conjunction with Fig. 4(d). At the point of exit, the strip surface is accelerated to keep up with the center, and the position of the flow-dividing plane is still a sensitive indicator of frictional conditions.

Friction from Forward Slip. The roll and strip move at the same speed at the neutral plane, and any reduction in thickness between the neutral and exit planes must result in a higher exit velocity v_1. Forward slip (S_f) is defined as:

$$S_f = \frac{v_1 - v}{v} \quad \text{(Eq 7)}$$

It increases with ϕ and thus with increasing friction. From the geometry of the pass, forward slip is approximately equal to:

$$S_f = \frac{1}{2}\phi^2\left[\frac{2R'}{h_1} - 1\right] \quad \text{(Eq 8)}$$

Forward slip can be simply measured by scratching one or more lines on the roll surface parallel with the roll axis; the imprints of the lines will be clearly visible on the rolled strip. At a given roll speed v, the rolls take a certain time to travel the carefully measured distance (l_0) between successive scratch lines; during the same time period, the strip travels at a higher speed v_1, and the distance between the imprints of the two marks is greater (l_1). Therefore:

$$S_f = \frac{l_1 - l_0}{l_0} \quad \text{(Eq 9)}$$

From Eq 8, the position of the neutral plane, and from Eq 5, the average external coefficient of friction μ can be calculated, with the advantage that $2k$ need not be known and that no special instrumentation is needed. Some precautions are necessary, though, when roll flattening is significant or when the shrinkage of the l_1 distance on cooling must be taken into account for rolling at elevated temperatures.

Forward slip values are comparable only for preset tensions; otherwise, l_1 must be extracted from Eq 6 or some similar formula. The techniques proposed in Ref 6 and 7 derive μ from simultaneously measured roll force, torque, and forward slip values. Their great advantage is that $2k$ need not be known, and the coefficient of friction is obtained even in the presence of tensions (Ref 6):

$$\mu = \frac{T}{PR'\left[1 - \dfrac{2S_f h_1}{h_0 - h_1}\right]} \quad \text{(Eq 10)}$$

where P is measured roll force, T is torque, S_f represents forward slip values, and R' is the locally increased radius of the rolls due to slight flattening under load.

For a constant μ, forward slip reaches a maximum when the angle of contact equals the friction angle. This would allow direct determination of μ by simply rolling with gradually increasing reductions until the maximum forward slip value is found. Unfortunately, this procedure is valid only if μ remains unchanged with increasing pass reduction, a condition seldom fulfilled.

Friction from Skidding. With increasing reductions, the neutral angle moves toward the exit, and, at some critical reduction when the neutral plane reaches the exit, skidding sets in. A value of μ or m can then be calculated (Ref 8). In the absence of tensions, and when α is small so that in Eq 4 sin $\alpha = \alpha$, the angle of entry at skidding is:

$$\alpha_{\text{skid}} = \left(\frac{h_0 - h_1}{R'}\right)^{1/2} = 2\mu \quad \text{(Eq 11)}$$

The onset of skidding also can be determined, in principle, from a single experiment if a wedge-shaped specimen is rolled (Ref 9). Once the critical reduction is reached, skidding sets in, and the entry angle is twice the friction angle.

Other Methods of Friction Determination. A method related to forward slip measurement has been proposed (Ref 10). Small irregularities, such as pickup on the roll surface, cause scratches to appear on the surface of a polished strip. If rolling is carried out slowly enough to stop the rolls with the strip in the gap, the length and direction of the scratches may be observed under a microscope. From the relative lengths of forward and reverse scratches, the position of the neutral plane can be determined, and, from Eq 5, μ can be derived. The scratches become hooks in the spread zone, allowing fully rolled strips to be examined, although some of the tailing scratches formed in the backward slip zone tend to be obscured by later contact. The position of the neutral plane can be determined also from backward slip, provided that some accurate means of measuring entry velocity v_0 and roll velocity v is available.

Strain Distribution

In the idealized case of flat rolling, a rectangular body is homogeneously deformed into another rectangular body of greater length. In reality, inhomogeneous deformation develops in the thickness, width, and length directions due to the effects of friction and the effects of pass geometry. The ideal of homogeneous deformation is approached only in rolling with a large L/h ratio and with low enough friction to maintain sliding along the whole arc of contact (Fig. 4a).

As soon as a zone of sticking develops, the formation of a dead-metal zone (Fig. 4b, c) leads to the extrusion effect already mentioned, causing a backward convexity of originally vertical planes in the entry zone and a forward curvature in the exit zone. Deformation may be likened to compression between inclined plates, and therefore, the angle of entry α, or, more realistically, the $(h_0 - h_1)/R$ ratio, affects the position of the flow-dividing plane. Friction determines, together with the L/h ratio, the extent of the sticking zone; therefore, it also affects inhomogeneity. Furthermore, frictional drag may cause a severe retardation of surface layers.

Through-Thickness Inhomogeneity. Inhomogeneity of through-thickness deformation may cause variations in properties such as hardness. One example of this is where an inhomogeneity factor is defined as the difference in hardness between the core and the surface, expressed as a percentage of core hardness (Ref 11). With this factor, inhomogeneity is shown to decrease with reductions in friction (Fig. 5), but, at very light reductions, inhomogeneity was evident even with the best lubricant, because the L/h ratio was small. At high L/h values, deformation became homogeneous irrespective of the lubricant used. Deformation can be inhomogeneous even in the rolling of very thin strip, if the pass reduction (or, rather, the L/h ratio) is small. This condition is intentionally induced in temper (skin pass) rolling of steel.

The homogeneity of deformation also affects the location at which new surfaces are generated. The absolute increase in surface area is a function only of pass reduction; thus, if a 50% reduction is taken, half of the surface is old (covered with adsorbed films, oxides, contaminants, etc.) and half of it is virgin metal. In well-lubricated rolling where sliding friction predomi-

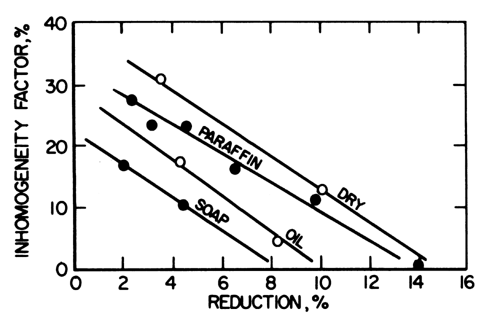


Fig. 5 Effect of lubrication on the inhomogeneity of deformation. Source: Ref 11

nates, the new surfaces are generated along the arc of contact, and, presumably, the lubricant effectively separates the two surfaces. With poor lubrication or in the absence of a lubricant, the surfaces break up just before entering the roll gap; virgin surfaces enter the contact zone, and metal-to-metal contact could occur. The consequences of this depend on the prevailing lubrication mechanism.

Occasionally, surface cracking is encountered in the hot rolling of difficult-to-work materials. The cause is a combination of low ductility (at less than optimal hot working temperatures) and secondary tensile stresses generated by cooling of the workpiece surface in contact with the rolls. Localized deformation associated with sticking friction contributes to these stresses. Lubricants alleviate the problem partly by increasing the homogeneity of deformation and partly through their heat-insulating properties.

When friction is higher on one roll than on the other, the larger elongation on the well-lubricated surface causes the rolled slab to curl around the roll with higher friction. This also happens when one of the rolls is smaller. Friction and inhomogeneity of deformation are also important factors in achieving bonding in roll cladding and in rolling of bimetallic strip. Localized deformation due to friction and pass geometry also can affect the rate of texture development in cold-rolled sheet.

Lateral Inhomogeneity. The assumption of plane-strain deformation implies that material flow is directed only in the longitudinal direction. In practice, it is inevitable that some lateral flow (spread) should also occur. One early experimental demonstration of this is outlined in Ref 12. Lateral flow is clearly evident, as is the resultant increase in width w (Fig. 6). Most of the spread develops in the zone of backward slip; as such, it is sensitive to the position of the neutral plane. Spread increases with increasing friction, but other factors enter, too. For a given pass reduction, spread is greater with a larger roll; for the same roll diameter, spread increases with the L/w ratio, because a longer arc of contact presents greater frictional resistance in the rolling direction. These observations have been repeatedly confirmed. A partial review is given in Ref 13. Confusion arises only when dry and lubricated experiments are conducted at a constant roll gap setting; lubricated rolling gives heavier reductions and therefore greater spread, creating the false impression that spread decreases with friction.

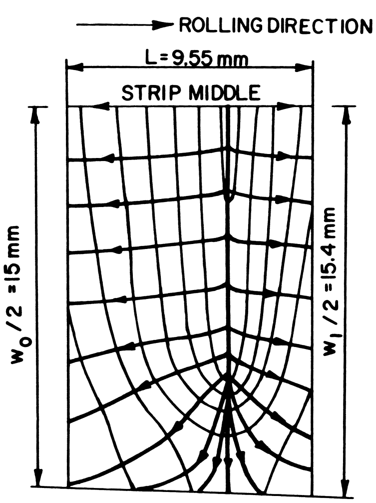


Fig. 6 Directions of material flow in dry rolling of aluminum strip (plan view of contact zone over half-width). Source: Ref 12

The friction hill shapes shown in Fig. 4 apply only to the central portion of a slab where plane strain prevails. Pressures gradually decay toward the spread zones, as already shown (Ref 12) (Fig. 7). At the edge, the normal pressure drops below $2k$, and yielding takes place in a stress state composed of the normal stress and a longitudinal tensile stress induced by the elongation of material adjacent to the edge (secondary tensile stress). Using a large number of pressure transducers embedded in the roll surface along a line parallel to the roll axis, it has been shown (Ref 14) that without a lubricant, the pressure drop-off is steep and limited to the edges, whereas in lubricated rolling, the pressure begins to drop off, very gradually, from the center of the strip.

Friction also affects the development of the side surfaces of the rolled material. In ideal homogeneous deformation, the edge profiles would remain straight (Fig. 8). However, friction restricts material flow at the interface in the width direction, too, and therefore, higher friction results not only in greater spread but also in greater barreling of the side faces (Fig. 8b). Under sticking conditions, some of the spread actually develops by folding over of the side surfaces (Fig. 8c). Single barreling is typical of heavy reductions taken on a relatively thin material—that is, with $L/h > 2$. Relatively light reductions on a thick workpiece ($L/h < 1$) lead to double barreling (Fig. 8d), which is only weakly dependent on friction.

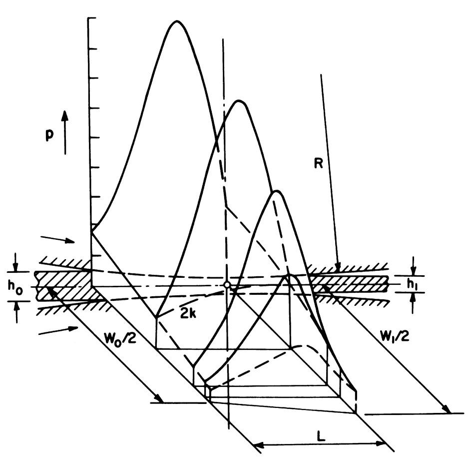


Fig. 7 Interface pressure distribution in dry rolling of aluminum strip on rough rolls. Source: Ref 12

Through its influence on the homogeneity of deformation, friction also affects the occurrence of rolling defects. Edge cracking is a result of longitudinal secondary tensile stresses generated at the edges of the rolled workpiece, where the roll pressure drops (Fig. 7). Secondary tensile stresses become especially large when the workpiece barrels severely; therefore, the danger of edge cracking increases with increasing friction. Friction, especially sliding friction, has also been suspected of contributing to opening up of the billet in the horizontal center plane (crocodiling or alligatoring), although no direct evidence or detailed analysis is available to support this assumption.

Strip Shape. Control of strip shape requires an understanding of the interactions of many variables. A flat product is rolled only if reductions and, consequently, elongations are uniform across the whole width of the sheet. If the strip was originally of uniform thickness, the roll gap must be parallel; if the strip is rolled with a slightly thicker center to ensure good tracking, the roll gap must be proportionally thicker in the middle. The roll gap referred to here is that defined by the rolls while the rolling load is applied; because the rolls deflect under the applied loads, cylindrical rolls would present a substantially larger gap in the middle of the strip. Thus, greater elongations would be imposed on the strip edges, making them longer and therefore wavy (Fig. 9).

These effects may be counteracted in several ways:

- The rolls may be ground to a barrel shape, with the crown (camber) calculated to give a parallel gap under load.
- Heat input into the rolls causes them to acquire a thermal camber, because much of the heat is extracted through the roll necks. Frictional heating contributes to the roll camber. Luckily, the larger thermal camber is at least partially compensated for by the larger roll deflection due to higher roll forces.
- The temperature profile and thus the thermal camber of the rolls are affected by the quantity, distribution, and heat capacity of the lubricant.
- The work rolls of a four-high or multiroll mill may be mechanically deflected to control the shape of the roll gap very rapidly.
- In addition to ground and thermal cambers and roll bending, strip shape is also affected

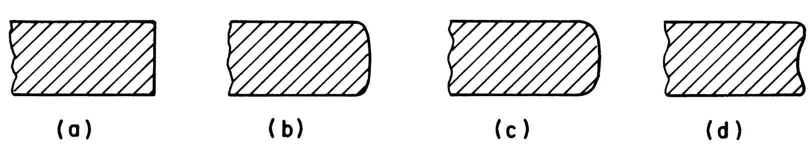


Fig. 8 Edge profiles of rolled slabs

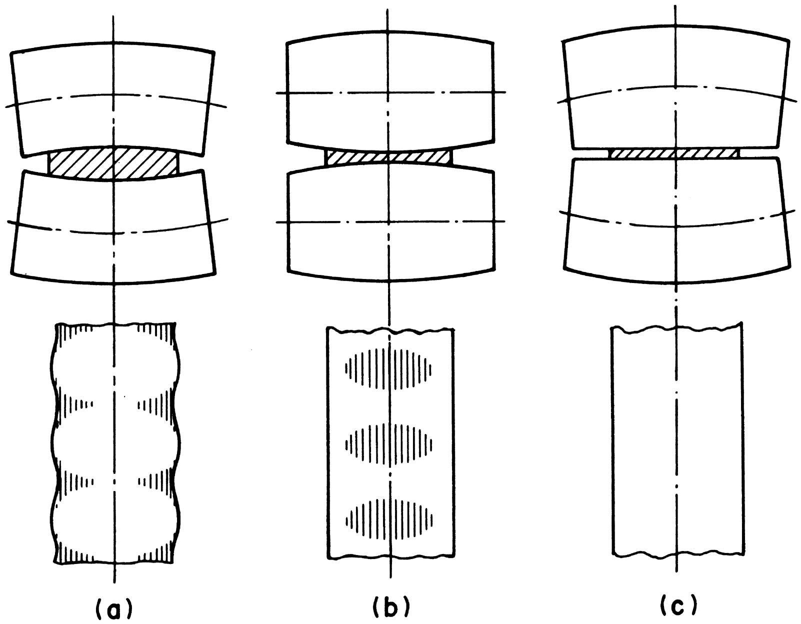


Fig. 9 Elastic deflection of rolls and its effect on strip shape. (a) Long, wavy shape at the strip edges due to inadequate crown (thermal or ground) of the rolls. (b) Long, wavy shape in the middle region due to excessive crown. (c) Parallel gap between rolls, and no waviness on the strip

by tensions, because, for a given reduction, roll forces are lower and roll deflections are smaller if higher tensions are applied.

The delicate balance may be upset by any change from steady-state conditions. For example, during mill start-up or when cooling by the lubricant is excessive, sufficient heating does not develop the thermal camber of the rolls, and the strip is rolled with a long (wavy) edge (Fig. 9a). If, on the other hand, insufficient coolant is applied or the lubricant is too warm, excessive thermal camber leads to rolling with a long (wavy) middle (Fig. 9b). When the waviness shifts off center, it is usual to speak of a quarter-wave. Highly localized poor shape may develop as a result of a blocked coolant nozzle, which allows local heating.

Lubrication

Friction is needed to draw the workpiece into the roll gap and, once the workpiece is accepted, to ensure its passage through the deformation zone. The minimum value of friction required for acceptance is twice that needed for continuous rolling. Lubricants are applied primarily to reduce friction and wear and to ensure temperature control. The relative importance of these functions depends on the process and the workpiece material. In cold rolling, the dominant mechanism is abrasive and adhesive wear accompanied by spalling, a form of fatigue wear.

In hot rolling, abrasive and thermal fatigue wear dominate. Lubrication then serves to reduce abrasive and adhesive wear and to minimize the thermal excursions that culminate in thermal fatigue. The purpose of lubrication in hot rolling depends on the adhesion between workpiece and roll materials. If adhesion is high, as in rolling of aluminum, lubrication is essential for the control of roll coating buildup. If adhesion is low, as in rolling of steel, the cooling and wear-reducing functions of the lubricant prevail. At high *L*/*h* ratios, lubricants aid in reducing roll force and power.

The most frequently used lubricants are tabulated in Table 1. For purposes of preliminary, approximate calculations, typical average coefficient of friction values are also shown, although it must be understood that, because of the prevalence of the mixed-film lubrication mechanism, these values change greatly with process conditions. Lubricants may be of an oil or a water base. Because of the predominance of the mixed-film mechanism, boundary and/or extreme pressure additives are essential. Oil-based lubricants are used, in general, only when water staining would develop, as in cold rolling of aluminum, or when an aqueous lubricant would prove inadequate in terms of reduction of friction or prevention of adhesion.

The lubricating power of emulsions can be increased by reducing their stability. It is also possible to combine the high lubricating power of a neat fat or oil with the high cooling power of water by applying the lubricant to the strip and the water to the roll, or by mechanically mixing the two at the point of application. The same principles apply to the hot rolling of sections (shapes) and tubes, except that the spatial configuration of the deformation zone becomes complex, and large speed differentials develop around the circumference of the section. Filling of the shape is generally hindered by friction, but friction is actually helpful in filling a leg formed by indirect draft.

In cold rolling, one of the main purposes of lubrication is that of reducing friction. Because of the need for some minimum friction, hydrodynamic lubrication is impractical, but hydrodynamic theory is of value in predicting the thickness of the film entrained for mixed-film lubrication. The gradually converging entry zone, periodic contact with the roll surface, and limited sliding in the contact zone make rolling responsive to lubrication with fluids. As expected from basic principles, friction in the mixed-film lubrication regime decreases with decreasing roll entry angle (smaller reduction

Table 1 Commonly used rolling lubricants and typical μ values

	Hot rolling		Cold rolling	
Material	**Lubricant(a)**	**μ(b)**	**Lubricant(a)**	**μ(c)**
Steels	None + cooling water	ST	Emulsion, 3–6% conc., of synthetic palm oil	MF
	Emulsion of fat (+ E.P.)	0.4		
	Fat (ester) (+ E.P.) + cooling water	0.3	Synthetic palm oil + water	MF
Stainless steels and Ni alloys	As for steel		M.O. (10–20) with Cl additive	MF
Al and Mg alloys	Emulsion, 2–15% conc., of M.O. (20–100) with <20% fatty acid, alcohol, ester	0.4	M.O. (4–20) with 1–5% fatty acid, alcohol, ester	MF
			As above, but synthetic mineral oil	MF
			Foil: as above, but (1.5–6)	MF
Cu and Cu alloys	Emulsion, 2–8% conc., of M.O. (80–400) with fat	0.3	Emulsion, 2–10% conc., of M.O. (80–400) with fat	MF
			M.O. (8–50) with fat (fatty acid)	MF
Ti alloys	None	ST	Oxidized surface, with:	
	Fat (+ water)	ST	Esters or soaps	0.2
			Castor oil (fatty oil)	0.2
			Compounded M.O. (4–10)	0.2
Refractory metals	Canning + lubricant for can material	0.4	M.O. with boundary and extreme pressure agents	MF
	Bare, dry	0.3		

(a) M.O., mineral oil; viscosity in centiStokes at 40 °C (105 °F) in parentheses. (b) ST, sticking friction. (c) MF, mixed-film lubrication; μ = 0.10 at low speeds, dropping to 0.03 at high speeds and viscosities and larger roll diameter), increasing viscosity, and increasing speed. Decreasing friction means lower forces and, for a given roll gap setting, thinner issuing gage. Thus, on acceleration of the mill, the strip thickness decreases; this speed effect is reinforced by changes in oil film thickness in mills equipped with full-fluid-film bearings.

With increasing pass reduction, lubrication shifts from predominantly hydrodynamic to predominantly boundary in character. Forward slip increases and roll forces rise; the rise in roll force becomes very steep once conditions of limiting reduction are approached. Very heavy reductions and/or high speeds thin the film to the point where breakdown occurs, resulting in heavy smearing and roll pickup with materials prone to adhesion (such as aluminum) and the development of a localized heat streak defect with other materials (such as steel). Increasing roll surface roughness generally shifts the mechanism toward the boundary regime but can be beneficial in delaying the onset of the heat streak defect.

Modeling of Strip Rolling

This section, adapted from Ref 2, provides a brief review of strip rolling theory and analysis. In strip rolling, the width of the strip is much greater than its thickness, and the deformation process is regarded as a problem of plane strain with sliding friction along the arc of contact (except at the neutral point). Models might also assume homogeneous deformation—that is, originally vertical sections remain vertical during rolling. This assumption restricts the validity of solutions to large L/h ratios and an ideal rigid-plastic material of constant flow strength.

The researcher von Karman (Ref 15) was the first to write the differential equations for the equilibrium of horizontal forces acting on a vertical section of thickness dx taken at a distance x from the exit plane (Fig. 10). Assuming that friction is low and x is small relative to the roll radius R, the radial roll pressure p may be considered to be approximately equal to the vertical component p_v, and equilibrium can be expressed as:

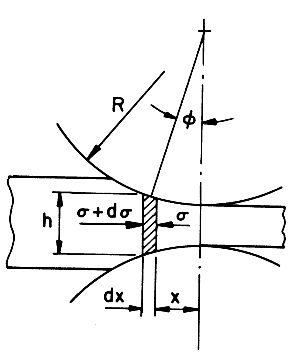


Fig. 10 Stresses acting on a vertical section of a rolled workpiece

$$\frac{d(\sigma h)}{dx} = p_v(\sin\phi \mp \mu\cos\phi) \quad \text{(Eq 12)}$$

where the minus sign refers to the zone of backward slip and the plus sign to the zone of forward slip. The horizontal compressive stress σ is assumed to be uniformly distributed over the height h of the section.

A more rigorous analysis (Ref 16) eliminated some of the unrealistic assumptions, notably, absence of strain hardening in the workpiece material, inhomogeneous deformation, and the presence of a sticking zone. The analysis was applied in the mid-1900s for calculation of rolling forces (Ref 17, 18). The roll-separating force and the roll torque can be estimated with various levels of approximations by such mathematical techniques as the slab method, the upper-bound method (Ref 19), or the slip-line method of analysis (Ref 20, 21). In more recent developments, computerized numerical techniques have been used to estimate metal flow, stresses, roll-separating force, elastic deflection of the rolls, thermal conditions, and strain distributions by viscoplastic analysis (Ref 19, 22–27). Viscoplastic analysis allows the construction of slip-line fields and gives information on pressure distribution and interface conditions. The basic formulation for a general viscoelastic material and constitutive equations are described in Ref 28, and the formulation of flow in rolling is introduced in Ref 29.

Of course, numerical solutions continue to become more refined and robust with advances in computer technology. Finite-element methods (FEM) are particularly important, because they are very flexible and can be applied for various situations ranging from constant friction, mixed sliding, and non-steady-state conditions. In FEM, the complex shape under the roll gap is divided into cells or elements with simple three-dimensional shapes. Through the analysis of these elements one at a time, the deformation pattern in a complex shape can be determined. Such calculations can take hours, depending on the complexity of the calculation and the speed of the computer, but new generations of computer hardware with greater speed continue to enhance the use of FEM. A more detailed review of FEM is beyond the scope of this chapter, but the following abstracts are a small sample of FEM applications in strip rolling analysis:

- "*An ALE Hydrodynamic Lubrication Finite Element Method With Application to Strip Rolling*" (Ref 30): A method that incorporates the hydrodynamic lubrication analysis into the arbitrary Lagrangian Eulerian (ALE) finite-element analysis is developed for steady-state strip rolling simulation. By employing the ALE formulation, only part of the workpiece, which is subjected to large plastic deformation within the roll-bite region, is modelled, so that the computational cost is substantially reduced. Two numerical examples, aluminium and steel strip rolling processes, are presented to demonstrate the merits of the proposed method.
- "*A Plane-Strain Elastoplastic Finite-Element Model for Cold Rolling of Thin Strip*" (Ref 31): Elastic roll deformation in cold rolling of thin strip is an important concern, and a coupled FEM is formulated, meshing a global strip-roll system with internal interface with sliding and friction. The model is two-dimensional and only analyzes roll flattening, but application to several kinds of rolling passes is examined (temper rolling, thin foil rolling), with special emphasis on roll-deformed shape and behavior of metal in the roll gap (sliding/sticking zones, elastic/plastic zones).
- "*A Finite Element Computer System for Analysis of Mechanical and Thermal Behaviors in Hot Strip Rolling*" (Ref 32): A two-dimensional, Eulerian finite-element-based model is presented for the prediction of heat transfer and metal flow occurring in the strip, and heat transfer and thermoelastic phenomena occurring in the work roll during hot strip rolling. Basic FEMs are described, with emphasis on combining each model to deal rigorously with the coupled aspect of the thermomechanical behaviors.
- "*A Three-Dimensional Finite Element Method for a Nonisothermal Aluminum Flat Strip Rolling Process*" (Ref 33): A three-dimensional coupled thermoelastic-plastic FEM of nonisothermal rolling is described for the analysis of strip curvature caused by the difference in the heat-transfer boundary conditions of the upper and lower rollers. In addition, the changes in shape, temperature field, and strain field of the strip during the various stages were analyzed to obtain the lateral plastic flow of the strip.
- "*Three-Dimensional Rigid-Plastic Finite Element Analysis of Strip Crown and Flatness During Rolling With Constraint Boundary Conditions Proposed by Shohet*" (Ref 34): Three-dimensional rigid-plastic FEMs and three-dimensional elastic FEMs were used to investigate the behavior of the contact boundary between the surface of a roll and a strip in strip rolling. The constraint boundary condition proposed by Shohet is incorporated into the whole system to evaluate both strip crown and flatness.

Simplified Method for Estimating Roll-Separating Force. The strip rolling process and symbols used in all subsequent equations are illustrated in Fig. 11. Because of volume constancy, the following relations hold:

$$w \cdot h_0 \cdot v_0 = w \cdot h \cdot v = w \cdot h_1 \cdot v_1 \quad \text{(Eq 13)}$$

where w is the width of the strip; h_0, h, and h_1 are the thicknesses at the entrance, in the deformation zone, and at the exit, respectively; and v_0, v, and v_1 are the velocities at the entrance, in the deformation zone, and at the exit, respectively.

In order to satisfy Eq 13, the exit velocity v_1 must be larger than the entrance velocity v_0. Therefore, the velocity of the deforming material in the x or rolling direction must steadily increase from entrance to exit. At only one point along the roll-strip interface is the surface velocity of the roll v_R equal to the velocity of the strip. This point is called the neutral point or neutral plane, indicated by N in Fig. 11.

The interface frictional stresses are directed from the entrance and exit planes toward the neutral plane, because the relative velocity between the roll surface and the strip changes its direction at the neutral plane. This is considered later in estimating rolling stresses.

An approximate value for the roll-separating force can be obtained by approximating the deformation zone, shown in Fig. 11, with the homogeneous plane-strain upsetting process. With this assumption, Eq 14 is valid; that is, the load per unit width (P/w) of the strip is given by:

$$\frac{P}{w} = \frac{2\bar{\sigma}}{\sqrt{3}}\left(1 + \frac{ml}{4h}\right)l \qquad \text{(Eq 14)}$$

However, in this case, the following approximations must be made:

- Average strip height $h = 0.5(h_0 + h_1)$.
- Average length of the deforming strip $l = R\alpha_D$, with $\cos \alpha_D = 1 - (h_0 - h_1)2R$. In the literature, it is often recommended that the value of the projection of strip length x_D (Fig. 11) be used for l; however, considering the effect of friction on the roll-strip interface length, $R\alpha_D$, it is more appropriate to use $l = R\alpha_D$.

To estimate average flow stress $\bar{\sigma}(\bar{\varepsilon}, \dot{\bar{\varepsilon}}, T)$ at a given rolling temperature T, the average strain $\bar{\varepsilon}$ is obtained from the thickness reduction; that is, $\bar{\varepsilon} = \ln(h_0/h_1)$. The strain rate $\dot{\bar{\varepsilon}}$ is given by:

$$\begin{aligned}\dot{\bar{\varepsilon}}_\alpha &= v_z / h = 2v_R \sin\alpha / h \\ &= [2v_R \sin\alpha] / [h_1 + 2R(1 - \cos\alpha)]\end{aligned} \qquad \text{(Eq 15)}$$

where v_z is the velocity at a given plane in the z-direction (Fig. 11), h is the thickness at a given plane (roll angle α) in the deformation zone, and v_R is the roll surface velocity. At the entrance plane:

$$v_z = 2v_R \sin \alpha_D;\ h = h_0$$

At the exit plane:

$$v_z = 0;\ h = h_1$$

By taking a simple average of these two limiting values, an approximate value of strain rate is obtained:

$$\dot{\bar{\varepsilon}} = [2v_R \sin\alpha_D / h_0 + 0] / 2 \qquad \text{(Eq 16)}$$

A more accurate value can be obtained by calculating an integrated average of $\dot{\bar{\varepsilon}}_\alpha$ (Eq 15) throughout the deformation zone. Then, an average approximate value is (Ref 35):

$$\dot{\bar{\varepsilon}} = \frac{v_R}{h_0}\left[\frac{2(h_0 - h_1)}{R}\right]^{1/2} \qquad \text{(Eq 17)}$$

The simplified (linear) stress roll pressure distribution in strip rolling is illustrated in Fig. 12. The maximum stress is at the neutral plane N. These stresses increase with increasing friction and length of the deformation zone X_D. Tensile stresses applied to the strip at entrance or exit have the effect of reducing the maximum stress (by an amount approximately equal to $\Delta\sigma_z$, in Fig. 12b) and shifting the position of the neutral plane. The analogy to plane-strain upsetting is illustrated in Fig. 12(a).

The stress distribution can be calculated by using the equations derived in most textbooks or by following the theory presented in Ref 16. However, these calculations are quite complex and require numerical techniques in order to avoid an excessive number of simplifying assumptions. A computerized solution, with all necessary details and the listing of the Fortran computer program, is also given in Ref 22.

For a numerical/computerized calculation of rolling stresses, the deformation zone can be divided into an arbitrary number of elements with flat, inclined surfaces (Fig. 13). The element, illustrated in this figure, is located between the neutral and exit planes, because the frictional stress τ is acting against the direction of metal flow. When this element is located between the entrance and neutral planes, τ acts in the direction of metal flow. The stress distribution within this element can be obtained by use of the slab method, as applied to plane-strain upsetting (Ref 36):

$$\sigma_z = \frac{K_2}{K_1}\ln\left(\frac{h_1}{h_0 + K_1 X}\right) + \sigma_{z1} \qquad \text{(Eq 18)}$$

where

$$K_1 = -2 \tan \alpha \qquad \text{(Eq 19)}$$

$$K_2 = -\frac{2\bar{\sigma}K_1}{\sqrt{3}} + 2\tau(1 + \tan^2 \alpha) \qquad \text{(Eq 20)}$$

$$\tau = m\bar{\sigma}/\sqrt{3} \qquad \text{(Eq 21)}$$

Following Fig. 13, for $x = \Delta x$, $h_0 + K_1 x = h_1$, and therefore, Eq 18 gives $\sigma_z = \sigma_{z1}$, the boundary condition at $x = \Delta x$, which is known. For $x = 0$:

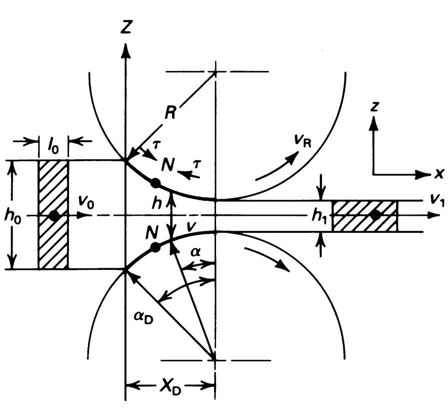


Fig. 11 Representation of strip rolling. The strip width w is constant in the y (width) direction.

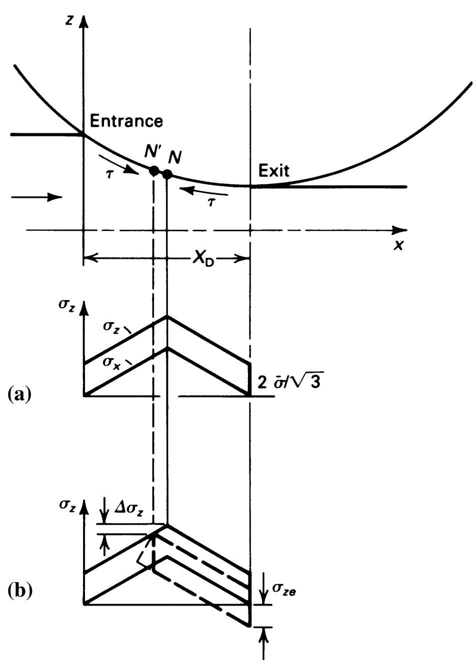


Fig. 12 Stress distribution in rolling. (a) With no tensile stresses at entry or exit. (b) With tensile stress σ_{ze} at exit

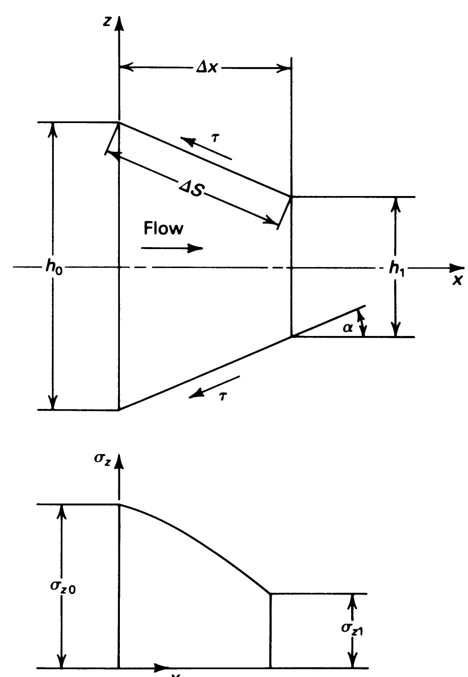


Fig. 13 Stresses in a deformation element used in computerized calculation of rolling stresses

$$\sigma_z = \sigma_{z0} = \frac{K_2}{K_1}\ln\left(\frac{h_1}{h_0}\right) + \sigma_{z1}$$

If the element shown in Fig. 13 is located between the entrance and neutral planes, then the sign for the frictional shear stress τ must be reversed. Thus, Eq 18 and 19 are still valid but:

$$K_2 = -\frac{2\bar{\sigma}}{\sqrt{3}}K_1 - 2\tau(1 + \tan^2\alpha) \quad \text{(Eq 22)}$$

In this case, the value of the boundary condition at $x = 0$, that is, σ_{z0}, is known, and σ_{z1} can be determined from Eq 18:

$$\sigma_{z1} = \sigma_{z0} - \frac{K_2}{K_1}\ln\left(\frac{h_1}{h_0 + K_{1\Delta x}}\right) \quad \text{(Eq 23)}$$

The stress boundary conditions at exit and entrance are known. Thus, to calculate the complete stress (roll pressure) distribution and to determine the location of the neutral plane, the length of the deformation zone X_D (Fig. 11, 12) is divided into n deformation elements (Fig. 14). Each element is approximated by flat top and bottom surfaces. Starting from both ends of the deformation zone, that is, entrance and exit planes, the stresses are calculated for each element successively from one element to the next. The calculations are carried out simultaneously for both sides of the neutral plane. The location of the neutral plane is the location at which the stresses, calculated progressively from both exit and entrance sides, are equal. This procedure has been computerized and extensively used in cold and hot rolling of sheet, plane-strain forging of turbine blades (Ref 37), and in rolling of plates and airfoil shapes (Ref 38, 39).

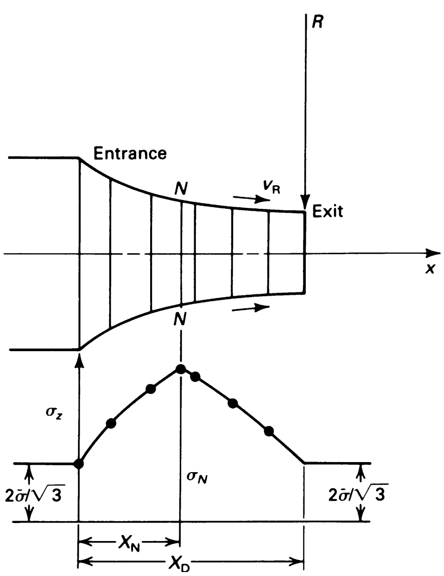


Fig. 14 Calculation of stress distribution by dividing the deformation zone into a number of tapered elements. In this case, tensile stresses in the strip are zero at both entrance and exit.

Roll-Separating Force and Torque. The integration of the stress distribution over the length of the deformation zone gives the total roll-separating force per unit width in strip rolling. In addition, the total torque is given by:

$$T = \int_0^{X_D} R\,dF \quad \text{(Eq 24)}$$

where X_D is the length of the deformation zone (Fig. 14), R is roll radius, and F is the tangential force acting on the roll. Assuming that all energy is transmitted from the roll to the workpiece by frictional force:

$$dF = \tau ds \quad \text{(Eq 25)}$$

$$ds = dx / \cos\alpha = \sqrt{1 = \tan^2 \alpha dx} \quad \text{(Eq 26)}$$

In the deformation zone, the frictional force is in the rolling direction between entry and neutral planes. It changes direction between the neutral and exit planes. Thus, the total roll torque per unit width is:

$$T = R\tau\left[\int_0^{X_N}(1 + \tan^2\alpha)dx - \int_{X_N}^{X_D}(1 + \tan^2\alpha)dx\right] \quad \text{(Eq 27)}$$

where τ equals $m\bar{\sigma}/\sqrt{3}$; R is roll radius; α is roll angle (Fig. 11); X_N is the x distance of the neutral plane from the entrance (Fig. 14); and X_D is the length of the deformation zone (Fig. 14).

Elastic Deflection of Rolls. During rolling of strip, especially at room temperature, a considerable amount of roll deflection and flattening may take place. In the width direction, the rolls are bent between the roll bearings, and a certain amount of crowning, or thickening of the strip, occurs at the center. This can be corrected by either grinding the rolls to a larger diameter at the center or by using backup rolls.

In the thickness direction, roll flattening causes the roll radius to enlarge, increasing the contact length. There are several numerical methods for calculating the elastic deformation of the rolls (Ref 23). A method for approximate correction of the force and torque calculations for roll flattening entails replacement of the original roll radius R with a larger value R'. A value of R' is suggested in Ref 40 and is referred to extensively in the literature. This is given as:

$$R' = R\left[1 + \frac{16(1 - \nu^2)p}{\pi E(h_0 - h_1)}\right] \quad \text{(Eq 28)}$$

where v is Poisson's ratio of the roll material, p is the average roll pressure, and E is the elastic modulus of the roll material.

It is obvious that R' and p influence each other. Therefore, a computerized iteration procedure is necessary for consideration of roll flattening in calculating rolling force or pressure. Thus, the value of p is calculated for the nominal roll radius R. Then R' is calculated from Eq 28. If $R'/R \neq 1$, the calculation of p is repeated with the new R' value, and so on, until R'/R has approximately the value of 1.

Hydrodynamic Lubrication (Adapted from Ref 41). Lubricant is applied at the entrance side of each pair (stand) of rolls. At the entrance side, the layers of lubricant that are in contact with the rolls or with the strip adhere to their respective metal surfaces and move inward toward the exit. An inlet entry zone, which is shaped as a wedge, also forms. The outer layers of this wedge move inward, and a return flow of lubricant (in the form of an eddy current) occurs between the surface layers (Fig. 15a). At low rolling speeds, the entry zone wedge is negligibly small. With increasing rolling speeds (or increasing liquid viscosity or increasing values of the Sommerfeld number), the wedge increases both in thickness and depth. The Sommerfeld number (S) is:

$$S = \frac{\text{Viscosity} \times \text{Strip exit velocity}}{\sigma_0 h_f} \quad \text{(Eq 29)}$$

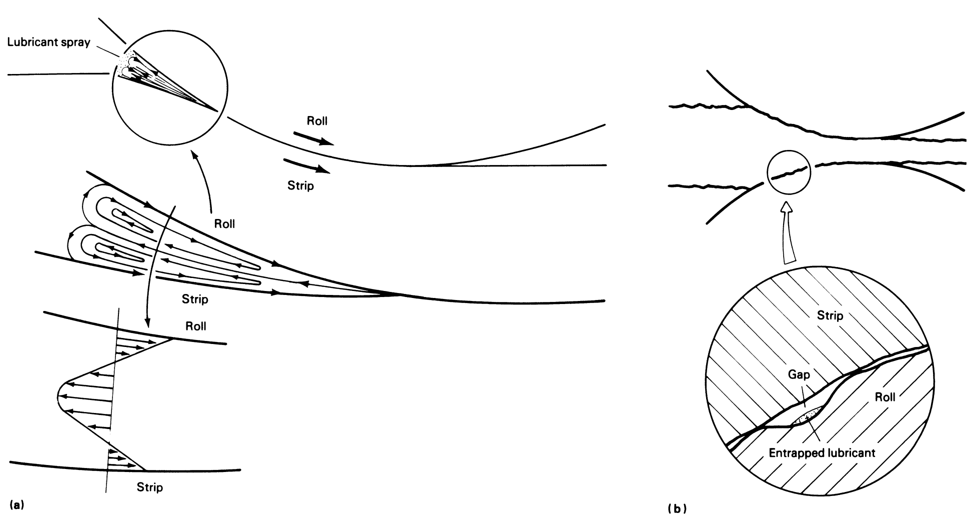


Fig. 15 (a) The lubrication entry zone and (b) surface irregularities associated with strip rolling

where h_f is the final (exit) strip thickness, and σ_0 is the flow stress. With increasing values of S, the point where metal-to-metal contact between the strip and the rolls is established moves further toward the exit.

The surfaces of both rolls and strip are not perfectly smooth surfaces in that they contain irregularities in the form of peaks (or crests) and valleys (cavities) (Fig. 15b). Some lubricant passes from the entrance to the exit side through the labyrinth of channels created by these irregularities. As the strip is deformed, the crests are flattened and the entrapped lubricant is pressurized in the diminishing volume of the cavities. At slow rolling speeds, the excess lubricant in the diminishing cavity space is squeezed to flow back into the entry zone. At higher rolling speed, the escape of excess fluid from the diminishing gap of the labyrinth between the rolls and the strip is relatively slower. The entrapped lubricant is then pressurized and causes partial separation between the rolls and the strip. At low speed, pressure is transmitted from the rolls to the strip through metal-to-metal contact. At higher speeds, more of the pressure is transmitted through the entry zone and through the entrapped lubricant. As the pressure transmitted through metal-to-metal contact is reduced with higher speeds, the friction is decreased. Eventually, at high speeds, no metal-to-metal contact exists.

When conditions for complete hydrodynamic lubrication are reached (Fig. 16), friction is at its minimum value. Therefore, with increased speeds, high shear rates are created in the liquid where the shear stress is proportional to the shear strain rate to the second power, and friction rises at a mild rate. Friction values are much lower when hydrodynamic lubrication prevails than when metal-to-metal contact is prevalent.

With increasing speed, the friction hill effect is also reduced, and roll-separation force and roll flattening become less pronounced. With reduced roll-separation force, the elastic stretching of the mill as well as roll bending and roll flattening are reduced, causing the gap between the rolls to decrease. With the increasing thickness of the lubricant film and reduced mill flexing, the thickness of the emerging strip reduces. The actual gap between the rolls is larger than the thickness of the strip by twice the film thickness. Because friction drag decreases with increasing rolling speed, the neutral point approaches the exit, and forward slip is reduced.

Two critical points may now be reached. The first is skidding due to insufficient friction drag, while the second is the establishment of hydrodynamic lubrication. If the strip is thin enough, hydrodynamic lubrication is reached first, and skidding will not develop. When speed continues to increase after hydrodynamic film lubrication is established, friction drag and forward slip, which have already reached their respective minimum points, begin to increase (lines 1 to 4 in Fig. 17). For increasing values of strip thickness (lines 2, 3, and 4 in Fig. 17), the forward slip decreases, and the critical value of roll speed at which the neutral angle and forward slip reach a minimum is increased. The values of the neutral angle and forward slip at which hydrodynamic lubrication begins increase with increasing thickness values.

At a critically high strip-thickness value, hydrodynamic lubrication commences when forward slip and the value of the neutral angle are zero (line 4 in Fig. 17). When the thickness of the strip is above the stated critical value, for lines 5 to 7, friction drops to below the minimum required for rolling before hydrodynamic lubrication can be established. Skidding then commences at the critical roll speed. It can be observed that at higher values of speed, referred to here as the minimum required speed, rolling with hydrodynamic lubrication may be reestablished. It should also be noted that the minimum required speed can be reduced by the use of lubricants of higher viscosity.

The following variations in the friction hill are expected with increasing speed:

- The peak of the friction hill gets lower and, together with the neutral point, shifts closer to the exit.
- The entry zone expands, and the corner of the liquid wedge moves further away from the entrance and closer to the exit. The meaningful slope of the friction hill on the entrance side starts at the area around the tip of the entry zone.
- The roll separation force gets lower.
- Roll bending/flattening and mill stretching are reduced.
- Strip thickness is reduced.
- For relatively thick strip (or low dry-friction value m_0), a critical roll speed of the first kind may be reached. When this occurs, friction becomes so low (below the critical value required) that the neutral point is at the exit, and skidding will commence (lines 5 and 6). At the moment that skidding begins, the strip stops.
- For relatively thin strip, another critical speed (the critical speed of the second kind) can be reached, even before skidding commences. When this occurs, the point of the entry zone will reach the exit, and hydrodynamic lubrication will commence. A further increase in speed will cause a thicker lubricating film to separate the rolls from the strip and make the rolling more stable (lines 1, 2, and 3 in Fig. 17).
- For relatively thick strip, even when a critical speed of the first kind commences first, a further increase in speed may ultimately produce a critical speed of the second kind and reestablish hydrodynamic lubrication (lines 4 to 7 in Fig. 17).

Changes in the entry zone, the friction hill, the neutral point, and strip thickness as a function of speed for thick and thin strip are shown schematically in Fig. 18.

Mechanics of Plate Rolling

In rolling of thick plates, metal flow occurs in three dimensions. The rolled material is elongated in the rolling direction as well as spread in the lateral or width direction. Spread in rolling is usually defined as the increase in width of a plate

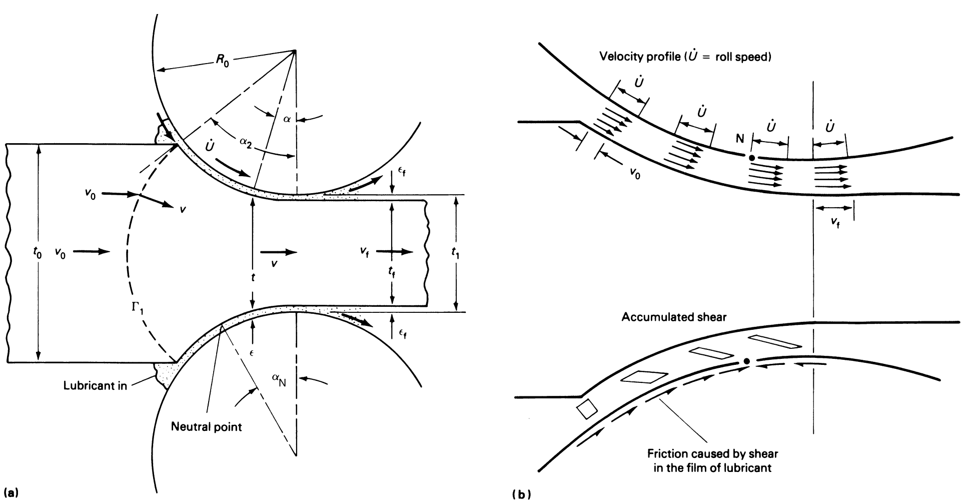


Fig. 16 Hydrodynamic lubrication during strip rolling. (a) Overall schematic. (b) Shear in the lubricant film

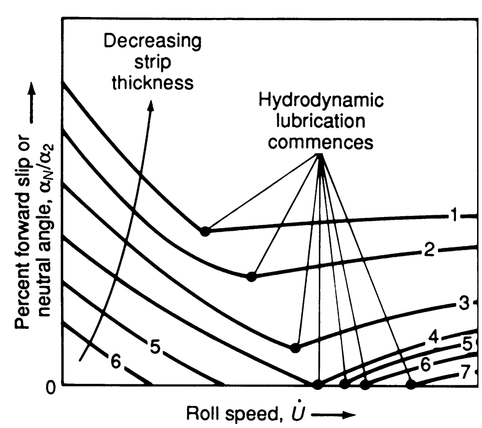


Fig. 17 Forward slip and position of the neutral point versus rolling velocity. See text for explanation of numbered curves.

or slab expressed as a percentage of its original width. The spread increases with increasing reduction and interface friction, decreasing plate width-to-thickness ratio, and increasing roll-diameter-to-plate-thickness ratio. In addition, the free edges tend to bulge with increasing reduction and interface friction. The three-dimensional metal flow that occurs in plate rolling is difficult to analyze. Therefore, most studies of this process have been experimental in nature. Several empirical formulas have been established for estimating spread (Ref 42–44), with some theoretical attempts to predict elongation or spread (Ref 45–47). Once the spread has been estimated, the elongation can be determined from the volume constancy, or vice versa.

An Empirical Method for Estimating Spread. Among the various formulas available for predicting spread, one formula (Ref 43) is used most extensively and is given as:

$$w_1/w_0 = abcd(h_0/h_1)^P \quad \text{(Eq 30)}$$

where w_1 and w_0 are the final and initial widths of the plate, respectively; h_1 and h_0 are the final and initial thicknesses of the plate, respectively; P equals $10^{(-1.269)}(w_0/h_0)(h_0/D)^{0.556}$; D is the effective roll diameter; and *a, b, c,* and *d* are constants that allow for variations in steel composition, rolling temperature, rolling speed, and roll material, respectively. These constants vary slightly from unity, and their values can be obtained from the literature (Ref 38, 43, 47).

An empirical formula for predicting spread, such as Eq 30, gives reasonable results within the range of conditions for the experiments from which the formula was developed. There is no formula that will make accurate predictions for all the conditions that exist in rolling. Thus, it is often necessary to attempt to estimate spread or elongation by theoretical means.

The theoretical prediction of spread involves a rather complex analysis and requires the use of computerized techniques (Ref 38, 45, 46). A modular upper-bound method has been used to predict metal flow, spread, elongation, and roll torque (Ref 38). The principles of this method are described subsequently. Figure 19 (Ref 38) illustrates the coordinate system, the division of the deformation zone into elements, and the notations used. The spread profile is defined in terms of a third-order polynomial $w(x)$, with two unknown coefficients, a_1 and a_2. The location of the neutral plane x_N is another unknown quantity. The following kinematically admissible velocity field (Ref 48) is used:

$$v_x = 1/[w(x)h(y)] \quad \text{(Eq 31)}$$

$$v_y = \frac{1}{h(x)}\frac{d}{dx}\left[\frac{1}{w(x)}\right] \quad \text{(Eq 32)}$$

$$v_z = \frac{1}{w(x)}\frac{d}{dx}\left[\frac{1}{h(x)}\right] \quad \text{(Eq 33)}$$

Using Eq 31 to 33, the upper-bound method can be applied to predict spread. A computer program, SHPROL, can be used for some steps in the analysis. More information on SHPROL is available in the section "Shape Rolling" in this chapter.

Prediction of Stresses and Roll-Separating Force. Once the spread (the boundaries of the deformation zone) has been calculated, this information can be used to predict the stresses and the roll-separating force. The computerized procedure used here is, in principle, the same as the method discussed earlier for predicting the stresses in strip rolling (Ref 38).

The deformation zone under the rolls is divided into trapezoidal slabs by planes normal to the rolling direction and along the stream tubes, as illustrated in Fig. 13 and 20 (Ref 38). The stresses acting on strips in the rolling and transverse directions are illustrated in Fig. 20(b) and (c), respectively. As expected from the slab analysis, the stress distributions are very similar to those illustrated for strip rolling in Fig. 12 to 14. By use of a numerical approach similar to that discussed for strip rolling, detailed predictions of stresses, in both the longitudinal and lateral directions, can be made. The stresses are calculated by assuming the frictional shear stress τ to be constant, as in the case of upper-bound analysis. Thus, the stress distribution at various planes along the width, or y, direction (Fig. 20) is linear on both sides of the plane of symmetry. The stress distribution in the rolling, or x, direction is calculated along the streamlines of metal flow (Fig. 19). At each node of the mesh, the lower of the σ_z values is accepted as the actual stress. Thus, a tentlike stress distribution is obtained (Fig. 21). Integration of the stresses acting on the plane of contact gives the roll-separating force.

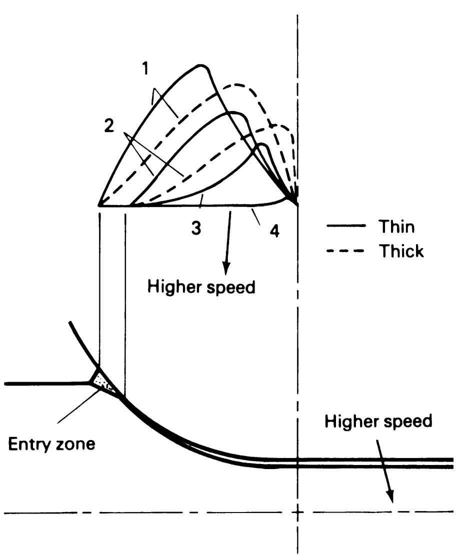


Fig. 18 Variations in the friction hill versus speed. The parameters m, R_0, σ_{xb}, σ_{xf}, and percent reduction are constant.

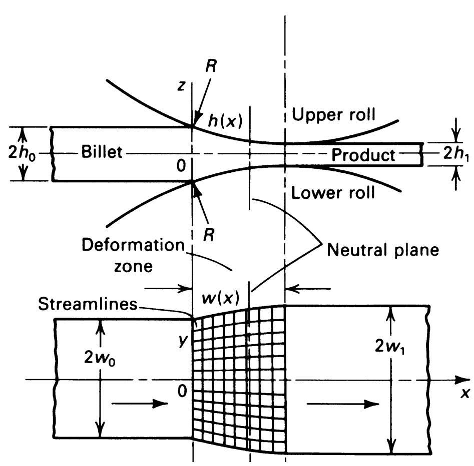


Fig. 19 Configuration of deformation and the grid system used in the analysis of the rolling of thick plates. Source: Ref 38

Shape Rolling

Rolling of shapes, also called caliber rolling, is one of the most complex deformation processes. A round or round-cornered square bar, billet, or slab is rolled in several passes into relatively simple sections, such as rounds, squares, or rectangles, or into complex sections, such as "L," "U," "T," "H," or other irregular shapes (Ref 49). For this purpose, certain intermediate shapes or passes are used, as shown in Fig. 22 for the rolling of angle sections (Ref 50). The design of these intermediate shapes, that is, roll pass design, is based on experience and differs from one company to another, even for the same final rolled section geometry.

The art of roll pass design, known since the late 17th century, has turned into a modern technology based on scientific foundations (Ref 51). The purpose of roll pass design is to ensure the production of the correct size and shape of a product with defect-free surface and intended mechanical properties, and at the same time ensure maximum output at lowest cost, ease the working conditions of the rolling crew, and minimize the roll wear. In the past, relatively few quantitative data on roll pass design have been available in the literature. Today, computer-aided roll design and FEMs are established for design of roll passes with computer numerical control machines for roll turning and computer-integrated roll manufacturing. In the future, roll pass design is expected to become more advanced for thermomechanical processing (Ref 51).

Basically, there are two methods for rolling shapes or sections. The first method is universal rolling (Fig. 23). The second method is caliber rolling (Fig. 22, 24). In universal rolling, the mill and stand constructions are more complex. However, in the rolling of I-beams or other similar sections, this method allows more flexibility than does caliber rolling and requires fewer passes. This is achieved because this method provides appropriate amounts of reductions, separately in webs and flanges.

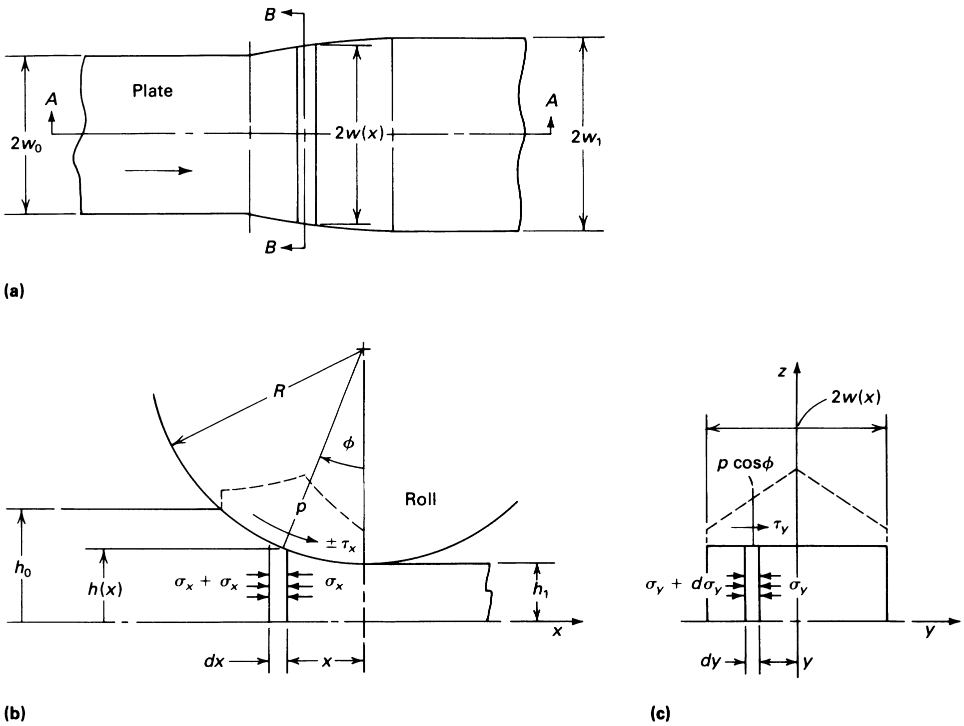


Fig. 20 Stress analysis of the rolling of plates. (a) Top view of the rolled plate. (b) Stresses in the rolling direction. (c) Stresses in the transverse direction. Source: Ref 38

For successful rolling of shapes, it is necessary to estimate the following for each stand: the roll-separating force and torque, the spread and elongation, and the appropriate geometry of the roll cavity or caliber. The force and torque can be estimated either by using empirical formulas or by approximating the deformation in shape rolling with that occurring in an equivalent plate rolling operation. In this case, the equivalent plate has initial and final thicknesses that correspond to the average initial and final thicknesses of the rolled section. The load and torque calculations can be performed for the equivalent plate, as discussed earlier in this chapter for plate rolling. The results are approximately valid for the rolled shape being considered.

Estimation of Elongation. During the rolling of a given shape or section, the cross section is not deformed uniformly, as can easily be seen in Fig. 24. This is illustrated further in Fig. 25 for a relatively simple shape. The reductions in height for zones A and B are not equal (Fig. 25a). Consequently, if these two zones, A and B, were completely independent of each other (Fig. 25b), zone B would be much more elongated than zone A. However, the two zones are connected and, as part of the rolled shape, must have equal elongation at the exit from the rolls. Therefore, during rolling, metal must flow from zone B into zone A so that a uniform elongation of the overall cross section is obtained (Fig. 25c). This lateral flow is influenced by the temperature differences that exist in the cross section because of variations in material thickness and heat flow.

To estimate the overall elongation, it is necessary to divide the initial section into a number of equivalent plates (A, B, C, and so forth), as shown in Fig. 25. The elongation for an individual section, without the combined influence of other portions of the section, can be estimated by using both the plate rolling analogy and the techniques discussed in this chapter. The combined effect can be calculated by taking a weighted average of the individual elongations. For example, if the original section is to be divided into an equivalent system consisting of two plate sections (A and B in Fig. 25) with individual cross-sectional areas A_a and A_b, then the following weighted-average formula can be used:

$$\lambda_m = \frac{A_0}{A_1} = \frac{A_{a0} + A_{b0}}{A_{A1} + A_{b1}} = \frac{A_{a1}\lambda_a + A_{b1}\lambda_b}{A_{a1} + A_{b1}} \quad \text{(Eq 34)}$$

where λ is the elongation coefficient (that is, the cross-sectional area at the entrance divided by the cross-sectional area at the exit); A is the cross-sectional area; m is a subscript denoting average; a and b are subscripts denoting section portions A and B; and 0 and 1 are subscripts denoting entrance and exit values, respectively.

Computer-Aided Roll Pass Design. Estimation of the number of passes and of the roll geometry for each pass is the most difficult aspect of shape rolling. To accomplish this, certain factors, discussed subsequently, must be considered.

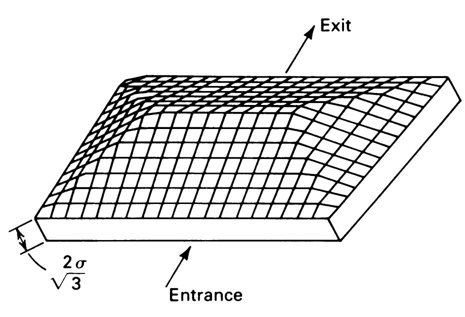


Fig. 21 The calculated stress (σ_z) distribution in plate rolling shown three-dimensionally. Source: Ref 38

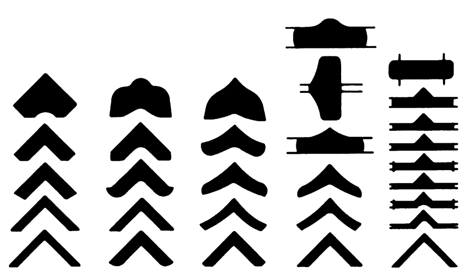


Fig. 22 Five possible roll pass designs for the rolling of a steel angle section. Source: Ref 50

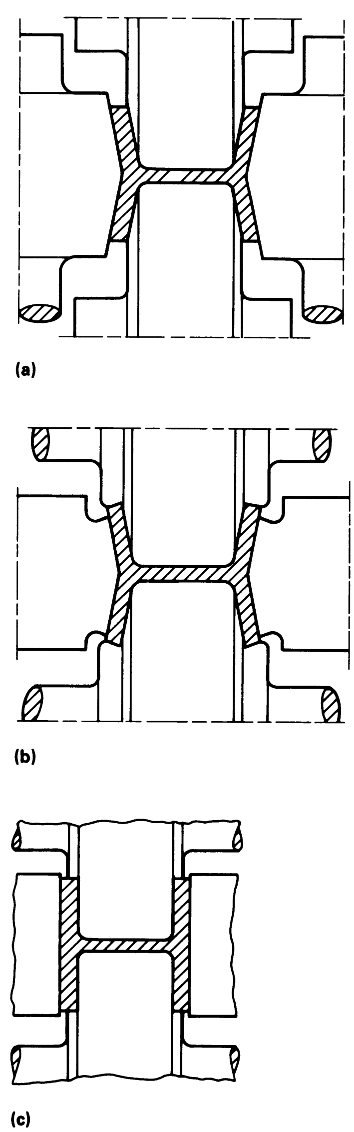


Fig. 23 Universal rolling of flanged beams. (a) Universal roll stand. (b) Edging stand. (c) Finishing stand

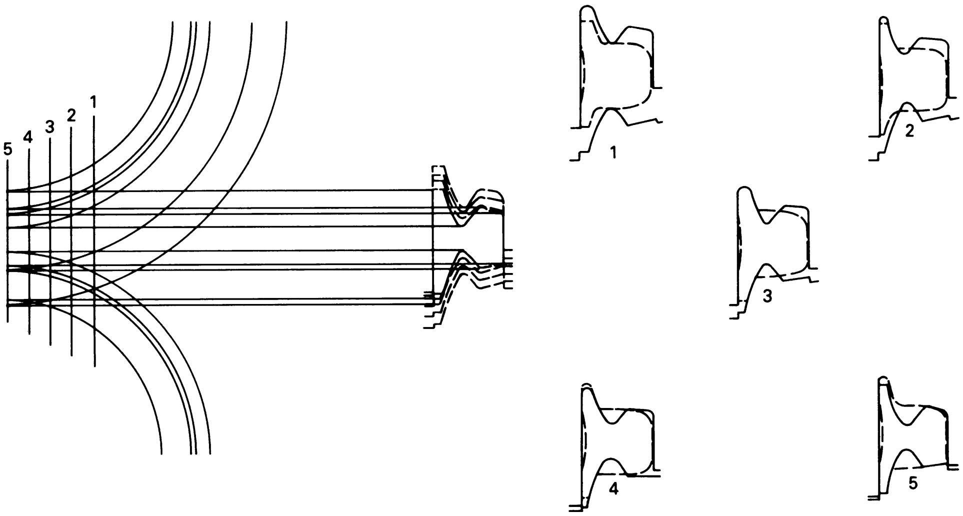


Fig. 24 Analysis of a roll stand used in rolling of rails. Sketches 1 through 5 illustrate the stock in broken lines and the roll in solid lines at various positions in the deformation zone.

The Characteristics of the Available Installation. These include diameters and lengths of the rolls, bar dimensions, distance between roll stands, distance from the last stand to the shear, and tolerances that are required and that can be maintained.

The reduction per pass must be adjusted so that the installation is used at a maximum capacity, the roll stands are not overloaded, and roll wear is minimized. The maximum value of the reduction per pass is limited by the excessive lateral metal flow, which results in edge cracking; the power and load capacity of the roll stand; the requirement for the rolls to bite in the incoming bar; roll wear; and tolerance requirements.

At the present stage of technology, the previously mentioned factors are considered in roll pass design by using a combination of empirical knowledge, some calculations, and some educated guesses. A methodical way of designing roll passes requires not only an estimate of the average elongation, as discussed earlier, but also the variation of this elongation within the deformation zone. The deformation zone is limited by the entrance, where a prerolled shape enters the rolls, and by the exit, where the rolled shape leaves the rolls. This is illustrated in Fig. 24. The deformation zone is cross-sectioned with several planes (for example, planes 1 to 5 in Fig. 24; 1 is at the entrance, 5 is at the exit). The roll position and the deformation of the incoming billet are investigated at each of these planes. Thus, a more detailed analysis of metal flow and an improved method for designing the configuration of the rolls are possible. It is evident that this process can be drastically improved and made extremely efficient by the use of computer-aided techniques.

In recent years, most companies that produce shapes have computerized their roll pass design procedures for rolling rounds or structural shapes. In most of these applications, the elongation per pass and the distribution of the elongation within the deformation zone for each pass are predicted by using an empirical formula. If the elongation per pass is known, it is then possible, by use of computer graphics, to calculate the cross-sectional area of a section for a given pass, that is, the reduction and the roll geometry. The roll geometry can be expressed parametrically (in terms of angles, radii, and so forth). These geometric parameters can then be varied to optimize the area reduction per pass and obtain an acceptable degree of fill of the roll caliber used for that pass.

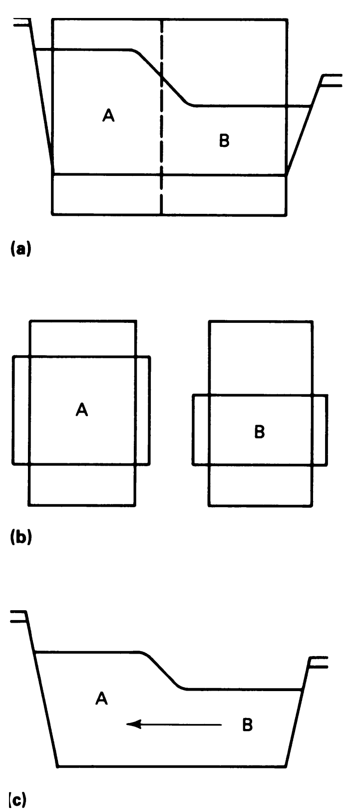


Fig. 25 Nonuniform deformation in the rolling of a shape. (a) Initial and final sections. (b) Two zones of the section considered as separate plates. (c) Direction of lateral metal flow

Computer-Aided Roll Pass Design of Airfoil Sections. Two examples of computer programs developed for the analysis of rolling airfoil sections are the SHPROL and ROLPAS programs (Ref 39). The first of these programs, SHPROL, uses upper-bound analysis in a numerical form to predict spread and roll torque. SHPROL is based on the following simplifying assumptions:

- The initial contact between the rolls and the entrance section can be approximated as a straight line. (This is only correct if the upper and lower surfaces of the initial section already have the shape of the rolls.)
- An airfoil shape can be considered as an aggregate of slabs, as shown in Fig. 26 (Ref 39).
- Plane sections perpendicular to the rolling direction remain plane during rolling. Thus, the axial velocity (velocity in the rolling, or *x*, direction) at any section perpendicular to the rolling direction is uniform over the entire cross section.
- The velocity components in the transverse, or *y*, direction and in the thickness, or *z*, direction are functions of *x* and linear in the *y* and *z* coordinates, respectively.

In Fig. 26, each element is considered to be a plate for which it is possible to derive a kinematically admissible velocity field. The total energy dissipation rate of the process $\dot{E}_T$ is:

$$\dot{E}_T = \dot{E}_P + \dot{E}_D + \dot{E}_F \qquad \text{(Eq 35)}$$

where $\dot{E}_P$ is the energy rate of plastic deformation and is calculated for each element by integrating the product of flow stress and the strain rate over the element volume, $\dot{E}_D$ represents the energy rates associated with velocity discontinu-

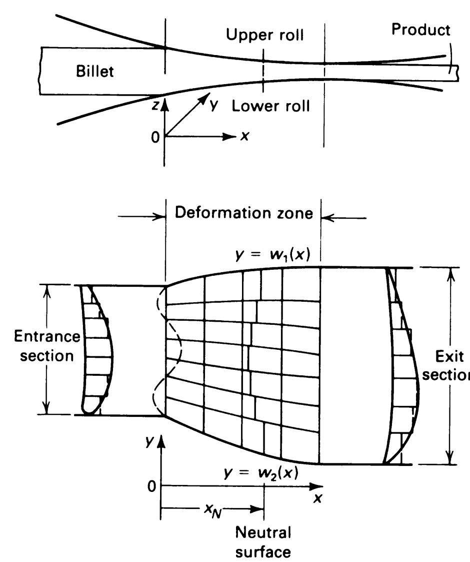


Fig. 26 Configuration of deformation zone in the application of numerical upper-bound analysis to the rolling of airfoil shapes. Source: Ref 39

ities and is due to internal shear between the elements, and $\dot{E}_F$ is the energy rate due to friction between the rolls and the deforming material.

The total energy dissipation rate $\dot{E}_T$ is a function of unknown spread profiles w_1 and w_2 (Fig. 26) and the location of the neutral plane x_N. As in the analysis discussed earlier for plate rolling, the unknown coefficients of w_1, w_2, and x_N are determined by minimizing the total energy rate. The computer program SHPROL uses the following as input data: roll and incoming shape geometry, friction, flow stress, and roll speed. SHPROL can predict the energy dissipation rates, the roll torque, and, most important, the amounts of elongation and spread within one deformation zone in the rolling of any airfoil shape.

The second program, called ROLPAS, uses interactive graphics and is capable of simulating the metal flow in the rolling of relatively simple shapes, such as rounds, plates, ovals, and airfoils (Fig. 27). ROLPAS uses the following as input: the geometry of the initial section, the geometry of the final section, the flow stress of the rolled material and the friction factor, and the variations in elongation and spread in the rolling direction, as calculated by the SHPROL program.

To simulate the rolling process, ROLPAS divides the deformation zone into a number of cross sections parallel to the roll axis (Fig. 19, 27). The simulation is initiated by considering the cross-sectional area, stresses present, and the roll-separating force and torque for the first section. These same analyses can then be performed on any succeeding section.

Computer-Aided Roll Pass Design for Round Sections. Computer-aided methods for designing caliber rolls for rod rolling have been used since the1960s and 1970s (Ref 52–55), and advances continue to be made in improving product quality and reducing dimensional variability. One example, based on a simplified finite-element-slab method, is discussed in the section "Modeling of Strip Rolling" in this chapter. Methods of experimental design are also used with the FEM of process simulation to arrive at robust roll pass designs (Ref 56).

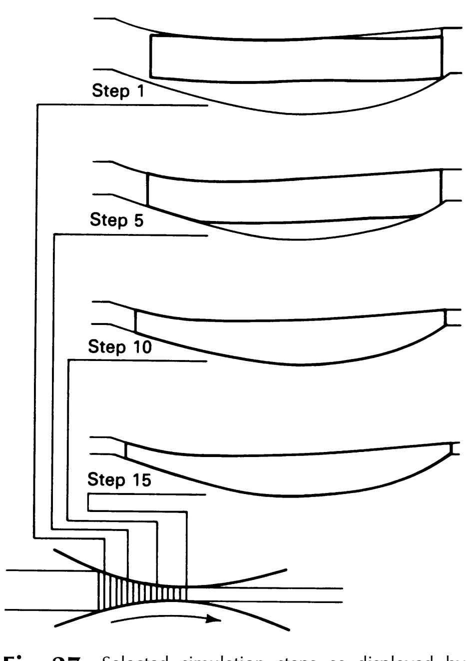


Fig. 27 Selected simulation steps as displayed by ROLPAS for a test airfoil shape cold rolled from rectangular steel stock

One example of a computer program called for establishing roll cross sections and pass schedules is RPDROD (Ref 57). RPDROD uses an empirical formula for estimating the variation of the spread in the roll bite and parametrically described alternative roll caliber designs. When using this program, the designer obtains an optimal roll pass schedule by evaluating a number of alternatives in which individual pass designs are selected from a variety of caliber shapes commonly used in rod rolling.

The computer program RPDROD consists of four modules, called Stock, Schedule, Groove, and Metal Flow. The Stock design module allows the user to design/specify the entry cross section for the first pass in the schedule. A square, rectangular, or round stock cross section can be defined. The Schedule design module allows the user to design the roll pass schedule by providing specific functions:

- Add a new pass to the roll pass schedule, by estimating alternative roll cross section dimensions from design data provided by spread/elongation calculations
- Delete pass design data from the schedule in order to investigate alternative pass designs
- Review and/or provide hard copy of existing pass design data

The Schedule design module allows the user to design an optimal roll pass schedule by investigating various alternative pass design and/or shape combinations. In principle, any roll cross-sectional shape considered by the program could be used for a given pass in the schedule. However, RPDROD has facilities for checking input data and thus for preventing the selection of an illogical pass design or the inappropriate selection of roll cross-sectional shape combinations.

The Groove design module can be used to change the initially suggested roll cross section dimensions, as the user deems appropriate. As in the Schedule module, input-checking facilities ensure that specified roll cross-sectional dimensions are consistent with the chosen roll cross-sectional shape and bar entry cross section.

The Metal Flow design module provides the user with details of metal-flow simulation, including the calculated cross sections of the deforming bar in the roll bite, stresses in the deforming material, roll-separating load, and roll torque. For this purpose, this module uses the ROLPAS program discussed earlier for the rolling of airfoil shapes.

Computer-Aided Roll Pass Design of Structural and Irregular Sections. Many companies use computer-aided graphics for the design and manufacture of the caliber shapes for the rolling of structural sections. Reference 58 gives an excellent summary of the practical use of computer graphics for roll caliber and roll pass design. In this case, the cross section of a rolled shape is described in general form as a polygon. Each corner or fillet point of the polygon is identified with the *x*- and *y*-coordinates and with the value of the corresponding radius (Fig. 28) (Ref 58). Thus, any rolled section can be represented by a sequence of lines and circles. This method of describing a rolled section is very general and can define a large number of sections with a single computer program. Lines or circles that are irrelevant in a specific case can be set equal to zero. Thus, a simpler section, with a smaller number of corner and fillet points, can be obtained. For example, in the rolling of the symmetric angle section shown in Fig. 22, several intermediate section passes are required. Such an intermediate section is shown in parametric representation in Fig. 29. In this figure, all the geometric variables can be modified to change the cross-sectional area and/or the amount of reduction per pass. These variables, which fully describe this section, are:

- SELA, length (of one leg) at centerline
- BETAG, angle at top corner
- RK, radius at top corner
- AL, length of straight portion at top
- RD, radius of leg at top
- PRST, projection of draft angle
- RRU, radius at lower tip of leg
- RH, radius at bottom corner

In establishing the final section geometry, the designer assigns desired values to the variables listed previously and, in addition, inputs the desired cross-sectional area and the degree of caliber fill, for example, the desired ratio of rolled section area versus section area on the caliber rolls. Thus, there is only one geometric variable

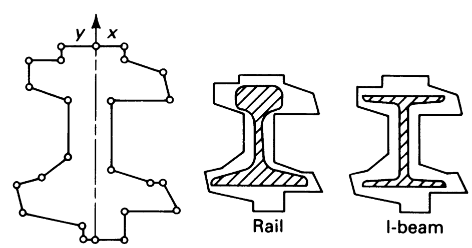


Fig. 28 Geometric representation of a rolled section as a polygon. Source: Ref 58

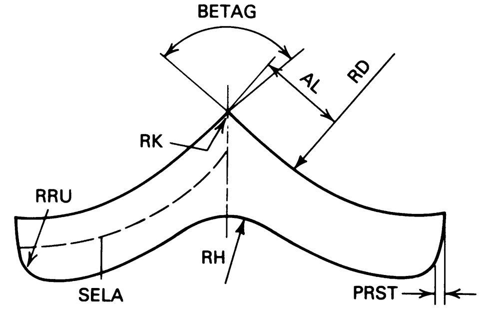


Fig. 29 Parametric representation of an intermediate-shaped rolling pass for a symmetric angle section. See also Fig. 30. Source: Ref 58

that is calculated by the computer program, and that is the thickness of the leg of the angle section. In the example shown in Fig. 30(a) (Ref 58), the leg thickness is calculated to be 18.2 mm (0.72 in.). The designer compares this section geometry (Fig. 30a) with the caliber geometry of the next pass that has been generated in a similar way. Assume that the section shown in Fig. 30(a) appears to be too long; that is, SELA is 67.5 mm (2.66 in.) and should be reduced to 65 mm (2.56 in.) without modifying the other variables. The interactive program is rerun with the new value for SELA. The modified section, shown in Fig. 30(b), is slightly thicker than the original section in order to maintain the same cross-sectional area.

This interactive graphics program does not involve any analysis of metal flow or stresses. Nevertheless, it is extremely useful to the designer for modifying section geometries quickly and accurately, calculating cross-sectional areas, and cataloging all this geometrical information systematically. The program also automatically prepares engineering drawings of the sections and the templates for quality control as well as tapes for numerically controlled milling of the templates and the graphite electrical discharge machining electrodes used in manufacturing the necessary cutting tools for roll machining (Ref 58).

Finite-element modeling (FEM) is used in the analysis of three-dimensional (3-D) metal flow and of stress and strain distributions in shape rolling; the laborious task of designing rolls for complex shapes can be simplified by FEM. Continued improvements in the speed of computer calculations enhance the use of FEM, and computer programs have been developed from the simplification of 3-D numerical methods to simulate the shape-rolling process. For example, shape rolling can be modeled by a finite- and slab-element method (FSEM), which uses a rigid-plastic two-dimensional FEM for the generalized plane-strain condition combined with the slab method (Ref 59). This reduces the computational effort without losing much accuracy obtained in the 3-D computer simulation of the shape-rolling process. The FSEM has been used to develop a computer program, called TASKS, for 3-D analysis of shape rolling as a kinematically steady-state process. The TASKS program has been used to simulate the metal flow and the bulge profile in flat rolling of slabs, the shape rolling of a simple "H" section, and the rolling of a practical H-beam section. The effect of roll pass sequence design on the roll forces and the quality (distribution of strain) of the rolled product can be studied. With reduced simulation time per pass, the code has potential use in the redesign of roll pass schedules for improved product properties (Ref 60).

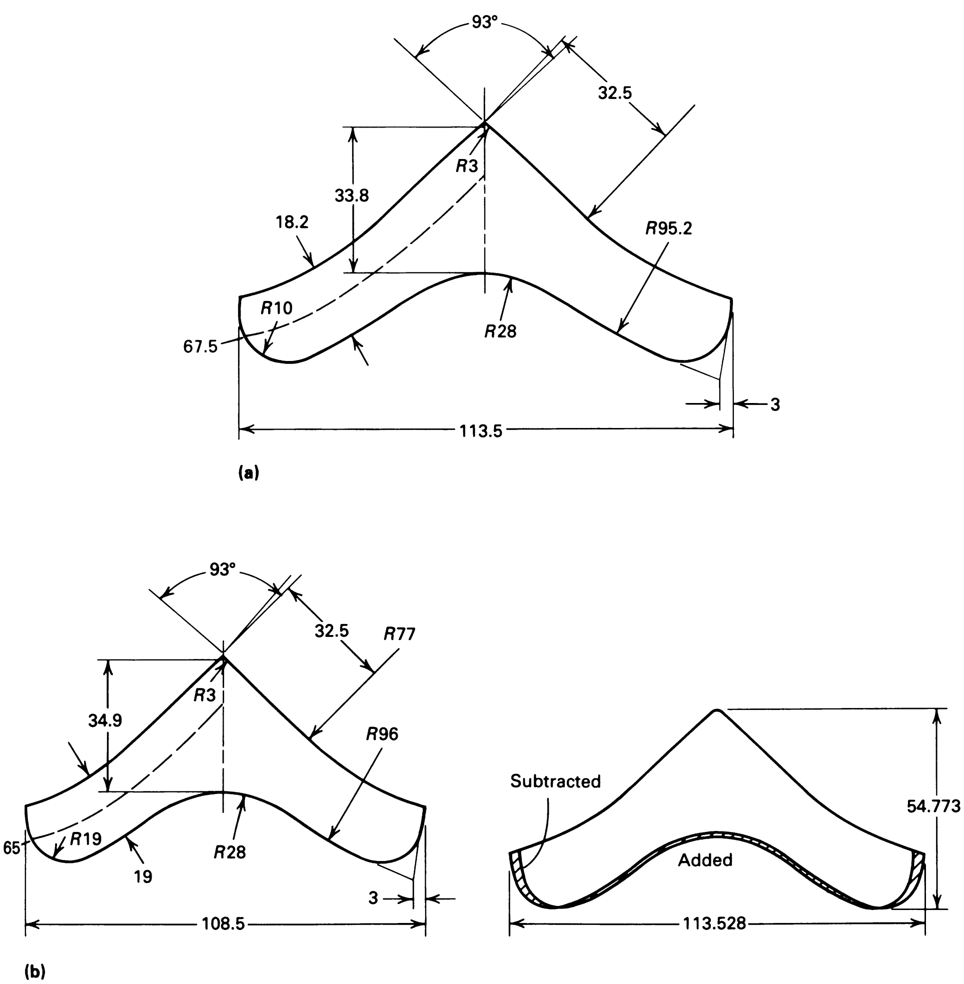


Fig. 30 (a) Alternative design for an intermediate-angle pass obtained by assigning values to the parameters given in Fig. 29. (b) Modified design of the intermediate-angle pass shown in (a), with new dimensions (left) and added and subtracted areas (right). Source: Ref 58

Summary

This chapter briefly introduces the mechanics of rolling with descriptions of plane-strain rolling, strain distributions, and lubrication. Some friction is needed to draw the pieces into the roll gap, and frictional stresses increase in the deformation zone until interface pressure reaches an apex roughly at the neutral point where the surface speed of the workpiece and roll are equal. The mechanics of this process are briefly described along with the factors that influence the homogeneity of deformation in strip, plate, and shape rolling. Friction, together with pass geometry, determines the homogeneity of deformation in several ways, as discussed in this chapter. The idealized case of plane-strain rolling is introduced first, and strip-rolling theory is described in more detail. Plate rolling mechanics and shape rolling are also discussed.

ACKNOWLEDGMENTS

This article was adapted from *Tribology in Metalworking*, by John Schey, American Society for Metals, 1983, p 249–333 and from "Flat, Bar, and Shape Rolling," by G.D. Lahoti and S.L. Semiatin, *Forming and Forging,* Volume 14 of the *ASM Handbook,* ASM International, 1988, p 343–360.

REFERENCES

1. J. Schey, Rolling, *Tribology in Metalworking*, American Society for Metals, 1983, p 249
2. G.D. Lahoti and S.L. Semiatin, Flat, Bar, and Shape Rolling, *Forming and Forging,* Vol 14, *ASM Handbook,* 1988, p 343–360
3. J. Schey, Friction in Metalworking, *Tribology in Metalworking*, American Society for Metals, 1983, p 15
4. S. Ekelund, *Steel,* Vol 93 (Aug 21), 1933, p 27–29
5. H. Ford, F. Ellis, and D.R. Bland, *J. Iron Steel Inst.*, Vol 168, 1951, p 57–72
6. W.L. Roberts, *Cold Rolling of Steel,* Dekker, 1978
7. H. Inhaber, *J. Eng. Ind. (Trans. ASME) B,* Vol 88, 1966, p 421–429
8. W. Evans and B. Avitzur, *J. Lubr. Technol. (Trans ASME) F,* Vol 90, 1968, p 72–80
9. W. Dobrucki and E. Odrzywolek, *J. Mech. Work. Technol.,* Vol 4, 1980, p 263–284 (in French)

10. J.M. Capus and M.G. Cockroft, *J. Inst. Met.,* Vol 92, 1963–1964, p 31–32
11. B.B. Hundy and A.R.E. Singer, *J. Inst. Met.,* Vol 83, 1954–1955, p 401–407
12. E. Siebel and W. Lueg, *Arch. Eisenhüttenwes.,* Vol 15, 1933, p 1–14 (in German)
13. G.D. Lahoti, N. Akgerman, S.I. Oh, and T. Altan, *J. Mech. Work. Technol.,* Vol 4, 1980, p 105–119
14. M. Vater, G. Nebe, and J. Petersen, *Stahl Eisen,* Vol 86, 1966, p 710–720 (in German)
15. T. von Karman and Z. Angew, *Mat. Mech.,* Vol 5, 1925, p 139–141
16. E. Orowan, The Calculation of Roll Pressure in Hot and Cold Flat Rolling, *Proc. Inst. Mech. Eng.,* Vol 150, 1943, p 140
17. J.T. Hockett, Calculation of Rolling Forces Using the Orowan Theory, *Trans. ASM,* Vol 52, 1960, p 675
18. R.B. Sims, The Calculation of Roll Force and Torque in Hot Rolling Mills, *Proc. Inst. Mech. Eng.,* Vol 168, 1954, p 191
19. B. Avitzur, An Upper-Bound Approach to Cold Strip Rolling, *J. Eng. Ind. (Trans. ASME),* Feb 1964, p 31
20. E.G. Thomsen, C.T. Yang, and S. Kobayashi, in *Mechanics of Plastic Deformation in Metal Processing,* Macmillan, 1965, p 373
21. G.W. Rowe, in *Principles of Industrial Metalworking Processes,* Edward Arnold, 1968, p 208
22. J.M. Alexander, On the Theory of Rolling, *Proc. R. Soc. (London) A,* Vol 326, 1972, p 535
23. G.D. Lahoti, S.N. Shah, and T. Altan, Computer Aided Analysis of the Deformations and Temperatures in Strip Rolling, *J. Eng. Ind. (Trans. ASME),* Vol 100, May 1978, p 159
24. H. Ford and J.M. Alexander, Simplified Hot-Rolling Calculations, *J. Inst. Met.,* Vol 92, 1963–1964, p 397
25. D.J. McPherson, Contributions to the Theory and Practice of Cold Rolling, *Metall. Trans.,* Vol 5, Dec 1974, p 2479
26. G.J. Li and S. Kobayashi, *J. Eng. Ind. (Trans. ASME) B,* Vol 104, 1982, p 55–64
27. P.F. Thomson and J.H. Brown, *Int. J. Mech. Sci.,* Vol 24, 1982, p 559–576
28. R.H. Wagoner and J.-L. Chenot, in *Fundamentals of Metal Forming,* John Wiley & Sons, 1997
29. R.H. Wagoner and J.-L. Chenot, *Metal Forming Analysis,* Cambridge University Press, 2001, p 210–212
30. V.-K. Hu and W.K. Liu, An ALE Hydrodynamic Lubrication Finite Element Method with Application to Strip Rolling, *Int. J. Numer. Methods Eng.,* Vol 36 (No. 5), March 1993, p 855–880
31. P. Gratacos, C. Fromholz, J.L. Chenot, and P. Montmitonnet, A Plane-Strain Elasto-plastic Finite-Element Model for Cold Rolling of Thin Strip, *Int. J. Mech. Sci.,* Vol 34 (No. 3), March 1992, p 195–210
32. C.G. Sun and S.M. Hwang, A Finite Element Computer System for Analysis of Mechanical and Thermal Behaviors in Hot Strip Rolling, *Second International Conference on Modelling of Metal Rolling Processes,* 9–11 Dec 1996 (London), Institute of Materials, 1996, p 84–91
33. Z.-C. Lin and C.-C. Shen, A Three-Dimensional Finite Element Method for a Nonisothermal Aluminum Flat Strip Rolling Process, *J. Mater. Eng. Perform.,* Vol 5 (No. 4), Aug 1996, p 452–461
34. K. Narita and K. Yasuda, Three-Dimensional Rigid-Plastic Finite Element Analysis of Strip Crown and Flatness During Rolling with Constraint Boundary Conditions Proposed by Shohet, METEC Congress 94, *Second European Continuous Casting Conference and Sixth International Rolling Conference,* Vol 2, 20–22 June 1994, Verein Deutscher Eisenhuttenleute (Dusseldorf), 1994, p 430–435
35. G.E. Dieter, in *Mechanical Metallurgy,* McGraw Hill, 1961, p 488
36. T. Altan and R.J. Fiorentino, Prediction of Loads and Stresses in Closed Die Forging, *J. Eng. Ind. (Trans. ASME),* May 1971, p 477
37. N. Akgerman and T. Altan, Application of CAD/CAM in Forging Turbine and Compressor Blades, *J. Eng. Power (Trans. ASME) A,* Vol 98 (No. 2), April 1976, p 290
38. G.D. Lahoti et al., Computer Aided Analysis of Metal Flow and Stresses in Plate Rolling, *J. Mech. Work. Technol.,* Vol 4, 1980, p 105
39. N. Akgerman, G.D. Lahoti, and T. Altan, Computer Aided Roll Pass Design in Rolling of Airfoil Shapes, *J. Appl. Metalwork.,* Vol 1, 1980, p 30
40. J.H. Hitchcock, "Roll Neck Bearings," research committee report, American Society of Mechanical Engineers, 1935; cited in L.R. Underwood, *The Rolling of Metals,* Vol I, John Wiley and Sons, 1950, p 15–16
41. B. Avitzur, Friction during Metal Forming, *Friction, Lubrication, and Wear Technology,* Vol 18, *ASM Handbook,* 1992, p 63–65
42. S. Ekelund, in *Roll Pass Design,* VEB Deutscher Verlag, 1963, p 48 (in German)
43. Z. Wusatowski, Hot Rolling: A Study of Draught, Spread and Elongation, *Iron Steel,* Vol 28, 1955, p 69
44. L.G.M. Sparling, Formula for Spread in Hot Rolling, *Proc. Inst. Mech. Eng.,* Vol 175, 1961, p 604
45. S.I. Oh and S. Kobayashi, An Approximate Method for Three-Dimensional Analysis of Rolling, *Int. J. Mech. Sci.,* Vol 17, 1975, p 293
46. R. Kummerling and H. Lipmann, On Spread on Rolling, *Mech. Res. Commun.,* Vol 2, 1975, p 113
47. H. Neumann, *Design of Rolls in Shape Rolling,* VEB Deutscher Verlag, 1969 (in German)
48. R. Hill, A General Method of Analysis for Metalworking Processes, *Inst. J. Mech. Sci.,* Vol 16, 1974, p 521
49. R.E. Beynon, *Roll Design and Mill Layout,* Association of Iron and Steel Engineers, 1956
50. A. Schutza, Comparison of Roll Pass Designs Used for Rolling Angle Sections, *Stahl Eisen,* Vol 90, 1970, p 796 (in German)
51. S.-E. Lundberg, Roll Pass Design: The Key Function in Control of Shape, Dimension and Mechanical Properties of Hot Rolled Products, *Scand. J. Metall.,* Vol 26 (No. 3), June 1997, p 102–114
52. P.S. Raghupathi and T. Altan, "Roll Pass Design in Shape Rolling," unpublished review of German literature, Battelle Columbus Laboratories, 1980
53. H. Neumann and R. Schulze, Programmed Roll Pass Design for Blocks, *Neue Hütte*, Vol 19, 1974, p 460 (in German)
54. H. Gedin, Programmed Roll Pass Design for Quality Steels, *Kalibreur,* Vol 11, 1969, p 41 (in German)
55. U. Suppo, A. Izzo, and P. Diana, Electronic Computer Used in Roll Design Work for Rounds, *Kalibreur,* Vol 19, Sept 1973, p 3
56. K. Yoshimura, S. Kini, and R. Shivpuri, A Technique for Robust Design of Roll Passes for Consistent Rod Quality, *38th Mechanical Working and Steel Processing Conference Proceedings,* Vol 34, USA, 13–16 Oct 1996 (Cleveland, OH), Iron and Steel Society/American Institute of Mining, Metallurgical and Petroleum Engineers, 1997, p 241–250
57. K.F. Kennedy, G.D. Lahoti, and T. Altan, "Computer Aided Analysis of Metal Flow Stresses and Roll Pass Design in Rod Rolling," *Iron Steel Eng.,* Vol 60 (No. 6), June 1983, p 50–54
58. F. Schmeling, Computer Aided Roll Pass Design and Roll Manufacturing, *Stahl Eisen,* Vol 102, 1982, p 771 (in German)
59. N. Kim, S. Kobayashi, and T. Altan, Three-Dimensional Analysis and Computer Simulation of Shape Rolling by the Finite and Slab Element Method, *Int. J. Mach. Tools Manuf.,* Vol 31 (No. 4), 1991, p 553–563
60. W. Shin, S.M. Lee, R. Shivpuri, and T. Altan, Finite-Slab Element Investigation of Square-to-Round Multi-Pass Shape Rolling, *J. Mater. Process. Technol.,* Vol 33 (No. 1–2), Aug 1992, p 141–154

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 248-257
DOI:10.1361/hwpd2003p248

Chapter 17

Thermomechanical Processing by Controlled Rolling

THERMOMECHANICAL PROCESSING (TMP) is a term referring to a variety of metal forming processes that involve careful control of thermal and deformation conditions to achieve advantageous microstructures and improvements in properties (particularly the refinement of grain size to improve strength and toughness). Perhaps the most common type of TMP application is the controlled rolling of microalloyed steels. Controlled rolling involves the careful conditioning of austenite during hot deformation so that the austenite transforms to a fine-grain ferrite in the final as-rolled product. Similar concepts also apply to steel bar and forgings, although TMP applications for these types of microalloying steel products have lagged behind that of flat-rolled steel products. The basic objective of TMP, regardless of form, is to ensure and/or improve properties through the control of microstructural changes during hot deformation. The concept of TMP also applies to nonferrous systems, such as the forging of titanium alloys and nickel-base superalloys (e.g., see Chapter 11, "Design for Deformation Processes").

This chapter focuses on the controlled rolling of steel, which is probably the most significant form of TMP application. The application of TMP methods for microalloyed steel bar and forgings, which involve different methods and justifying objectives than flat-rolled product, is briefly addressed in Chapter 11, "Design for Deformation Processes." In addition, Chapter 3, "Evolution of Microstructure during Hot Working," describes the general effects of thermomechanical processing on microstructure evolution with attention on the key processes that control microstructure evolution—that is, dynamic and static recovery and recrystallization and grain growth. These topics are addressed in more detail in this chapter, as they relate to the methods of controlled rolling. By careful process control (and the use of microalloying), controlled rolling of steels results in improved properties and advantages in subsequent operations. For example, improvement in strength from controlled rolling may allow reduction in carbon contents, which thus enhances the weldability of the product. Improved mechanical properties of as-rolled products also may eliminate the need for subsequent heat treatment.

Fundamentals of Controlled Rolling (Ref 1)

In general terms, *controlled rolling* refers to rolling processes designed for strict temperature and deformation control to obtain specific objectives for austenite conditioning. Obviously, all hot-rolling practices occur under some sort of temperature and deformation control, but not all rolling practices are designed to manipulate the condition of austenite prior to transformation. The general difference between conventional hot rolling (CHR) and various forms of thermomechanical processing is perhaps most simply described using Fig. 1. The method denoted as A in Fig. 1 is conventional hot rolling, which shows that the reheating, rough rolling, and finish rolling all occur at the highest possible temperatures. The primary goal of conventional hot rolling is to optimize productivity.

In contrast to CHR methods, controlled rolling involves special methods to control the microstructural condition of austenite microstructure during hot rolling. For example, if the hot rolling takes place below the recrystallization stop temperature (T_{RXN}), the austenite grains become highly elongated and, at a sufficiently large strain, become filled with intragranular defects such as deformation bands and twins. This technique, which perhaps represents the most well-known and common variant of thermomechanical processing, is referred to as *conventional controlled rolling* (CCR). In the CCR method, finish rolling is typically done below the recrystallization stop temperature (method B in Fig. 1) so that deformation results in very elongated (pancake-like) austenite grains with intragranular crystalline defects, which then transform into very fine ferrite grain sizes during cooling.

In another method of controlled rolling, if the rolling temperatures are high to allow recrystallization, then deformed equiaxed grains of austenite recrystallize into a different set of

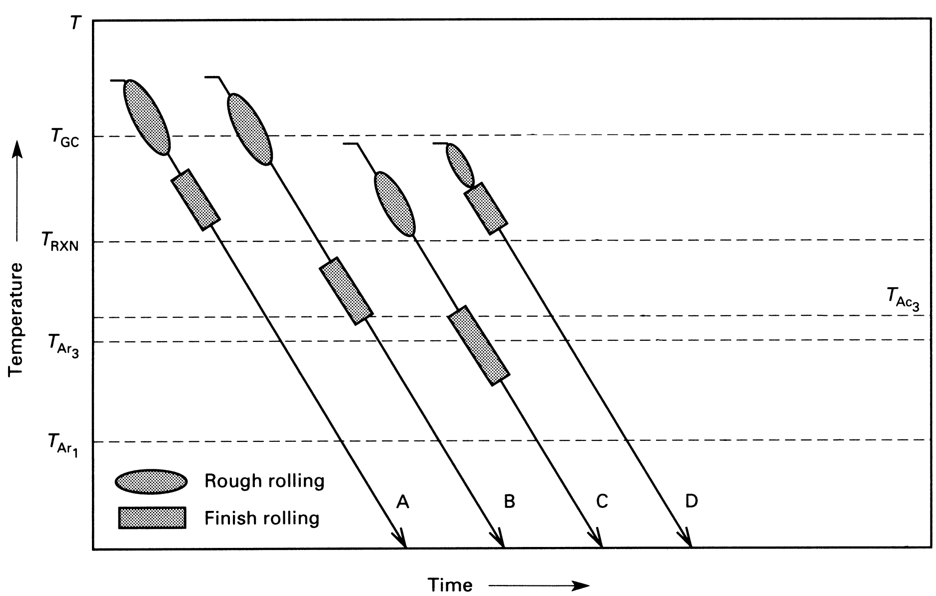


Fig. 1 Comparison of selected thermomechanical treatments based on critical austenite temperatures, transformation temperatures, and rough and finish rolling operations. A, conventional hot rolling; B, conventional controlled rolling; C, intensified controlled rolling; D, recrystallization controlled rolling

equiaxed grains that differ from the original chiefly in size. This type of hot-rolling process, which leads directly to fine equiaxed grains, is referred to as *recrystallization controlled rolling* (RCR). This form of thermomechanical processing involves repeated recrystallization of austenite by both rough rolling and finish rolling above the recrystallization stop temperature (method D in Fig. 1). The success of this technique depends not only on achieving a fine austenite grain size by repeated recrystallization but also on maintaining a fine grain by inhibiting grain-coarsening mechanisms. These different TMP techniques of controlled rolling are discussed in more detail after a brief review of the basic metallurgy of austenite conditioning. The principal goals are to obtain large amounts of fine ferrite grains and small amounts of lower temperature transformation products that contain cementite (such as pearlite or upper bainite), which degrade toughness.

Microalloying. Thermomechanical processing of ferritic steel product involves control of austenite microstructure, but it is important to note that improvements in the strength and toughness depend on the synergistic effect of microalloy additions and thermomechanical conditions. This synergism, recognized from early research on controlled rolling, is evidenced by the following observations:

- It is very difficult to improve the mechanical properties by thermomechanical processing without microalloying elements.
- Steels with microalloying elements, but without thermomechanical processing, display a lower toughness than their plain-carbon steel counterparts do.

The importance of microalloying elements on austenite during thermomechanical processing can be understood by the three critical temperatures of austenite metallurgy:

- Grain-coarsening temperature, T_{GC}
- Recrystallization stop temperature, T_{RXN}, below which recrystallization of austenite ceases
- Transformation temperatures, such as the critical temperature when austenite begins to transform to ferrite upon cooling (Ar_3), or the bainite-start temperature (B_s) above which austenite will not transform to bainite upon cooling

Critical transformation temperatures are an important part of microalloying, but they are not discussed in this chapter because they are related more to hardenability concepts and heat treatment. For more details, see Ref 1.

Solubility, which is the extent that an element can be maintained in solid solution, is a key aspect of how microalloying elements behave in austenite during TMP. Microalloying elements during the deformation of austenite may remain in solid solution, or they may chemically bond with another alloying element to form a precipitate, or second-phase, compounds. The extent and conditions of this partitioning of solutes between solid solution and precipitates in austenite have a strong influence on the migration and movement of austenite grains during hot deformation. This is explained in more detail in the section "Restoration Processes." The ability of an element to remain a solute or precipitate is directly controlled by its solubility product, which is a function of temperature and alloying concentrations. Elements such as niobium, titanium, and vanadium have been shown to have solubility products as to make them particularly useful for TMP conditioning of austenite. However, further discussion of solubility products of these microalloying elements is a subject of physical metallurgy that is beyond the scope of this chapter. The intent is to describe only briefly the importance of microalloying in the TMP conditioning of austenite.

Comparison of the various precipitation systems is done through the use of the appropriate solubility relations as a function of temperature. Titanium precipitates are very effective for high-temperature control (for example, during reheating), while niobium precipitates are very effective for intermediate temperatures (for example, control of recrystallization stop temperature). Vanadium, on the other hand, displays supersaturation at the very lowest range of hot-rolling temperatures and can be retained easily in solid solution in the austenite for the eventual precipitation (VC and VN) hardening of the transformation product.

Austenite Conditioning. The metallurgical condition of the austenite at the transformation temperature depends on a complex interplay between the stored energy of deformation and the restoration processes of recovery and recrystallization. These restoration processes drive the metallurgical state back to a condition that is energetically comparable to or lower than the condition existing before deformation. However, various thermal and metallurgical factors may also retard or limit the restoration processes that are driven, in part, by the stored energy of deformation. Therefore, the austenite conditioning depends on a complex interplay of thermomechanical and metallurgical factors that drive or retard the restortion processes. This interaction is described in a little more detail in the section "Restoration Processes."

In general, as-rolled austenite can show three types of microstructures:

- Fully recrystallized
- Partially recrystallized
- Fully unrecrystallized

While the grain size of fully recrystallized austenite is a simple matter to comprehend and determine, the determination of a fully unrecrystallized microstructure is not. In order to specify the grain size of fully recrystallized austenite, only an average grain diameter is required; this fully describes the grain boundary area per unit volume. In the case of fully unrecrystallized austenite, both the starting grain size and the aspect ratio of the deformed grains are required to determine the grain boundary surface per unit volume. In general, because the ferrite grain size is related to the number of sites for ferrite nucleation, and because this number is related to the austenite grain boundary area per unit volume, the ferrite grain size will decrease as this boundary area increases. Furthermore, finer austenite grains reduce the chances of forming upper bainite during transformation because the austenite grain volume can be completely consumed by the ferrite and pearlite reactions. Although both of the issues were qualitatively recognized in the 1960s (for example, ferrite grain size was reduced by lowering the finish rolling temperature of austenite), it was not until the mid-1970s that effects such as these could be quantitatively assessed.

The understanding of the hot deformation of austenite in general and its influence on final ferrite grain size was improved dramatically through research (Ref 2–6) in the mid-1970s. A summary of this work is presented in Ref 6. The key to understanding the process by which austenite links alloying and hot rolling with the final ferrite grain size was based on quantitatively assessing the number of pre-existing or strain-induced heterogeneities introduced into the austenite during rolling that could act as sites for ferrite nucleation (Ref 2). Earlier research (Ref 7) using quantitative metallography led to the definition of a parameter referred to as the *austenite interfacial area* (S_v), which attempts to account for all the near planar crystalline defects that may exist in a polycrystalline aggregate. The parameter S_v has units of mm^{-1} or equivalently area (mm^2) per unit volume (mm^3). Researchers (Ref 2) applied the parameter S_v, which they defined as the *total effective interfacial area per unit volume*, to the problem of quantitatively assessing the number of sites in thermomechanically processed austenite for potential nucleation of ferrite and developed the correlation shown in Fig. 2.

Attempts to quantify the influence of S_v on final ferrite grain size have shown that the ferrite grain size decreases continuously with increases in S_v (Ref 2, 8, 9). There is still some disagreement over the exact relationship, but the general form has been well defined (Fig. 3). Hence, from the standpoint of ferrite grain refinement, a well-conditioned austenite exhibits a large S_v. For example, if the rolling takes place above the recrystallization temperature of the austenite, then the deformed austenite grains are replaced with a different set of equiaxed, recrystallized grains that differ from the original chiefly in size. For equiaxed grains:

$$S_v = \frac{3}{D_\gamma} \quad \text{(Eq 1)}$$

where D_γ is the diameter of the austenite grains. A reduction in grain size by hot rolling thus leads to an increase in grain boundary area and, hence, an increase in S_v.

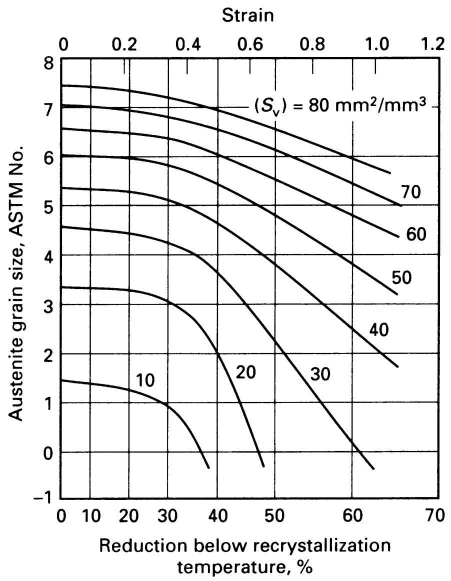


Fig. 2 Variation of effective austenite interfacial area (S_v) with initial grain size and amount of reduction below the recrystallization stop temperature for a 0.03 wt% Nb steel. Source: Ref 2

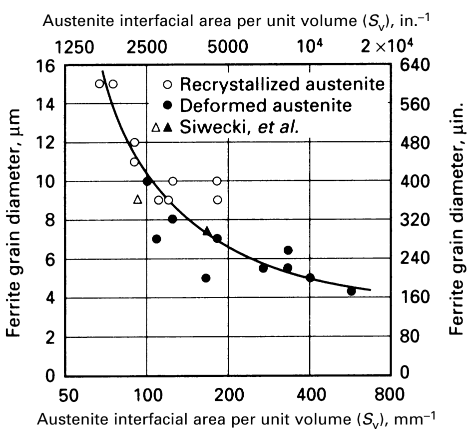


Fig. 3 Ferrite grain sizes produced from recrystallized and unrecrystallized austenite at various S_v values. Source: Ref 10

If the hot rolling takes place below the recrystallization temperature, the austenite grains become highly elongated and, at a sufficiently large strain, become filled with intragranular defects such as deformation bands and twins. In this case, the final S_v comprises two terms, one describing the grain boundary surface area per unit volume and the other describing the surface area of the deformation bands and twins contained within the elongated grains such that:

$$S_v = S_v(\mathrm{GB}) + S_v(\mathrm{IPD}) \quad \text{(Eq 2)}$$

where S_v(GB) is the total effective area per unit volume from the grain boundary contribution, and S_v(IPD) is the contribution from intragranular planar defects. The important parameters include the extent of deformation below the recrystallization temperature and the grain size of the austenite prior to pancaking (Ref 10, 11, 12).

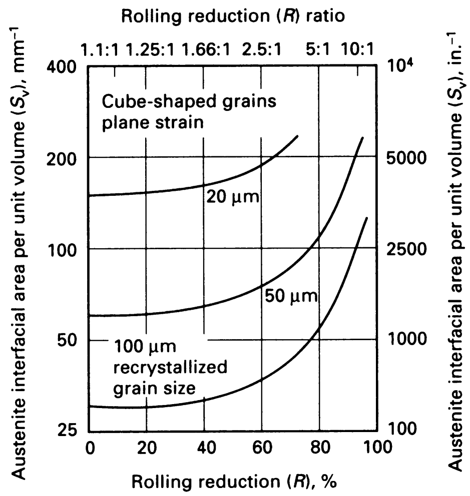


Fig. 4 Effect of rolling reduction on S_v for cube-shaped austenite grains. Source: Ref 10

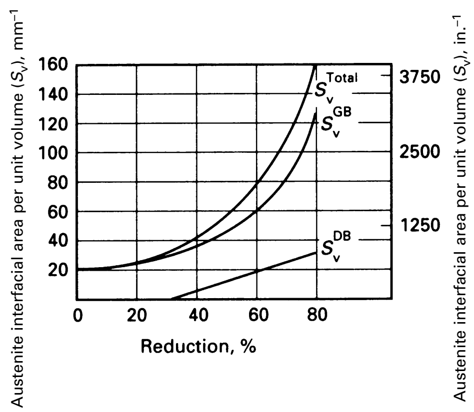


Fig. 5 Increase in S_v because of deformation bands. Austenite grain diameter, D_γ = 100 µm (0.004 in.); S_v^{GB} grain boundary planar-linear model; S_v^{DB} deformation band. $S_v^{Total} = S_v^{GB} + S_v^{DB}$. Source: Ref 10

This work has been summarized (Ref 10) and is illustrated in Fig. 4 and 5. The variation of S_v was calculated for a hypothetical array of cube-shaped grains of variable initial size that were subjected to various plane strain deformations, all to simulate the behavior of grains during controlled rolling (Ref 10).

Restoration Processes (Ref 1)

The metallurgical condition of the austenite at the transformation temperature is dependent on a complex interplay between the stored energy of deformation, restoration processes, and the forces (or factors) that retard restoration. The following discussions briefly describe the forces (or factors) that drive and retard the restoration processes of grain growth, recovery, and recrystallization. In general, restoration processes involve the movement of grain boundaries so that stored energy tends to become more dispersed. However, the mechanisms of grain-boundary movement also can be retarded by solutes and precipitates that act on the grain boundaries.

Grain Coarsening

Grain coarsening is driven by the minimization of grain boundary energy (or area) per unit volume. Like surface energy, grain-boundary energy tends to minimize itself when possible by decreasing the total grain-boundary area of a system. Thus, one grain grows at the expense of another to reduce the overall grain-boundary area per unit volume. Grain boundaries that have sharp corners or bends tend to smooth and straighten, and small grains (with high surface-to-volume ratios) become absorbed by larger grains.

Grain growth can be complicated (or slowed) when the movement of grain boundaries is restricted by metallurgical conditions such as:

- Solute atoms that cause grains to grow more slowly
- Second-phase particles or precipitates that "pin" or restrict the movement of grain boundaries

These two mechanisms, which are key factors in the restriction of grain growth, are described briefly here. They are not, however, the exclusive mechanisms. For example, another reason for slow grain growth may be the presence of a strongly preferred crystal orientation or texture. With respect to other restoration processes, grain growth tends to occur much more slowly than recrystallization, but this depends on the temperature and the amount of stored energy that may drive recrystallization processes. In particular, inhibition of grain growth is important in RCR methods, where finish rolling is done at relatively higher temperatures (i.e., above T_{RXN}).

Solute Drag. Solutes can cause grains to grow more slowly through the stabilization of grain boundaries. Solute atoms lower the grain-boundary energy and therefore prevent the grain boundary from moving as rapidly as it would in the absence of a solute atom. This is referred to as *solute drag*. The effect of solutes in retarding grain growth depends on the solute involved; elements that distort the lattice the most impede grain growth the most. The effectiveness of solute drag has been shown to be related to differences in size and valence between the solute and solvent atoms (Ref 13, 14, 15) and has been discussed extensively in the literature (Ref 16, 17, 18, 19, 20). However, unlike the effects of particle pinning, quantitative treatment of solute-drag effects are not established.

Grain-Boundary Pinning by Particles. The effects of particle pinning on grain growth have been expressed in more quantitative terms. The first attempt to treat the retarding effects of second-phase particles on a migrating grain boundary was made by Zener (Ref 21). He indicated that when particles were present in the vicinity of a grain boundary, the effective energy of the grain boundary would be lowered. The reduction in grain-boundary energy occurs because the surface area of second-phase particles would replace part of the grain boundary. Therefore, the movement of a grain boundary away from particles would require work as the effective grain-boundary area is increased (Ref 22). Zener's original theory was later quantified with the assumption of the rigid motion of grain boundaries through a regular array of spherical particles (Ref 23). It was shown that the pinning force for each particle (F) and the particle radius (r) were related by:

$$F = 4r\sigma_I \quad \text{(Eq 3)}$$

where σ_I is the interfacial energy per unit area of boundary. An expression was derived subsequently for the critical particle size (r_c) below which grain boundaries are pinned (Ref 23):

$$r_c = \frac{6R_0 f_v}{\pi} \cdot \left(\frac{3}{2} - \frac{2}{z}\right)^{-1} \quad \text{(Eq 4)}$$

The relationship shown by Eq 4 was obtained by equating the rates of the grain-boundary energy increase during grain growth and the particle pinning force. In this expression, R_0 is the mean radius of the matrix grain, and f_v is the volume fraction of second-phase particles. The term z represents the ratio of the radii of growing and matrix grains. It was deduced (Ref 24) that z can range in value between $\sqrt{2}$ to 2 throughout the grain-growth process. The important point to be realized from Eq 4 is that a high volume fraction of small particles inhibits boundary migration.

Equation 3 can be expanded to account for the total pinning force (F_{PIN}) that a number of particles per unit area (N_s) exert on a migrating boundary. This total pinning force is expressed as:

$$F_{PIN} = 4r\sigma_I N_s \quad \text{(Eq 5)}$$

Grain Coarsening Temperature (T_{GC}). The grain coarsening temperature (T_{GC}) is defined to be that temperature above which grain coarsening by secondary recrystallization commences and refers to the temperature above which the undissolved precipitates can no longer suppress grain growth. The influence of various microalloying elements on grain growth during reheating is shown in Fig. 6 (Ref 10). While it does not appear that the final average austenite grain size is very dependent on the reheated grain size prior to multipass deformation (Ref 25), it does appear likely that the distribution of grain sizes about the average is much smaller when the reheating temperature is kept below the grain-coarsening temperature (Ref 26).

During reheating, a slab is slowly heated from room temperature to the reheating temperature, where it is soaked for an appropriate period of time and after which it is available for hot rolling. During this heating and soaking period, events that ultimately control the grain size and composition of austenite prior to hot rolling occur. The ultimate response of a given steel to subsequent processing will, to a large extent, be strongly influenced by the reheating stage. The major events that may occur during reheating include:

- Nucleation, growth, and possible coarsening of austenite grains
- Elimination of inhomogeneities in solute distribution
- Dissolution of microalloy precipitates that were inherited from the original microstructure

Austenite first forms at ferrite-carbide (cementite) interfaces (Ref 27), and the increase in the amount of austenite is accompanied by a decrease in the amount of cementite, at least in the early stages of austenite formation (Ref 27). Also, although the kinetics of austenite formation are very rapid at normal reheating temperatures, the elimination of heterogeneities in solute distribution is very sluggish (Ref 27).

Once the austenite grains have formed from the ferrite-pearlite starting structure, they will coarsen with time and temperature. This coarsening behavior has been reviewed in some detail (Ref 28). Reheating controls not only the grain size, but also the initial composition of the austenite (Ref 25, 28). Niobium is not fully taken into solution until a steel is heated to approximately 100 °C (180 °F) above its grain-coarsening temperature (Ref 25).

Similar results have been found in a separate work (Ref 29). The factors that control the composition of the austenite at various stages of processing have been reviewed (Ref 30).

Recovery and Recrystallization

Recovery and recrystallization are driven by the stored energy of deformation. The concentrations of carbon, nitrogen, microalloying elements (niobium, titanium, vanadium), the degree of strain, the time between passes, the strain rate, and the temperature of deformation all influence recrystallization during hot working.

The driving force for recovery and recrystallization has been shown to be equal to the difference in dislocation density that is present between strained and strain-free local volumes. This driving force has been estimated and, based upon the area under the flow curve (Ref 31), is expressed by:

$$\Delta\sigma = 0.2\mu \mathbf{b}\sqrt{\Delta\rho} \quad \text{(Eq 6)}$$

where $\Delta\sigma$ is the increment in flow stress due to work hardening, μ is the shear modulus, $\mathbf{b}$ is the Burgers vector, and $\Delta\rho$ is the excess dislocation density.

Recrystallization Stop Temperature (T_{RXN}). Perhaps the most important effect of microalloying elements is in controlling the recrystallization stop temperature of austenite (Ref 31). Austenite recrystallization ceases at temperatures below T_{RXN}, and the influence of several microalloying elements on the recrystallization stop temperature is shown in Fig. 7. As shown in Fig. 7, niobium has the most profound effect on

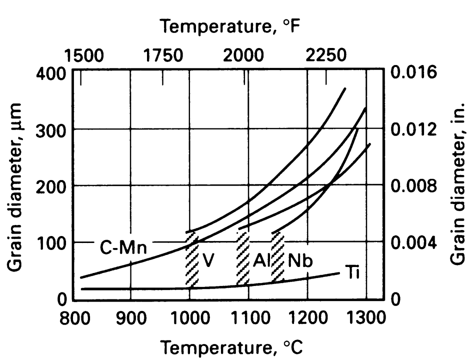


Fig. 6 Austenite grain-coarsening characteristics in steels containing various microalloying additions. Source: Ref 10

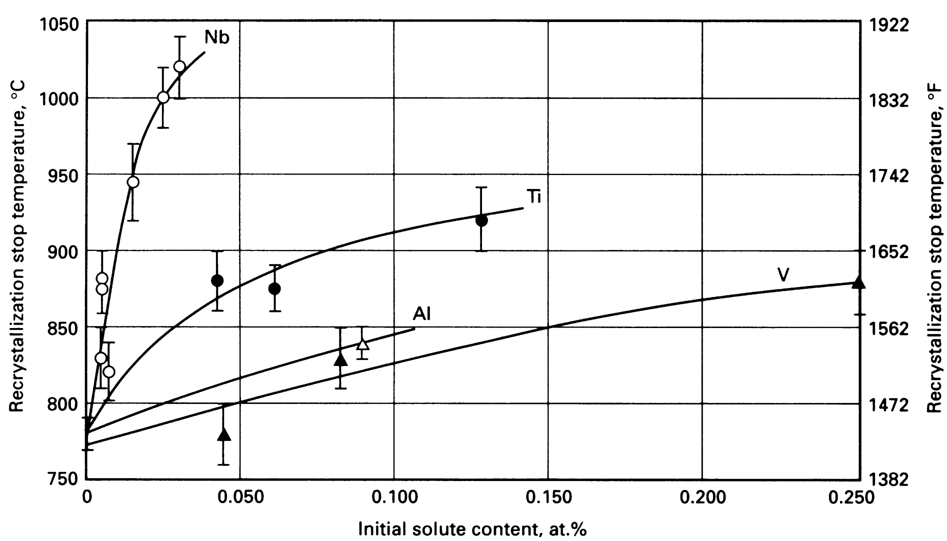


Fig. 7 Effect of microalloying additions on the recrystallization stop temperature of austenite. Source: Ref 31

T_{RXN}, and this accounts for its selection as the primary microalloying element in modern high-strength, low-alloy (HSLA) steels. Some early steels used vanadium in conjunction with nitrogen additions to produce a moderate effect in retarding austenite recrystallization, but such practices are not compatible with current ideas for limiting free nitrogen concentrations, nor were they particularly effective. Consequently, the predominant use of vanadium is for precipitating vanadium carbide in the low-temperature transformation products. Similarly, titanium is not considered useful for retarding austenite recrystallization but is highly effective in preventing grain growth (Fig. 6). The grain-refining benefits of titanium have another primary use in welding that is unrelated to the benefits during austenite processing, and this has led to widespread use of approximately 0.015 wt% Ti in plate and pipeline steels. This approach is not readily accessible when nitrogen concentrations are above 0.008 wt% N or when the ingot casting route is involved.

The data shown in Fig. 6 and 7 reveal that the various microalloying elements all show the same general behavior. The differences between them are predominantly the temperature range over which they exert their influence and the level of intensity of that influence. This difference in behavior is due chiefly to differences in the solubility of the various precipitates in austenite. The difference in the temperature range of precipitation in austenite is shown clearly in Fig. 8 for vanadium and niobium carbonitrides (Ref 32, 33). In Fig. 8, the deformation temperature is plotted against the deformation time required to initiate dynamic recrystallization (i.e., time-to-peak stress, t_p), which is one form of austenite restoration. When no strain-induced precipitation forms, the curves are only slightly offset from the base steel. This is indicative of the weak retarding effect of solute drag as a retarding force. However, when precipitation accompanies the early stages of deformation as it does in the vanadium plus nitrogen steel at low temperatures and the two niobium steels at higher temperatures, there is a drastic retardation of dynamic recrystallization, and the curves display a plateau. Notice that when precipitation occurs, the curves are virtually identical. They are only displaced on the temperature scale.

Models of Recrystallization Suppression. Three models have been postulated to explain how microalloy precipitates suppress austenite recrystallization (Ref 16, 22, 34). All of these models are based on the general form of Eq 5. Their differences arise from the method by which N_s is calculated. The first of these models was an extension of earlier research (Ref 23). This model is often referred to as the *rigid boundary model* because it assumes the motion of a rigid grain boundary, which is capable of interacting with those particles lying within $\pm r$ of the boundary plane. Under this set of conditions, N_{sR} is defined by:

$$N_{sR} = 2r\,N_v \qquad \text{(Eq 7)}$$

where N_v is the number of particles per unit volume. Assuming spherical particles having radius r and volume fraction f_v, N_v is defined by the following expression:

$$N_v = \frac{3f_v}{4\pi r^3} \qquad \text{(Eq 8)}$$

Combining Eq 7 and 8 yields the number of particles per unit area assuming a rigid boundary model, N_{sR}:

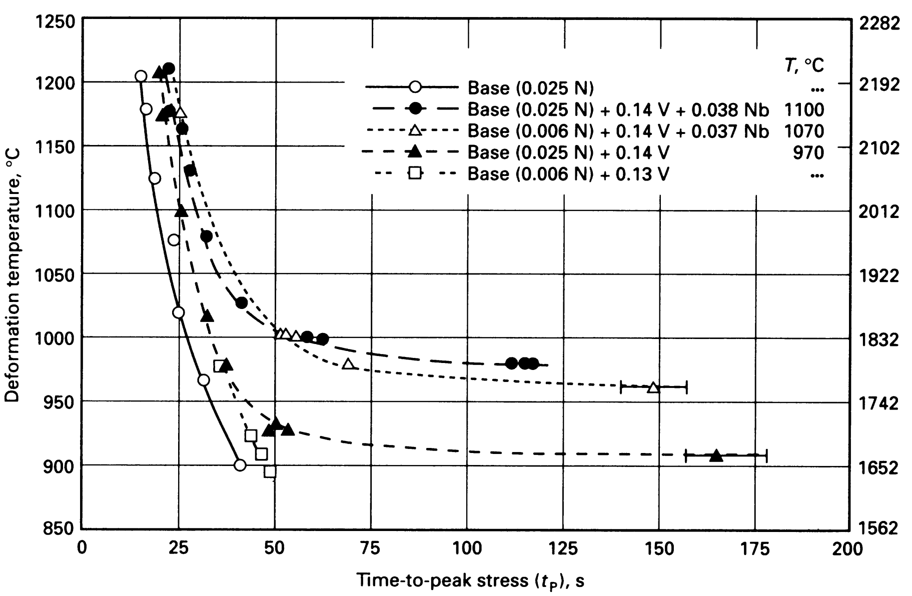


Fig. 8 Dependence of the time-to-peak stress (t_p) on deformation temperature at a low strain rate ($\dot{\varepsilon}$ is 0.006 s^{-1}) for selected microalloyed steels containing two levels of nitrogen (0.006 and 0.025% N). The 0.025% N steel is a reference steel. Source: Ref 32

$$N_{sR} = \frac{3f_v}{2\pi r^2} \qquad \text{(Eq 9)}$$

The rigid boundary model, however, was rather simple and unrealistic since it did not account for any flexibility of the austenite grain boundaries (Ref 22).

A model at the opposite extreme of the rigid boundary model was proposed (Ref 31). This model was termed the *flexible boundary model*. It assumed that an infinitely flexible boundary is capable of interacting with every particle in the three-dimensional array until it is fully pinned (Ref 31). This led to the relationship:

$$N_{sF} = \frac{rN_v}{(f_v)^{1/3}} \qquad \text{(Eq 10)}$$

which, when combined with Eq 8, yields the number of particles per unit area assuming a flexible boundary model, N_{sF}:

$$N_{sF} = \frac{3(f_v)^{2/3}}{4\pi r^2} \qquad \text{(Eq 11)}$$

The most realistic of these models was the one proposed in Ref 34 and 35. This model considered the effect of a precipitate distribution that could exist on austenite subgrain boundaries prior to the start of recrystallization. Assuming that the average subgrain intercept is l, the surface area per unit volume for such subgrain boundaries would be $2/l$, and the number of particles per unit subgrain boundary area would be given by (Ref 34):

$$N_{sS} = \frac{lN_v}{2} \qquad \text{(Eq 12)}$$

As before, combining Eq 8 and 12 gives the number of particles per unit area assuming a subgrain boundary model, N_{sS}:

$$N_{sS} = \frac{3lf_v}{8\pi r^3} \qquad \text{(Eq 13)}$$

Substitution of the expressions for N_{sR}, N_{sF}, and N_{sS} (Eq 9, 11, 13) into Eq 5 gives the respective pinning force for each model.

Conventional Controlled Rolling

Conventional controlled rolling was the first type of controlled rolling to come into regular commercial use. About 8 to 10% of the total steel tonnage rolled annually in the 1990s was produced in this way. This process was originally developed for the production of plate grades for the manufacture of oil and gas pipelines with required minimum yield strengths ranging from 350 to 490 MPa (50 to 70 ksi). Because of the need for good weldability, low concentrations of carbon and carbon equivalents were specified. These were obtained readily by reducing carbon concentrations to 0.06 or 0.07%; small amounts of niobium (about 0.04%), in combination with vanadium (up to 0.1%) and molybdenum (up to 0.30%), were added for higher-strength grades.

As noted previously, the goal of conventional controlled rolling is to produce very fine ferrite grain sizes by conditioning the austenite through extensive rolling in the non-recrystallization region of austenite. During roughing operations, the coarse reheated austenite grains in a slab are first refined by repeated recrystallization, bringing the grain sizes down to about 20 mm (0.8 mil) or less. The transfer bar can then cool below the T_{RXN} during transfer from roughing to the finishing facilities. When rolling is restarted or continued below the T_{RXN}, recrystallization is no longer possible, and the austenite structure is progressively flattened in an operation known as *pancaking*. For pancaking to be successful, the accumulated reductions applied in this temperature range must add up to at least 80%. Finally, when the flattened austenite grains go through their transformation to ferrite, the ferrite produced has a very fine grain structure because of the large number of nucleation sites available on the expanded surfaces of the pancaked austenite grains. This leads to ferrite grain sizes in the range of 5 to 10 mm (ASTM grain size numbers 10 to 12). The fine-grain ferrite is responsible for the attractive combination of good toughness and strength.

Suppression of Recrystallization. Austenite pancaking is only possible in the absence of recrystallization. The three different forms of recrystallization that might take place during multipass hot rolling are (Ref 36, 37):

- Static recrystallization that might occur in the time period between deformations
- Dynamic recrystallization that might occur during deformation in the roll gap
- Metadynamic recrystallization that, if it occurred, would begin in the roll gap and go to completion in the time period between deformations

Static recrystallization is the more dominant form of recrystallization during industrial processing because the time required for substantial amounts of dynamic recrystallization is in orders of magnitude longer than the times actually available in the roll gap (Ref 38–41). Suppression of static recrystallization thus is a critical factor, and its arrest is caused by the copious precipitation of Nb(C,N) during delays between mill passes. Because the precipitation of Nb(CN) in the austenite during hot rolling retards recrystallization and raises the recrystallization-stop temperature, a broader temperature range is possible for hot working the steel to produce highly deformed austenite.

Niobium is the most effective alloying element for grain refinement by conventional controlled rolling. The optimum amount of niobium to suppress recrystallization between passes can be as little as 0.02%. During the rolling reductions at temperatures below 1040 °C (1900 °F), the niobium in solution suppresses recrystallization by solute drag or by strain-induced Nb(C,N) precipitation on the deformed austenite slip planes. The strain-induced precipitates are too large to affect precipitation strengthening but are beneficial for two reasons: they allow additional suppression of recrystallization by preventing migration of austenite grain subboundaries, and they provide a large number of nuclei in the deformed austenite for the formation of fine ferrite particles during cooling.

Titanium, zirconium, and vanadium are not as effective as niobium in raising the recrystallization stop temperature. Titanium and zirconium nitride formed during solidification and upon cooling of the slab do not readily dissolve upon reheating to hot-rolling temperatures. Although these nitrides may prevent grain coarsening upon reheating, they are not effective in preventing recrystallization because insufficient titanium or zirconium remains in solution at the rolling temperature to precipitate on deformed austenite boundaries *during* hot rolling and thus suppress austenite recrystallization. Vanadium, on the other hand, is so soluble that precipitation does not occur readily in the austenite at normal hot-rolling temperatures.

Effect of Niobium Precipitation on Ar_3. When microalloyed austenite undergoes multipass hot rolling at temperatures where large supersaturations exist, various amounts of strain-induced precipitation may occur depending on factors such as the finishing temperature and the interpass holding time. The effect on the amount of precipitation caused by variations in deformation temperature, extent of recrystallization, and interpass holding time is shown in Fig. 9. Figure 9 shows that the amount of strain-induced precipitation, which attends hot rolling, is a variable, and hence, the amount of niobium in solution can also vary. Because it is known that the Ar_3 temperature varies with the amount of niobium in solution, it should be expected that the Ar_3 temperature may also vary with the details of the rolling practice. This is shown in Fig. 10, which plots the dependence of Ar_3 versus the amount of deformation that takes place below the recrystallization stop temperature (Ref 43).

Intensified Controlled Rolling. A variation in conventional controlled rolling practice, known as *intensified controlled rolling*, is shown in Fig. 1. Note that intensified controlled rolling differs from conventional controlled rolling in that finish rolling not only occurs below the T_{RXN} but also extends to temperatures below the Ar_3. The goal of intensified controlled rolling is to increase the strength and toughness of the as-rolled microalloyed steel over what could be achieved through conventional controlled rolling. The combination of the lower reheat and

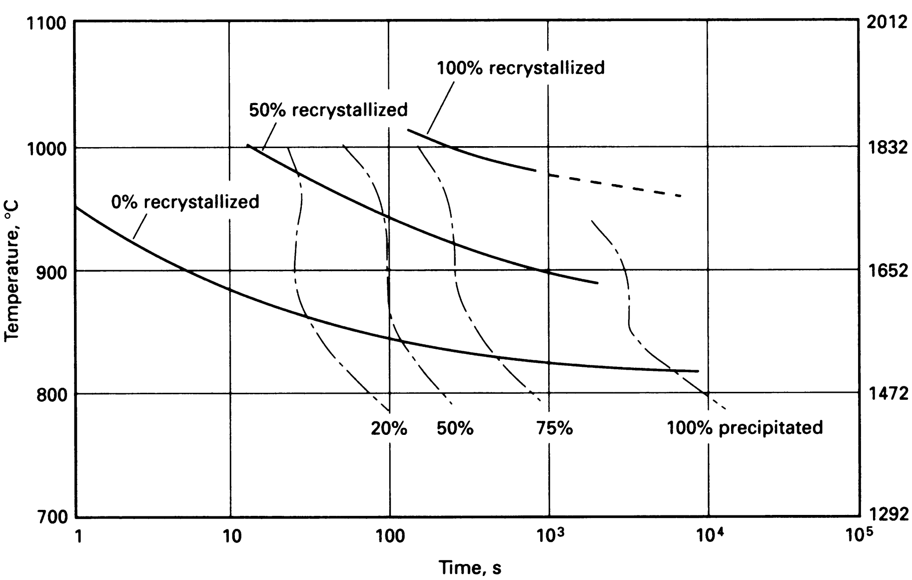


Fig. 9 Isothermal recrystallization and precipitation in a niobium steel after 50% deformation. The steel composition was 0.10C-0.99Mn-0.04Nb-0.008N. Source: Ref 42

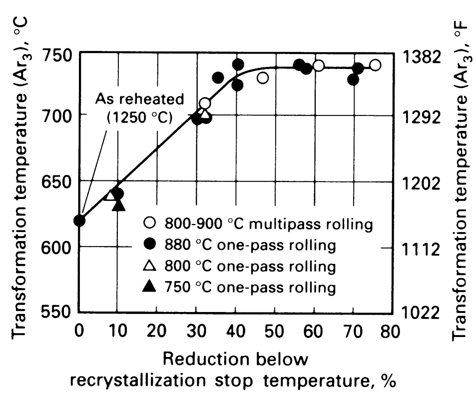


Fig. 10 Increase in Ar_3 temperature as a result of controlled rolling. Source: Ref 43

lower rolling temperatures leads to finer as-rolled austenite. Furthermore, because a portion of the rolling occurs in the intercritical or two-phase (α+γ) region, a certain amount of proeutectoid ferrite would be deformed during this rolling. This deformed ferrite is partially responsible for the higher strengths observed with intensified controlled rolling.

Recrystallization Controlled Rolling

Recrystallization controlled rolling is an attractive alternative to conventional controlled rolling when low-temperature working is impractical for certain products (hot forgings or thick plate) or with the equipment limitation of an underpowered mill. Conventional controlled rolling is generally based on the use of low finishing temperatures (that is, in the vicinity of 800 to 900 °C, or 1470 to 1650 °F). However, such finishing at such low temperatures is inappropriate for certain products, such as heavy plates and thick-walled seamless tubes, due to excessive rolling loads. For such applications, it is desirable to condition austenite at a higher rolling temperature in the crystallization regime. By carefully controlling the recrystallization of austenite and arranging for it to occur at successively lower temperatures during finish rolling, recrystallization controlled rolling achieves a fine, equiaxed austenite.

The success of recrystallization controlled rolling depends on two key factors. First, the recrystallization process should not be sluggish or suppressed so that the times required for recrystallization are not too long. The second factor is the inhibition of grain growth. Rapid grain coarsening can occur during the time between rolling passes, and so steels processed at the higher temperatures of the RCR method require microalloying to inhibit grain coarsening. For this purpose, sufficient titanium is added to have about 0.01% available for the formation of fine particles of TiN during cooling after continuous casting. When this dispersion has an appropriate size and frequency distribution, it can almost completely prevent grain growth of the austenite after each cycle of recrystallization.

In general, methods of grain-growth inhibition can take one of two forms. The most common method relies on the pinning force (Eq 5) from stable particles. As noted previously, the driving force for normal grain coarsening is typically smaller than for recrystallization, depending on the temperature and deformation conditions. The driving force for grain growth can be three orders of magnitude smaller than that for recrystallization (Ref 22). Because grain growth is slower, the pinning force needed for inhibition of grain growth is less than that needed for suppression of recrystallization. The particle pinning force also varies with f_v/r, and so particle distributions with rather small f_v and large r can still be effective barriers to grain coarsening. This, of course, would not be true in the suppression of recrystallization, where only large, local volume fractions of fine particles are effective. In this regard, particles such as titanium nitride have been shown to be particularly well suited for the suppression of grain coarsening. These particles must have sufficient stability to resist Ostwald ripening (solution-crystallizer phenomenon in which small crystals, more soluble than large ones, dissolve into larger particles).

An alternate grain-coarsening inhibition mechanism involves intense solute drag as the critical element. It has been found that high levels of soluble molybdenum and niobium can be very effective in retarding boundary motion at high reheating temperatures (Fig. 11). Hence, these solute effects act to retard grain coarsening both during reheating and after post-deformation static recrystallization.

Dynamic Recrystallization Controlled Rolling (DRCR) (Ref 44). When the interpass time is short, as in the case of rod, hot strip, and certain other rolling processes, insufficient time is available for conventional recrystallization during the interpass delay. The amount of carbonitride precipitation that can take place is also severely limited. As a result, an alternative form of recrystallization, known as *dynamic recrystallization*, is initiated. Dynamic recrystallization involves the nucleation and growth of new grains during (as opposed to after) deformation (Ref 45, 46). This also requires the accumulation of appreciable reductions, of the order of 100%, to enable the recrystallization process to spread completely through the microstructure inherited from the roughing process. Austenite grain sizes as small as 10 mm (0.4 mil) can be achieved with DRCR (Ref 23).

Low-temperature finishing by DRCR has the advantage of producing finer ferrite grain sizes after transformation than CCR does; that is, 3 to 6 mm (0.12 to 0.24 mil), as opposed to 5 to 8 mm (0.2 to 0.3 mil) for the latter process (Ref 45). However, such low-temperature finishing increases the rolling load, and it can also make mill control more difficult because of the load drop associated with the initiation of dynamic recrystallization. It is important to note that under industrial rolling conditions, CCR, RCR, and DRCR can all occur to different degrees during a given operation. This can happen when the processing parameters have not been optimized so as to favor only strain-induced precipitation and austenite pancaking in the case of CCR, conventional recrystallization in the case of RCR, and dynamic recrystallization in the case of DRCR.

Thermomechanical Processing of Microalloyed Bar (Ref 47)

Plate-rolling operations lend themselves to considerable control of the thermomechanical treatment. The slab reheat temperature can be reduced if desired. In fact, some rolling strategies involve only reheating to 960 °C (1760 °F) prior to rolling. Delays can be built into the rolling operation (although with some penalty in productivity), and a considerable range of finishing temperatures can be achieved. The operation can accommodate the most severe controlled-rolling schedules, including the deformation of austenite-ferrite mixtures.

In contrast to the plate-rolling process, the possible thermomechanical treatment on a modern bar mill (Fig. 12) is somewhat limited in scope. For example, the temperature-time profile for the rolling of 44 mm (1 in.) diam bar shown in Fig. 13 can be compared with the plate-rolling profile shown in Fig. 14. There are

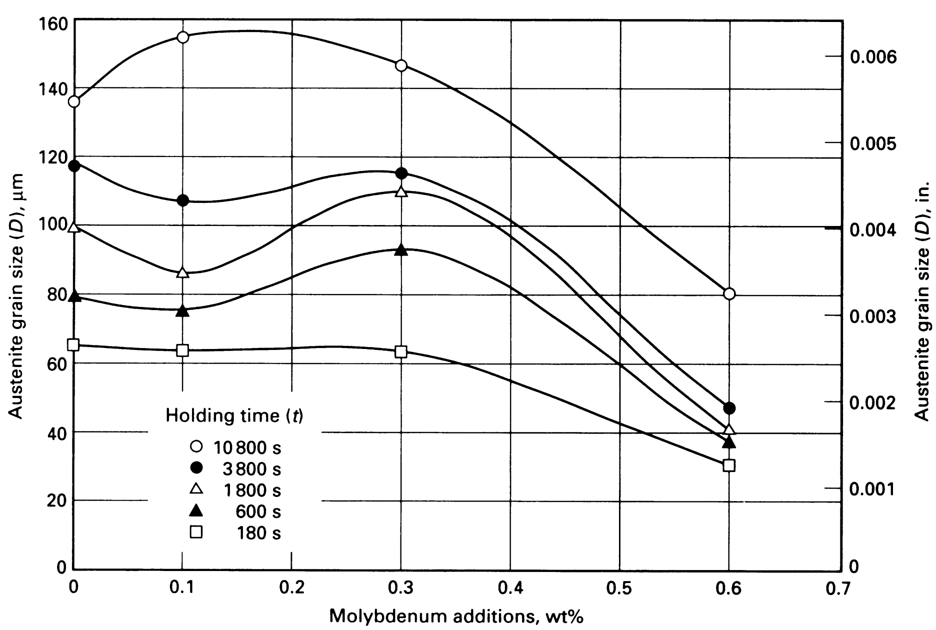


Fig. 11 Effect of molybdenum addition on austenite grain size after reheating at 1150 °C (2100 °F) and holding at this temperature for various times. Courtesy of C.I. Garcia and A.J. DeArdo, University of Pittsburgh, 1991

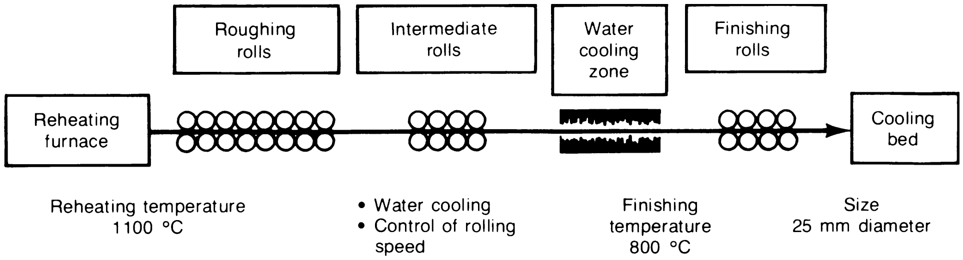


Fig. 12 Controlled-rolling process for microalloyed steel bar

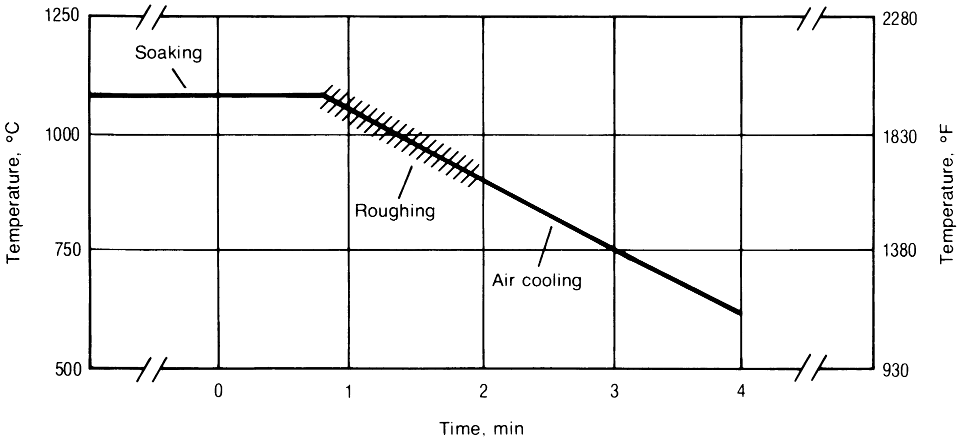


Fig. 13 Temperature-time profile for the controlled rolling of 44 mm (1$^3/_4$ in.) diam microalloyed steel bar. Compare with Fig. 14.

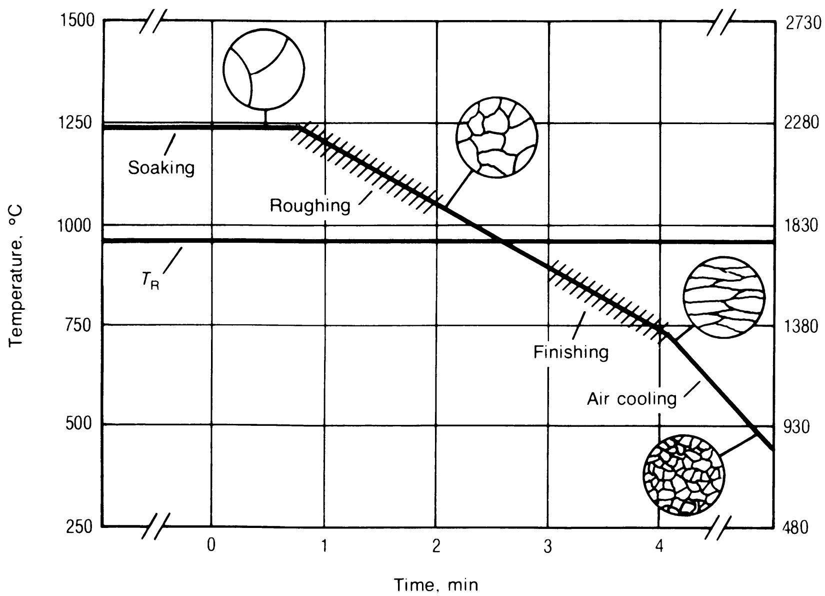


Fig. 14 Temperature-time profile for controlled rolling of 19 mm ($^3/_4$ in.) thick microalloyed steel plate. T_R, recrystallization temperature

clear differences between these rolling processes:

- Lower reheat temperatures, typically in the range of 1100 to 1200 °C (2010 to 2190 °F), are used in bar rolling. This lower temperature, in combination with the generally higher carbon levels in bar products, limits the amount of niobium that can be dissolved upon reheating. For example, in a 0.20% C steel, only about 0.01% Nb is soluble at 1100 °C (2010 °F). In contrast, vanadium is still readily soluble at bar reheat temperatures. Consequently, in HSLA bar grades, vanadium is the microalloying element commonly used to obtain the highest possible strength levels.
- Even though the lower reheat temperatures typical of bar products place some limitations on the use of different microalloying elements, these lower temperatures do provide for a finer as-reheated austenitic grain size than is typical of slabs reheated for conversion to plate. With a small titanium addition and continuous casting, as-reheated austenitic grain sizes of 50 to 60 mm (2 to 2.4 mils) can be achieved in billets destined for bar.
- Finishing temperatures in bar rolling are relatively high, even with the use of interstand cooling.

Consequently, recrystallization controlled rolling becomes quite important in bar rolling, and the rolling strategy must be designed to produce the finest possible recrystallized austenitic grain size. Subsequent control of the austenite-to-ferrite transformation range is still important to maximize ferritic grain refinement. Nevertheless, the ferritic grain size that can be produced on transformation from a recrystallized austenite is limited compared with the grain size that can be produced on transformation from austenitic grains that have been flattened by rolling below the recrystallization temperature. Thus, while moderate grain refinement can be achieved in an as-rolled microalloyed bar, this grain size will be somewhat coarser than the grain size of controlled-rolled HSLA plates.

Because the degree of ferritic grain refinement possible in as-rolled microalloyed bar steels is somewhat limited, and because substructural strengthening is not possible, alternative strengthening mechanisms are employed to reach yield strength levels comparable to those of plate grades. For example, in the alloy design of microalloyed bar steels, precipitation and pearlite strengthening must be relied on to a greater extent than in the design of plates. In view of the limited solubility of niobium or titanium at the reheat temperatures used in bar processing, vanadium is usually used to obtain the required level of precipitation strengthening in HSLA bar grades. Precipitation of V(C,N) during or after transformation can provide significant strengthening increments. In this regard, nitrogen level is also of importance. Judicious selection of both the vanadium and nitrogen levels is required to produce the desired level of precipitation strengthening. Similarly, an increase in the carbon content and thus the pearlite volume fraction of a bar steel can also be used to increase strength. Of course, these two strengthening mechanisms have very deleterious effects on toughness. Toughness levels currently available in commercial microalloyed bar steels may be adequate for many applications, but there is considerable interest in improving the toughness of microalloyed bar grades.

Dynamic Recrystallization in Bar Rolling. Similar to the controlled rolling of plate, dynamic recrystallization can be a factor in the rolling of bar, rod, or tube. At the finishing end of hot-rolling processes such as rod and bar rolling, the stock moves quickly, and the interpass intervals are of the order of 0.010 to 0.1 s. Under these conditions, there may not be enough time for significant amounts of conventional static recrystallization to take place, particularly when low finishing temperatures were employed. The strain, therefore, accumulates from pass to pass until the retained strain reaches the critical level required for the initiation of dynamic recrystallization (DRX), followed immediately upon unloading by postdynamic static recrystallization (PDSR). This has important consequences for the rolling loads and grain size evolution. The influences of the deformation parameters and chemical composition on PDSR are described in Ref 48.

In another analysis (Ref 49), data from pass-by-pass evolution of austenite grain size during the rod rolling of plain carbon steel was organized into a commercial computer spreadsheet program with previously developed equations relating grain size and hot-working parameters. By considering the substanial "redundant" strains developed during rod rolling, the analysis revealed that metadynamic recrystallization (MRX) is the dominant microstructural process. The spreadsheet was also employed to examine the effects of modifications to rod rolling practice aimed at refining the austenite grain size. Little benefit is obtained by increasing the strain rate or by reducing the distance from the final rolling pass to the laying head. On the other hand, increasing the cooling rate on the forced-air cooling deck should lead to measurable grain refinement. Lowering the temperature during rolling can lead to even more refinement, and various strategies to achieve this refinement are discussed in Ref 49.

Summary and Acknowledgment

Thermomechanical processing of steel by controlled rolling is an ongoing activity of research and development because property improvements in as-rolled steels products can have important engineering and economic implications. Bibliographic searches of current literature on this topic provide many examples, and thorough coverage is beyond the scope of this chapter. By necessity, this chapter only briefly describes some of the basic concepts and methods of controlled rolling.

Substantial portions of this chapter were adapted from Ref 1 as noted in the section headings. The editors extend their thanks for the reuse of this material.

REFERENCES

1. A.J. DeArdo, C.I. Garcia, and E.J. Palmiere, Thermomechanical Processing of Steels, *Heat Treating,* Vol 4, *ASM Handbook,* ASM International, 1991, p 237–255
2. I. Kozasu, C. Ouchi, T. Sampei, and T. Okita, *Microalloying 75,* M. Korchynsky et al., Ed., Union Carbide Corporation, 1977, p 120
3. T. Tanaka, N. Tabata, T. Hatomura, and C. Shiga, *Microalloying 75,* M. Korchynsky et al., Ed., Union Carbide Corporation, 1977, p 107
4. H. Sekine and T. Maruyama, *Trans. Iron Steel Inst. Jpn.,* Vol 16, 1976, p 427
5. W.B. Morrison and J.A. Chapman, *Philos. Trans. R. Soc. (London) A,* Vol A282, 1976, p 289
6. T. Tanaka, *International Metallurgical Reviews,* 1981, p 185
7. E.E. Underwood, *Quantitative Metallography,* McGraw-Hill, 1968, p 77
8. W. Roberts, H. Lindefelt, and A. Sandberg, *Hot Working and Forming Processes,* The Metals Society, London, 1980, p 38
9. L.J. Cuddy, *Metall. Trans. A,* Vol 15, 1984, p 87
10. G.R. Speich, L.J. Cuddy, C.R. Gordon, and A.J. DeArdo, *Phase Transformations in Ferrous Alloys,* A. Marder and J. Goldstein, Ed., The Metallurgical Society of AIME, 1984, p 341
11. Y. Zheng, G. Fitzsimons, and A.J. DeArdo, *HSLA Steels: Technology and Applications,* American Society for Metals, 1984, p 85
12. M.R. Blicharski, C.I. Garcia, S. Pytel, and A.J. DeArdo, *Processing, Microstructure and Properties of HSLA Steels,* A.J. DeArdo, Ed., The Metallurgical Society of AIME, 1988, p 317
13. R.A. Oriani, *Acta Metall.,* Vol 7, 1959, p 62
14. J.L. Walter and C.G. Dunn, *Trans. AIME,* Vol 215, 1959, p 465
15. E.L. Holmes and W.C. Winegard, *Acta Metall.,* Vol 7, 1959, p 411
16. R.A.P. Djaic and J.J. Jonas, *Metall. Trans. A,* Vol 4, 1973, p 621
17. R.A. Petkovic, M.J. Luton, and J.J. Jonas, *Can. Metall. Q.,* Vol 14, 1975, p 137
18. M.J. Luton, R. Dorvel, and R.A. Petkovic, *Metall. Trans. A,* Vol 11, 1980, p 411
19. G.L. Wang and M.G. Akben, *Processing, Microstructure and Properties of HSLA Steels,* A.J. DeArdo, Ed., The Metallurgical Society of AIME, 1988, p 79
20. L.J. Cuddy, J.J. Bauwin, and J.C. Raley, *Metall. Trans. A,* Vol 11, 1980, p 381
21. C. Zener, *Trans. AIME,* Vol 175, 1949, p 15
22. M.F. Ashby, *Recrystallization and Grain Growth of Multi-Phase and Particle Containing Materials: 1st RISO International Symposium on Metallurgy and Material Science,* N. Hansen, A.R. Jones, and T. Leffers, Ed., RISO National Laboratory, Roskilde, Denmark, 1980, p 325
23. T. Gladman, *Proc. R. Soc. (London) A,* Vol 294, 1966, p 298
24. M. Hillert, *Acta Metall.,* Vol 13, 1965, p 227
25. L.J. Cuddy, *Metall. Trans. A,* Vol 12, 1981, p 1313
26. A.J. DeArdo, University of Pittsburgh, unpublished research, 1983
27. C.I. Garcia and A.J. DeArdo, *Metall. Trans. A,* Vol 12, 1981, p 521
28. A.J. DeArdo, J.M. Gray, and L. Meyer, *Niobium,* H. Stuart, Ed., The Metallurgical Society of AIME, 1984, p 685
29. E.J. Palmiere, Ph.D. thesis, University of Pittsburgh, 1991
30. J.M. Gray and A.J. DeArdo, *HSLA Steels: Metallurgy and Applications,* J.M. Gray et al., Ed., American Society for Metals, 1986, p 83
31. L.J. Cuddy, *Thermomechanical Processing of Microalloyed Austenite,* A.J. DeArdo, G.A. Ratz, and P.J. Wray, Ed., The Metallurgical Society of AIME, 1982, p 129
32. I. Weiss et al., *Thermomechanical Processing of Microalloyed Austenite,* A.J. DeArdo, G.A. Ratz, and P.J. Wray, Ed., The Metallurgical Society of AIME, 1982, p 33
33. K. Tiitto, G. Fitzsimons, and A.J. DeArdo, *Acta Metall.,* Vol 31, 1983, p 1159
34. S.S. Hansen, J.B. Vander Sande, and M. Cohen, *Metall. Trans. A,* Vol 11, 1980, p 387
35. J.G. Speer and S.S. Hansen, *Metall. Trans. A,* Vol 20, 1989, p 25
36. J.J. Jonas, C.M. Sellars, and W.J. McG. Tegart, *Met. Rev.,* Vol 14, 1969, p 1
37. H.J. McQueen and J.J. Jonas, *Plastic Deformation of Materials,* R.J. Arsenault, Ed., Academic Press, 1975, p 393
38. J.D. L'Ecuyer and G. L'Esperance, *Acta Metall.,* Vol 37, 1989, p 1023
39. J.J. Jonas, *International Conference on Physical Metallurgy of Thermomechanical Processing of Steel and Other Metals,* I. Tamura, Ed., The Iron and Steel Institute of Japan, 1988, p 59
40. O. Kwon and A.J. DeArdo, Interactions Between Recrystallization and Precipitation in Hot-Deformed Microalloyed Steels, *Acta Metall. Mater.,* Vol 39 (No. 4), April 1991, p 529–538
41. A.J. DeArdo, *Mathematical Modelling of Hot Rolling of Steel,* S. Yue, Ed., CIM, 1990, p 220
42. T.M. Hoogendoorn and M.J. Spanraft, *Microalloying 75,* M. Korchynsky et al., Ed., Union Carbide Corporation, 1977, p 75
43. C. Ouchi, T. Sampei, and L. Kozasu, *Trans. Iron Steel Inst. Jpn.,* Vol 22, 1982, p 214
44. R.I.L. Guthrie and J.J. Jonas, Steel Processing Technology, *Properties and Selection: Irons, Steels, and High-Performance Alloys,* Vol 1, *ASM Handbook*, ASM International, 1990, p 117
45. F.H. Samuel, S. Yue, J.J. Jonas, and K.R. Barnes, *Effect of Dynamic Recrystallization of Microstructural Evolution During Strip Rolling,* ISIJ International, Vol 30 (No. 3), March 1990, p 216–225
46. L.N. Pussegoda, S. Yue, and J.J. Jonas, Laboratory Simulation of Seamless Tube Piercing and Rolling Using Dynamic Recrystallization Schedules, *Metall. Trans. A,* Vol 21 (No. 1), Jan 1990, p 153–164
47. Bulk Formability of Steels, *Properties and Selection: Irons, Steels, and High-*

Performance Alloys, Vol 1, *ASM Handbook,* ASM International, 1990, p 587

48. J.J. Jonas; T.M. Maccagno, and S. Yue, The Role of Dynamic Recrystallization in Industrial Hot Working, *Conf. Proc.: Phase Transformations During the Thermal/Mechanical Processing of Steel,* Vancouver, 20–24 Aug 1995, Canadian Institute of Mining, Metallurgy and Petroleum, 1995, p 179–193
49. T.M. Maccagno, J.J. Jonas, and P.D. Hodgson, *Spreadsheet Modelling of Grain Size Evolution during Rod Rolling.* ISIJ International, Vol 36 (No. 6), 1996, p 720–728

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 258-277
DOI:10.1361/hwpd2003p258

Chapter 18

Workability and Process Design in Rolling

John G. Lenard, University of Waterloo

THE MANUFACTURE of successful products requires knowledge about the three basic elements of any metal forming operation: the machine, the workpiece, and the interface where the transfer of mechanical and thermal energies takes place. These three elements determine both the workability conditions and the process-design factors for successful manufacture of a part. In the overall context of this book, workability can be defined in general terms as any conditional limit on the successful creation of a part. Likewise, this is closely related to process design, where the three elements are properly configured in technological terms to manufacture products with acceptable quality in terms of the mechanical, metallurgical, and geometrical requirements of the customer.

In this chapter, the subject of workability and process design is reviewed for metal rolling based on some current studies, as well as important studies conducted earlier. Substantial portions of this chapter are due to Schey's original publication on this topic (Ref 1), and many of the comments and explanations of the issues are as valid now as they were in 1984. Since the publication of the original review of workability in rolling, numerous examinations of the topic have been presented in the technical literature. While many of these considered physical simulations and the development of failure criteria, a significant portion of the studies involved the application of mathematical models to the prediction of the occurrence of defects. In what follows, the phenomenological issues of rolling are discussed first. Mathematical models of various complexities are presented next. These will then be applied to workability in general terms and then in reference to rolling. A presentation of process design in rolling follows. The chapter closes with some general and specific conclusions.

The Rolling Process

As noted, the three basic elements of a forming process are the machine, the workpiece, and the interface where the transfer of mechanical and thermal energies takes place. In the flat or shape rolling process, the machine refers to the rolling mill, including the driving motor, the work rolls, the back-up rolls, the screw-down system, the mill frame, the bearings, the drive spindles, the scale breakers, the coilers, loopers, and the control system of the mill. The parameters of relevance that may affect workability include the available power and speed of the driving motor, the mill stretch and stiffness, the resonant frequencies, the lubricant, emulsion and the coolant delivery systems, the manufacture of the roll, material, strength, hardness, diameter and crown, the furnaces, and the responsiveness of the control systems.

The interface indicates the contact between the work roll and the rolled metal, and the parameters describing its mechanical and tribological behavior include the roughness and hardness of the roll and the strip, the presence of a lubricant, an emulsion, a coolant, and the presence of oxides and/or scales. The workpiece is the rolled metal, which may be flat—that is, of rectangular cross section—or may have a particular, nonrectangular cross section. The attributes and parameters of significance refer to the resistance of the workpiece to deformation as a function of mechanical and metallurgical variables. It is argued that while the list is certainly incomplete, it includes the most significant parameters. Further, it is well understood that no realistic analysis of the metal rolling system can take all of these factors into consideration, and mathematical models for acceptance and use by rolling mill engineers invariably contain simplifying assumptions and approximations. A fairly recent addition to the family of mathematical models is the use of artificial intelligence techniques, which, if well designed, may in fact take all parameters and possibilities into consideration and may be continuously upgraded using the most current data.

This section briefly reviews some phenomenological aspects of rolling, which is also introduced in Chapter 16, "Rolling." The flat and shape rolling processes, given their highly developed economic significance, are also reviewed in detail in a large number of publications. Some of the books devoted exclusively to rolling include Ref 2 to 11. Texts concerned with the theory of plasticity and metal forming include chapters on the rolling process. In addition to the traditional texts (Ref 12–18), the books cited in Ref 19 to 21 should be mentioned. Publications in various technical journals, concerned with mechanical and metallurgical aspects of flat or shape rolling, are numerous and the list is much too long to be given here.

The Forces, Torques, and Speeds in Flat Rolling (Ref 22). While the flat rolling process is well known, the sequence of events experienced by the rolling mill and the metal indicate its complexity, and the fundamental ideas described next (adapted from Ref 22) briefly consider the forces acting in the roll-metal system.

The strip enters the deformation zone because of the friction forces exerted by the work rolls on it, and as a result, it first experiences elastic deformation. The limit of elasticity is reached soon after entry. The permanent deformation regime is thus in existence through most of the roll gap region, followed by the elastic unloading regime. These regimes are illustrated in Fig. 1(a) to (c), which show a schematic of a two-high mill and a strip in various stages of the process. In Fig. 1(a) the strip is ready to be rolled and entry is imminent. In Fig. 1(b) it is rolled partway through to roll gap and in Fig. 1(c) the rolling process is continuous and the forces and torques acting on the work roll are also shown. The conditions shown describe either a laboratory situation where no front and back tensions exist, or a single stand, reversing roughing mill.

In Fig. 1(a), where the strip is about to make contact, if the coefficient of friction is larger than the tangent of the bite angle, a relationship that is often used to determine the minimum friction necessary to start the rolling process, the strip enters the deformation zone. In a laboratory mill, the usual practice is to push the strip, placed on the delivery table, toward the work rolls and allow the friction forces to cause entry; under certain circumstances it is necessary to taper the leading edge of the strip to facilitate the bite. In a strip mill, edge rolls force the strip into the roll gap in the first stand, and the momentum

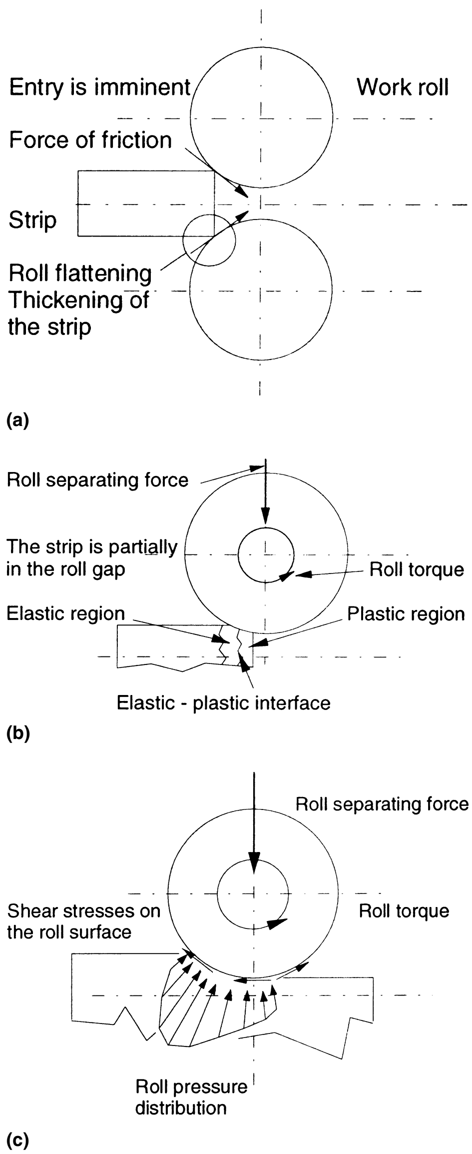


Fig. 1 (a) Schematic of the entry of the strip into the roll gap. (b) The strip is partially in the deformation zone. (c) Free body diagram of the work roll and the rolled strip

of the strip exiting from there carries it into the deformation zone of the next stand. In either case, entry creates some longitudinal compression at the leading edge of the strip, where there will be some initial thickening as well. This is accompanied by local, elastic deformation of the work rolls, indicating that the usual simplification about the entry point—located where the edge of the strip encounters the undeformed, perfectly cylindrical roll—does not represent reality very well. Relatively little has been done to analyze the exact entry conditions of the strip into the roll gap, indicated by the circle in Fig. 1(a). In one of the attempts, high-speed photography was used (Ref 23) to allow visualization of the entry conditions and the length of contact in the hot-rolling process. A doctoral dissertation (Ref 24) on the topic has also been produced; however, the writer is not familiar with publications that may have resulted from the research.

Figure 1(b) shows the strip about halfway through the deformation zone. As previously mentioned, the entering metal first experiences elastic deformation, and when and where the yield criterion is first satisfied, plastic flow is observed. These two regimes are separated by an elastic-plastic boundary, the location of which should be determined by an analysis of the rolling process. In the elastic region, the theory of elasticity describes the deformation. In the permanent deformation region, the criterion of yielding, the associated flow rule and the condition of incompressibility describe the situation.

The rolls are further flattened. The magnitude of the roll stresses should not exceed the yield strength of the roll material. The theory of elasticity is to be used to determine the roll distortion and the corresponding changes of the length of contact.

In Fig. 1(c), the leading edge of the rolled metal has exited and the process is continuing. The figure shows the pressures, forces, and torques acting on the roll and strip. These include the roll pressure distribution and the interfacial shear stress, the integrals of which over the contact length lead to the roll separating force and the roll torque. These are the variables the models are designed to determine. If front and back tensions are present, as would be the case under industrial conditions, their effect on the longitudinal stresses at the entry and exit should be included in the definitions of the boundary conditions.

The surface velocities of the roll and the strip should also be considered. It may be assumed that the rolls rotate at a constant angular velocity, even though there may be some slow down under high loads. The strip usually enters the roll gap at a surface velocity less than that of the roll. The friction force always points in the direction of the relative motion, and on the strip it acts to aid its movement. As the compression of the strip proceeds, its velocity increases and it approaches that of the surface of the roll. When the two velocities are equal, the no-slip region is reached, often referred to as the neutral point, or more correctly, a neutral region, where the strip and the roll move at the same speed (see Chapter 16, "Rolling"). If the neutral point is between entry and exit, beyond it the strip experiences further compression, and its surface velocity exceeds that of the roll. Between the neutral point and the exit, the friction force on the strip has changed direction and is now retarding its motion.

Figure 2 illustrates the side-view of the flat rolling process, showing L, the projected contact length, and φ, the independent variable in the rolling direction. It is generally accepted that in strip rolling the field variables—stress, strain,

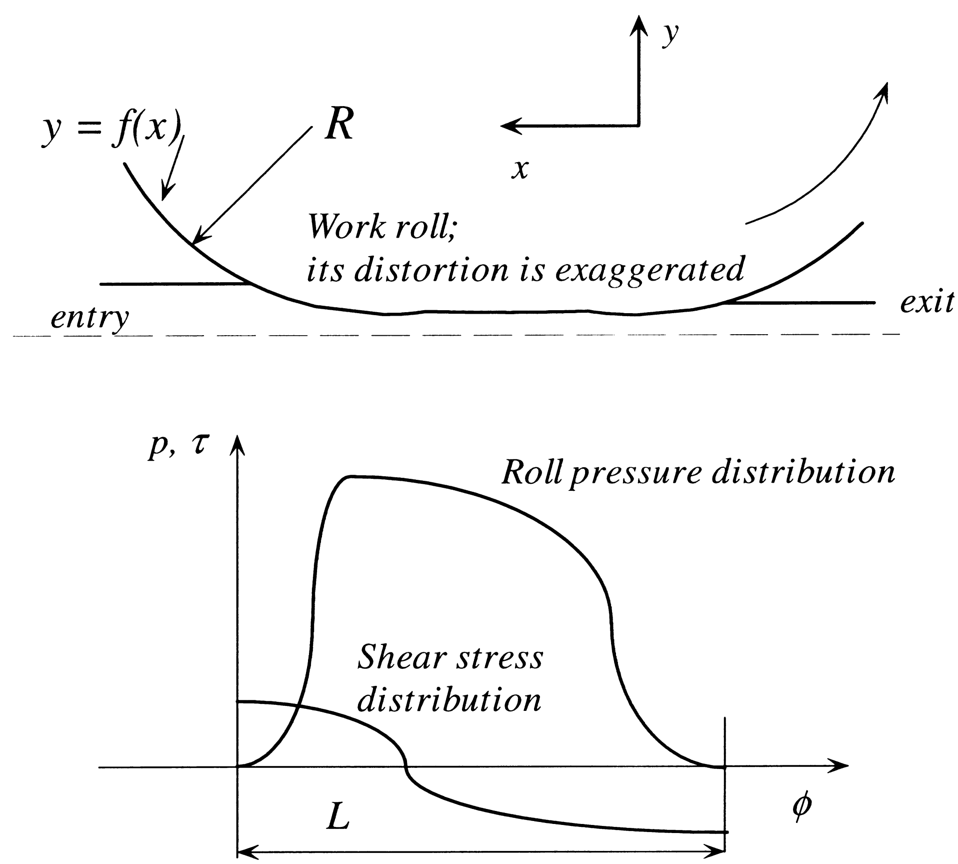


Fig. 2 Schematic of the flat rolling process, showing the distorted work roll, the roll pressure, and the interfacial shear stress distributions

rate of strain, and temperature—don't vary in the direction perpendicular to the plane shown. This is a consequence of the very low straining the rolled metal experiences in the width direction and makes the assumption of plane-strain deformation realistic. A similar schematic of the rolling process appeared in the first edition (Ref 1), but in more idealized form with undeformed work rolls and the traditional friction hill, in which the location of the saddle point is taken as the location of the no-slip region (Fig. 3). The sharp saddle point, or cusp, in Fig. 3 results from one-dimensional analysis of the abrupt change in the direction of frictional resistance at the neutral point. If a more realistic friction model is used to relate relative slip between the roll and strip, then the sharp cusp in the traditional friction hill is replaced by a smooth-hill top.

Figure 2 indicates the deformation of the work roll, albeit in a highly exaggerated fashion and gives its shape as $y = f(x)$, to be determined as part of any mathematical analysis. The roll pressure distribution is also different, depicting a more realistic, smooth shape, not the sharp saddle-point obtained in simplified one-dimensional analyses. It is recalled that in the traditional, one-dimensional analyses of the flat rolling process (Ref 25–27), the roll pressure distribution is obtained by integrating the differential equations of equilibrium from the entry and from the exit, and the location of the neutral point is taken as the place where the two curves meet. However, it is also well known that the one-dimensional analysis of the friction hill approach can fail when interface friction assumptions fail. One such case occurs when large reductions of thin, hard strips using large diameter rolls are analyzed.

The shear stress distribution in Fig. 2 shows the location of the no-slip point, which does not necessarily coincide with the location of the maximum roll pressure, as demonstrated experimentally (Ref 28–30). Also, when strips are rolled at higher speeds to higher reductions and lubricants of higher viscosities are used, no neutral point is located, the strip exits at a velocity less than the surface velocity of the work roll, and the forward slip becomes negative. These differences are of importance when cold rolling of thin strips and foils are considered. They are of less importance in the study of the hot rolling of strips and plates when the resistance of the rolled metals is substantially lower.

Homogeneous Versus Non-Homogeneous Deformation. The usual assumption of homogeneous deformation, that is, planes remaining planes during a rolling pass, made in most one-dimensional mathematical models of the flat rolling process, represents the actual conditions very well, provided the ratio of the roll diameter (D) to the strip thickness (h) is significantly larger than unity. This condition is well satisfied when cold rolling of strips and foils is considered, in which case the D/h ratio may approach 10^3. When hot rolling strips in the finishing train of strip mills, this ratio is smaller but the deformation is not too far from being homogeneous. The deformation may well deviate, in a significant manner, from the homogeneous condition during rough rolling of strips, plate rolling, or bar rolling, however. The important parameter is the shape factor (Ref 1), defined as the ratio of the average thickness of the rolled metal to the projected contact length, h/L. The description of the role of this ratio, given in Ref 1 is most relevant and in what follows, it is repeated essentially unchanged. When $h/L > 1$, deformation becomes somewhat similar to that observed in the indentation of a slab with two opposing narrow tools. At high values of h/L, deformation is concentrated in zones adjacent to the rolls (Fig. 4) and does not penetrate the full thickness of the workpiece. Deformation against the restraint presented by the rigid (or only elastically deformed) central portions of the workpiece necessitates a rise in interface pressures (Ref 31) for forging and then recognized as relevant for rolling (Ref 32). Because rolling may be regarded as a progression of indentation (Ref 33), the interface pressure can be calculated by analogy to the plane-strain indentation of a slab, either from the curve given by Hill from slip-line field analysis (Ref 12), or from the experimental relationship determined by Smirnov and reported in Ref 9:

$$\frac{p}{2k} = \left(\frac{h}{L}\right)^{0.4} \leq 2.57$$

where p is the average interface pressure and $2k$ is the plane-strain flow stress of the workpiece material. It is immaterial whether the interface is lubricated; the pressure increase is due to geometric restraint and not to friction.

The transition from almost homogeneous to highly inhomogeneous deformation is, of course, not abrupt. At intermediate values of h/L, the deformation zone may be extremely complex (Ref

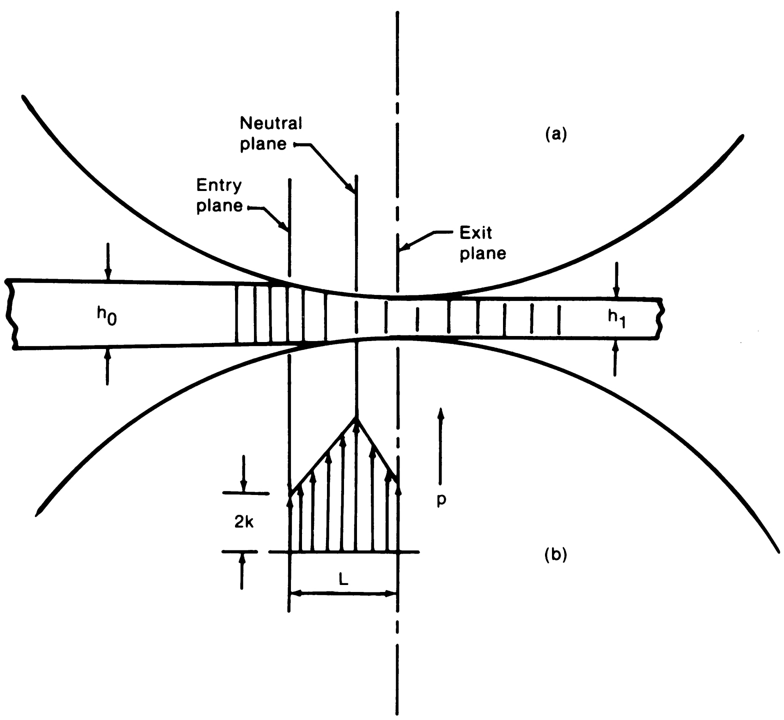


Fig. 3 Idealized view of homogeneous plane-strain rolling without deformation of work rolls. Source: Ref 1, p 270

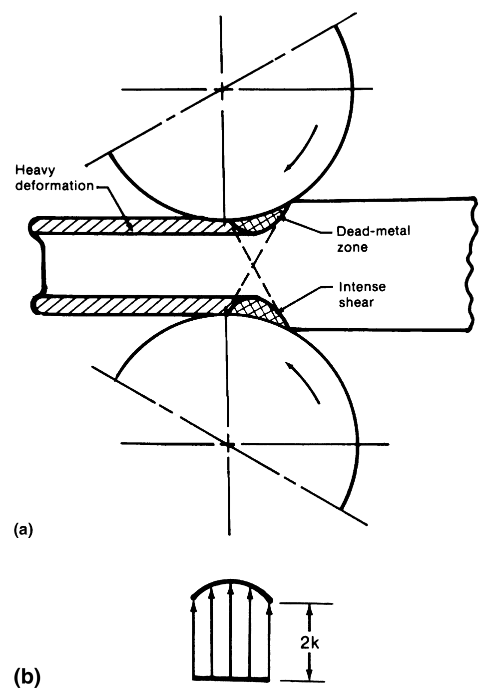


Fig. 4 Highly inhomogeneous plane-strain deformation at large h/L ratios. (a) Deformation. (b) Pressure distribution in arc of contact. Source: Ref 1, p 272

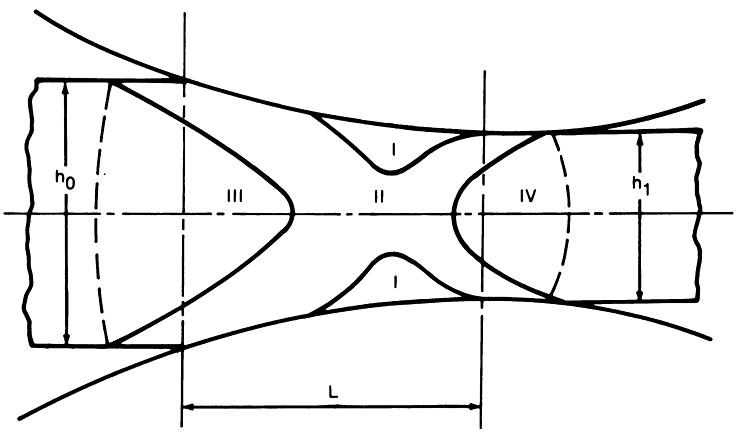


Fig. 5 Deformation zones at intermediate *h/L* ratios. (I) Dead-metal zone; (II) Intense deformation. (III and IV) Indirect deformation zones. Source: Ref 1, p 273

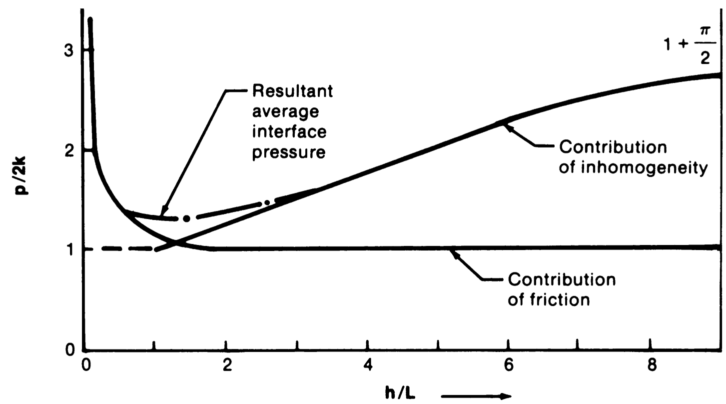


Fig. 6 Pressure-multiplying factor as a function of *h/L* ratio. Source: Ref 1, p 274

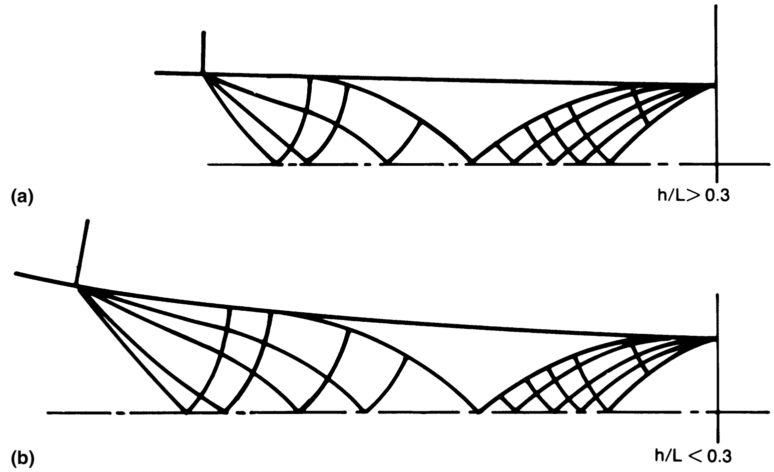


Fig. 7 Slip-line fields for hot rolling. (a) Single dead-metal zone extending from entry to exit. (b) Relative motion between strip and roll surface at entry and exit. Source: Ref 1, p 274

11). Rigid (dead-metal) zones extend partway into the body of the workpiece from the sticking zone on the roll surface (zones I in Fig. 5). Intense deformation occurs between these two zones, whereas minor plastic deformation takes place outside the entry and exit where the workpiece actually thickens slightly while contracting longitudinally; this is a consequence of the interaction of the deforming zone with the rigid parts of the outlying portions of the workpiece.

The combined effects of inhomogeneous deformation and friction on roll forces can be readily calculated, at least approximately, by combining the effects of the *h/L* ratio and the friction hill; the resulting curve shows (Fig. 6) a minimum of the ratio $p/2k$ at around $h/L = 1$. Such a trend is indeed found when roll pressures calculated from numerous production roll-force measurements are analyzed (Ref 34). Much more difficult is a theoretical treatment of the inhomogeneous deformations themselves, but the deformation-zone geometry and the relative velocities for several geometries typical of hot rolling were identified. It was discovered that a single dead-metal zone (Fig. 7a) extends from entry to exit at combinations of relatively low pass reduction and roll-radius to strip-thickness ratio. Their relationship can be reduced to the statement that a single, fully extended dead-metal zone exists at $h/L > 0.3$ (Ref 36), or at $h/L = 0.5$. At lower *h/L* values, the rigid zone is bounded by two intense deformation zones in which relative motion occurs at the strip/roll interface (Fig. 7b). This solution puts the transition from full to partial sticking at a lower *h/L* value than that found in the experiments reported in Ref 11. In these latter experiments, unusually smooth rolls were used for unlubricated hot rolling of aluminum and copper, but an approach to full sticking was found only at $h/L \geq 2$. The distributions of stresses, strains, rates of strains and temperatures, supporting the above conclusions, can be determined by the use of mathematical models.

Mathematical Models of the Flat Rolling Process. The effects of process, material, and geometrical parameters on the rolling mill and the rolled material may be predicted using mathematical models of various sophistication. The output of the models depends on their mathematical rigor and the assumptions made during the derivations. They may include the roll separating force and the roll torque; the radius of the deformed roll, the roll pressure and the interfacial shear stress; or, if two- or three-dimensional finite element or finite difference approaches are used, the distributions of the displacements, velocities, strains, strain rates, stresses, temperatures and grain sizes. Further sophistication also allows the calculation of the post-rolling attributes: the yield strength, the hardness and the size of the austenite/ferrite/pearlite/bainite grains.

The models, discussed in brief subsequently, may be used to shed light into various mechanisms that operate during the rolling pass and from that point of view, are of immense use when workability limits are analyzed. Further, the models may be used for the predictions of the mechanical parameters and material attributes, and in this context their contributions are of importance in the design of the hardware or of the rolling process. The models may be "on-line" or "off-line". While in the former, the time taken for computation must be short, in the latter the time required for number crunching may be long due to the iterative calculations.

The models of the rolling process, for potential on-line use, include both empirical and semi-empirical types (Ref 27, 37, 38). Simple relations for the roll separating force and the roll torque, for use during hot rolling of carbon steels, are given in Ref 37. Simplified formulae for the same parameters are presented in Ref 27 and 38.

The curves presented in Ref 39 may also be used, but some method of transferring the values, predicted by the curves, to a computer needs to be devised. A more elaborate, iterative one-dimensional model (Ref 25) includes the effect of roll flattening, the contributions of the elastic entry and exit regions, and strain hardening as the strip or plate is traveling through the roll gap. The models presented in Ref 26 and 27 are both obtained from the Orowan model (Ref 25) through various simplifications. A statistical analysis of the predictive capabilities of Orowan's and Ford and Alexander's (Ref 37) models indicated that the standard deviation of the difference between measurements and predictions decreased with the increasing sophistication of the model (Ref 40).

The one-dimensional models, most of which assume that planes remain planes, that is, the rolled strip experiences homogeneous compression, are useful for the prediction of the roll separating forces, as long as two conditions are satisfied. The first is a good understanding of the tribological conditions at the contact surfaces, allowing one to use an appropriate value for the coefficient of friction. The second is a mathematical limitation; the roll diameter to strip thickness ratio should not be excessive since otherwise the models will not converge. Further, since the calculations of the roll torque are performed by integrating over the contact surface and the true extent of that is not known accurately, torque predictions are seldom accurate and consistent.

If more accuracy *and* consistency of predictions are required, a two-dimensional rigid-plastic finite element model (FEM) may be necessary. Many of these models (Abaqus, LS-Dyna, MARC; Deform; Forge, for example) are commercially available. Many of them are based on the minimization of the required energy, and they determine the roll torque from the power thus resulting in improved prediction of both the roll separating force and the torque. For proper functioning, they need an appropriate description of the boundary conditions, and they also need appropriate values of the coefficients of friction and heat transfer, both of which may be difficult to provide.

The rigid-plastic formulation is considered adequate for the simulation of hot rolling. For suitable analysis of cold rolling, an elastic-plastic model is necessary. Rigorous analysis of shape rolling needs three-dimensional models. Models, capable of predicting the mechanical and metallurgical attributes after hot rolling, are also available (Ref 22).

Stress, Strain, Strain Rate and Temperature Distribution in the Flat Rolling Process. As just stated, the states of stress, strain, rate of strain and temperature in the rolled sample will determine the limits of workability of the rolled metal. These distributions are shown in Fig. 8, 9, and 10, obtained by a two-dimensional rigid-plastic finite element approach, using the commercially available program Elroll (Ref 22). Hot rolling of a C-Mn steel plate, reduced by 40% in

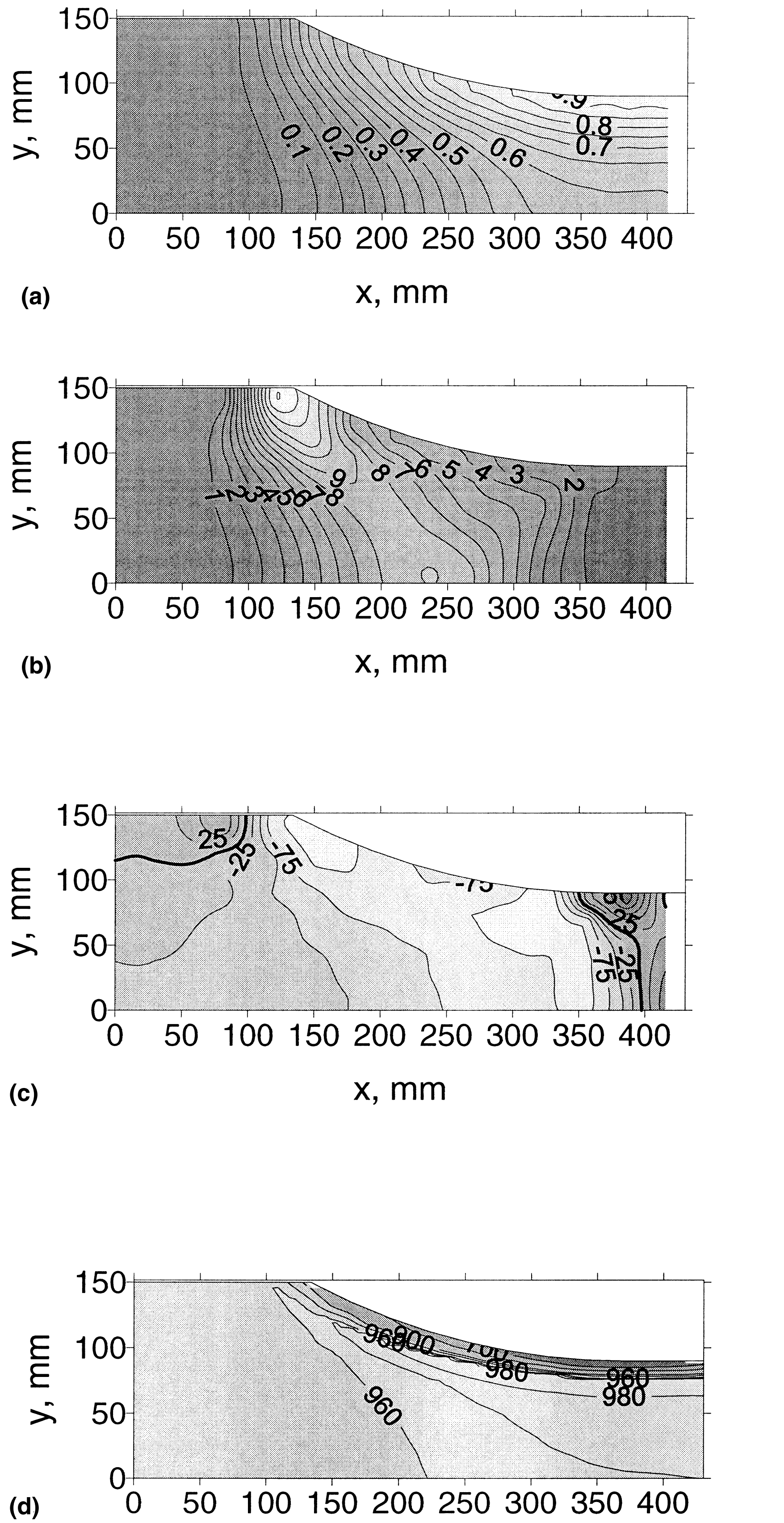


Fig. 8 (a) The effective strain distribution; h_{entry} = 300 mm (12 in.); h/L = 0.98. (b) The effective strain rate distribution; h_{entry} = 300 mm (12 in.); h/L = 0.98. (c) The effective stress distribution; h_{entry} = 300 mm (12 in.); h/L = 0.98. (d) The temperature distribution; h_{entry} = 300 mm (12 in.); h/L = 0.98

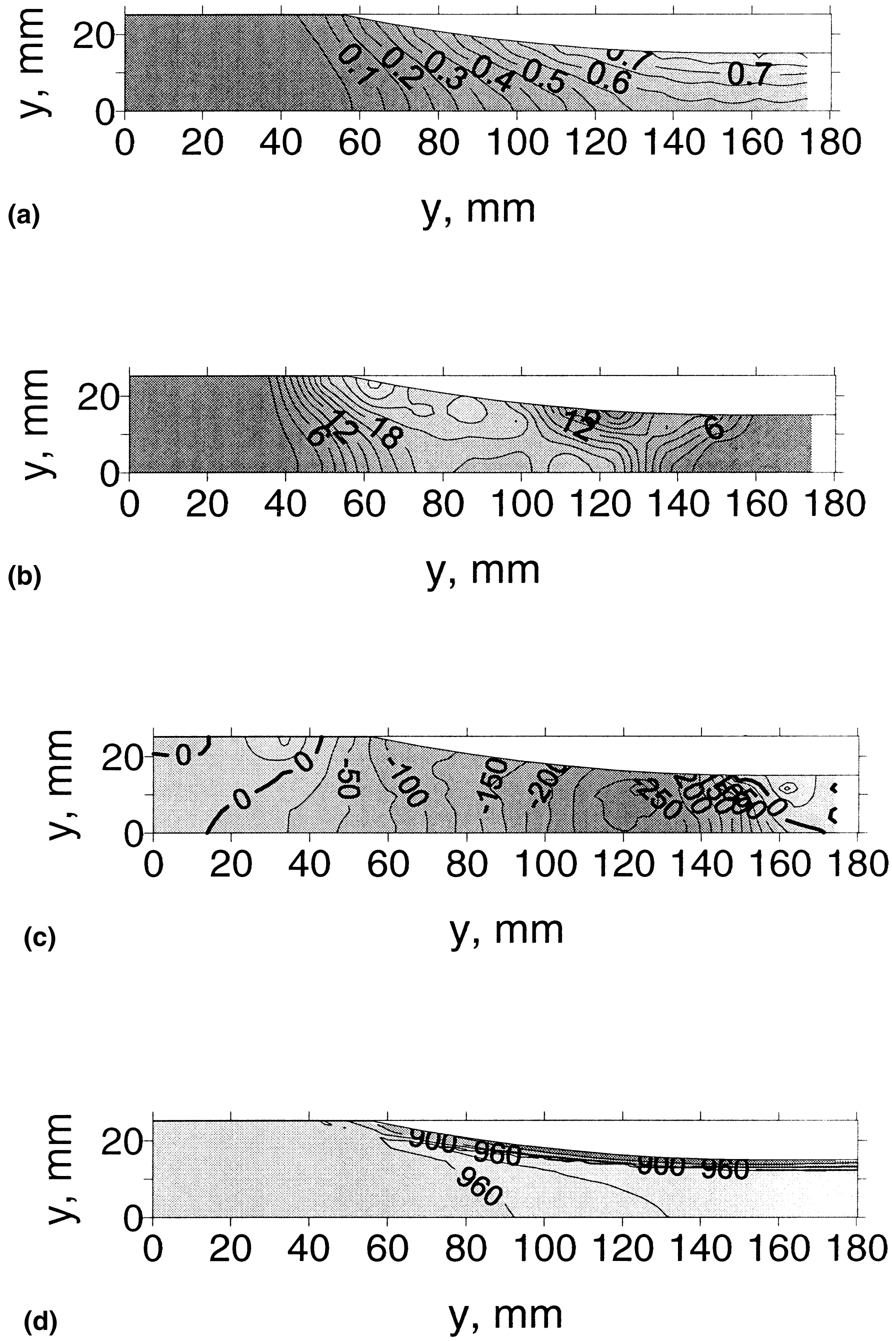


Fig. 9 (a) The effective strain distribution; h_{entry} = 50 mm (2 in.); h/L = 0.40. (b) The effective strain rate distribution; h_{entry} = 50 mm (2 in.); h/L = 0.40. (c) The effective stress distribution; h_{entry} = 50 mm (2 in.); h/L = 0.40. (d) The temperature distribution; h_{entry} = 50 mm (2 in.); h/L = 0.40

one pass is presented, similar to a pass in the roughing stand of a hot strip mill. The calculations were performed and the figures prepared by Prof. M. Pietrzyk, (Akademia Górniczo-Hutnicza, Krakow, Poland) using Elroll, a 2D finite element program.

The slab is heated to a uniform temperature of 950 °C (1740 °F) before rolling. The process parameters are as follows. The roll radius is 500 mm (20 in.), the roll velocity is 50 rpm, and the friction coefficient is 0.4. The equations in Ref 41 describing the flow strength as a function of the carbon content, temperature, strain rate and strain, are used as the constitutive model, with the carbon equivalent taken as 0.3. The FE model also requires the coefficient of heat transfer, and in the figures that follow, 50,000 W/m^{2}K was used. In all calculations, the roll was assumed to remain rigid, not accurate, but not unreasonable, in the hot-rolling process.

Figure 8 demonstrates the distributions of the strains (a), the rates of strains (b), the stresses (c) and the temperatures (d), considering a thick plate of 300 mm (12 in.) entry thickness, having a shape factor of h/L = 0.98. In Fig. 9, the same distributions are given, but the entry thickness is reduced to 50 mm (2 in.), leading to an intermediate shape factor of 0.4. Finally, in Fig. 10, the distributions are repeated, and a thin strip of 2 mm (0.08 in.) entry thickness is considered. In this case, the mean thickness to projected contact length ratio is 0.08.

The expected nonhomogeneity of the deformation is clearly observable in Fig. 8(a), (b), (c), and (d). The largest effective strains are located near the exit plane, at the contact surface in between the work roll and the rolled strip. The strains reduce toward the centerline, and the strain gradient in the through-thickness direction is high, pointing to the localization of the strains near the exit. The nonhomogeneity of the deformation is also observable at the entry. The rates of strain are also distributed in a nonhomogeneous manner, as noted in Fig. 8(b). The maximum strain rates, however, are predicted to be near the entry point, and they diminish toward the exit. The distribution of the stresses (Fig. 8c), indicates the highest concentration near the entry, as was the case with the rates of strain. This is expected, since at the temperatures considered, the resistance of the steel to deformation is mostly dependent on strain rates. The stress maxima are observed in the contact zone where intense temperature gradients are also found, as shown in Fig. 8(d). The distributions of the temperatures are also highly nonhomogeneous and the surface of the rolled strip is predicted to cool to a low of 700 °C (1300 °F), creating a material with sharply reduced ductility.

The nonhomogeneity is reduced, but still noticeable, when the shape factor is reduced to 0.4, as indicated in Fig. 9(a) to (d), showing the results when the entry thickness of the hot strip was reduced to 50 mm (2 in). The distributions of the effective strains and strain rates are not significantly different from that shown in Fig. 8(a) and (b); the largest strains are near the exit and the zone of contact and the largest rates of strain are near the point of entry. Both indicate a certain measure of nonhomogeneity. The largest stresses are now found near the possible location of the maximum roll pressure and nearer the centerline of the rolled strip, demonstrating the reduced nonhomogeneity. The magnitudes of the strain rates and, hence, the stresses are larger than those shown in Fig. 8(b) and (c) for the thicker strip, necessitated by the creation of a larger portion of stressed metal. Not much difference in the temperature distributions is observed when Fig. 8(d) and 9(d) are compared; the coolest part is again near the surface of contact. The loss of the temperature is reduced; however, the surface is warmer and correspondingly more ductile. Almost complete homogeneous compression is observed in Fig. 10(a) to (d) where the shape factor is 0.08. The major differences concern the very much larger rates of strain and the very high stresses as a result.

These stress and strain distributions may be used to indicate the potential locations of ductile fracture, based in the flow localization

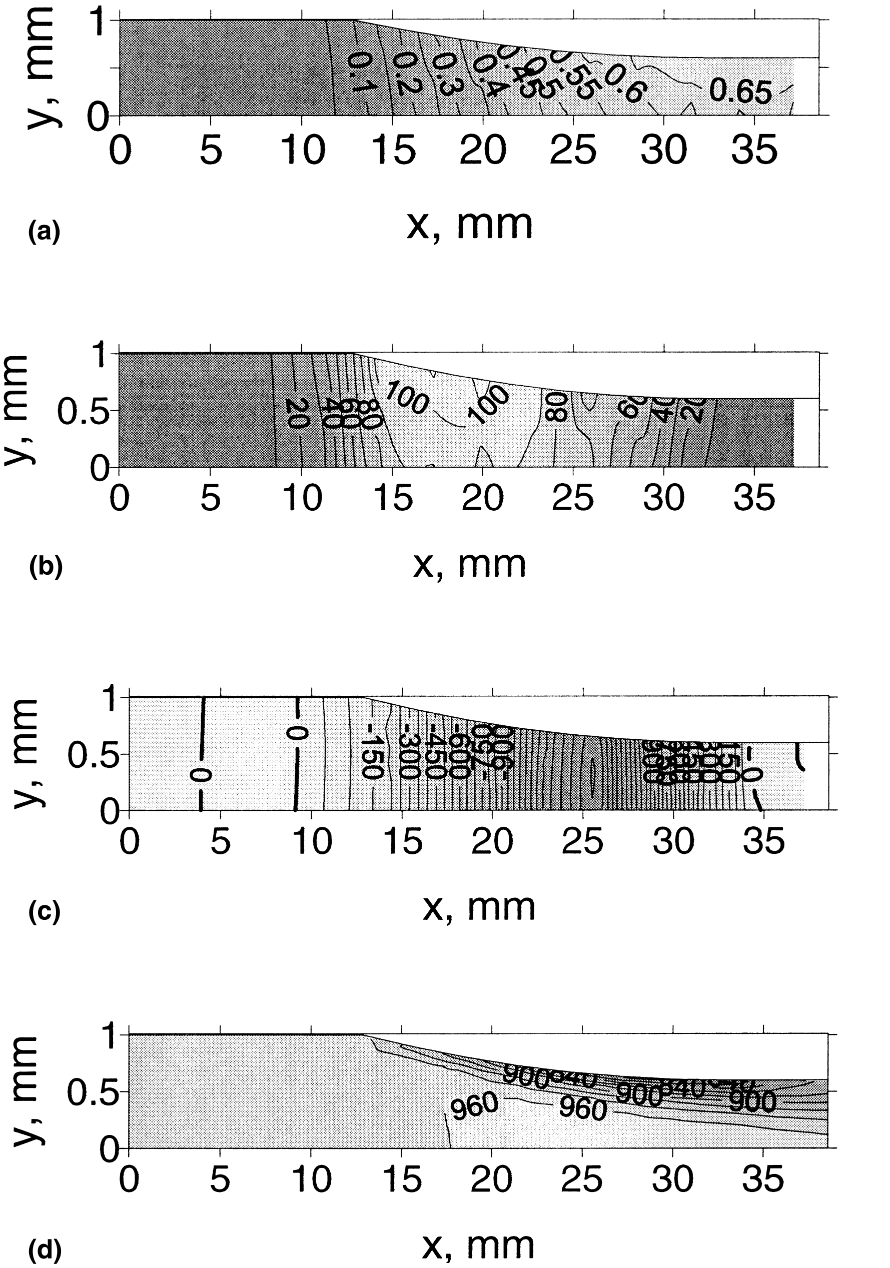


Fig. 10 (a) The effective strain distribution; h_{entry} = 2 mm (0.08 in.); h/L = 0.08. (b) The effective strain rate distribution; h_{entry} = 2 mm (0.08 in.); h/L = 0.08. (c) The effective stress distribution; h_{entry} = 2 mm (0.08 in.); h/L = 0.08. (d) The temperature distribution; h_{entry} = 2 mm (0.08 in.); h/L = 0.08

technique. It is realistic to conclude that fracture occurs (depending on the chosen failure criterion) where the stresses, strains, and rates of strains are the highest and where their gradients are high. These, combined with the loss of ductility resulting from roll chilling, are usually responsible for the types of defects that cause the limit of workability to be breached. Another approach, not used often to predict failures, is the application of the three-dimensional Mohr's circle, which allows the visualization of potential surfaces of fracture. If the state of stress can be determined, the surface of fracture may be predicted when a failure criterion is defined in conjunction with the shear stresses (e.g., see Chapter 12, "Workability Theory and Application in Bulk Forming Processes").

Workability in Rolling

The limits introduced by the rolling mill refer to its load, torque, and power carrying capability, including its deformation and dynamic response, and these are best considered when designing the rolling process, discussed in some detail in the section "Process Design." The limits contributed by the workpiece are twofold. The first set refers to mechanical phenomena and is concerned with the ability of the rolled metal to deform as the rolling process demands (Ref 52). These involve the nonhomogeneity of the variables— strains, rates of strain, stresses, and the temperatures— discussed previously. The second limitation is metallurgical in nature and is caused by the often-observed ductility trough. This is most significant when hot rolling of plates and strips is considered. The limits indicate the reaction of the material to the thermal-mechanical treatment it is receiving, including the effect of the treatment on ductility.

The limits associated with the interface, the contact zone between the rolled metal and the roll, affect the quality of the surfaces. These include friction, heat transfer, and the development of the roughness of the rolled metal, all of which depend on tribological phenomena and are influenced by the transfer of mechanical and thermal energies. These factors are not directly related to fracture, and they are not often considered in the usual treatments of workability. However, the interface has much to do with fracture. The heat transfer at the interface and the nature of surface roughness (at the interface) can affect surface fracture. Indirectly, friction also affects internal stresses and even the formation of the edge profile, which can have a large influence on edge cracking.

Workability Limits due to Nonhomogeneity

In the idealized situation of homogeneous deformation under conditions of plane-strain rolling, straight vertical sections in the workpiece remain almost straight (Fig. 3). No secondary tensile stresses are developed during rolling, and no residual stresses remain in the product. This idealized result from elementary plasticity theory assumes that an ideal rigid-plastic workpiece is rolled on relatively large-diameter rigid rolls at heavy reductions. Of course, this idealization is not completely realistic in terms of deformations. Even under conditions of plane-strain (approximated under conditions of thin products with large widths and lengths), deformation can be frequently inhomogeneous in the through-thickness direction. At large h/L ratios, for example, secondary tensile stresses may develop in the centerplane and could lead to center bursts. Conversely, with decreasing h/L ratios, the danger of inhomogeneous deformation now shifts to the surface. Wherever the actual point of transition from homogeneous to inhomogeneous deformation may be, it is evident that a variety of situations can develop that may be harmful in rolling of materials, as described subsequently and in Chapter 12, "Workability Theory and Application in Bulk Forming Processes."

A detailed, systematic study of the limits of workability during hot rolling of aluminum slabs was conducted, using samples of rectangular and tapered cross sections. The state of stress in the hot rolled slabs was examined and the possible planes of fracture were identified using the three-dimensional Mohr circle analysis (Ref 43). Two alloys were selected for this investigation: a 5182 alloy, containing 0.06% Cu, 0.37% Mn and 4.65% Mg, which has a susceptibility to alligatoring and edge cracking during the early stages of hot rolling, following reheating, that is known to be very high, and a 7075 alloy containing 1.45% Cu and 2.5% Mg, which is normally extruded rather than rolled, but possesses hot ductility that is very different from that of the 5182 alloy. During the process, the work rolls were covered with a thin layer of mineral seal oil to eliminate sticking. The speed of rolling was 50 rpm, corresponding to a surface speed of approximately 0.4 m/s (1.5 ft/s). The average shape factor in the tests was 0.6, near the shape factor of 0.4, analyzed previously and for which the distributions are given in Fig. 9(a) to (d). Several types of workability limits were discovered, including surface cracking, alligatoring, edge cracking, and a combination of these. These limits are discussed below, first considering the rectangular cross-section samples, followed by those with tapered edges.

Surface Cracking. The cracking of the surfaces is caused by the combination of roll chilling, as indicated in Fig. 8(d), 9(d), and 10(d), and the corresponding loss of metal ductility. The temperature gradients shown in Fig. 8(d), 9(d), and 10(d) indicate the severity of the cooling of the surfaces. This phenomenon, combined with the possible loss of lubricants, either due to lubricant breakdown as the temperatures rise or due to improper lubricant delivery, results in the dramatic increase of the interfacial shear stresses, approaching or surpassing the metal's resistance to deformation (Ref 44). The cracks usually occur in a direction parallel to the roll's axis of rotation.

Surface cracking is demonstrated in Fig. 11, showing a sample of the 5182 alloy, rolled at 560 °C (1040 °F) to a strain of 0.58. It is interesting to observe that the surface cracks are deepest near the edges and are located near the leading edge of the sample. They appear to vanish as the rolling pass nears completion. The resistance of the metal to deformation at the temperature of the pass is highly dependent on the rate of strain, and, as shown in Fig. 8(b), 9(b), and 10(b), the localization of the strain rates is highest at the entry. The rates drop as the exit is approached. Contributing to this is the possibly incomplete coverage of the roll's surface by the layer of the mineral seal oil, some of which may have burnt off as the pass was nearing completion. The potential loss of lubrication is overwhelmed by the strain rate localization, and the trailing portion of the strip, where strain rates are more uniform, does not demonstrate surface cracking. Once sufficient lubrication was available, no cracks were produced. The 7075 alloy, rolled at 460 °C (860 °F), also suffered some minor surface cracking, as in Fig. 12. While the cracks appear to cover all of the surface of the strip, from entry to exit, they don't extend over the complete width. It is possible that surface lubrication was more effective in the test shown in Fig. 12. The strain rates in the two tests were similar, so the superior ability of the 7075 alloy to avoid surface cracking is demonstrated.

Cracking of the surfaces during the hot rolling of steel is influenced by the development of the layer of tertiary scales. In the finishing stands of the hot strip mill, the thickness of this layer is in the order of 10 to 20 μm and is likely quite hard with limited ductility (Ref 45).

A schematic of the cross section of the hot steel is reproduced here as Fig. 13, indicating the potential surface problems that may be encountered as a result. As the scale layer cracks, the islands become separated as the strip is elongated. Hot metal then extrudes between the islands and sticks to the rolls while the sliding islands move farther apart and promote tensions applied to the sticking portion, thereby reducing the load. At the same time, simultaneous smoothing and roughening of the rolled surface occurs, which is visible after pickling and may well cause difficulties in subsequent cold rolling passes (Ref 47). The surface of a carbon steel strip, rolled at 877 °C (1610 °F) shows no surface cracking, as in Fig. 14. When the temperature was 842 °C, severe cracking was observed as shown in Fig. 15, again indicating the chilling and the loss of surface ductility (Ref 48) in addition to the very strong dependence of workability on the temperature.

Reference 1 describes the tensile stresses that may be imposed on the surface and may cause surface cracks; similar conclusions are presented in Ref 35. Surface cracking may also be due to stress corrosion (Ref 16) or by sudden changes of tribological conditions as the rolling pass proceeds (Ref 44). This was also experienced by the author of this article in an experiment while attempting to overcome the minimum friction necessary to cause bite. A tapered leading edge was prepared and kept dry. Neat oil was applied only some distance away from the edge. Successful entry was achieved, but the steel strip tore violently as the lubricated portion was reached. The effect of changing surface conditions on surface cracking was also observed by others (Ref 44, 49, 50). Surface cracking during direct rolling of continuously cast low-alloy steel slabs was studied by means of hot tensile tests (Ref 51). The loss of ductility was explained in terms of dynamic precipitation of carbides and nitrides.

Edge cracking. As stated in Ref 1, the most severe problem in the loss of workability is the cracking of the edges. This cracking is attributed

Fig. 11 Surface cracking of a 5182 Al alloy, rolled at 560 °C (1040 °F) to a strain of 0.58. Source: Ref 43

Fig. 12 Surface and edge cracking of a 7075 Al alloy, rolled at 460 °C (860 °F). Source: Ref 43

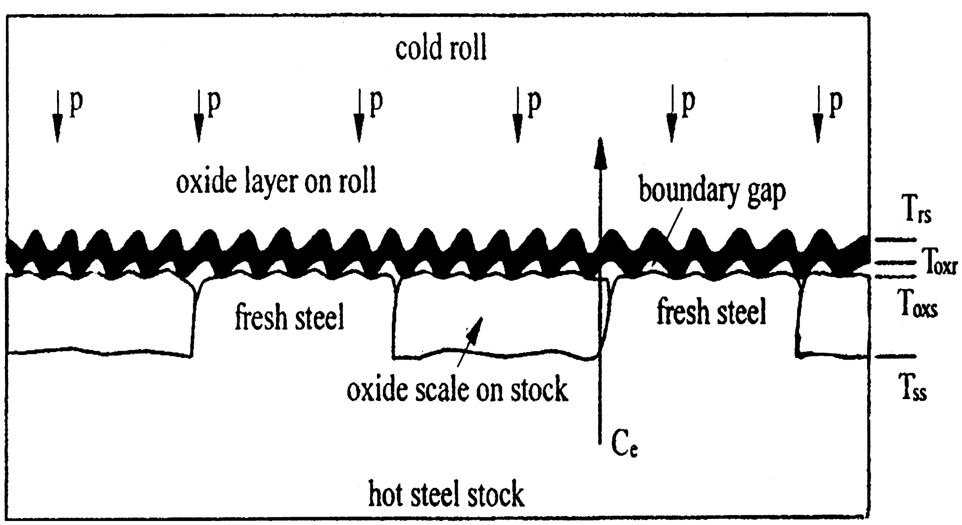


Fig. 13 Schematic of the cross section of a hot steel strip, indicating the cracked layer of scale. T_{rs}, depth of roller scale; T_{oxr}, thickness of oxide on roller; T_{oxs}, thickness of oxide on stock; T_{ss}, depth of scale on stock. Source: Ref 46

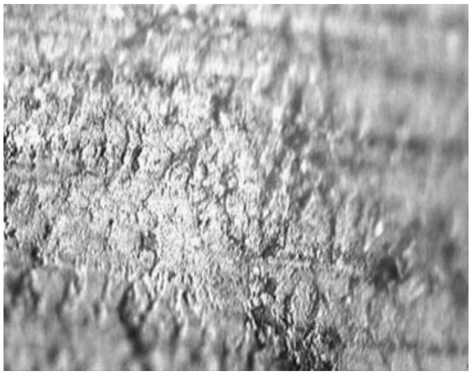


Fig. 14 No surface cracking on a strip of low carbon steel, rolled at 877 °C (1611 °F). Source: Ref 48

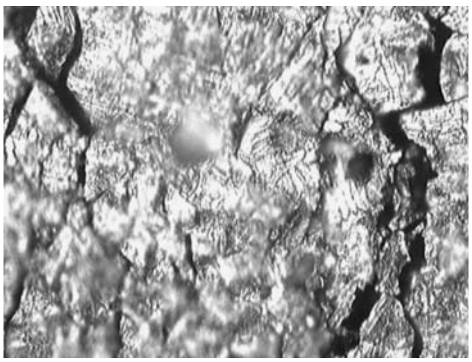


Fig. 15 Cracked surface of a hot steel strip, rolled at 842 °C (1548 °F). Source: Ref 48

Fig. 16 No ductility problems of a 5182 Al alloy, rolled at 520 °C (968 °F). Source: Ref 43

Fig. 17 Edge cracking of the 7075 Al alloy, rolled at 460 °C (860 °F). Source: Ref 43

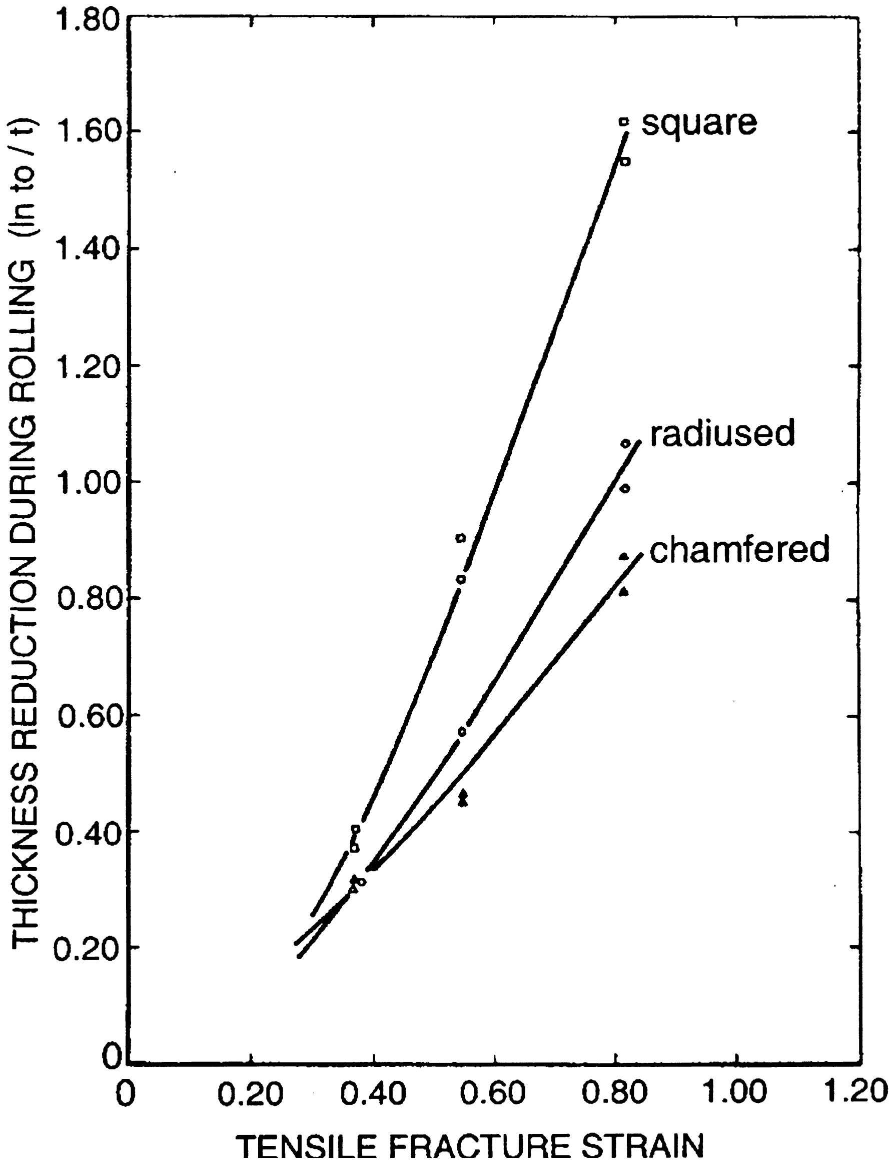


Fig. 18 The relationship between thickness reduction and tensile fracture strain. Source: Ref 56

to the secondary tensile stresses, developed as a result of nonhomogeneity of deformation in slabs of finite widths. The effect of the tensile stresses is enhanced by the temperature losses and the added loss of ductility near the edges, which may well be up to 100 °C (212 °F) cooler than the centers of the rolled metal. Schey (Ref 1) also describes the effects of double barreling at high *h/L* ratios, found when the shape factor is greater than 1.8 (Ref 61–63). The workability of Al alloy 7075-T6 in upsetting and rolling at room temperature was investigated (Ref 54) with reference to the occurrence of free-surface cracks. A workability chart was constructed by combining the fracture criterion, the complete solution for various friction conditions, and several specimen dimensions. With this theory it was possible to predict workability in multipass rolling with reference to edge cracking. The limiting reductions in thickness were given as functions of workpiece dimensions for various friction values. Experiments demonstrated that the predictions were very good. The hot workability of the Al alloy was established by tensile and torsional testing (Ref 55). Three-dimensional diagrams, showing the effects of preliminary deformation and temperature on workability and the effects of the stress state and deformation rate on workability show that the optimum rolling temperature range for this alloy is between 440 and 500 °C (820 and 930 °F). The workability is minimal at 30 to 50% hot reduction.

A slab of the 5182 alloy, rolled at 520 °C (970 °F) to a true strain of 0.53, showed no ductility problems (Fig. 16). The 7075 alloy exhibited some edge cracking when rolled at 460 °C to a strain of 0.6 (Fig. 17). Reference 56 discussed the effect of edge cracking, which progresses to the center of the square-edged rolled strip, made of Al 2024 T351 alloy. The relationship in between the reduction of thickness and the tensile fracture strain is shown in Fig. 18. The figure indicates the dependence of the possible thickness reduction on the manner of edge preparation. The state of stress was found to be most severe when chamfered edge strips were rolled and least severe when square edges were used.

The criterion from Ref 57 was applied successfully (Ref 58) to correlate crack formation and stresses, computed by an FEM of the rolling process. A three-dimensional model with thermal-mechanical coupling, using flow formulation and the Norton-Hoff viscoplastic material description, was developed. Experiments in flat and shape rolling were conducted, using plasticine as the model material. The crack opening criterion function was defined as:

$$C = \int_{\text{pass}} \max[(\sigma n_c) n_c, 0] d\bar{\varepsilon}$$

where σ is the Cauchy stress tensor, n_c is the unit normal, and $\bar{\varepsilon}$ is the equivalent strain.

Edge cracking of low carbon steel sheets, produced in a mini-mill, was examined (Ref 59, 60). It is recognized that while the rolling of thinner gages in mini-mills enhances the industry's competitive advantage, cracking of the edges was found to cause difficulties. The reason for the problem was identified as embrittlement at high temperatures. The authors concluded that the increase in manganese content and smaller amounts of FeS helped in the suppression of edge cracking. Isothermal preheat treatment also helped in reducing the cracking of the edges. Reference 61 considers defects due to porosity or shrinkage, which are present in all steels, following solidification. Process conditions under which the internal stresses are able to eliminate porosity are defined. It was concluded that when the reduction was in excess of 30% all pores closed.

Centerburst defects are more commonly found in the extrusion process but can occur in slab rolling when the slab product thickness is large, such as for some ship plates or pressure vessel plates. There are few public reports on the subject because it has not been treated by academic institutions, and so it is not surprising that the search of the literature yielded no recent publications concerning the occurrence of centerburst during rolling of metals. As noted in Ref 1, internal defects due to secondary tensile stresses generated in the center of the slab may occur at early stages of hot rolling when *h/L* is very large, but the temperature is still high enough to ensure reasonable ductility in all but the most brittle materials. Defects that form also have a chance of being healed by forge welding. Most practical hot rolling is performed in the finishing stands, and as shown in Fig. 10, this causes little inhomogeneity. It is conceivable, however, that this danger exists more generally during the rolling of superalloys, refractory-metal alloys and tool steels, particularly if restricted mill power limits the attainable reductions and thus keeps *h/L* high. Centerburst of a steel beam during rolling was demonstrated (Ref 62).

Alligatoring. This defect is also related to nonhomogeneity in the rolling process, in the through-thickness direction. It affects metals of lower ductility but not exclusively. The shape factor at which alligatoring is observed varies broadly. Reference 1 quotes *h/L* values of 1.3 (Ref 53). While rolling aluminum (Al-8 Mg) billets, alligatoring appeared at shape factors of 0.5 to 0.7 (Ref 63). In the study quoted in Ref 43, the 7075 alloy showed alligatoring when rolled at 510 °C (950 °F) (Fig. 19). An upper bound approach (Ref 64), which successfully predicted the occurrence of central burst and alligatoring during cold rolling of copper and aluminum strips in terms of the shape factor, was developed. It was concluded that as long as the shape factor, defined as the ratio *h/L*, is less than 0.75, neither defect will occur, somewhat contradictory to the conclusions reached in Ref 43. The results were presented in plots of the relative reduction versus the relative thickness and the safe and nonsafe zones identified. The results are shown in Fig. 20, obtained for a perfectly plastic material. The figure gives the safe zone, the central burst zone, and the zone where both defects are expected in terms of the relative reduction and the relative thickness. These defects tend to be promoted by small roll radii and large initial workpiece thickness.

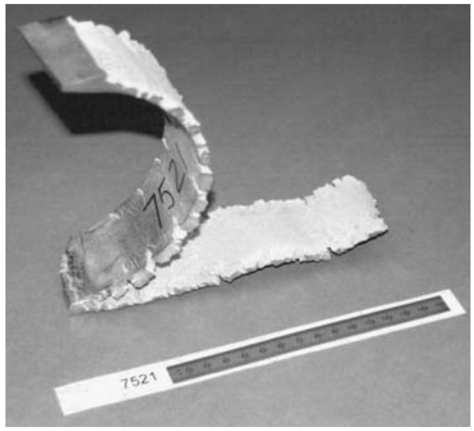


Fig. 19 Alligatoring of the 7075 Al alloy, rolled at 510 °C (950 °F). Source: Ref 43

Limits of Workability While Rolling Tapered-Edge Samples. The study of workability limits of aluminum samples also considered the effect of tapered edges (Ref 43), which created complex stress states, and as a result, a combination of edge and surface cracking in addition to alligatoring was observed. Surface cracking and alligatoring were observed when the 7075 alloy was rolled at 460 °C (860 °F); the edges were tapered at 15° from the vertical (Fig. 21). Recall that the same alloy, having rectangular edges and rolled at 460 °C (860 °F), suffered minor cracking of the edges (Fig. 17). Similar behavior is observed when the 5182

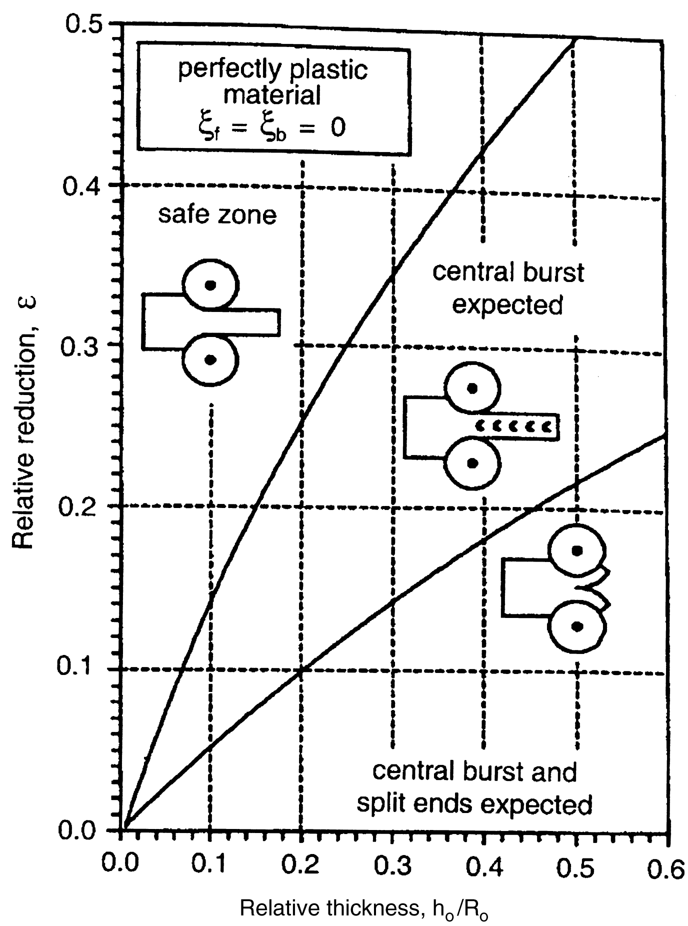


Fig. 20 The safe, central burst, and alligatoring zones during bar rolling. Ref 64

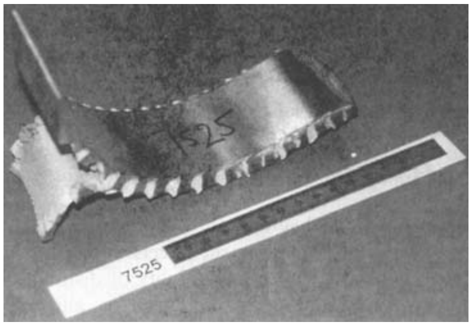


Fig. 21 Alligatoring of the 7075 Al alloy, rolled at 460 °C (860 °F). Source: Ref 43

alloy, with edges making 45° with the vertical, was rolled at 558 °C (1036 °F) (Fig. 22). The tapered edges produced a significant reduction of the workability, reinforcing the conclusions reached in Ref 56. Chamfered, sheared, or fully rounded edges also affect workability in cold rolling. Data obtained in a rolling mill (Ref 65) with rolls of 250 mm (10 in.) diameter and rolling strips of 2.5 by 63 mm (0.10 by 2 in.), reported in Ref 1, are repeated in the table below:

	Reduction at fracture (%)			
Edge	1.3% C steel	1.5% W steel	Al-5Mg	Al-7Mg
Square, ground	58–64	53	90	60
Fully rounded	15–30	10	80	48
30° chamfer	14–17	. . .	. . .	. . .
Sheared	31–52	. . .	. . .	. . .
As-received	26–31	. . .	. . .	. . .

Roll Surface Roughness. The effect of the direction of the surface roughness of the work roll on the attainable reduction was examined (Ref 66). The researchers cold rolled Al-Mg-Cu strips, using rolls with surfaces prepared by the traditional grinding technique and compared the results with those obtained when the surface of the roll was prepared by electrical discharge machining (EDM), which produced a random roughness direction. Using the ground roll, the roll force increased exponentially with increasing reductions, and the strip tore at a reduction of about 45%. Rolling with the EDM surface allowed reductions of 78% with no tearing. The beneficial and efficient distribution of the lubricant by the randomly placed open and closed pores was credited with the increased rollability.

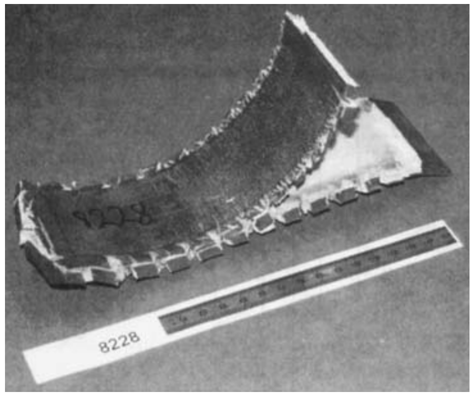


Fig. 22 Alligatoring of the 5182 Al alloy, rolled at 558 °C (1036 °F). Source: Ref 43

The state of stress can be derived from simple mechanical considerations at different locations of the sample passing through the roll gap (Fig. 23). The stress in the longitudinal direction is mostly tensile. The stress in the transverse direction is generally compressive, but in the tapered edge or in a bulged side (location 5), it may vanish. The shear stresses are most severe near the surface at the edge (location 2), where they act in two directions. The elements located at the vertical centerline do not experience shearing (location 3), except at the roll/sample interface (location 1). The stresses in the direction of rolling and in the transverse direction are practically identical near the longitudinal centerline, whereas near the edge, the magnitude of the transverse stress decreases. The shear stresses are approximately half of the normal stresses differences.

According to Tresca's criterion for yielding, deformation will occur when the maximum shear stress reaches the yield strength of the metal in pure shear. Further, the most probable surface of fracture may be located by searching for the plane on which shear stress of the greatest magnitude is acting. Thus, the description of maximum shear stresses, as determined by 3D Mohr's circle, gives the planes on which fracture is likely to propagate (but not necessarily the initiation of fracture). In alligatoring, for example, the most likely cause is excessive shearing stresses at the roll/metal interface, creating a state of stress such as the one shown in Fig. 24(a), near the exiting portion of the rolled strip (location 2). Separation of the rolled aluminum begins on the plane of absolute maximum shear stress, defined in this figure. As to the edge, the stress in the transverse direction reduces from a fairly large magnitude to zero. There are two planes within the element at location 5 in Fig. 23 on which the shearing stresses are at their maximum level, as in Fig. 24(b). It is the resistance of the material to deformation at that location that will determine on which of these planes fracture may occur.

This simple approach neglects many parameters, such as material anisotropy, material resistance to deformation, friction effects, temperature distribution, or thermal history, all of which have potentially important effects. For instance, even though the temperature distribution within the furnace is uniform, the heating rate is not, with the surfaces reaching the furnace temperature sooner than the central portion. The portions of the metal near the surface are kept therefore at the furnace temperature for longer time periods. If the materials attributes are sensitive to the rate of heating, as should be the case for partly homogenized materials for instance, the balance between the edge cracking and alligatoring will be affected. Edges, especially those tapered to 15 and 30°, are also at considerably lower temperatures than those at the center of the workpieces.

Adequate modeling of the failure mechanisms observed during rolling tests thus requires a quantitative estimate of the thermomechanical history of the different zones of the samples, a precise estimate of the deformation path, taking into account materials properties and friction, and the use of fracture criteria (for crack initiation) and stress analysis to identify the probable planes of crack propagation. Stress analysis may be very simple, such as Tresca's criterion, or more complex. In addition to Tresca's criterion, many existing criteria of fracture also show that the hydrostatic stress state has a dominant effect on the probability of fracture along planes of

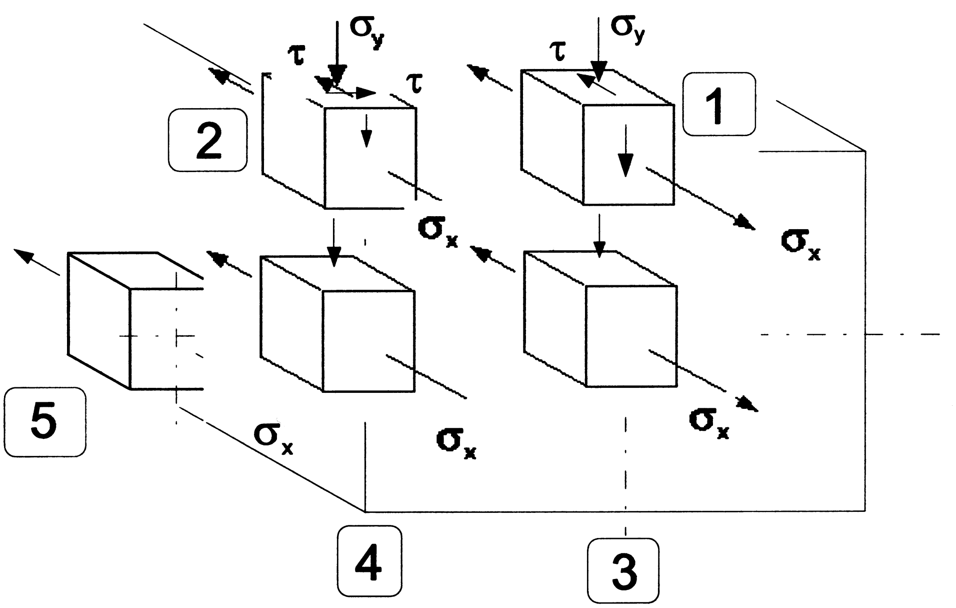


Fig. 23 The state of stress in a rolled sample. Source: Ref 43

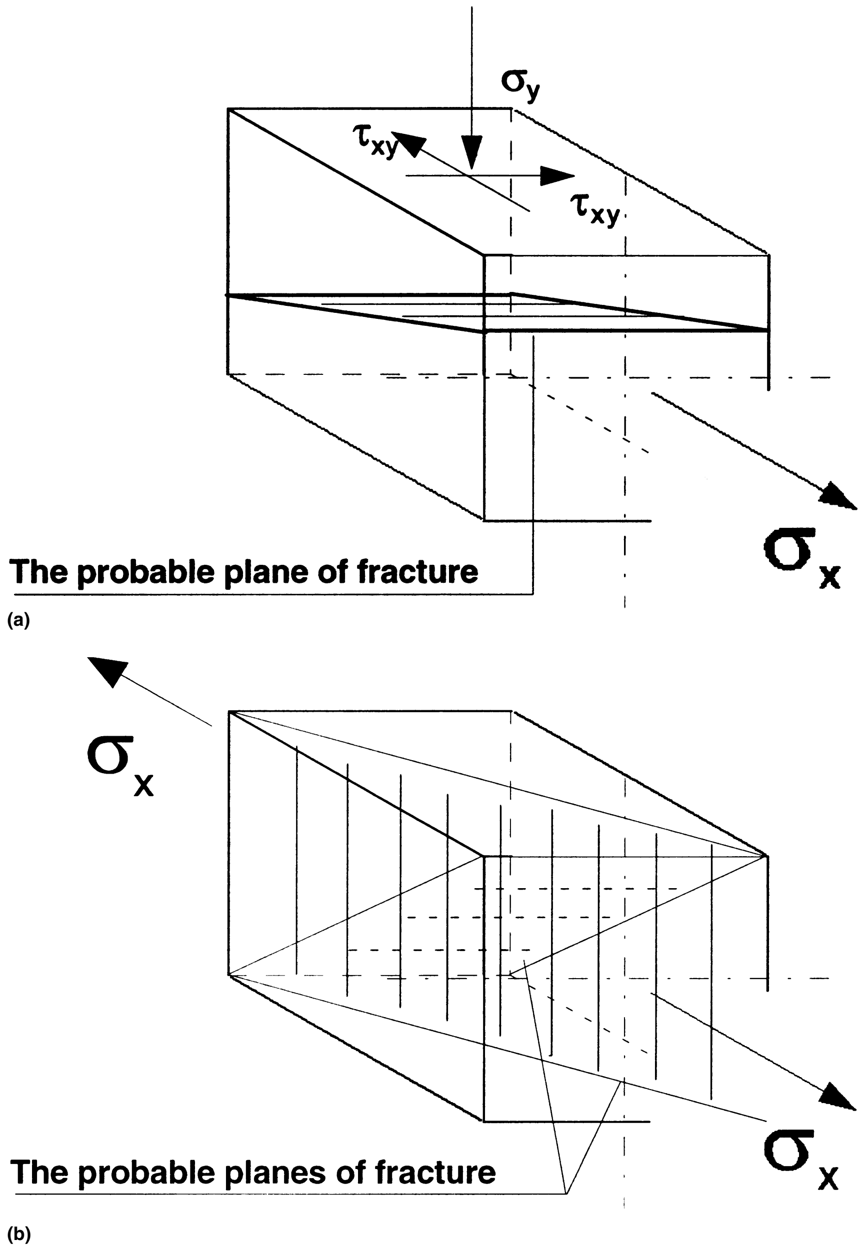


Fig. 24 Potential surfaces of fracture. (a) Near the surface and strip edge shown as location 2 of Fig. 23. (b) Interior shown as location 5 of Fig. 23. Source: Ref 43

Fig. 25 The split ends of a hot rolled bar. Source: Ref 64

maximum shear. This problem can be addressed properly only with the assistance of two-dimensional or, most preferably, three-dimensional, FEM. Additional information on fracture criteria is given elsewhere in this book (e.g, see Fig. 25–29 in Chapter 12, "Workability Theory and Application in Bulk Forming Processes," for edge cracking in bar rolling).

Workability in Shape Rolling. By necessity, rolling of sections and shapes generates large secondary tensile stresses because a more complex, nonrectangular shape is gradually developed from a simple starting cross section. Various portions of the cross section are subjected to different reductions, and the differential elongation leads to cracking in the less-deformed zones, whether they are on the surface or in the center. The purpose of roll-pass design is the equalization of these strains.

Uneven cooling of thinner portions can be especially troublesome in hot rolling, and edge cracking due to cold corners is avoided by designing roll passes with more rounded corners. In general, it is not meaningful to speak of a particular *h/L* ratio in shape rolling because both the contact surface and the deformed zone are very complex in shape.

Workability depends on the chemical composition, structural state, and temperature of the material, in addition to the state of stress in the deformation zone (Ref 67). The rolling of a rectangular cross-sectioned bar by a two-arc oval pass and by a circular pass were considered. A stress state index was defined as:

$$n_\sigma = \frac{\sigma_m}{\sigma_i}$$

where σ_m is the mean stress and σ_i is the flow stress. The index is related to the limit strain:

$$\varepsilon_{iL} = K_1 \exp(-K_2 n_\sigma)$$

where K_1 and K_2 are to be determined and two tests—a tension and a torsion test—are needed. The actual and predicted values of the limit strain intensity were found to be close. In a discussion of the effect of shape factor on internal defects (Ref 64), the split ends of a steel bar, which initiates a crack, are shown (Fig. 25).

Workability Limited by Metallurgical Phenomena

Charles writes in a keynote address about processability (Ref 68) that "the behavior of the material during hot processing by mechanical work is governed by its intrinsic structural strength and ductility as a function of the microstructure and by the presence of gross imperfections. . . ." He classifies the problems encountered as surface defects, porosity, segregation, and inclusions. Reference 69 discusses the metallurgical causes of ductile fracture. It describes cavity-nucleated fractures, shear bands, cavitation, and grain anisotropy. Reference 70 presents five phenomena and the events that control them, all of which affect formability:

- Ductility, controlled by the state of stress, temperature, inclusions, segregation of residuals
- Anisotropy, controlled by texture, thermal-mechanical history, elongated inclusions, second phase particles, stacking fault energy
- Strength level, controlled by chemical composition, grain size, precipitates, solutes

- Work hardening, controlled by temperature, solutes, strain gradients, degree of deformation
- Surface properties, controlled by oxide films, lubricants, surface plastic deformation, hard particles

The hot workability of a large number of materials are compared and metallurgical mechanisms that affect the limits are identified in Ref 71. As the researchers write, plastic deformation occurs by dislocation glide. Reduction in the resistance of the metals to deformation is observed when slip and climb mechanisms cause dislocation annihilation. The restoration processes that result—recovery and recrystallization, dynamic, static or metadynamic—affect ductility in a significant manner. High levels of dynamic recovery prevent triple-point cracking. At lower levels grain-boundary fissuring may be noted. Poor ductility may also result in the prevention of dynamic recrystallization or its retardation by low temperatures. When particles pin grain boundary movement, the metal may experience intergranular fracture. The ductility troughs demonstrated by C-Mn, C-Mn-Al, C-Mn-Nb-Al, C-Mn-V-Al, and C-Mn-Ti-Al steels, all of which fall in the range of temperatures used in the finishing stands of hot strip mills, are shown in Fig. 26. The effect of minor elements on hot workability are discussed in Ref 73. It was concluded that in Nb-added two-phase stainless steels reduction of the phosphorus content and the additions of yttrium and/or cerium improve ductility. According to Ref 74, while the ductility of ferritic stainless steels is little dependent on the temperature, that of austenitic grades 304 and 316 are strongly temperature dependent. Boron and titanium were found to increase the ductility of the 304 and 316 grades, respectively, while carbon, nitrogen, chromium, nickel, molybdenum, sulfur, manganese, and silicon lowered it.

The workability of cast pure chromium produced by an induction-slag-melting process was investigated through the study of rolling and isothermal upsetting (Ref 75). The effect of rolling, which was performed to change the coarse cast structure to a finer one, on the ductile-to-brittle transition temperature (DBTT) was examined. The results are summarized as follows. The DBTT of the as-cast specimen tested in tension at a strain rate of 2.4×10^{-2} s^{-1} was approximately 500 K. A minimum ductility and a maximum flow stress appeared at approximately 973 K, owing to dynamic strain aging. The forming limit in upsetting, the percent reduction of height at the beginning of cracking on the specimen wall, increased to 80% at 773 K, and no crack appeared at temperatures over 873 K. At all temperatures, there occurred cracks in the as-cast specimen deformed by conventional rolling. Cracks occurred slightly in the specimen even in sandwich rolling by which a sintered chromium was successfully deformed. Cracking in the as-cast specimen was mainly due to the coarse structure with large crystal grains. Once the structure was changed to a finer one by upsetting, the specimen was easily rolled to a thin sheet without crack occurrence. The DBTT of the specimen rolled to 80% and, followed by annealing at 1273 K, was higher by approximately 200 K than that of the as-cast specimen, while annealing at 1473 K yielded a slightly lower value than that of the as-cast material.

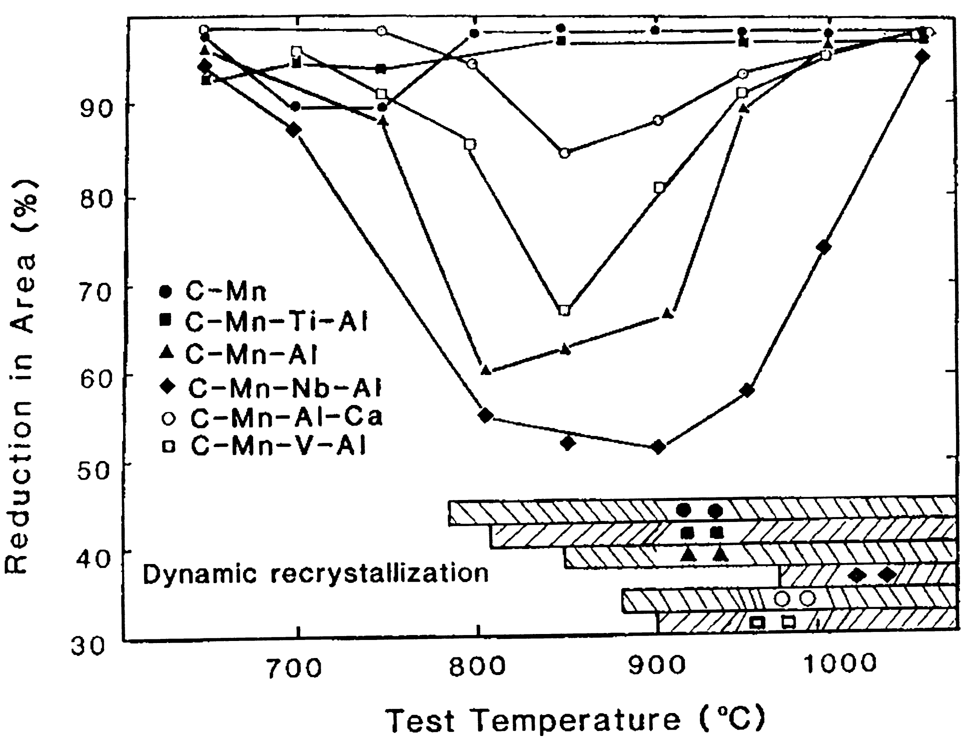


Fig. 26 The ductility trough for several alloy steels. Source: Ref 72

A typical industrial hot-rolling process is simulated by means of an interrupted torsion test, using 5083 aluminum alloy as the test material (Ref 76). In addition, the effect of soaking time on the hot workability of the material is explored, where it is shown that the resistance to deformation of the alloy when soaked for 24 hours is as much as 20% less than that when soaked for 4 hours.

Improvement of hot workability is shown to depend upon the reduction of the denuded zone around the dendritic cast structure and on the refining of the precipitated second phases that occur with increasing deformation. The flow and fracture of MP35N (35Co, 35Ni, 20Cr, 10Mo) have been studied by uniaxial compression and plane strain bending in the temperature range 1000 to 1200 °C (1800 to 2200 °F) and strain rate range 0.01 to 10 s^{-1}, which cover the normal bar rolling production conditions (Ref. 77). The strain to fracture in plane strain bending was found to increase with increasing strain rate, roughly coinciding with the increase of the strain to the peak stress in the flow curves. Within most of the temperature and strain rate ranges investigated, and under plane strain bending deformation conditions, microvoid nucleation was found to be concurrent with or greatly enhanced by the onset of dynamic recrystallization. A small addition of boron improves hot workability of austenitic stainless steels (Ref 78), and it is expected that boron should have the same effect on ultra-high-strength metastable austenitic stainless steels. Examination of hot workability and ductility at room temperature using hot rolling, hot tensile, notched tensile, and Erichsen tests showed that the mechanical properties are related to boron content. Hot workability is improved with increasing boron addition. However, excessive boron content has a deleterious effect, causing reduced room temperature ductility in steels aged at various temperatures after cold rolling. This significant loss of ductility is attributed mainly to borides that precipitate after annealing.

Mathematical Models for Workability Prediction

The physical phenomena of the loss of workability may be translated into mathematical terms. This section reviews some of these mathematical models. While some of these models are applied directly to rolling, some address the general problem of bulk forming. A general review of prediction of damage in materials processing is given in Ref 79.

The model of Montmitonnet has been referred to (Ref 58) in the context of correlating crack formation and the stress field. As mentioned, a finite element program that uses a viscoplastic material model and a steady state approach with free surface updating by minimizing the material flux through the surface was developed. Hot workability is presented from the systems point of view (Ref 80). Based on a polar reciprocity relationship between the stress and strain rate spaces, a measure of hot workability, known as *hardening power index*, is defined. Deformation processing maps of this parameter are plotted for a titanium alloy IMI685 (Ti-6Al-5Zr-0.2Mo-0.1Si) using flow data from high-temperature axial compression tests in the temperature range 800 to 1000 °C (1470 to 1830 °F) and strain rate range 10^{-3} to 10^2 s^{-1}. Maps plotted at different strain values reveal the optimum processing conditions and indicate the probable occurrence of instabilities.

Reference 81 describes a new method of preform design in multistage metal forming processes considering workability, limited by ductile fracture. The FE simulation combined with the ductile fracture criterion has been performed in order to predict ductile fracture. The artificial neural network using the Taguchi method has been implemented for minimizing objective functions relevant to the forming process. The combinations of design parameters used in FE simulation are selected by an orthogonal array in the statistical design of experiments. The orthogonal array and the result of simulation are used as training data for artificial neural networks. The cold heading process is taken as an example of designing preforms, which do not form any fracture in the finished product. The results of analysis to validate the proposed design method are presented. The use of ductile fracture criteria in conjunction with the FEM for predicting failures in cold bulk metal forming is described in Ref 82.

Four previously published ductile fracture criteria are selected, and their relative accuracy for predicting and quantifying fracture initiation sites is investigated. Experiments with ring, cylindrical, tapered, and flanged upset samples are performed to investigate the validity of the workability criteria under conditions of stress and strain similar to those usually found in bulk metal forming processes. The implementation of ductile fracture criteria into a rigid-plastic FE computer program is also presented. Local stress and strain distributions throughout the deformation are computed and compared with experimental measurements. In general, a good agreement is found. However, only two of these workability criteria have successfully predicted the location at which fracture initiates for all the upset tests performed in this work. The paper concludes with a discussion of the importance of the critical damage at fracture which is to remain independent of the technological processes.

In the model described in Ref 83, irreversible thermodynamics and the orthogonal flow rule are used in formulating a microdamage evolution equation, taking advantage of the specific free energy and plastic potential with internal variables. Furthermore, by analogy with the microdamage evolution equation, a macrodamage evolution equation is derived according to the principle of minimum strength. Afterward, predictions are made for proportional loading and nonproportional loading, and the experimental data available in the relevant literature are used to verify these predictions. It can be concluded that the new damage model can successfully reveal the ductile damage evolution during the plastic deformation process. Thus, it opens the way to predict product quality and materials workability for metal-forming processes. In Ref 84, plasticity theory is modified for compressible materials and the upper bound theorem or the second extremum principle, which incorporates a normal velocity discontinuity, considered to be a measure of fracture. The theorem is applied to predict the occurrence of fracture and central bursting in extrusion or drawing. The occurrence of fracture appears to depend on the value of the parameters in the theorem; these parameters may be related to the ductility or the workability of the material in question. The upper bound technique is then applied to multistage extrusion of a carbon steel for which the values of the parameters are given with some assumptions. It is thus shown that the stage at which central bursting occurs agrees with experimental results and that it apparently depends on the extrusion condition. Reference 85 presents a ductile-damage model based on the precise measurement of relative density changes. The rationality of the proposed model is verified through observation of the change in the volume fraction of voids during the tensile test, by the use of scanning electron microscopy. An example to demonstrate the application of the model is given. Combining the proposed model with a proper method of the analysis of metal-forming processes, it is possible to predict the evolution of ductile damage in the interior of the workpiece and thus the workability of the material.

Reference 86 examines the behavior of a longitudinal V-shaped crack on the surface of a continuously cast steel slab during hot rolling. The analysis is carried out by means of the commercial FE-code LS-DYNA3D. Process parameters obtained from industry are used as a reference. The slab of initial width of 1000 mm (39 in.) and 220 mm (9 in.) thickness is rolled down to 30 mm (1.2 in.). It is assumed that the material can be treated as rigid-perfectly plastic and that the cracks do not propagate. The latter assumption is in agreement with industrial observations for a steel grade similar to that analyzed here. The aim of the study is to investigate the possibility of controlling the plastic deformation so that the cracks disappear or so that their deteriorating effects are minimized. The analysis is focused upon the influence of friction, roll radius, and rolling schedule on the change in the shape of a crack of initial depth 20 mm (0.8 in.) and a crack angle of 6°. The reliability of the simulations is checked by pilot-plant experiments using aluminum as the model material for steel. The results indicate that it is not possible to prevent the bottom side surfaces of the crack from coming into contact, especially not for small reduction-pass and small roll radii. The influence of friction was found to be marginal.

In another study (Ref 87), researchers incorporated ductile damage mechanisms in their elastic-plastic finite strain model, based on irreversible thermodynamics. Their model accounts for the degradation of elastic properties, plastic dilatation caused by the growth of voids and microcracks and strain softening. They successfully predicted the occurrence of defects in the process of extrusion. They emphasize the importance of stress and strain history in addition to materials parameters. Reference 88 discusses and compares damage models and underlines two aspects of the fracture mechanisms in metal forming processes. When progressive ductile damage development is the mechanism, material models that include an account of voids or cracks are needed. When fracture is occasioned by plastic instabilities, increasing void volume fraction is not an appropriate measure. The author uses a perturbation technique for the detection of instabilities. The authors of Ref 89 present an optimization technique for process design with damage minimization as the main objective. Others (Ref 90) used the Oyane criterion (Ref 91) and a three-dimensional FEM to predict the occurrence of surface defects during hot rolling of heavy ingots. The Oyane criterion, based on the critical void volume, is defined in terms of the Oyane parameter, B, given by:

$$\int_0^{\bar{\varepsilon}} \left(1 + A\frac{\sigma_m}{\sigma_{eq}}\right) d\,\bar{\varepsilon} = B$$

where A and B are materials characteristics, depending on thermal conditions; and σ_m and σ_{eq} are the mean stress and the equivalent stress, respectively. The authors state that while the Cockroft and Latham criterion could not fit their data, the Oyane criterion did.

Process Design

In Ref 92, a simplified description is given of the "Taguchi methods" (Ref 93) as applied to quality engineering, a topic of which process design is an essential part. Off-line quality control consists of two stages: product design and process design (Ref 93). Process design is, in fact, the essence of manufacturing engineering (Ref 92). The process specifies the equipment and sets work standards as well as workable specifications.

When the rolling process, or in industry parlance, the draft schedule, is designed, several questions need to be answered. These questions concern the three components of the rolling system: the mill, the workpiece, and their interface. As far as the rolling mill is concerned, these questions may include:

- Are the mill stands of appropriate stiffness, providing satisfactory mill stretch?
- Are the drive spindles capable of transmitting the necessary torque?
- Are the driving motors of sufficient power?
- Is the roll crown satisfactory?
- Is the speed of the mill sufficient?
- Are the roll cooling facilities appropriate?
- Is the surface roughness of the work roll adequate?

- Is lubricant/emulsion delivery suitable?
- Are the furnaces adequate?

To include the attributes of the metal to be rolled, the designer may need to know:

- What is the ductility of the metal, especially the ductility vs. temperature curve?
- What is the constitutive relation of the metal?
- What are the metallurgical attributes—i.e., the hardening and recovery processes—of the metal?

When the interface is considered, the limits concern three events:

- The ability of the two metal surfaces—that of the roll and of the rolled metal—to successfully entrain the lubricant or the emulsion
- After entrainment, how well the lubricant is distributed in the contact zone
- The boundary conditions as a result of the contact of the rolls and the metal; note that these usually refer to the coefficients of friction and heat transfer at the contact zone

When the process of the rolling of shapes is designed, all of these considerations are of some importance. Of prime significance, however, is the process of roll pass design, which, in spite of a large number of technical publications, is most often an experience-driven, trial-and-error procedure.

Example

As the preceding discussion indicates, these phenomena are interdependent, making process design a difficult and crucial undertaking. A well-designed draft schedule for a hot strip mill, a schematic of which is shown in Fig. 27, includes considerations of the following.

The Temperature of the Reheat Furnace. While the objective of reheating the slab prior to rough rolling is to allow all alloying elements to enter into solid solution of the matrix, the usual practice is to use a preset temperature, arrived at most often by a trial-and-error procedure. Some of the alloying elements dissolve at relatively low temperatures, while others (for example, titanium and vanadium) require heating in excess of 1300 °C (2370 °F). The usual temperature of the reheat furnace is between 1200 and 1250 °C (2190 and 2280 °F), and the fact that not all elements are dissolved is simply accepted.

Time in the reheat furnace is another variable of the process, which has been established by long experience. The usual time slabs spend in the reheat furnace is in the order of 2 to 3 hours, which is believed to be sufficient for them to reach a uniform temperature distribution.

Descaling and Rough Rolling Practice. In the roughing mill the slab's thickness is reduced from approximately 200 to 300 mm (8 to 12 in.) to about 25 to 30 mm (0.98 to 1.2 in.) in several steps. Descaling before and after the rough rolling passes, which is necessary to avoid poor surfaces, is accomplished by high-pressure water jets. These practices also have been used in steel companies for some time, and they are not easily changed. The flying shears remove the distorted edges from the rolled metal.

Time on the Transfer Table or in the Coil Box. In older mills the bar waits to enter the finishing train on the transfer table. In more recently built mills, the coil box, placed between the rougher and the finishing train and in which the steel is coiled up on exit from the rougher and uncoiled in order to enter the finishing stands, allows the development of some temperature uniformity and leads to improved homogeneity of the finished product. The temperature of the entry of the transfer bar to the first stand of the finishing train is one of the most important parameters because it controls the metallurgical attributes of the finished product. There are some compromises to be made in deciding the entry temperature. Values that are too high reduce the loads on the mill but may create a strip that is too soft. Temperatures that are too low would produce higher strength and hardness but may overload the drive train and increase roll wear.

The Draft Schedule on the Finishing Train. This part of the hot strip mill contributes most to the attributes of the product. The final geometry of the strip, its surface quality, and its metallurgy and the attendant mechanical properties are determined by the reductions and their sequence on the five, six or seven stands of the finishing train. While the speed of the passes has a major impact on the development of the properties of the product, engineers may freely decide only the speed of the first stand. The speeds of the subsequent stands must follow the laws of mass conservation, and if these are disobeyed, tearing or cobble may occur.

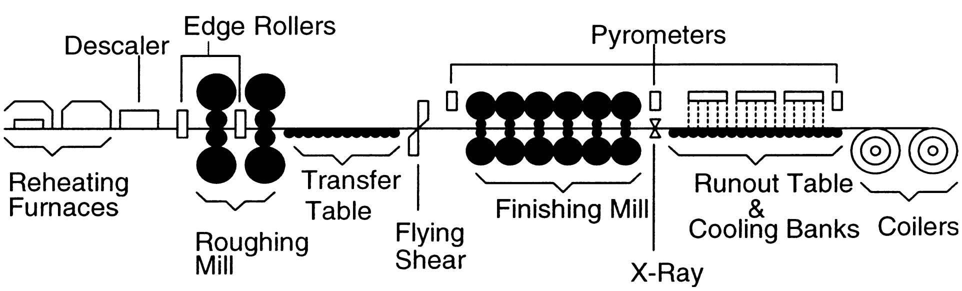


Fig. 27 Schematic of a hot strip mill

Cooling Tables. After exiting the last stand of the finishing train, the strip is cooled at a predetermined rate to a final temperature before it is coiled for shipment. The cooling process has a very significant effect on the final product properties and there are a large number of designs in use in the steel mills.

Coilers. Here the strips are coiled, allowed to cool, and are prepared for shipment.

Thermomechanical Processing

Of the items in the preceding section, one important topic that has received the most attention of researchers is the development of the draft schedule on the finishing stands of a hot strip mill. Most of the advances in process design originate in the realization that the attributes of the hot-rolled product are influenced by both the mechanical and the thermal treatment it receives in the finishing train of the hot strip mill. This is referred to as *thermomechanical processing,* or in the case of rolling, *controlled rolling* (see Chapter 17, "Thermomechanical Processing by Controlled Rolling").

The design of thermal-mechanical treatment on the finishing train of hot strip mills has been considered in detail (Ref 94). A schematic of several possibilities is shown in Fig. 28. The figure shows four different designs for the draft schedules. The first one indicates normal rolling of carbon steels, showing that both the roughing and the finishing processes are performed at high temperatures, well above that required for recrystallization. In the second example, still for carbon steels, finish rolling is performed at significantly lower temperatures, but they are still within the dynamic recrystallization range. The third and the fourth schedules are for microalloyed steels, containing various amounts of niobium. In both cases, finish rolling is completed at temperatures below that required for dynamic recrystallization, taking advantage of the resulting elongated γ grains.

Thermal-mechanical working, in which the effects of hot working on the kinetics of recrystallization are used to produce a fine, uniform microstructure, is industry's favored technique toward creating the required attributes in the rolled metal. The alloys that respond best are among the family of microalloyed steels, most often containing strong carbide formers such as niobium, vanadium, molybdenum, titanium, boron or aluminum, singly or in combinations. Of these, niobium is used most, and its principal advantage is its contribution to strengthening through precipitation (Ref 95). The other major benefits are grain refinement and the retardation of recovery and recrystallization of austenite. Vanadium is often added as a second alloying element with the objective of providing increased strength. In addition, the extent of retardation of the recovery processes, with multiple additions of alloys, is expected to be greater than the sum of the individual components. For automotive applications, yield strengths of up to 700 MPa

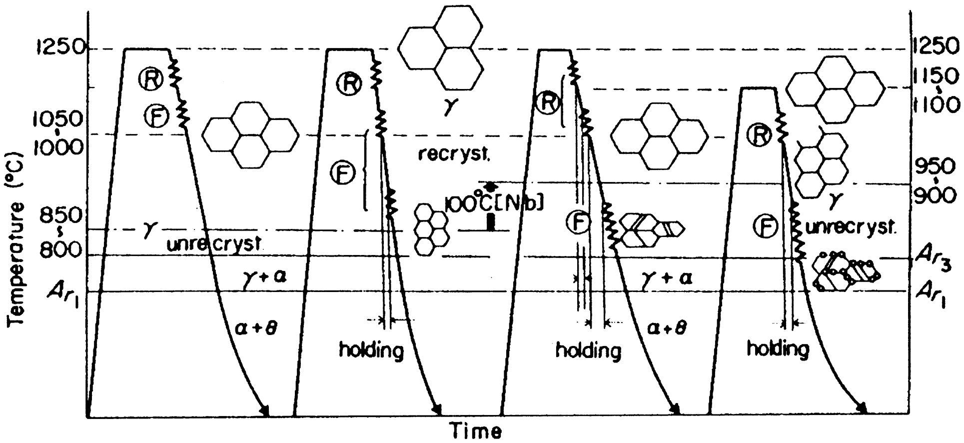


Fig. 28 Thermal-mechanical treatment. Source: Ref 94

(120 Ksi) may be required, establishing the need for bainitic microstructures, obtained through the control of transformation kinetics (Ref 95). The technical literature contains a large number of publications that detail the effects of these elements on the hot, as well as the room temperature, strength of the steels. A detailed examination of the chemical compositions of the metals tested, however, reveals that most of the work concerned steels with niobium contents under 0.06%.

Other researchers (Ref 96) worked with several niobium steels, containing up to 1.03%, by mass, of the alloy. They listed the important features of thermal-mechanical processing, which are the control of the initial austenite grain size and the volume of carbides and nitrides in the solution; the mechanical treatment; the composition; and the interpass time. Increasing the amount of niobium refined the austenite grains at all temperatures up to 1200 °C (2190 °F). Beyond the stoichiometric Nb:C ratio, slight coarsening of the grains was noted. Solute drag appeared to be the mechanism of retardation of the restoration processes at higher temperatures, while at lower temperatures, strain-induced precipitation of carbonitrides caused the delay. At the lower temperatures, the strengthening was caused by the precipitation of fine, planar Nb(CN), coherent with the ferrite matrix, the carbides having been in solution in the austenite grains during the reheating process. Thermomechanical treatment was by hot rolling at temperatures of 900 to 950 °C (1650 to 1740 °F) and 1200 to 1250 °C (2190 to 2280 °F), 20 to 50% reduction per pass, followed by quenching in ice brine. Rolling the 0.16% Nb steel at 950 °C (1740 °F) to 50% reduction caused only partial recrystallization.

Using a 0.073% Nb steel, researchers (Ref 97) noted sharply increased retardation of dynamic recrystallization at lower rates of strain, caused by precipitation. At a strain rate of 13 s^{-1}, the addition of 0.14% V caused further retardation of the beginning of dynamic recrystallization. In an attempt to optimize plate rolling schedules, steels containing a large variety of microalloys, including niobium, of up to 0.118% by weight were considered (Ref 98). It was found that in multistage plane strain compression, carried out at temperature ranges that varied from one that would produce complete recrystallization to one that would cause pancaking only, most of the reduction of the grains would occur in the first stage. In a later study (Ref 99) steels with the niobium content up to 0.112% by weight were considered. Niobium contents up to 0.16% were examined (Ref 100). The alloys were hot rolled in several passes at various temperatures, and it was observed that the roll separating force remained independent of the niobium content in the first three passes. Beyond a total reduction of 70% however, the loads increased with the amount of niobium, in an apparent contradiction to the conclusions reached in Ref 96. Another researcher (Ref 101) also indicated that there are beneficial effects to be gained from increased niobium content beyond 0.08% by weight. Reference 102 discusses the role of microalloying elements in the modification of the softening behavior of the steels, mentioning the three possibilities: solute effects, strain-induced precipitation, and a combination of both. Studying single alloy additions, it was concluded that the most likely effect is the combination of the two mechanisms. One of the steels tested contained 0.035% Nb. When deformed at 900 °C (1652 °F), strain-induced precipitation began in approximately 10 s. The effect of high-temperature deformation on the size distribution of Nb(CN) was examined (Ref 103). In addition, the researchers studied the retardation of dynamic recrystallization due to precipitation. The interaction of manganese with niobium and vanadium has been studied (Ref 104, 105) with attention focused on dynamic precipitation. Increasing the manganese content from 1.25% by weight to 1.9% decreased the rate of precipitation of niobium carbonitrides. While dynamic precipitation started earliest in the single-alloy niobium steel, the addition of other microalloying elements retarded the kinetics of precipitation.

New steels are being developed constantly and information about their response to thermal-mechanical treatment is always required by the producers. A very informative approach is to subject samples of the steel to compression under closely controlled conditions in which the rates of strain are kept constant during the process. Study of the resulting true stress-strain curves leads to conclusions that may, in turn, lead to the efficient scheduling of strip mill operations.

The Web site of the American Iron and Steel Industry AISI includes up-to-date information concerning the description and the processing of recently developed steel alloys. As given in the site, advanced high-strength steels (AHSS) in use in the automotive industry include the dual phase steels (DP), the microstructure of which includes ferrite and up to 20 and 70% volume fraction of martensite. While the use of bainite helps to enhance the capability to resist stretching on a blanked edge, the ferrite phase leads to high ductility and creates high work hardening rates that give the DP steels higher tensile strength than conventional steels. Further, transformation induced plasticity (TRIP) steels are also used, the microstructure of which consist of a ferrite matrix containing a dispersion of hard second phases—martensite and/or bainite in addition to retained austenite in volume fractions greater than 5%. During deformation, the hard second phases create a high work hardening rate, while the retained austenite transforms to martensite, increasing the work hardening rate at higher strain levels. The complex phase (CP) steels consist of a very fine microstructure of ferrite and a higher volume fraction of hard phases, which are strengthened further by fine precipitates. In the martensitic (MART) steels, the austenite that exists during hot rolling or annealing is transformed almost entirely to martensite during quenching on the run-out table or in the cooling section of the annealing line. All AHSS are produced by controlling the cooling rate from the austenite or austenite plus ferrite phase, either on the runout table of the hot mill (for hot-rolled products) or in the cooling section of the continuous annealing furnace (continuously annealed or hot dip coated products). Advanced high-strength cooling patterns and resultant microstructures are schematically illustrated on the continuous cooling-transformation diagram, available for examination on the AISI Web site (www.steel.org). The cooling patterns are designed on the bases of mathematical models, which attempt to predict the structures and properties resulting from the processing technique. Recent research concentrated on the development of these models.

According to Ref 106, most important models, developed for all aspects of the hot rolling of steel, are those that attempt to predict the evolution of the microstructure—including the properties after hot rolling and cooling in addition to rolling load prediction. The models are now used, offline, to design the process of hot rolling of steels. Several of these models were reviewed and discussed (Ref 22). The models include those found in Ref 107 to 118. In these reports, various closed form equations are presented, describing the processes of recrystallization and grain growth. Their predictive capabilities under laboratory and industrial conditions have also been discussed and, in general, found reasonable; see for example, the experiments and calculations of the size of the ferrite grains after hot rolling experiments described in Ref 118. Niobium-carrying steels were hot rolled, cooled in air, and the resulting ferrite grain sizes were determined. As observed in the following table (reproduced from Ref 22), the comparison of the measurements and the predictions is good.

Measured and calculated ferrite grain size

Test, °C (°F)	Center		Surface	
(Start temperature)	Measured	Calculated	Measured	Calculated
1187 (2169)	7.3	8	5.5	6.8
1168 (2134)	8.4	8.2	6.4	6.7
1163 (2125)	8.4	7.4	7.0	6.5
1209 (2208)	8.4	7.4	7.0	6.5
1169 (2136)	7.4	6.1	6.3	5.0
1187 (2169)	4.8	6.1	3.3	4.6
1169 (2136)	6.3	7.7	3.6	6.0

The Symposium on Emerging Steel Processing Technologies, held in conjunction with the 43rd Mechanical Working and Steel Processing Conference (October, 2001), considered the application of these approaches to various steel grades. Bake hardening and aging of TRIP steels are considered in Ref 119. The development of texture in dual phase steels and of TRIP steels were the topics of Ref 120 and 121, respectively. In each of these studies the effect of the thermal treatment on the resulting properties was examined. Higher intercritical holding temperatures resulted in higher strength, higher elongation, and higher formability index. The annealing cycles used in industrial conditions were presented in Ref 122, and the best mechanical properties were obtained at an overaging temperature of 400 °C (752 °F). The effect of different prior austenite conditions on transformation kinetics and on the resulting properties of TRIP steels was examined in Ref 123. The thermal schedule was given and it was concluded that the volume fraction of the retained austenite is not strongly dependent on the refinement of ferrite. Further, the effect of the retained austenite on the mechanical properties was more significant than the ferrite grain refinement. Strain-induced transformation was used to create ultra-fine ferrite grains of 1 to 3 μm size (Ref 124). The success in producing these very small grains depended on the fine tuning of the thermal-mechanical process. The rate of strain was found to be one of the most important parameters.

The state-of-the-art of thermomechanical processing is discussed in Ref 125. The authors divided the techniques into those that involve control of the deformation process and those that control the post-deformation cooling history. They also describe processes that create ultra-fine ferrite grains either through strain-induced transformation or by creating new plant configurations. Various processes and their details are listed in a comprehensive table, including conventional and recrystallization controlled rolling, accelerated cooling, direct quenching, quenching and self-tempering, rod cooling, warm forming, and intercritical rolling. They conclude by writing that thermomechanical processing has not changed in a marked fashion over the past few decades and that most of the recent research was devoted to the development of mathematical models of the process.

Conclusions

While the limits of workability in the rolling process depend on a large number of materials and process parameters, the most important contributor appears to be the geometrical parameter—the shape factor—defined as the ratio of the average thickness of the rolled metal to the projected contact length of the pass. When the shape factor is larger than unity, the distributions of the field variables, i.e., the stresses, strains, strain rates, and temperatures, are highly nonhomogeneous and will result in flow localization. The probability of reaching the limits of workability is very high. This phenomenon, combined with a possible loss of ductility, is the likeliest cause of ductile fracture, whether in the form of edge cracking, alligatoring, or surface cracking.

When the shape factor is much less than unity, such as in the cold-rolling process or in the finishing stands of the hot strip mill, the most likely cause of failure is probably the tensile cracking of the edges of the rolled strip. In the rolling of shapes of non-rectangular cross sections, the definition of the shape factor is of little use. Still, failure is most likely caused by flow localization and the loss of ductility.

Mathematical models, attempting to predict the limits of workability, are also available in the technical literature. It appears, however, that the Cockroft and Latham criterion (Ref 57) is the model of choice because it is the one whose predictions have been most often substantiated.

Research publications concerning process design in rolling are few, probably because of the proprietary nature of the processes. When the draft schedule for a particular steel has been developed and has been shown to be successful, the details are understandably guarded closely by the companies. Experimental simulation of the hot-rolling process of microalloyed steels shows the success of thermal-mechanical treatment in producing high-quality, high-strength steels.

ACKNOWLEDGMENT

Substantial portions of this chapter are due to Schey's original publication (Ref 1) on this topic, and several portions of his text are reproduced, essentially unchanged.

REFERENCES

1. J.A. Schey, Workability in Rolling, *Workability Testing Techniques,* G.E. Dieter, Ed., ASM International, 1984, p 269–314
2. W.L. Roberts, *Cold Rolling of Steel,* Marcel Dekker Inc., 1978
3. W.L. Roberts, *Hot Rolling of Steel,* Marcel Dekker Inc., 1983
4. W.L. Roberts, *Flat Rolling of Steel,* Marcel Dekker Inc., 1978
5. V.B. Ginzburg, *High-Quality Steel Rolling: Theory and Practice,* Marcel Dekker, 1993
6. Z. Wusatowski, *Fundamentals of Rolling,* Pergamon, Oxford, 1969
7. E.C. Larke, *The Rolling of Strip, Sheet and Plate,* 2nd ed., Chapman and Hall, 1965
8. C.W. Starling, *The Theory and Practice of Flat Rolling,* University Press, London, 1962
9. A. Tselikov, *Stress and Strain in Metal Rolling,* MIR, Moscow, 1967
10. L.R. Underwood, *The Rolling of Metals*, John Wiley & Sons, 1950
11. I. Ya. Tarnovskii, A.A. Pozdeyev, and V.B. Lyashkov, *Deformation of Metals during Rolling,* Pergamon, Oxford, 1965
12. R. Hill, *The Mathematical Theory of Plasticity,* Oxford University Press, 1950
13. O. Hoffman and G. Sachs, *Introduction to the Theory of Plasticity for Engineers,* McGraw-Hill, 1953
14. B. Avitzur, *Metal Forming: Processes and Analysis,* McGraw-Hill, 1968
15. G.W. Rowe, *Principles of Industrial Metalworking Processes,* Edward Arnold, London, 1977
16. W.A. Backofen, *Deformation Processing,* Addison-Wesley, 1972
17. W.F. Hosford and R.M. Caddell, *Metal Forming: Mechanics and Metallurgy,* 2nd ed., Prentice Hall, Englewood Cliffs, 1993
18. W. Johnson and P.B. Mellor, *Plasticity for Mechanical Engineers,* Van Nostrand, London, 1962
19. J. Lubliner, *Plasticity Theory,* Macmillan, New York, 1990
20. E.M. Mielnik, *Metalworking Science and Engineering,* McGraw-Hill, 1991
21. R.H. Wagoner and J.-L. Chenot, *Fundamentals of Metal Forming,* John Wiley & Sons, 1996
22. J.G. Lenard, M. Pietrzyk, and L. Cser, *Mathematical and Physical Simulation of the Properties of Hot Rolled Products,* Elsevier, 1999
23. D. Kobasa and R.A. Schultz, *Experimental Determination of the Length of the Arc of Contact in Cold Rolling,* Iron and Steel Eng., 1968, p 1–6
24. J.R. Yow, "Behavior of the Elastic Entrance Region in Cold Strip Rolling," Ph.D. dissertation, North Carolina State University, Raleigh, NC, 1970
25. E. Orowan, The Calculation of Roll Pressure in Hot and Cold Flat Rolling, *Proc.*

Inst. Mech. Eng., Vol 150, 1943, p 140–167
26. D.R. Bland and H. Ford, The Calculation of Roll Force and Torque in Cold Strip Rolling with Tensions, *Proc. Inst. Mech. Eng.,* Vol 159, 1948, p 144–153
27. R.B. Sims, The Calculation of Roll Force and Torque in Hot Rolling Mills, *Proc. Inst. Mech. Eng.,* Vol 168, 1954, p 191–200
28. L.S. Lim and J.G. Lenard, A Study of Friction in Cold Strip Rolling, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 106, 1984, p 139–146
29. A.N. Karagiozis and J.G. Lenard, Temperature Distribution in a Slab During Hot Rolling, *J. Eng. Mater. Technol. (Trans. ASME),* Vol 110, 1988, p 17–21
30. B. Hum, H.W. Colquhoun, and J.G. Lenard, Measurements of Friction During Hot Rolling of Aluminum Alloys, *J. Mater. Process. Technol.,* Vol 60, 1996, p 331–338
31. L. Prandtl, Anwendungsbeispiele zu einem Henckyschen Satz über das plastische Gleich-gewicht, *ZAMM,* Vol 3, 1926, p 401–406
32. E. Orowan and K.J. Pascoe, A Simple Method of Calculating Roll Pressure and Power Consumption in Hot Flat Rolling, Special Report 34, Iron and Steel Institute, 1946, p 124
33. F. Körber and E. Siebel, "Über die Beanspruchungsverhältnisse beim Schmieden und Walzen," Mitt. Kaiser-Wilh.-Inst. Eisenforsch., Vol 10, 1928, p 15–22
34. O. Pawelski, *Stahl und Eisen,* Vol 83, 1963, p 1440–1451
35. F.A.A. Crane and J.M. Alexander, *J. Inst. Met.,* Vol 96, 1968, p 289–300
36. P. Dewhurst, I.F. Collins, and W. Johnson, *J. Mech. Eng. Sci.,* Vol 15, 1973, p 439–447
37. H. Ford and J.M. Alexander, Simplified Hot-Rolling Calculations, *J. Inst. Met.,* Vol 92, 1963–64, p 397–403
38. J.A. Schey, *Introduction to Manufacturing Processes,* 3rd ed., McGraw-Hill, 2000
39. P.M. Cook and A.W. McCrum, *The Calculation of Load and Torque in Hot, Flat Rolling,* British Iron and Steel Research Association, 1958
40. A. Murthy and J.G. Lenard, "Statistical Evaluation of Some Hot Rolling Theories," *J. Eng. Mater. Technol.,* Vol 104, 1982, p 47–52
41. S. Shida, "Effect of Carbon Content, Temperature and Strain Rate on Compression Flow Stress of Carbon Steels," Hitachi Research Report, 1974, p 1–9
42. J.A. Schey, Fracture in Rolling Processes, *J. Appl. Metalwork.,* Vol 1, 1980, p 48–59
43. D. Duly, J.G. Lenard, and J.A. Schey, Applicability of Indentation Tests to Assess Ductility in the Hot Rolling of Aluminium Alloys, *J. Mater. Process. Technol.,* Vol 75, 1998, p 143–151
44. Discussion: Edge Cracking, Lamination and Surface Cracking in Hot and Cold Rolling, *J. Inst. Met.,* Vol 92, 1964, p 254–256
45. D.T. Blazevic, Tertiary Rolled in Scale—the Hot Strip Mill Problem of the 1990's, *37th MWSP Conf. Proc.,* ISS-AIME, XXXIII, 1996, p 33–38
46. Y.H. Li and C.M. Sellars, Modelling Surface Temperatures during Hot Rolling of Steel, *Modelling of Metal Rolling Processes 3,* J.H. Beynon, M.T. Clark, P. Ingham, P. Kern, K. Waterson, Ed: IOM Communications, London, 1999, 178–186
47. M. Krzyzanowski and J.H. Beynon, The Tensile Failure of Mild Steel Oxides under Hot Rolling Conditions, *Steel Res.,* Vol 70, 1999, p 22–27
48. J.B. Tiley, J.G. Lenard, and Y. Yu, Roll Bite Deformation of the Thin Scale Layer on a Plain Carbon Steel during Hot Rolling, *42nd Mechanical Working and Steel Processing Conf.* (Toronto), Vol 38, Iron and Steel Society, 2000, p 215–222
49. A. Jones and B. Walker, *Met. Technol.,* Vol 1, 1974, p 310–315
50. E.A. Leach, P. Gregory and R. Eborall, *J. Inst. Met.,* Vol 83, 1954–55, p 347–353
51. Y. Maehara, K. Nakai, K. Yasumoto, and T. Mishima, Hot Cracking of Low Alloy Steels in Simulated Continuous Casting—Direct Rolling Process, *Trans. Iron Steel Inst. Jpn.,* Vol 28, 1988, p 1021–1027
52. W. Dahl, E. Wildschütz, and W. Schiffgen, *Arch. Eisenhüttenwes,* Vol 39, 1968, p 501–509
53. F. Kasz and P.C. Varley, *J. Inst. Met.,* Vol 76, 1949–50, p 423–428
54. S. Kobayashi and S.I. Oh, Workability of Aluminum Alloy 7075-T6 in Upsetting and Rolling, *J. Eng. Ind. (Trans. ASME B),* Vol 98, 1976, p 800–806
55. V.K. Vorontsov, A.I. Baganov, B.V. Kucheryaev, V.V. Brinza, and V.P. Gusev, Study of Workability of AK4-1ch Alloy, *Tekhnol. Legk. Splavov,* Vol 7, 1981, p 15–19
56. A L. Hoffmanner, *Metal Forming: Interrelation Between Theory and Practice,* Plenum, 1971
57. M.H. Cockroft, and D.J. Latham, Ductility and Workability of Metals, *J. Inst. Met.,* Vol 96, 1968, p 33–39
58. P. Montmitonnet, J.L. Chenot, C. Bertrand-Corsini, C. David, T. Iung, and P. Buessler, A Coupled Thermomechanical Approach for Hot Rolling by a 3D Finite Element Method, *J. Eng. Ind. (Trans. ASME),* Vol 114, 1992, p 336–344
59. J.-H. Kwak, J.-H. Chung, and K.M. Cho, Optimum Mn/S Ratio and Heat Treatment of Low Carbon Steels in Mini-Mill Process to Prevent Edge Cracking, *J. Korean Inst. Met. Mater.* (South Korea), Vol 38, 2000, p 213–218
60. J.-H. Kwak, J.-H. Chung, and K.M. Cho, A Study on the Edge Cracking of Low Carbon Steel Sheets Manufactured by Mini-Mill Process, *42nd Mechanical Working and Steel Processing Conf. Proc.,* Vol 38, Iron and Steel Society, 2000, p 311–320
61. A. Wang, P.F. Thomson, and P.D. Hodgson, A Study of Pore Closure and Welding in Hot Rolling Process, *J. Mater. Process. Technol.,* Vol 60, 1996, p 95–102
62. S. Turczyn and Z. Malinowski, Split Ends and Central Burst Defects in Rolling, *Materials Processing Defects,* S. K. Ghosh and M. Predeleanu, Ed., Elsevier, p 401–416
63. J.A. Schey, *J. Inst. Met.,* Vol 94, 1966, p 193–200
64. S. Turczyn, The Effect of the Roll-Gap Shape Factor on Internal Defects in Rolling, *J. Mater. Process. Technol.,* Vol 60, 1996, p 275–282
65. W.C.F. Hessenberg and L. Bourne, BISRA Report MW/A/51/49, 1949
66. S. Sheu, L.G. Hector and O. Richmond, Tool Surface Topographies for Controlling Friction and Wear in Metal Forming Processes, *J. Tribology (Trans. ASME),* Vol 120, 1998, p 517–527
67. P. Macura and J. Petruska, Numerical and Experimental Simulation of Pass Rolling, *J. Mater. Process. Technol.,* Vol 60, 1996, p 55–60
68. J.A. Charles, Hot Processability and the Optimization of Hot-Worked Properties, *Hot Working and Forming Processes,* C.M. Sellars and G.J. Davies, Ed., The Metals Society, 1979, p 87–98
69. D.V. Wilson, Influences of Metallurgical Structure on Formability, *Formability and Metallurgical Structure,* A.K. Sachdev and J.D.Embury, Ed., The Metallurgical Society, Warrendale, PA, 1987, p 3–32
70. J.D. Embury, Microstructural Aspects of Formability, *Formability and Metallurgical Structure,* A.K. Sachdev and J.D. Embury, Ed., The Metallurgical Society, Warrendale, PA, 1987, p 101–116
71. H.J. McQueen, and D.L. Bourell, Hot Workability of Metals and Alloys, *J. Met.* 1987, p 28–35
72. D.N. Crowther, Z. Mohamed, and B. Mintz, Influence of Micro-Alloying Additions on the Hot Ductility of Steels Heated Directly to the Test Temperature, *Trans. Iron Steel Inst. Jpn.,* Vol 27, 1987, p 366–375
73. H. Kobayashi, S. Yamaguchi, and M. Endo, Effect of Minor Elements on Hot Workability of Two-Phase Stainless Steel, *Hot Working and Forming Processes,* C.M. Sellars and G.J. Davies, Ed., The Metals Society, 1979, 133–139
74. S.R. Keown, Hot Ductility of Wrought Austenitic Stainless Steels, *Hot Working and Forming Processes,* C.M. Sellars and G.J. Davies, Ed., The Metals Society, 1979, 140–147
75. Y. Harada, M. Ohmori, and S.-I. Ohnishi, Workability in Rolling and Upsetting of Cast Chromium, *J. Jpn. Inst. Met.,* Vol 54, 1990, p 473–479

76. M. Ueki, S. Horie, and T. Nakamura, Simulation of the Hot Working of 5083 Aluminum Alloy By Means of the Torsion Tests, *J. Mech. Work. Technol.,* Vol 11, 1985, p 365–376
77. G. Fitzsimons and H.A. Kuhn, Flow and Fracture of a Multiphase Alloy MP35N for Study of Workability, *Metall. Trans.,* Vol 15A, 1984, p 1837–1848
78. K. Tomimura and K. Miyakusu, Effect of Boron on Hot Workability and Ductility in Ultra-High-Strength Metastable Austenitic Stainless Steels, *Nisshin Steel Tech. Rep.* (Japan), Vol 80, 2000, p 10–16
79. M. Predeleanu, Damage Mechanics Approach for Predicting Material Processing Defects, *Surface Treatment V: Computer Methods and Experimental Measurements (USA),* 2001, p 129–139
80. T. Rajagopalachary and V. V. Kutumbarao, Model for the Evaluation of Hot Workability of Metallic Materials. *J. Mater. Process. Technol.,* Vol 24, 1990, p 333–342
81. D.-C. Ko, D.-H. Kim, and B.-M. Kim, Application of Artificial Neural Network and Taguchi Method to Preform Design in Metal Forming Considering Workability, *Int. J. Mach. Tools Manuf.,* Vol 39, 1999, p 771–785
82. B.P.P.A. Gouveia, J.M.C. Rodrigues, and P.A.F. Martins, Fracture Predicting in Bulk Metal Forming, *Int. J. Mech. Sci.,* Vol 38, 1996, p 361–372
83. M. Zheng, Z.J. Luo, and X. Zheng, New Damage Model for Ductile Materials, *Eng. Fract. Mech.,* Vol 41, 1992, p 103–110
84. S. Shima and Y. Nose, Development of Upper Bound Technique for Analysis of Fracture in Metal Forming, *Ingenieur-Arch.,* Vol 60, 1990, p 311–322
85. Z.J. Luo, W.H. Ji, N.C. Guo, X.Y. Xu, Q.S. Xu, and Y.Y. Zhang, Ductile-Damage Model and Its Application to Metal-Forming Processes, *J. Mater. Process. Technol.,* Vol 30, 1992, p 31–43
86. E. Ervasti and U. Stahlberg, Behavior of Longitudinal Surface Cracks in the Hot Rolling of Steel Slabs, *J. Mater. Process. Technol.,* Vol 94, 1999, p 141–150
87. J.C. Gelin and M. Predeleanu, Finite Strain Elasto-Plasticity Including Damage—Application to Metal Forming Problems, *NUMIFORM 89,* E.G. Thompson, R.D. Wood, O. Zienkiewicz, and A. Samuelsson, Ed., A.A. Balkema, 1989, p 151–157
88. J.C. Gelin, Modelling of Damage in Metal Forming Processes, *J. Mater. Process. Technol.,* Vol 80–81, 1998, p 24–32
89. P. Picart, O. Ghouati, and J.C. Gelin, Optimization of Metal Forming Process Parameters with Damage Minimization, *J. Mater. Process. Technol.,* Vol 80–81, 1998, p 597–601
90. O. Branswyck, C. David, C. Levaillant, J.L. Chenot, J.P. Billard, D. Weber, and J.P. Guerlet, Surface Defects in Hot Rolling of Copper Based Alloys Heavy Ingots: Three Dimensional Simulation Including Failure Criterion, *Computational Methods for Predicting Material Processing Defects,* M. Predeleanu, Ed., Elsevier, 1987, p 29–38
91. M. Oyane, T. Sato, K. Okimoto, and S. Shima, Criteria for Ductile Fracture and Their Applications, *J. Mech. Work. Technol.,* Vol 4, 1980, p 65–81
92. M.P. Groover, *Fundamentals of Modern Manufacturing,* 2nd ed., John Wiley and Sons, 2002
93. G. Taguchi, E.A. Elsayed, and T.C. Hsiang, *Quality Engineering in Production Systems,* McGraw Hill, 1989
94. I. Tamura, H. Sekine, T. Tanaka, and C. Ouchi, *Thermomechanical Processing of High-Strength Low-alloy Steels,* Butterworth and Co., 1988
95. G. Tither, *Proc. Int. Conf. on HSLA Steels: Processing, Properties and Applications,* G. Tither and S. Zhang, Ed., The Minerals, Metals, and Materials Society, Warrendale, PA, 1992, p 61–68
96. R.K. Amin and F.B. Pickering, *Proc. Int. Conf. on Thermomechanical Processing of Microalloyed Austenite,* A.J. DeArdo, G.A. Ratz, and P.J. Wray, Ed., AIME, Pittsburgh, PA, 1981, p 1–33
97. I. Weiss, G.L. Fitzsimons, K. Mielityinen-Tiitto, and A.J. DeArdo, *Proc. Int. Conf. on Thermomechanical Processing of Microalloyed Austenite,* A.J. DeArdo, G.A. Ratz, and P.J. Wray, Ed., AIME, Pittsburgh, PA, 1981, p 33–59
98. L.J. Cuddy, *Metall. Trans. A,* Vol 12A, 1981, p 1313–1320
99. L.J. Cuddy, *Proc. Int. Conf. on Thermomechanical Processing of Microalloyed Austenite,* A.J. DeArdo, G.A. Ratz, and P.J. Wray, Ed., AIME, Pittsburgh, PA, 1981, p 129–140
100. Y.C. Hirsch and B.A. Parker, *Proc. Int. Conf. on Advances in the Physical Metallurgy and Application of Steels,* The Metals Society, Liverpool, 1981, p 26
101. P.E. Repas, *Proc. Int. Conf. on HSLA Steels—Technology and Applications,* (Philadelphia, PA), ASM International, 1984, p 203–208
102. H.L. Andrade, M.G. Akben, and J.J. Jonas, *Metall. Trans. A,* Vol 14A, 1983, p 1967
103. I. Weiss and J.J. Jonas *Metall. Trans. A,* Vol 11A, 1980, p 403
104. M.G. Akben, I. Weiss, and J.J. Jonas *Acta Metall.,* Vol 29, 1981, p 111
105. M.G. Akben, B. Bacroix, and J.J. Jonas *Acta Metall.,* Vol. 31, 1983, p 161
106. S. Yue, private communication, 2001
107. C.M. Sellars, Modelling Microstructural Development During Hot Rolling, *Mater. Sci. Technol.,* Vol 6, 1990, p 1072–1081
108. C.M. Sellars, An Internal State Variable Approach to Modelling Microstructural Evolution during Thermomechanical Processing, *Proc. THERMEC '97,* T. Chandra, and T. Sakai, Ed., Wollongong, 1997, p 3–11
109. A. Laasraoui and J.J. Jonas, Prediction of Temperature Distribution, Flow Stress and Microstructure during the Multipass Hot Rolling of Steel Plate and Strip, *ISIJ Int.,* Vol 31, 1991, p 95–105
110. A. Laasraoui and J.J. Jonas, Recrystallization of Austenite after Deformation at High Temperature and Strain Rates—Analysis and Modelling, *Metall. Trans. A,* Vol 22A, 1991, 151–160
111. P. Choquet, P. Fabregue, J. Giusti, B. Chamont, J.N. Pezant, and F. Blanchet, Modelling of Forces, Structure and Final Properties during the Hot Rolling Process on the Hot Strip Mill, *Proc. Mathematical Modelling of Hot Rolling of Steel,* S. Yue, Ed., Hamilton, 1990, p 34–43
112. P.D. Hodgson, and R.K. Gibbs, A Mathematical Model to Predict the Mechanical Properties of Hot Rolled C-Mn and Microalloyed Steels, *ISIJ Int.,* Vol 32, 1992, p 1329–1338.
113. H. Yada, Prediction of Microstructural Changes and Mechanical Properties in Hot Strip Rolling, *Proc. Symp. Accelerated Cooling of Rolled Steel,* G.E. Ruddle and A.F. Crawley, Ed., Pergamon Press, Winnipeg, 1992, p 105–119
114. J.H. Beynon, and C.M. Sellars, Modelling Microstructure and Its Effects during Multipass Hot Rolling, *ISIJ Int.,* Vol 32, 1992, p 359–367
115. T. Sakai, Dynamic Recrystallization Microstructures under Hot Working Conditions, *J. Mater. Process. Technol.,* Vol 53, 1995, p 349–361.
116. R. Kuziak, Y.-W. Cheng, M. Glowacki, and M. Pietrzyk, "Modelling of the Microstructure and Mechanical Properties of Steels during Thermomechanical Processing," NIST Technical Note 1393, Boulder, 1997
117. C. Devadas, I.V. Samarasekera, and E.B. Hawbolt, The Thermal and Metallurgical State of Steel Strip during Hot Rolling, Part III: Microstructural Evolution, *Metall. Trans.,* Vol 22A, 1991, p 335–349
118. M. Pietrzyk, C. Roucoules, and P.D. Hodgson, Modelling the Thermomechanical and Microstructural Evolution during Rolling of a Nb HSLA Steel, *ISIJ Int.,* Vol 35, 1995, p 531–541
119. M. De Meyer, A.K. De, L. Tosal-Martinez, and B.C. De Cooman, The Bake Hardening and Ageing Behaviour of Cold Rolled TRIP Steels, *43rd Mechanical Working and Steel Processing Conf. Proc.,* Vol 39, Iron and Steel Society, 2001, p 349–358
120. T. Waterschoot, L. Kestens, S. Vandeputte, and B.C. De Cooman, Development of the Texture in MnCr Hot Rolled Dual Phase Steel: Influence on the r-Value, *43rd Mechanical Working and Steel Processing*

Conf. Proc., Vol 39, Iron and Steel Society, 2001, p 373–383

121. M. De Meyer, L. Kestens, L. Tosal-Martinez, and B.C. De Cooman, Texture Development in Intercritically Annealed TRIP Steels, *43rd Mechanical Working and Steel Processing Conf. Proc.,* Vol 39, Iron and Seel Society, 2001, p. 333–347
122. A. Pichler, S. Traint, H. Pauli, H. Mildner, J. Szynyur, M. Blaimschein, P. Stiaszny, and E. Werner, Processing and Properties of Cold Rolled TRIP Steels, *43rd Mechanical Working and Steel Processing Conf. Proc.,* Vol 39, Iron and Steel Society, 2001, p 411–433.
123. S. Jiao, F. Hassani, R.L. Donaberger, E. Essadiqi, and S. Yue, The Effect of the Prior Heat Treatment on a Cold Rolled and Annealed Mo-Nb Microalloyed TRIP Steel, *43rd Mechanical Working and Steel Processing Conf. Proc.,* Vol 39, Iron and Steel Society, 2001, p 385–395
124. T. Yokota, K. Sato, M. Miyake, and M. Niikura, Ultra Refinement in Austenite Grain Size Through the Spontaneous Reverse Transformation Due to Adiabatic Deformation Heating, *43rd Mechanical Working and Steel Processing Conf. Proc.,* Vol 39, Iron and Steel Society, 2001, p 479–488
125. P.D. Hodgson and M.R. Barnett, The Thermomechanical Processing of Steels—Current Status and Future Trends, *Metal Forming 2000,* M. Pietrzyk, J. Kusiak, J. Majta, P. Hartley, and I. Pillinger, Ed., Balkema, 2000, p 21–28

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 278-290
DOI:10.1361/hwpd2003p278

Chapter 19

Drawing of Wire, Rod, and Tube*

IN THE DRAWING PROCESS, the cross-sectional area and/or the shape of a rod, bar, tube, or wire is reduced by pulling through a die. One of the oldest metalforming operations, drawing allows excellent surface finishes and closely controlled dimensions to be obtained in long products that have constant cross sections. The deformation is accomplished by a combination of tensile and compressive stresses that are created by the pulling force at the exit from the die and by the die configuration (Fig. 1).

In drawing, a previously rolled, extruded, or fabricated product with a solid or hollow cross section is pulled through a die, sometimes at exit speeds as high as several thousand feet per minute. The die geometry determines the final dimensions, the cross-sectional area of the drawn product, and the reduction in area. The purpose of drawing might be simply to tighten tolerances, improve surface finish, or increase the strength of the product; in this case, a single draw (pass) at room temperature might suffice. More frequently, the end product is of a smaller cross section than can be produced by hot working. In this case, a sequence of passes may be employed, usually at cold working temperatures, with interpass process anneals if required. Occasionally, material of limited ductility is drawn warm or hot. Drawing is usually conducted at room temperature using a sequence of passes or reductions through consecutively located dies. An important exception is the warm drawing of tungsten to make incandescent lamp filaments. Annealing might occasionally be necessary after several drawing passes before the drawing operation is continued.

Wire and bar drawing are basically the same. In general, drawing is a steady-state process from both the mechanical and tribological viewpoints. The workpiece is long and is continuously drawn into the die. Even though non-steady-state conditions exist during acceleration, they are significant in only a few instances. Unless the lubricant breaks down, forces and temperatures attain steady levels. Proper lubrication is essential in rod, tube, and wire drawing. In contrast to rolling, little or no friction is needed for wire drawing, tube sinking, and tube drawing on a fixed plug. However, some minimum friction is essential for drawing with a floating plug, and friction is helpful on the tube/bar interface in drawing on a bar. Therefore, if at all possible, the lubricant is chosen to give lowest friction and minimum wear. It is essential, though, that the heat generated be extracted, especially in high-speed drawing; if this is not done, the lubricant may break down, and the properties of the drawn product may suffer.

Drawing of Bar and Wire

Significant quantities of bar and wire materials are produced by drawing. In wire or rod drawing, the section is usually round but could also be a shape. In the cold drawing of shapes, the basic contour of the incoming shape is established by cold-rolling passes that are usually preceded by annealing. After rolling, the section shape is refined and reduced to close tolerances by cold drawing. Again, several steps may be necessary to eliminate the effects of strain hardening, that is, to reduce the flow stress and to increase the ductility.

A typical carbide drawing die is shown in Fig. 2. During drawing, the workpiece is pulled through a stationary die with a converging channel, where reduction of the cross-sectional-area from A_o to A_1 takes place. The wire or rod makes

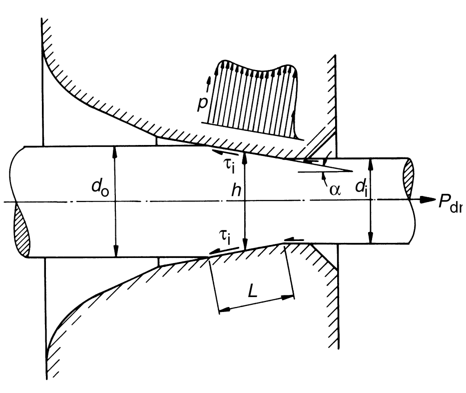


Fig. 1 The drawing process. Source: Ref 1, p 345

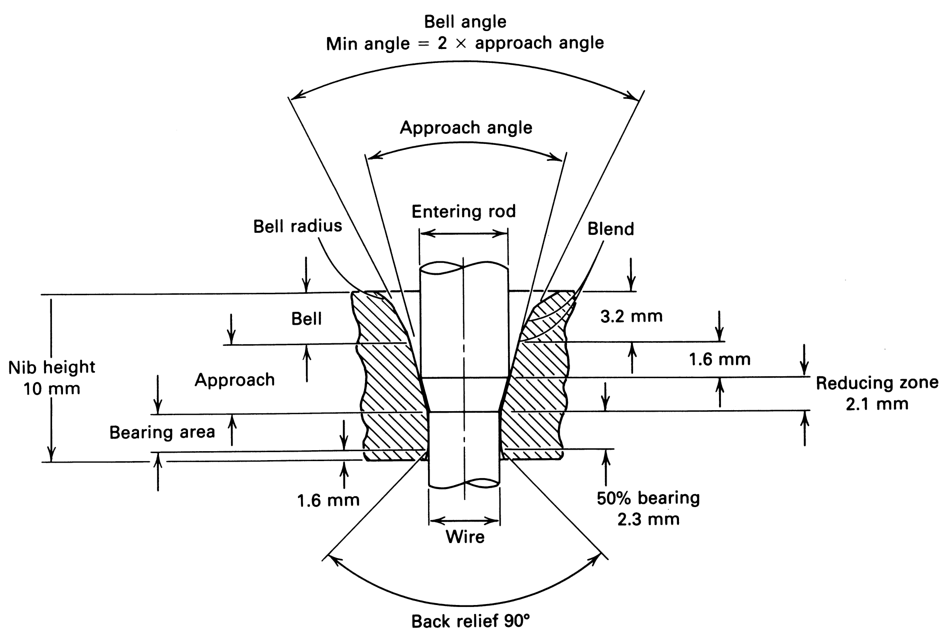


Fig. 2 Cross section of a typical carbide die for wire drawing, 5.5 mm (0.218 in.) diam rod to 4.6 mm (0.180 in.) diam wire (17% reduction per pass)

*Portions adapted from *Tribology in Metalworking: Friction, Lubrication and Wear* by John Schey, American Society for Metals book, 1983, pages 343–393 (Ref 1) and from "Drawing of Rod and Wire," by Roger N. Wright, *Encyclopedia of Materials Science and Engineering,* Volume 2, M.B. Bever, Ed., Pergamon Press and The MIT Press, 1986, pages 1227–1231 (Ref 2).

contact in the drawing cone along the approach angle and is reduced to the dimensions of the drawing cone exit. To prevent sudden decompression of lubricant film and to avoid a rapid increase in diameter on die wear, a parallel land (bearing) is always provided. The bearing region involves no further reduction and allows the die to be refinished without a change in the exit dimensions of the drawing cone. The back relief ensures a gradual drop in stresses in the die and reduces the amount of abrasion that takes place if the drawing stops or if the die is out of alignment. Lubricant is introduced at the bell portion of the die and is pulled into the die/wire interface by the moving wire.

Reduction is defined as the engineering compressive strain

$$r = \frac{A_0 - A_1}{A_0} \quad \text{(Eq 1)}$$

or

$$r = \frac{A_0 - A_1}{A_0} \times 100(\%)$$

For purposes of calculation, the true strain ε is preferred:

$$\varepsilon = \ln \frac{A_0}{A_1} \quad \text{(Eq 2)}$$

Because the volume of the wire must remain constant, the wire elongates. If drawn at a speed v_1, it enters the die at a lower v_0 speed, which can be calculated from constancy of volume:

$$v_0 A_0 = v_1 A_1 \quad \text{(Eq 3)}$$

However, in contrast to rolling, there is no neutral plane here, and the wire always slides over the whole contact zone. To a close enough approximation, it can be assumed that the interface sliding velocity increases from entry to exit linearly in plane-strain drawing and parabolically in axially symmetrical drawing in a conical die. Sliding generates a frictional stress opposing the movement of the wire (Fig. 1), and the contribution of this stress to the draw force is of immediate concern. Deformation occurs under the combination of the longitudinal (draw) stress and the indirect compressive stress generated in the die. Friction is equivalent to a back tension and thus lowers the interface pressure, but at the expense of higher draw forces.

Stresses and Forces (Ref 2)

The pulling force, or drawing stress, cannot exceed the strength of the wire or rod being drawn (otherwise, fracture or unstable deformation would occur). In fact, practical considerations often limit the drawing stress to about 60% of the as-drawn flow stress. This sets a limit on the attainable reduction, which decreases with increasing friction and height-to-length ratio (h/L), of approximately 55% under the best conditions. However, no wire can be perfect, and drawing with critical reductions would lead to frequent wire breaks. Therefore, practical reductions are limited to 20 or 30% per pass and rarely greater than 30 to 35%. A particularly common reduction is that of an American Wire Gauge of 1, or about 20.7%. Multiple reductions or drawing passes are needed to achieve a large overall reduction, or larger reductions can be achieved in a single operation with extrusion. Alternatively, drawing can be used to generate larger quantities of small-diameter product (for example, 0.01 mm, or 0.0004 in.) with excellent dimensional control (assuming proper die maintenance).

The force required for drawing can be calculated by a variety of methods. By analogy, all solutions can be brought to the general form:

$$p = \hat{\sigma}_0 \left[f\left(\varepsilon; \frac{\tau_i}{\alpha}; \frac{h}{L} \right) \right] = \hat{\sigma}_0 Q \quad \text{(Eq 4)}$$

where $\hat{\sigma}_0$ is the mean flow stress, and p is the applied drawing pressure. The bracketed term (pressure-multiplying factor Q) is a function with three factors that define the three types of work performed by the drawing force. The first term is the portion of the drawing force that does work by pure (homogeneous) deformation of the worked material. Pure deformation work occurs from the homogeneous strain (ε), which is directly proportional to the flow stress σ_0 (or, in cold working, the mean flow stress, $\hat{\sigma}_0$) of the wire material.

The other two terms in the bracketed function of Eq 4 are the force components that overcome friction and expend work in the form of inhomogeneous deformation (known as redundant work). Friction and redundant work, which are always additional to the work required for pure deformation, depend on the die geometry and work interface between the die and drawn material. The contribution of friction increases with increasing interface shear stress (τ_i) and, for a given reduction, decreases with increasing die half angle α (Fig. 3) because of a decrease in sliding length. The interface is usually described by a constant coefficient of friction (μ) even though evidence indicates that friction varies along the contact surface with the die. Friction on the die land also contributes to the draw force.

The contribution of redundant work is the additional work expended due to inhomogeneous deformation. As in other processes, the amount of redundant work increases with an increasing h/L ratio. If the contact length (L) is smaller than the mean thickness or diameter, deformation will be inhomogeneous, as the center would deform less than would be expected. Because deformation proceeds against the restraining effect of the less-deformed center, deformation stresses, forces, and power requirements increase. Usually, a more important consequence of inhomogeneous deformation is the generation of tensile stresses in the less-deformed parts of the workpiece. If these secondary tensile stresses reach high values, a workpiece material of low ductility may suffer internal fracture (called an arrowhead defect). Even if there is no failure during deformation, residual stresses may remain in the workpiece, which can lead to subsequent distortion or to delayed failure due to stress corrosion. Any inhomogeneously deformed workpiece also undergoes differential strain hardening and, if subsequently annealed, will show large grain-size and property variations.

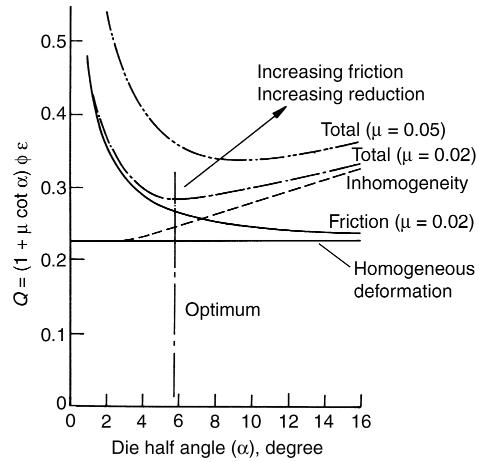


Fig. 3 Components of draw force and optimal die half angle (for 20% reduction in area). Source: Ref 1, p 346

In many processes, friction effects and inhomogeneity induced by process geometry occur simultaneously, sometimes reinforcing and sometimes opposing each other. For a given reduction in wire drawing, the friction contribution decreases and the contribution of redundant work increases with die angle. Therefore, an optimal half angle can be found, as discussed further in the next section on approach angle. The important point is that the effects of friction cannot be artificially separated from the effects of inhomogeneous deformation or redundant work.

Friction effects and redundant work are described in more detail below, but a simple and useful formula for the drawing pressure is:

$$p = \hat{\sigma}_0 Q = \hat{\sigma}_0 (1 + \mu \cot \alpha) \phi \varepsilon \quad \text{(Eq 5)}$$

where ϕ allows for inhomogeneous deformation. Its value can be taken, on the basis of experiments, for round wire as:

$$\phi = 0.88 + 0.2 \frac{h}{L} \quad \text{(Eq 6a)}$$

and for plane-strain drawing as:

$$\phi = 0.8 + 0.2 \frac{h}{L} \quad \text{(Eq 6b)}$$

There is no general agreement on an appropriate formula; the above formulas can give underestimates, and other experiments give better agreement with different formulas.

For a given workpiece material and process geometry, a larger draw force generally indicates higher friction. As in rolling, a back pull (back tension) can be applied to the wire. The resultant drop in interface pressure can be beneficial in terms of the lubrication mechanism and wear, but at the expense of greater draw force.

The approach angle is perhaps the most important feature of the die for most drawing applications. As previously noted, an optimal half angle also can be found, because the frictional component of Eq 4 decreases with an increase of the die angle, while the amount of redundant work increases with an increase of the die angle. The optimal half angle is usually between 5 and 7°, where draw force is minimum. Lower angles are permissible only with very low friction. With an optimal die angle, the contributions of friction and inhomogeneous deformation are of the same order of magnitude. With increasing reduction, die pressures drop, deformation becomes more homogeneous, and the efficiency of the drawing process (the ratio of total work to pure deformation work) increases. The optimal angle shifts to higher values. Because of this, the trumpet-shape dies used in the past actually gave a better approximation of the optimal die profile.

The approach angle also cannot easily be considered independent of section reduction and metal flow. In drawing theory, a Δ parameter is used:

$$\Delta = (\alpha/r)[1 + (1 - r)^{1/2}]^2 \qquad \text{(Eq 7)}$$

where α is the approach semiangle (one-half the included angle) in radians and r is the fractional drawing reduction, given by Eq 1 as $r = 1 - A_1/A_0$,where A_0 and A_1 are the starting and finishing cross-sectional areas, respectively. Commercial die design often involves approach semiangles in the range of 6 to 10° and drawing reductions of about 20%. The corresponding Δ values typically range from 2 to 3, with higher values corresponding to lower reductions and higher die angles, and lower values corresponding to higher reductions and lower die angles. For simple compression, $\Delta = h/l$, but at an oblique angle, then Δ reduces to Eq 7.

Effect of Friction. Basically, low Δ values may involve excessive frictional work between the wire and the drawing cone, and high Δ values involve redundant work or plastic strain beyond that calculable from the reduction in area of the pass. Some degree of redundant work exists for Δ > 1, with redundant work increasing as Δ increases, much as frictional work can increase as Δ decreases. The net effect is that some intermediate value of Δ involves the minimum work and therefore the minimum drawing force, because the drawing force multiplied by the drawing velocity is the work consumed per unit time. Similarly, the drawing stress equals the work per unit volume of wire drawn. The Δ for minimum drawing stress can be approximated by:

$$\Delta_{\min} \cong 4.9\left[\frac{\mu}{\ln(1/1-r)}\right]^{1/2} \qquad \text{(Eq 8)}$$

where μ is the coefficient of friction between the wire and the drawing cone. The drawing stress σ_d can be usefully approximated as:

$$\sigma_d = \hat{\sigma}_0\left[\frac{3.2}{\Delta + 0.9}\right](\alpha + \mu) \qquad \text{(Eq 9)}$$

where $\hat{\sigma}_0$ is the yield or flow stress of the wire during the drawing pass.

Determination of Friction. In principle, the magnitude of friction should be obtainable by various techniques. In practice, however, there are several problems. Because the contributions of friction and inhomogeneous deformation to the total force are of similar magnitude when an optimal approach angle is used, back-calculation of the magnitude of friction from the total draw force is difficult. Simultaneous measurement of draw force and normal force requires more sophisticated equipment and instrumentation, but gives a direct method of measuring frictional forces. (Ref 1, p 347).

The interface shear strength cannot be determined from the total frictional force in drawing over a plug or a bar because there is no assurance that μ_d and μ_p (or μ_d and μ_b) are equal. Measurement of die pressure presents the same difficulties as in bar drawing, although strain gages have been successfully applied to the die circumference. Much work has, therefore, been conducted with plane-strain drawing simulation (see Ref 1, p 204). Plane-strain drawing simulation tests are designed to evaluate specific tribological aspects of deformation processes such as rolling, wire drawing, and sheet drawing. The incentive for developing these test methods is the desire to have better control over process variables, and tests are usually designed to facilitate quantitative measurements of friction and wear. Their common feature is that interface shear stress (τ_i) (or μ) can be obtained without knowing flow stress (σ_f) and without recourse to theory. A common feature of these tests is that new surfaces are generated by means of deforming the specimen through its entire thickness.

Redundant Work of Shear Deformation. Redundant work refers to the additional energy expended due to inhomogeneous (shear) deformation, in addition to the work expended for pure deformation and friction. As previously noted, redundant work increases with an increasing h/L ratio. The ratio can be calculated from various formulas, but, as shown in Fig. 1, it can be simply taken as the mean diameter (or, in plane-strain drawing, the mean thickness) of the workpiece divided by the contact length L. For a given reduction, the ratio increases with increasing α (Fig. 3). Redundant work is also expressed in terms of the redundant work factor or the ratio of total plastic deformation work to the work implied by dimensional change. Experimental studies suggest that the redundant work factor Φ can be estimated to be:

$$\Phi \approx \Delta/6 + 1 \qquad \text{(Eq 10)}$$

The effects of redundant work are considered in great detail by Blazynski (Ref 3) and Avitzur (Ref 4). Inhomogeneous deformation also has several other consequences besides the effect on the drawing pressure. These include differential strain hardening, periodic centerbursts, and bulge formation.

Heat Generation

The management of heat is of great concern in drawing. Much heat is generated directly by the plastic deformation, and this heat is only partially removed by interpass cooling; practical cold-drawing operations can involve wire-temperature increases of a few hundred degrees Kelvin. Most of the total work expended in wire drawing is converted into heat. The die is of small mass and is in continuous contact with the wire, and thus heating presents an even greater danger than in rolling. The dies extract little heat under commercial conditions and become very hot.

At high drawing speeds, most of the heat is retained in the wire; still, the die also heats up, and in multihole drawing the entry temperature of the wire also rises. The surface temperature of the wire is higher because of concentration of redundant work and frictional heating. From the tribochemical view point, peak temperatures are important. On emergence of the wire from the die, temperatures rapidly equalize, and the main concern is the overall temperature level. Therefore, cooling becomes an important function of the lubricant.

Under adiabatic conditions, the temperature increase (ΔT_d) associated with plastic deformation in a single pass is approximately:

$$\Delta T_d = \frac{\Phi\hat{\sigma}_0 \ln[1/(1-r)]}{C\rho} \qquad \text{(Eq 11)}$$

where C and ρ are the heat capacity and density of the wire, respectively. Additional heat generation is associated with frictional work. This heat is concentrated at the die/wire interface and can lead to diminished lubrication, further heating, and catastrophic lubricant breakdown. Accompanying problems include poor wire-surface quality and metallurgical changes near the wire surface. If the coefficient of friction is not influenced by Δ, frictional heating is aggravated by low Δ processing. Fortunately, there is a tendency for low approach angles (and thus low Δ) to foster hydrodynamic lubrication and a reduced coefficient of friction.

Flow through Conical Dies

Relative Drawing Stress and Power Consumption (Ref 5). Wire drawing can be modeled in terms of relative drawing stress and the relative amount of power consumed by internal deformation, redundant work, and friction. The symbols used in this formulation are shown in Fig. 4 for wire drawing through a converging die where section size is reduced from R_0 to R_f. As it passes through the die, the wire rubs against the conical and cylindrical surfaces of the die and encounters friction resistance. The characteristics of the die and the flow patterns in Fig. 4 are common to wire drawing, open-die extrusion, and hydrostatic extrusion.

Under steady-state conditions, a typical solution (Ref 4) expresses the relative drawing stress by the ratio of the drawing stress (or front tension, σ_{xf}) and the flow stress (σ_0) such that:

$$\frac{\sigma_{xf}}{\sigma_0} = \frac{\sigma_{xb}}{\sigma_0} + \dot{W}_i + \dot{W}_s + \dot{W}_f \qquad \text{(Eq 12)}$$

where σ_{xb} is the back tension (Fig. 4), $\dot{W}_i$ is the relative (dimensionless) portion of power consumed by internal deformation, $\dot{W}_s$ is the portion of power consumed by shear or redundant work, and $\dot{W}_f$ is the power portion consumed by friction. The values of these dimensionless components for relative portion of power consumption are calculated from the geometry as follows.

The relative portion of power consumed by internal deformation ($\dot{W}_i$) is expressed as:

$$\dot{W}_i = 2f(\alpha)\ln\frac{R_0}{R_f} \qquad \text{(Eq 13)}$$

where the function of the semicone angle (α) is:

$$f(\alpha) = \frac{1}{\sin^2\alpha}\left\{1 - \cos\alpha\sqrt{\left(1 - \frac{11}{12}\sin^2\alpha\right)} + \frac{1}{\sqrt{(11\cdot 12)}}\ln\left[\frac{1+\sqrt{11/12}}{\sqrt{11/12}\cdot(\cos\alpha) + \sqrt{(1-[11/12]\cdot\sin^2\alpha)}}\right]\right\}$$

The relative portion of power consumed by the shear (or redundant power) component ($\dot{W}_s$) is:

$$\dot{W}_s = \frac{2}{\sqrt{3}}\left(\frac{\alpha}{\sin^2\alpha} - \cot\alpha\right) \qquad \text{(Eq 14)}$$

The relative portion of power consumed by friction ($\dot{W}_f$) along the conical surface of the die is:

$$\dot{W}_f = \frac{2}{\sqrt{3}} m\cdot(\cot\alpha)\ln\frac{R_0}{R_f} \qquad \text{(Eq 15)}$$

where m is an interface friction factor (see Eq 50 in Chapter 2, "Bulk Workability of Metals"). Relative drawing stress is plotted in Fig. 5 versus the semicone angle (α) of the die for various values of the friction factor (m). Each curve in Fig. 5 demonstrates a minimum at some optimal angle. For angles smaller than the optimal angle, friction losses ($\dot{W}_f$) are excessive. Friction losses drop with increasing die angles, but redundant power losses increase with increasing die angle. Beyond the optimal angle, excessive distortion occurs with increasing die angles, and the drawing stress increases. The optimal die angle that minimizes the drawing stress increases with the increase in reduction and friction. With higher friction it is advisable to use larger die angles.

More expressions have been derived for the determination of process limitations due to tearing, dead zone formation, and shaving (Ref 5). When the phenomenon of central burst is analyzed, one finds that increasing friction deters central burst during extrusion, but promotes it during drawing (Fig. 6).

Strain Distribution. Inhomogeneous strain distribution is governed by the h/L ratio, but it also increases with friction. Draw forces are affected by inhomogeneity of deformation—as expressed by the h/L ratio. Inhomogeneous strain distribution also involves some other factors, as described in this section.

Residual Stresses. The first consequence of inhomogeneous strain is that residual stresses are set up in the wire. At low h/L ratios the surface residual stress is tensile. For a given die angle α, stresses decrease with increasing pass reduction and improved lubrication, both of which serve to make deformation more homogeneous. Consequently, the service properties of the wire also change. In a bar of large enough diameter, inhomogeneity is also reflected in hardness variations. Hundy and Singer (Ref 6) found this variation to increase with friction for a 20° die angle, but to be independent of friction for a 30° angle, presumably because inhomogeneity had an overwhelming effect. The surface deformation associated with very high h/L ratios can be beneficial because residual compressive stresses impart greater fatigue resistance.

Die Geometry. The simple conical die used in practice does not ensure the most homogeneous deformation, and ideal die profiles can be theoretically calculated for more uniform strain or strain-rate distribution while also minimizing friction. However, the improvement is seldom sufficient to justify the cost of making dies with convex, sigmoidal, or other complex shapes.

Centerbursts. Surface compressive stresses are balanced by central tensile stresses. At high h/L ratios, triaxial tensile stresses are set up in the center. The density of wire decreases because of void formation, and discontinuity in material flow can result in periodic centerburst defects* in wire of moderate ductility, especially if it is subjected to a succession of high h/L

*For an example of periodic centerburst defects in extended product, see Fig. 5 in Chapter 21.

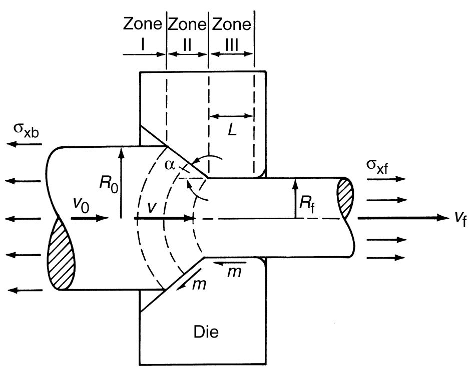


Fig. 4 Wire drawing or extrusion processes. Source: Ref 5, p 59

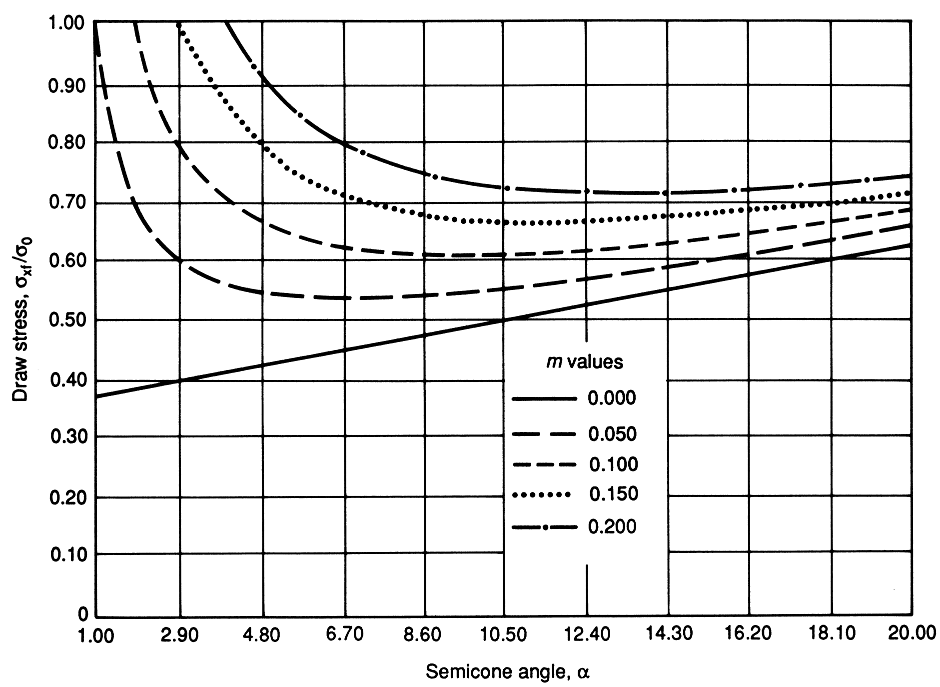


Fig. 5 Relative drawing stress as a function of semicone angle and friction factor (m). Source: Ref 5, p 61

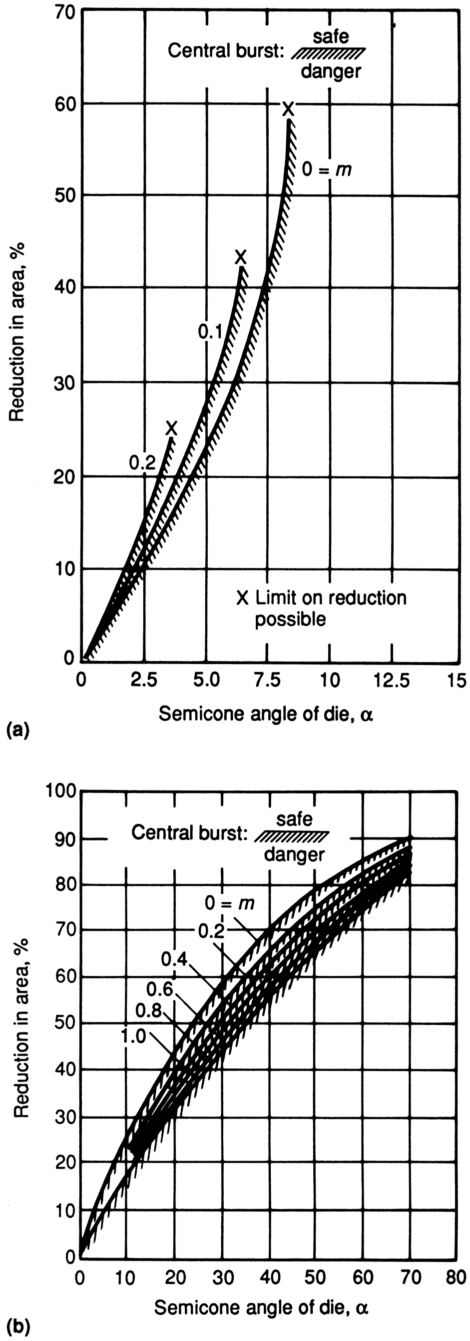


Fig. 6 Criteria for central burst in (a) drawing and (b) extrusion. The safe region is above or to the left of each line. Source: Ref 5, p 61

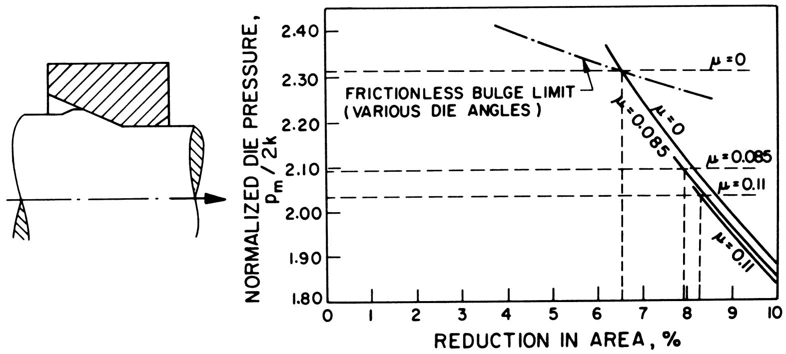


Fig. 7 Condition for bulge formation in drawing with light reductions. Source: Ref 9

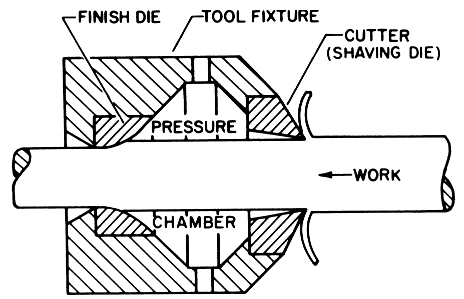


Fig. 8 Pressure-lubricating shaving/drawing die. Source: Ref 10

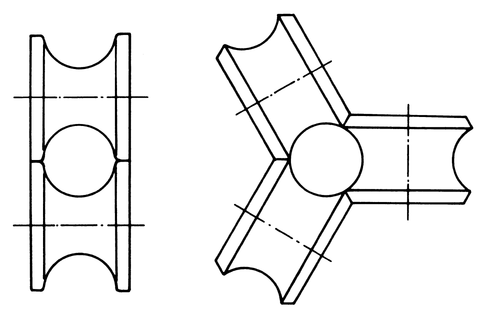


Fig. 9 Configurations of roller dies. Source: Ref 1, p 352

passes. Because friction increases the back tension, it aggravates the situation, and the theoretical safe zone for drawing is reduced (Fig. 6a). Of course, this represents only a necessary but not a sufficient criterion for the onset of centerburst: it is also necessary that the material have low ductility. A model based on the volume fraction of voids has been proposed (Ref 7).

External Deformation and Shaving. High *h*/*L* ratios affect the external deformation, too. First, a bulge forms at the die entry; for a given die angle, bulging sets in at a lower reduction when friction is high (Fig. 7) (Ref 8, 9). At yet higher angles the surface of the wire is shaved off by a process related to cutting with a negative-rake-angle tool. Such shaving is an important step in making wire free of surface defects inherited from the hot processing stages. In one version of pressure die for lubrication, the inlet sealing is replaced with a shaving die (Fig. 8).

Drawing of Profiles. An increasingly important activity is the drawing of bars with more-or-less complex cross sections, often to close tolerances, and in difficult-to-draw materials. Drawing stresses are somewhat higher, but the major concern is uniformity of material flow, which can be promoted with convex dies. Drawing of profiles presents some special lubrication difficulties; as material flow is less uniform, local stress concentrations arise at corners and grooves, and surface expansion can be much greater than in drawing of a round to a round. Lubricants may be evaluated by drawing profiles representing varying degrees of difficulty, and the order of merit of lubricants may change relative to wire drawing.

Roller Dies and Rolling. Many of the lubrication problems in wire drawing can be bypassed by using roller dies or by switching to cold rolling of sections. Rectangular sections with varying width-to-thickness ratios can be drawn through turks heads in which four rollers are mounted adjustably, with their axes in the same plane.

Roller dies are usually constructed with two or three rollers (Fig. 9) mounted in a frame with their axes in a single plane. A consideration of the three-dimensional geometry will show that the bar is fully enclosed only in the plane of the axes, and that gaps open farther back. This limits the reduction attainable without flash formation, and successive die sets must be set at a different angular position so that any incipient flash is rolled back into the body of the bar. Some spread is unavoidable, and the pass must be slightly opened to accommodate it. The repeated deformation places greater demands on workpiece ductility. However, sliding friction is replaced by the much lower rolling friction, and the tribological demands are much less stringent: the danger of die pickup and die wear are much reduced. Lubrication practices are similar to those of cold rolling.

From a mechanical view point, the process is equivalent to rolling with a high enough front tension to obviate the need for the rolls to be driven. This simplifies construction, but limits the attainable reduction. This limit is removed when the rolls are driven. Three-roll arrangements can become quite complex, but are feasible even for small-diameter wire (Ref 11). It is not actually necessary to have profiled rolls. As pointed out by Sayer and Moller (Ref 12), ductile wires can be reduced in a succession of flat rolls and then shaped in the last rolls or finished by drawing.

Tube Drawing

With four exceptions, the methods and equipment used for cold drawing tubes in straight lengths are basically identical to those used for bar drawing. The four exceptions are:

- Some tubes require more than one drawing pass.
- Tubes are usually longer than bars. Drawbenches for tubes are usually correspondingly longer, some permitting drawn lengths of more than 30 m (100 ft).
- Tube diameters are generally larger than bar diameters, ranging to about 305 mm (12 in.). The bigger tube drawbenches have larger components than do bar drawbenches.
- Tubes require internal mandrels or bars for simultaneous working or support of the interior surface during drawing. Tube drawbenches are usually equipped with one of several available devices, usually powered, for ready assembly of the cleaned, coated, and pointed workpiece onto internal bars or rod-supported mandrels. If rod-supported mandrels are used, they are usually air-operated so that the mandrel can be placed and maintained in the plane of the draw die after pulling starts. Butt or electric-welded tubes are sometimes drawn to smooth the weld seams and tube walls.

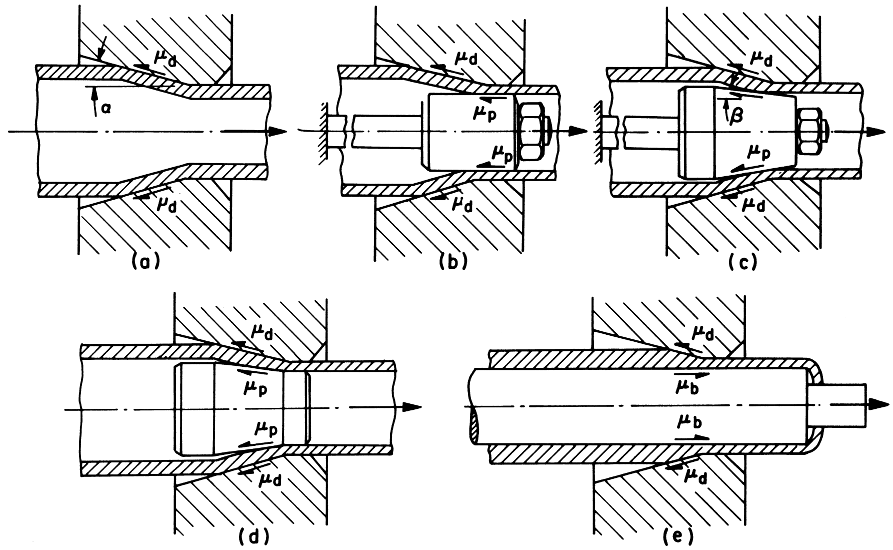


Fig. 10 Tube drawing processes. (a) Tube sinking. (b) Drawing with a fixed cylindrical plug. (c) Drawing with fixed conical plug. (d) Floating plug. (e) Drawing on a bar (or mandrel). Source: Ref 1, p 353

Tubes, particularly those having small diameters and requiring working only of their outer surfaces, are produced from cold-drawn coils on machines that straighten the stock and cut it to required lengths. As with bars, however, most tubes are produced from straight lengths rather than coiled stock. The surface finish of a tube blank can be improved, and its wall thickness and/or diameter reduced, by cold rolling on a pilger mill or, more frequently, by cold drawing.

In tube drawing both the diameter and wall thickness change. In calculating strain from Eq 1 and 2, A_o and A_1 now are cross-sectional areas of the tube before and after drawing, respectively. In all forms of tube drawing, inhomogeneity of deformation arises when h/L is high; here, h stands for mean wall thickness. As in bar drawing, most of the total work is transformed into heat, and the temperature of the tube rises.

Seamless tubes are mostly produced by one of the hot tube-rolling or tube-extrusion techniques. Only the Ehrhardt process utilizes a technique related to drawing. Hot, heavy-wall, closed-ended tube blanks are pushed through a series of roller dies in mechanical push benches to reduce the tube wall. For example, the Ehrhardt Push-Bench Method is effective in producing seamless steel pipes, as described Ref 13 for production of seamless pipe with outside diameters of 165 to 950 mm (6.5 to 37 in.) for power plants, nuclear power plants, chemical plants, and construction purposes.

Sinking. Figure 10 illustrates the basic types of tube drawing. Sinking (Fig. 10a) is closest to wire drawing in that the tube is drawn through a die without a mandrel (or plug). In tube sinking, the tube is initially pointed to facilitate feeding through the die; it is then reduced in outside diameter while the wall thickness and the tube length are increased. The magnitudes of thickness increase and tube elongation depends on the flow stress of the drawn part, die geometry, and interface friction. Its diameter is reduced while the wall slightly thickens and the inside surface roughens. Thickening of the wall is a function of friction and of the tube wall-thickness-to-diameter ratio (Fig. 11) (Ref 14). The effects of friction are the same as in bar drawing, but interface pressures are lower and optimal die angles are larger. Dévényi (Ref 15) observed optimal die angles of about 12° at 10% reduction and about 20° at 40% reduction in drawing aluminum tubes.

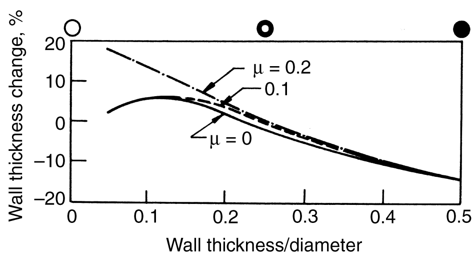


Fig. 11 Change in wall thickness in tube sinking with 15% reduction in diameter, in dies with 9° half angle. Source: Ref 14

Drawing with a fixed plug (Fig. 10b, c) is used for drawing large-to-medium diameter straight tubes. The plug, when pushed into the deformation zone, is pulled forward by the frictional force created by the sliding movement of the deforming tube. Therefore, the plug must be held in the correct position by a bar. The internal diameter and surface finish are controlled when drawing is done on a plug. In addition to the diameter (sink) being reduced, the wall thickness (draw) is also reduced to a dimension defined by the diameter of a cylindrical plug (Fig. 10b) or the position of a conical plug (Fig. 10c). The sink-to-draw ratio affects interface pressures, sliding velocities, and surface extension. The severity of operation increases with increasing draw. The additional friction, μ_p, on the plug increases the draw force. This effect can be allowed for in Eq 5 by replacing $\mu \cot \alpha$ with the term:

$$(\mu_d + \mu_p)/(\tan \alpha - \tan \beta) \qquad \text{(Eq 16)}$$

The optimal die half angle is usually 12 to 15° in cooperation with β at 10 to 11° for conical plugs (Fig. 10c).

Tube Drawing with Floating Plug. In drawing long and small-diameter tubes, the plug bar might stretch and even break. In such cases, it is advantageous to use a floating plug (Fig. 10d). This process can be used to draw any length of tubing by coiling the drawn tube at speeds as high as 10 m/s (2000 ft/min). If the plug is designed so that the frictional force keeps it in the deformation zone, the holding bar can be omitted. Such a floating plug allows drawing of long lengths on a drum (bull block), and heavier reductions can be taken. Plugs can be designed from basic principles if friction can be characterized. Here, too, there is an optimal angle.

Tube Drawing with Moving Mandrel. In drawing on a moving bar (also called a mandrel), the tube material is forced to slide over a bar (Fig. 10e). The mandrel travels at the identical speed at which the section exits the die. This process, also called ironing, is widely used for thinning the walls of drawn cups or shells in, for example, the production of beverage cans or artillery shells. A frictional stress is set up that transfers the drawing stresses from the tube wall to the bar (Fig. 7). Because friction on the bar opposes friction on the die, the stress in the tube wall can be approximated by substituting in Eq 5 for $\mu \cot \alpha$ the term:

$$(\mu_d - \mu_b)/\tan \alpha \qquad \text{(Eq 17)}$$

If $\mu_d < \mu_b$ (a quite likely circumstance), the stress imposed on the tube wall drops below that prevailing in frictionless drawing. Of course, the force necessary to overcome friction on the bar has to be provided. After drawing, the tube has to be reeled (cross rolled) off the bar, as in the Ehrhardt process.

Lubrication

Proper lubrication is essential in rod, tube, and wire drawing. In contrast to rolling, no friction is needed at all for wire drawing, tube sinking, and tube drawing on a fixed plug. However, some minimum friction is essential for drawing with a floating plug, and friction is helpful on the tube/bar interface in drawing on a bar. Thus, if at all possible, the lubricant is chosen to give lowest friction and minimum wear. It is essential, though, that the heat generated should be extracted, especially in high-speed drawing, because otherwise the lubricant can fail and the properties of the wire suffer.

Dry and Wet Drawing

These lubrication and cooling requirements can be satisfied by one of two techniques:

- Dry drawing, where the lubricant is chosen for its tribological attributes and the wire is cooled while it resides on the internally cooled draw drums (capstans)
- Wet drawing, where the lubricant is chosen both for its tribological attributes and for its cooling power

A transition between the two techniques is sometimes employed, particularly in low-speed drawing of bar and tube. A high-viscosity liquid or semisolid is applied to the workpiece and/or die. Drawing speeds range from 0.3 to 2 m/s (59 to 390 ft/min) for bar and heavy tube, through 7 m/s (1380 ft/min) for heavy (say, 5 mm, or 0.2 in.) wire, to 40 m/s (7870 ft/min) for fine steel wire and 70 m/s (13,780 ft/min) for nonferrous metals. The limit is often set by die life and coil length, which would force a die change or rethreading at unreasonably short intervals. Dove (Ref 16) provides additional details on the lubrication of ferrous wire.

In dry drawing, the lubricant is usually a dry soap powder, placed in a die box and picked up by the wire surface upon its passage through the box. This technique is used for steel wire larger than 0.5 to 1 mm (0.02 to 0.04 in.) in diameter, for which the relatively rough surface produced is acceptable. For the most severe draws and for tubes, the soap is often preapplied from a solution, if necessary, over a conversion coating; the soap must be allowed to dry. Additionally, external air cooling of the wire coil and water cooling of the die holder are possible. If water is applied to the wire at all, it must be totally removed before the wire enters the next die.

These techniques are customarily described as dry drawing (a process that has nothing to do with dry working in the sense of an unlubricated interface). With high-strength materials such as steels, stainless steels, and high-temperature alloys, the surface of the rod or wire can be coated either with a softer metal or with a conversion coating. Copper or tin can be chemically deposited on the surface of the metal. This thin layer of softer metal acts as a solid lubricant during drawing. Conversion coatings may consist of sulfate or oxalate coatings on the rod; these are then typically coated with soap, as a lubricant. Polymers are also used as solid lubricants, such as in the drawing of titanium.

In the case of steels, the rod to be drawn is first surface treated by pickling. This removes the surface scale that could lead to surface defects and therefore increases die life. For the dry drawing of ferrous metals, four types of nonslip continuous machines are in general use:

- Accumulating-type machines
- Double-block accumulating-type machines
- Controlled-speed machines
- Straight-through machines

An accumulating-type multiblock continuous wire-drawing machine is equipped with electromagnetic block clutches. Photocells sense high and low wire accumulation on each block and disengage or engage appropriate block clutches. Double-block accumulating machines have individually driven blocks. Wire is transferred from the first drawing block by means of an intermediate flyer sheave that reverses the direction of the wire (without twisting it) onto a coiling block mounted immediately above the first drawing block. The wire is then held temporarily in storage until demanded by the second drawing block.

On controlled-speed machines, the wire follows an essentially flat path from block to block with a constant, unvarying amount of wire storage without twisting and slipping. A tension arm between the blocks, activated by a loop of the wire being drawn, regulates the speed of the adjustable-speed direct-current motor on the preceding block. Straight-through machines, without tension arms, are also available. The spindles are often canted from the vertical axis to accommodate wire buildup on the blocks and to provide unimpeded, straight entry into the succeeding die; this is usually done when large-size workpieces are required. Skillful operators are necessary because torque adjustments may need to be altered at each block when stringing up the machines in order to make the electrical system function properly.

In wet drawing, the lubricant is chosen both for its tribological attributes and for its cooling power and can be either oil-based or aqueous. The lubricant/coolant can be applied to the die inlet, to the wire, and often also to the capstan, or the entire machine can be submerged in a bath. When the machine operates with slip, the lubricant must reduce wear of the capstan while maintaining some minimum friction. This wet-drawing practice is typical of all nonferrous metals and of steel wires less than 0.5 to 1 mm (0.02 to 0.04 in.) in diameter.

The continuous drawing of nonferrous rod and wire, as well as some intermediate and fine sizes of ferrous wire, is generally done on wet-drawing slip-type machines. On these machines, the surface speed of the capstans, except for the final (pullout) capstans, exceeds the speed of the wire being drawn, thus creating slip of the wire on the capstans. Brighter surface finishes are generally produced with these machines, but the machines are limited to smaller reductions per pass than with dry-drawing nonslip continuous machines.

With wet-drawing slip-type machines, the drawing operation is generally confined to an enclosed chamber, with the lubricant bathing the dies and wire as it is being drawn. These machines are less complicated electrically than nonslip machines, and only one drive system is employed. They are designed with either tandem or cone-type configurations, usually with horizontal spindles, but sometimes with a vertical spindle for the finishing capstan. Cone capstans have drawing surfaces (usually hardfaced) that are stepped outward to provide increasing peripheral speeds. This compensates for the elongation and increasing speed of the wire as it is reduced in diameter during drawing.

Plastohydrodynamic (Full-Fluid-Film) Lubrication

As in other processes, mixed-film lubrication is often encountered, but the high speeds of drawing—coupled with a favorable process geometry—make plastohydrodynamic (PHD) lubrication possible and even desirable, at least for early draws in which a rougher surface can be tolerated. In principle, a thick film can develop with any lubricant if other conditions are conducive to it. Under some very specific conditions, it is possible to maintain a continuous (or full-fluid) film between the die and workpiece in plastic deformation processes.

Early development of hydrodynamic lubrication to wire drawing is described in Ref 17 and 18. More detailed information on the phenome-

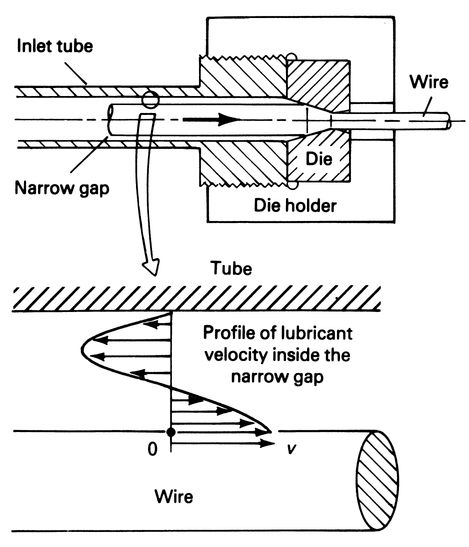


Fig. 12 Inlet tube for hydrodynamic lubrication. See Ref 5, p 62

non of hydrodynamic lubrication during flow through converging dies can be found in Ref 1, 4, 5, and 19. The analysis of hydrodynamic lubrication by Hillier (Ref 20) offers a simple and applicable treatment. In early studies of the equipment design for analysis of hydrodynamic lubrication, a long tube with a narrow gap between the tube and the wire was firmly attached at the entrance side of the die (Fig. 12). The lubrication adhered to the wire and was dragged into the clearance between the tube and wire. At about 3 m/s (590 ft/min), the pressure that built up at the approach to the die reached 70 to 275 MPa (10 to 40 ksi), and the liquid formed a film between the wire and the die (Ref 21).

Treatment of PHD lubrication has many similarities to elastohydrodynamic (EHD) lubrication, and much of the work came from research on EHD lubrication. In its simplest form (with the die and workpiece assumed to be rigid prior to entry in the deformation zone), the entrained film thickness (h_0) is:

$$h_0 = \frac{6\eta v}{\sigma_0 \tan\theta} \qquad \text{(Eq 18)}$$

where η is the viscosity, v is the mean surface speed, σ_0 is the flow strength of the workpiece material, and θ is the angle between the converging surfaces (Fig. 13a). The entrained film thickness increases with increasing viscosity (η) and surface speed (v) or with diminishing wedge angle θ ($=\alpha$ for a conical die) and with diminishing flow stress σ_f (or, more relevantly, with die pressure p_m, which is governed by reduction, die angle, friction, and back tension). The film thickness increases in steady-state processes if:

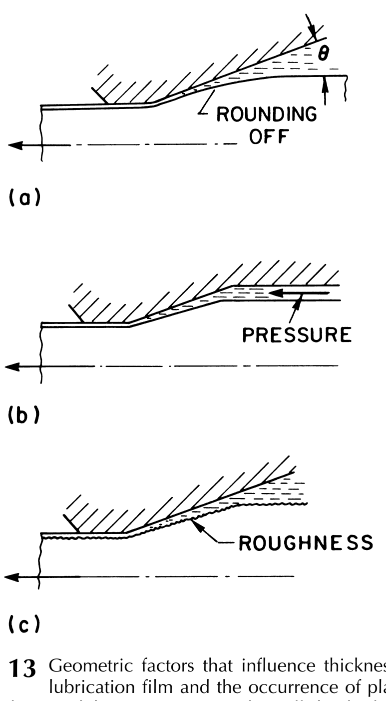


Fig. 13 Geometric factors that influence thickness of lubrication film and the occurrence of plastohydrodynamic lubrication. (a) Rounding off. (b) Fluid pressure. (c) Surface roughness. Source: Ref 1, p 69

- The angle θ or the entire entry geometry changes by deformation (rounding) of the workpiece just before entry into the die (Fig. 13a).
- The fluid pressure at the die entry is increased (as in wire drawing, hydrostatic extrusion, or deep drawing) (Fig. 13b).
- The angle θ is reduced by elastic deflection of the die.
- The workpiece surface roughness helps carry lubricant into the deformation zone (Fig. 13c).

Viscosity is also a function of pressure and temperature. The magnitude of the viscosity-pressure coefficient (φ) is an influential factor in determining film thickness, as follows:

$$h_0 = \frac{6\eta_0 \varphi v}{\tan\theta[1-\exp(-\varphi\sigma_f)]} \qquad \text{(Eq 19)}$$

where η_0 is the viscosity at atmospheric pressure.

In high-speed wet drawing, there is little doubt that film thickness increases with velocity, although this is not necessarily proof of a full-fluid-film mechanism. The hydrodynamic effect does not depend entirely on the speed, viscosity, or chamber/tube length. As noted, hydrodynamic lubrication can be accomplished at lower velocities and with lower fluid viscosities when higher pressure is applied. Pressurized chambers (Fig. 14) have been used, whereby the lubricating liquid is supplied to the chamber at high pressure by external means (Ref 22). There is experimental evidence to suggest that a full film can be generated with a high-viscosity lubricant in drawing of soft metals. Felder (Ref 23) observed thick (15 μm) films (as judged from weight differential) and no electrical contact in drawing soft aluminum and copper wires with a chlorinated oil of 200 Pa · s viscosity, at speeds from 0.5 to 5 m/s (98 to 980 ft/min). The presence of a bamboolike defect (a periodic variation of diameter) also suggested full-fluid-film lubrication on soft copper. However, only short pieces of wire were drawn, and it is not clear whether fluid-film conditions could have been maintained on long lengths of wire after attaining a thermal equilibrium.

Hydrodynamic Lubrication in Dry Drawing. In high-speed dry drawing, there is more opportunity for forming a fully separating film with a solidlike substance. Hydrodynamic lubrication through a pressure box is shown in Fig. 15. The pressure chamber is inserted into the soap bin of a conventional wire-drawing bull block. The bin preceding the entrance die to the chamber is filled with powdered soap. The wire running through the powder supply drags powder into the chamber through the narrow gap between the incoming wire and the approach die.

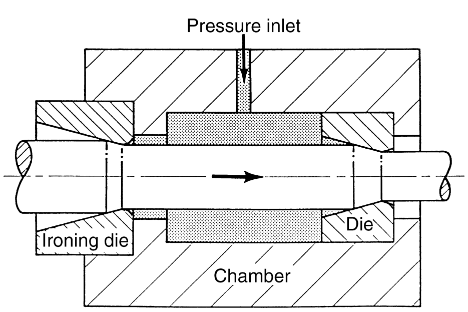


Fig. 14 Wire drawing through a pressurized chamber. Source: Ref 5, p 62

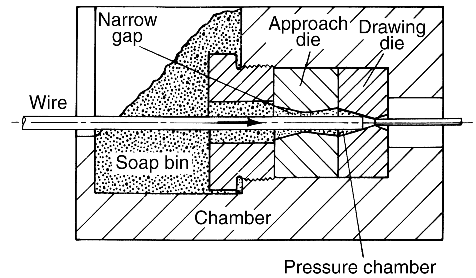


Fig. 15 Setup for hydrodynamic lubrication with dry soap. Source: Ref 5, p 62

When the wire is pulled through the soap, some soap is dragged into the chamber. In the small clearance gap, the soap powder shears, heats because of this shear, and then melts. Liquid lubricating soap then enters the pressure chamber. The faster the wire is drawn, the higher are the temperatures and pressures introduced in the chamber causing hydrodynamic lubrication. A chemical bond between the wire and the molten soap is also produced. As soon as the wire exits the die, its temperature drops and the soap freezes and forms a layer on the wire. There is now a layer of lubricant performing the duties of hydrodynamic lubrication, that is, keeping full separation between the deforming wire and the next five or six dies. The pressures produced in these chambers, without any external or auxiliary agent, are in excess of 275 MPa (40 ksi).

One difficulty associated with the use of a powdered-soap bin is in maintaining steady drag of powder into the chamber. Inconsistencies in the particle size and dryness of the powder vary the effectiveness of its adhesion to the wire and the quantity of powder dragged. At the extreme, a hollow channel in the powder surrounds the incoming wire and no soap is dragged in. Greater uniformity of the powder and constant agitation (directly or through the box) may improve the performance. Proper surface preparation also may be called for. A recent development utilizes the conventional powder spray system with electrostatically charged particles.

With phosphating, the most popular method, a predetermined layer of a spongy phosphate coating is applied to the surface of the wire electrolytically. This sponge can absorb and retain a large volume of lubrication liquid. Even without the lubricant, the phosphate sponge provides an effective separation between the die and the workpiece, thereby minimizing friction and wear.

Soaps in Hydrodynamic Lubrication. The rheological behavior of soaps is highly effective in allowing hydrodynamic lubrication of metal-

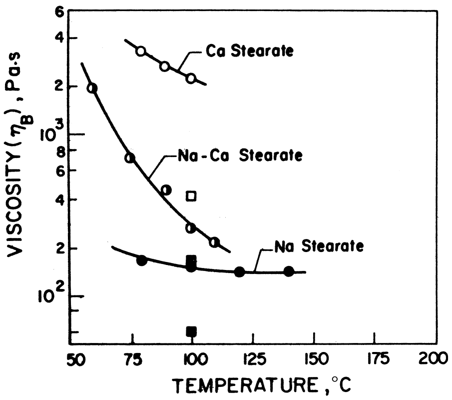


Fig. 16 Apparent viscosity of various soaps versus temperature. Source: Ref 24

drawing processes. The viscosity of soap is reduced at high temperature (Fig. 16) and a higher drawing speed is then required for full-film lubrication. The effect of speed is twofold; first it increases film thickness, but then heat generation causes a drop in viscosity. Thus, the industrial practice of using leaner (higher lime-content) soaps for heavier gages and harder alloy is justified.

In addition to lime (CaO), other thickeners are also incorporated. Borax and soda ash (Na_2O) are used instead of or in addition to lime. Iron sulfate replaces the hydrated iron oxides found on sull-coated wire, and imparts a brown color to the drawn wire. Graphite is added for a black finish, and TiO_2 (which also acts as a polishing agent) for white color. Mica and talc, even though of lamellar structure, act only as parting agents. Often, MoS_2 is added to a sodium soap, sometimes in company with graphite and lime, mostly for difficult-to-draw, strong alloys. Sulfur added to steel-drawing soaps acts as an extreme-pressure agent. Evidently, experience has demonstrated that film thickness is frequently insufficient to ensure complete die/workpiece separation. It is not clear whether or not such additives affect lubricant rheology to a significant degree. They can, however, interfere with film formation. Occasionally, an additive, such as graphite, can be used as a coating applied prior to drawing. On nonferrous wire, sodium and potassium soaps are used as coatings without carriers, but their rheology is most likely affected by the changes due to pressure, temperature, and reactions with the substrate.

In considering the rheology of soap, one must remember that the soap undergoes changes in service. Reactions with the wire, and especially reactions with oxides remaining after descaling and with borax, are also possible. Of course, soap viscosity and composition matter little if the wire coating and/or wire roughness are inadequate to carry soap into the die. The soap must be of controlled particle size; beaded (granulated) soaps help avoid tunneling and reduce dusting. Vibration of the die box or agitation of the soap is necessary when the wire runs free of vibration or whipping that would aid feeding. Melting point is important in controlling feeding. Excess soap picked up by the wire is rejected by the die; if it flows back in a molten form, the clumps thus formed deny further access of soap to the die. Sodium soaps are prone to caking because of their hygroscopic nature.

Tube Drawing. The possibility of increasing the film thickness in tube drawing has also attracted attention. In tube sinking, a pressure tube or multiple dies can offer benefits (Ref 25). In drawing on a plug, the critical problem is to increase lubricant supply to the plug/inner-tube interface. Pumping of the lubricant through a hollow plug is feasible, but in itself does not help to build up a lubricant film. For this, a double-plug arrangement, with supply of pressurized lubricant to the enclosed space, is needed. Some improvement in plug drawing is achieved simply by using a suitable plug profile, as reported by Rees (Ref 26). The oscillating floating plug (Fig. 17a) is designed to trap lubricant; the pressure in the lubricant is assumed to force the shoulder back against the spring, creating an oscillating-pumping action. The adjustable floating plug (Fig. 17) creates a lubricant cavity and combines the advantages of straight and tapered plugs. The fluid pressure generated in the cavity prevented galling in drawing of a limited number of stainless steel tubes (Ref 24).

Mixed-Film Lubrication

Mixed-film lubrication occurs when there is an interruption in the conditions for the continuous (full-film) lubrication. Average film thickness has little meaning under these conditions. Nevertheless, it is a useful concept for visualizing process conditions. As indicated by Eq 18, film thickness increases with decreasing entry angle. Indeed, it has been repeatedly observed that lubricant flow increases greatly with very low (about 2°) half angles. Even in low-speed experiments, a transition from thick-film to mixed-film lubrication may be observed on increasing α to some critical value.

Die geometry is important in industrial practice. Schmidt (Ref 28) notes that in drawing of steel wire 5 or 6 mm (0.2 or 0.24 in.) in diameter at speeds greater than 8 m/s (1580 ft/min), the half angle must be kept at about 6° instead of at the previously used 8 to 10°; obviously, the lower angle is needed to maintain a sufficiently thick film at the higher temperatures generated by the higher speeds. For this reason, cooling must be improved as well. Land length is also important. An excessively long land strips the wire of its soap coating, whereas a short land wears rapidly. Schmidt (Ref 28) suggests that, for steel wire 2.0 to 6.5 mm in diameter, land length (in mm) = $5/v$, where v is drawing speed in m/s. A similar relation may well hold for other materials and lubricants.

Under mixed-film conditions, friction and wear may be difficult to characterize. There is no single value of coefficient of friction μ or other descriptor can possibly characterize the conditions existing for a wide range of speeds, geometries, and reductions. As Schey notes in Ref 1 (p 372): "It would be futile to compile lists of μ values; in addition to the effects of process variables, μ values are influenced also by the method of their determination (plane-strain drawing, split die, strain-gaged die) and by the method of calculation from measured draw force. Furthermore, many laboratory measurements are conducted at very low speeds and sometimes on lengths as short as 25 mm, and thus the results are quite unrepresentative of production practices." For predictive calculations, the values listed in Table 1 are suggested by Schey.

Solid-Film Lubrication

This section describes lubricants that do not need a hydrodynamic wedge action to develop a film.

Layer-Lattice Compounds. Graphite and MoS_2 are seldom used on their own, except for high-temperature work. Their friction is found to be independent of speed (Ref 29). The effect of pressure is controversial. With a truly continuous, well-developed film, μ should be constant. In sheet drawing, μ has been observed to drop or rise with pressure in different situations. It is conceivable that in some instances the film was burnished and improved, while it may have lost its continuity in other tests. Generally, μ rises

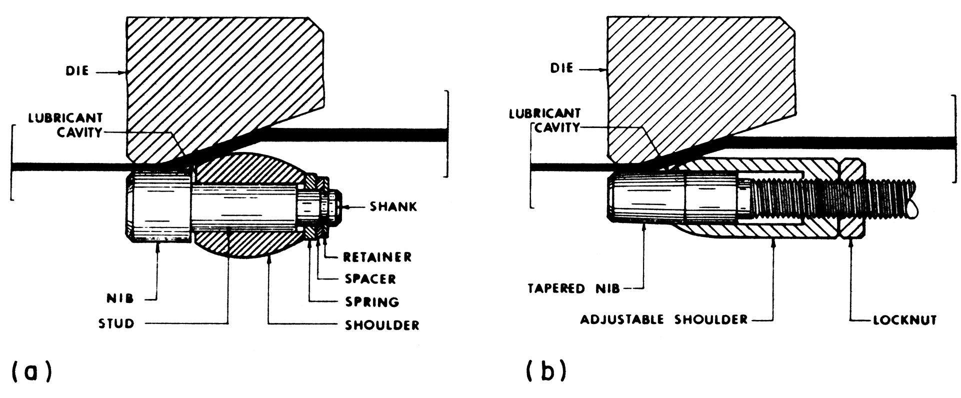


Fig. 17 Profiles of floating plugs with full-film lubrication. (a) Oscillating floating plug. (b) Adjustable floating plug. Source: Ref 27

Table 1 Commonly used drawing lubricants and typical μ values (lubricants listed according to increasing severity of conditions)

Material	Lubricant(a)	μ(b)
Wiredrawing		
Steels	Over 1 mm: dry (Ca-Na) soap on lime or borax	MF
	Over 1 mm: phosphate + soap	MF
	Under 1 mm: EM (MO + fat + EP)	0.07
	Under 1 mm: phosphate + EM	0.1
	Metal (Cu, Zn, brass) + EM	MF
Stainless steels and Ni alloys	EM (MO + Cl) (on lime)	0.1
	MO + Cl additive (on lime)	0.07
	Chlorinated paraffin, wax	0.05
	Oxalate + soap	0.05
	Metal (Cu) + EM (or oil)	MF
Al and Mg alloys	MO + fatty derivatives	MF
	Synthetic MO + fatty derivatives	MF
Cu and Cu alloys	EM (MO + fat) (+ EP)	MF
	Metal (Sn) + EM or MO	MF
Ti alloys	Oxidized + Cl oil (wax)	0.15
	Fluoride-phosphate + soap	0.1
	Metal (Cu or Zn) + soap or MO	0.07
Refractory metals	Hot: graphite coating	0.15
	warm (cold): graphite or MoS_2	0.1
	Oxidized + wax	0.15
	Metal (Cu) + MO	0.1
Bar and tube drawing		
Steels	Heavy oil, soap-fat paste, grease	MF
	Heavy oil, soap-fat, paste, grease (+ EP) (+ MoS_2, etc.)	MF
	Polymer coating + EP oil	0.07
	Phosphate + soap	0.05
	Metal + MO (or EM)	MF
Stainless steels and Ni alloys	MO + Cl additive	0.15
	Chlorinated wax	0.07
	Polymer (chlorinated) (+ MO)	0.07
	Oxalate + soap	0.05
	Metal (Cu) + MO	MF
Al and Mg alloys	MO + fatty derivatives	MF
	Soap coating	0.07
	Wax (lanolin) coating	0.05
	Polymer coating	0.05
Cu and Cu alloys	EM (fat)	MF
	MO (+ fat) (+ EP)	MF
	Soap film	0.05
Ti alloys	Polymer coating	0.07
	Oxidized + Cl oil (wax)	0.15
	Fluoride-phosphate + soap	0.1
	Metal + soap	0.07
Refractory metals	Hot: graphite coating	0.15
	Warm (cold): graphite or MoS_2	0.1
	Oxidized + wax	0.15
	Metal (Cu) + MO	0.1

(a) EM, emulsion of ingredients shown in parentheses. MO, mineral oil; of higher viscosity for more severe duties, limited by staining. EP, EP compounds (S, Cl, and/or P). (b) MF, mixed-film lubrication; μ = 0.15 at low speeds, dropping to 0.3 at high speeds. Source: Ref 1, p 373

with reduction, most likely because surface extension exposes more asperities.

Conversion Coatings. By definition, conversion coating involves the reaction of the surface atomic layers of the workpiece material with the anions of a selected medium. It is similar to extreme-pressure (EP) lubrication in that a controlled corrosion process produces a firmly bonded surface layer. The difference is that a conversion coating does not lubricate, but acts only as a parting agent and lubricant carrier. A superimposed and sometimes reacted lubricant layer is necessary to complete the lubrication system.

Conversion coatings are indispensable for tube drawing, low-speed bar drawing, and wire drawing with heavy reductions on materials that are otherwise difficult to lubricate. Coating types include sull coating, oxide coatings, sulfuration, chromating, and most significantly phosphating. Phosphating was originally developed for steel, but modifications for other metals are also available:

- Zinc and its alloys can be phosphated in solutions enriched with iron.
- Aluminum is phosphated mostly in proprietary solution containing chromates and fluorides in addition to phosphates.
- Titanium can be treated in a solution containing sodium orthophosphate, potassium fluoride, and hydrochloric acid.

Steels containing more than 5% Cr cannot be phosphated. Therefore, stainless steels, chromium-molybdenum steels, nickel-base superalloys, and other high-chromium alloys are oxalated.

Polymer Coatings. Deposition or bonding of polymer films provides lubrication in a wide variety of bearing or wear applications. In metalworking applications, only thermoplastic polymers with appropriate glass-transition temperature (T_g) are of interest. The film thickness of a solid coating deposited from a solvent can be controlled quite independently of process conditions; this is a great advantage in low-speed bar and tube drawing. The friction of coatings is independent of speed as long as no major temperature rise is encountered. Polytetrafluoroethylene (PTFE) has often been investigated, but other polymers, closely related to the coatings used on sheet have found use. Numerous polymers are useful in bearing applications, but for metalworking purposes only thermoplastic polymers of the appropriate glass-transition temperature are of interest.

Polyethylene (T_g = −120 °C, or −185 °F; melting temperature, or T_m = 137 °C, or 280 °F) is readily available in sheet form and possesses substantial elongation (100 to 300%). In combination with mineral oil, sheets 2.5 to 7.5 μm thick provide excellent lubrication. Polypropylene (T_g = −18 °C, or −0.5 °F; T_m = 176 °C, or 350 °F) has similar properties.

Polytetrafluoroethylene (T_g = −50 °C, or −60 °F; T_m = 327 °C, or 620 °F) sheets have limited ductility, but well-adhering films of controlled thickness and good extendability can be deposited from a tricholorotrifluorethylene dispersion of telomers (which have aliphatic end chains grafted onto the PTFE). Sputtering is also possible.

Polyvinylchloride (T_g = 87 °C, or 190 °F; T_m = 212 °C, or 415 °F) is available both in sheet form and as a deposited coating.

Polymethylmethacrylate (PMMA, T_g = 105 °C or 220 °F; T_m = 240 °C, or 465 °F) and acrylics in general have the advantage that they are readily deposited from either solvent solutions or emulsions, and they can be modified so as to endow them with boundary-lubrication properties.

Acetal resins contain the linkage $–CH_2O–CH_2O–$ and have found some limited application.

Polyimides formed with aromatic radicals have outstanding temperature stability.

Metal coatings are among the most powerful aids to lubrication. Even though thin metal coatings can provide lubrication by reducing adhesion or friction, they are almost always used in conjunction with a liquid or semisolid lubricant. Thus, they serve more as a lubricant carrier, with the lubricant chosen for the coating rather than the substrate metal. The metal film may improve performance of another (usually liquid) lubricant either by providing better entrapment through favorable surface configuration or by allowing chemical reactions. The film must be thick enough to ensure continuous coating, and the metal film can only be effective if it is well bonded to the substrate.

Surface Preparation

Any surface oxide or dirt carried into the die greatly accelerates die wear and lubricant deterioration. Therefore, the first step for successful lubrication is preparation of the surface. Both mechanical and chemical techniques are used. In drawing steel wire, blast cleaning with cast iron or cut-wire shot is used. Steel wires also can be descaled with the scale itself. Much but not all of the scale can be removed by reverse bending, provided that the sum of the wrap angles exceeds 180° or that a total elongation of 5 to 10% is achieved. Elongation of the wire reduces the scale residue to some 0.6%. Normally, the residue must be removed by pickling, but Bernot (Ref 27) found wet blasting (with scale in water) to be an ecologically and economically attractive alternative. Abrasive belts have been used on heat treated rod. Continuing efforts are made to increase the efficiency of pickling, not only by control of chemical composition but also by application of ultrasonics. Recycling of spent liquors is increasingly practiced. Even though these examples refer to steel, similar problems exist, to varying degrees, with other metals.

Prevention of oxidation eliminates the need for descaling. Steel wire rod cooled by water or nitrogen has a thin scale (4.5 μm max) consisting of the less-abrasive FeO and could be drawn, at least experimentally, without descaling, in a lubricant consisting of a stearate soap with 20% inorganic and 12% EP additives (Ref 30). Copper annealed in a protective atmosphere can be drawn without descaling.

Wear

One can compensate for wear in rolling by resetting the roll gap. No such opportunity exists in drawing through a die. Thus, wear results in an increase in the dimension of the product, and the tolerances set an absolute limit on the allowable wear. The geometry of the converging hole also changes, and thus the lubrication mechanism is affected, lubricant breakdown can occur, and the product finish suffers. Therefore, minimization of wear is a principal preoccupation of tribology in drawing.

In a typical die, wear occurs at three major points (Fig. 18):

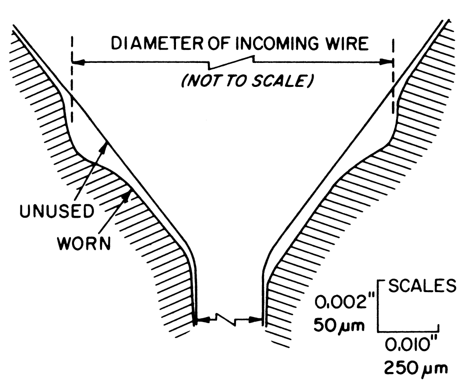


Fig. 18 Section of typical worn wire-drawing die. Source: Ref 31

- Severe wear at the point where the wire enters the die results in so-called ringing. The change in entry geometry changes the film thickness and thus the surface finish of the product.
- Wear of the draw cone changes the geometry of operation and affects film thickness.
- Wear of the die land is directly responsible for loss of tolerances and also affects the surface finish of the product. If wear of a wire die is nonuniform, the wire acquires a noncircular profile.

Wear Mechanisms

Deformation would proceed without wear, but only if the die and workpiece were separated by a thick film of nonreactive lubricant without any foreign particles. However, PHD lubrication is seldom achieved; lubrication is mostly of the mixed-film or boundary type. Thus, reactions with surface typically occur and result in progressive loss (wear) of the interface materials. Mild wear of the workpiece material may be considered normal and acceptable. Wear of the die, or severe wear of the workpiece—perhaps coupled with scoring of the workpiece surface—is obviously undesirable and must be prevented with proper lubrication.

Wear generally follows Eq 20 and increases linearly with the length (or weight) of wire drawn. Under steady-state conditions, the wear volume (V) inversely proportional to the materials hardness (H) and is proportional to the distance (l) or time of sliding with a normal load (P) such that:

$$\frac{V}{l} = K\frac{P}{H} \qquad \text{(Eq 20)}$$

where K is some constant of the wear situation. Compilations of the K coefficient are given by Rowe (Ref 32) and Rabinowicz (Ref 33), but minor variation in the interface conditions can have a much more marked effect on the wear rate than on friction and the K (for ostensibly identical conditions). The chief difficulty of wear theories is the interpretation and prediction of K for various wear situations and mechanisms.

In general terms, three general types of wear situations are common, as noted in Ref 34. The first occurs when two solid bodies are in contact and move relative to one another. This first situation is generally subdivided by the predominant nature of the relative motion, such as sliding, rolling, or impact. The second situation occurs when the wear is caused by a liquid moving relative to a solid surface. Wear in that situation is often called erosion or erosive wear. The third situation, which is generally referred to as abrasive wear, occurs when the wear is caused by hard particles. Erosion and abrasive wear situations can also be subdivided into more specific categories. Examples of these are cavitation erosion, solid-particle erosion, gouging abrasion, and slurry erosion.

The actual mechanisms of wear are complex and may vary in extent depending on the nature of the wear situation. For the different types of wear situations, there are four general ways (or types of mechanisms) by which material wear occurs:

- Adhesive processes
- Abrasive or deformation processes
- Fatigue or fatiguelike processes
- Oxidative or corrosive processes

With adhesive processes, wear occurs as a result of the bonding that takes place between two surfaces in contact. With subsequent separation of the two surfaces, material from either surface may be pulled out, resulting in wear. Adhesive wear processes normally occur only with sliding. However, adhesive wear processes can occur under nominal impact and rolling conditions, because of the slip that is often present in those situations (Ref 34).

Abrasive wear processes are those fracture, cutting, and plastic deformation processes that can occur when a harder surface engages a softer surface. These mechanisms tend to produce machining-chip-like debris. Fatigue or fatiguelike wear processes are those associated with crack initiation and propagation or progressive deformation as a result of repeated contact, such as a ratchet mechanism. Corrosive wear processes are those associated with the loss of wear of in situ formed reaction product (e.g., oxide layers).

The relative wear rates of these mechanisms can differ by several orders of magnitude, depending on the wear situation. Table 2 shows the observed range of a dimensionless wear coefficient for each of these mechanisms under sliding conditions. This coefficient is obtained by dividing the amount of wear by the distance of sliding and the load and multiplying by the hardness of the wearing surface. The value for each of the mechanisms is based on data obtained from tribosystems in which that mechanism is considered to be the predominant one. The higher the value, the more severe the wear situation. Generally, engineering applications require that this coefficient be in the range of 10^{-5} or less. In some cases, very small values are required. Table 2 shows that the most severe forms of wear are associated with adhesive and abrasive wear mechanisms.

All wear mechanisms can be active in a given situation, and the foregoing mechanisms are not mutually exclusive. They can coexist and interact to form more complex wear processes, and when worn surfaces are examined, features indicative of more than one mechanism are usually found. However, in most cases, one type of mechanism

Table 2 Sliding-wear coefficients

Wear mechanism	Sliding wear coefficient(a)
Adhesive wear	10^{-7} to 10^{-1}
Abrasive wear	10^{-6} to 10^{-1}
Corrosive war	10^{-7} to 10^{-2}
Fatigue wear	$\leq 10^{-6}$

(a) $(V \times H)/(P \times S)$, where V is the volume of wear, H is the hardness (load per unit area), P is the applied load, and S is the sliding distance. Source: Ref 34

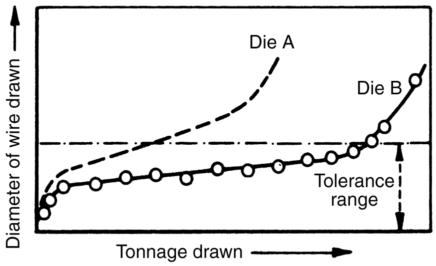


Fig. 19 Effect of wear rate on die life for a given tolerance range. Source: Ref 36

tends to predominate and ultimately be the controlling one. In wire drawing, wear mechanisms have been studied extensively by Wistreich (Ref 31), and the applications of wear theory have been reviewed by Shatynski and Wright (Ref 35). There is a phase of early rapid wear (Fig. 19), which Papsdorf (Ref 36) attributed to removal of the asperities in a new die, and an acceleration in the last stage, which is associated with heavy ringing.

In general, the following wear mechanisms are most active:

- *Abrasive wear is unavoidable.* Most metals carry a harder oxide; even if the heavy oxide inherited from hot working is removed, fresh oxide always forms. Some lubricants and coatings are mildly abrasive; others contain abrasive constituents (e.g., SiO_2 or Al_2O_3 in lime). Wear debris in wet drawing, unless continuously removed, creates a three-body wear condition. In a simulation test with a copper disk rotating against a diamond flat, Harper et al. (Ref 37) found wear to double when a new lubricant was recirculated and to increase sevenfold when Cu_2O particles were added. A used lubricant, in which fines were entrapped in the soap-fat emulsion, gave similarly high rates of wear.
- *Adhesive wear is likely* whenever metal-to-die contact occurs, although the severity of wear is highly dependent on the specific workpiece-die combination.
- *Surface fatigue wear* has been suspected as the main culprit in ringing. The high stress gradient, even if loading is continuous, is damaging in itself and is aggravated by vibration of the wire. It has also been found that entering of the wire at an angle greatly increases wear rates; therefore, accurate alignment and elimination of vibration by guide pulleys reduce wear.
- *Thermal fatigue wear* can lead to crazing of steel dies but is, generally, not a major problem.

Catastrophic failure (fracture) of dies is usually due to overloading. The danger is greater with lower die angles, and brittle, hard dies must be encased in steel rings.

Die Wear

Because die wear is proportional only to the length of wire drawn and is not related to wire diameter, a given amount of absolute wear represents a greater percentage increase for a thinner wire. This means that more wear-resistant die materials must be found for finer wires if tolerances are to be held. Wear rates are expressed as the length of wire that can be drawn before the increase in wire diameter reaches 1 μm. Typical values are: 1 to 2 km (0.6 miles) for low-carbon steel with die steel dies, 10 km (6 miles) for patented wire, and up to 500 km (310 miles) for low-carbon steel wire with tungsten carbide dies. Lives are 10 to 200 times longer for diamond dies (e.g., 1.6×10^5 km, or 100,000 miles, of copper wire may be drawn before repolishing). All die materials have their places; discussions of various types are given in the steel-wire handbooks.

Die life is also increased by minimizing the presence of abrasive oxides and removing debris from the lubricant. The surface finish obtained on finish polishing is equally important and must be measured. Early polishing is important (particularly for the prevention of incipient ringing), and checking the die profile is needed to ensure lubricant film formation. Cast silicone rubber replicas allow inspection without special equipment. A drastic improvement is obtained by the use of roller dies—an improvement that is equivalent to a tenfold to 100-fold increase in die life. Wear often begins as an angular ring on the approach angle of the die. Die life may be increased by as much as 200% if the die is removed and repolished at the first appearance of this ring (Ref 38).

Table 3 (Ref 38) gives recommended materials for wire-drawing dies. For round wire, dies made of diamond or cemented carbides are always recommended without regard to composition or quantity of the metal being drawn. For short runs or special shapes, hardened tool steel is less costly, although cemented carbides give superior performance in most applications. Die steels are still used for sections and large-diameter tubes and bars, as they are for roller dies and rolls. They could, no doubt, be improved with coatings, but some of the incentive for doing so has disappeared with the ready availability of large carbide dies. Hard-chromium plating is commonly used for plugs. Boronizing has been reported to give better results. Ceramic coatings, particularly ZrO_2, look promising for drawing of stainless steel. Pure zirconia plugs have also been used.

Conclusions

No friction between workpiece and die is needed in wire drawing and tube drawing on a fixed plug; some moderate friction is required on the plug in drawing tubes with a floating plug, and friction on the bar is beneficial in drawing on a bar.

The frictional force is additional to the force required to perform the deformation (work of homogeneous deformation) and the force arising from the inhomogeneity of deformation (redundant work). The frictional force increases, whereas the force for inhomogeneous deformation decreases, with increasing die angle, and therefore an optimal die angle is found. Because the contributions of friction and inhomogeneous deformation to the total force are of similar magnitude, back-calculation of friction from the total draw force is difficult. However, for a given workpiece material and process geometry, a larger draw force is always indicative of higher friction.

Yielding takes place under the combined effects of axial tension and radial compression, and interface pressures are below the flow stress of the material except at low die angles, when they can rise to higher values. Even though pressures are, normally, not excessive, the process

Table 3 Recommended materials for wiredrawing dies

	Wear size		Recommended die materials for:	
Metal to be drawn	mm	in.	Round wire	Special shapes
Carbon and alloy steels	<1.57	<0.062	Diamond, natural or synthetic	M2 or cemented tungsten carbide
	>1.57	>0.062	Cemented tungsten carbide	
Stainless steels; titanium, tungsten, molybdenum and nickel alloys	<1.57	<0.062	Diamond, natural or synthetic	M2 or cemented tungsten carbide
	>1.57	>0.062	Cemented tungsten carbide	
Copper	<2.06	<0.081	Diamond, natural or synthetic	D2 or cemented tungsten carbide
	>2.06	>0.081	Cemented tungsten carbide	
Copper alloys and aluminum alloys	<2.5	<0.100	Diamond, natural or synthetic	D2 or cemented tungsten carbide
	>2.5	>0.100	Cemented tungsten carbide	
Magnesium alloys	<2.06	<0.081	Diamond, natural or synthetic	
	>2.06	>0.081	Cemented tungsten carbide	

Source: Ref 39

presents severe tribological conditions because sliding prevails over the whole contact zone, and heating due to deformation and friction raises the temperature. In multihole drawing the high temperature of the entering wire aggravates the situation.

Lubricants are applied to reduce friction, wear, and temperature. A good lubricant reduces temperature by reducing friction. Nevertheless, in high-speed multihole drawing the lubricant must also have adequate cooling capacity to remove heat from the product. The method of coolant application becomes most important. Wear is brought within tolerable limits, in part, by the selection of highly wear-resistant die materials, primarily tungsten carbide and diamond.

Drawing without a lubricant would result in immediate pickup. So-called dry drawing is conducted with a nonliquid lubricant, usually a soap, whereas wet drawing is performed with viscous oils or aqueous emulsions. The distinction is practical rather than fundamental: under the conditions prevailing at the interface, soap can be modeled as a viscous fluid of temperature- and strain-rate-sensitive viscosity, or as a Bingham substance.

Drawing is a steady-state process with a favorable entry geometry for the entrainment of lubricants. Full-fluid-film lubrication can be encouraged by special die constructions that aid in developing hydrodynamic films or that allow the introduction of externally pressurized fluids. The most practicable of these is the pressure die for soaps; it builds up sufficient pressure by purely hydrodynamic action to cause yielding of the wire prior to entry into the die. The surface finish of the workpiece produced under full-fluid-film conditions is rough.

Most practical drawing is conducted under mixed-film conditions. Initial draws may be made with a dry lubricant, but brighter, smoother surfaces can be produced only by deliberately thinning the film to the point where boundary lubrication predominates. This is accomplished by increasing the die angle (although this increases also the inhomogeneity of deformation and could lead to internal arrowhead fracture) and by drawing with a low-viscosity lubricant. The potential for die pickup necessitates the use of boundary or EP additives.

Solid film lubrication is of greatest importance in warm and hot drawing and in cold drawing of adhesion-prone alloys at low speeds, in which an effective fluid film cannot be maintained.

Tube drawing on a fixed or floating plug presents special difficulties of lubricant application to ensure effective lubrication of the inner tube surface.

Many problems of lubrication and wear can be bypassed by replacing the draw die with a rotary die composed of two to four rollers, particularly in drawing of more ductile materials.

Lubricant recommendations are summarized in Table 1. For preliminary calculations, coefficient of friction values are given with the understanding that their magnitude is a function of drawing speed, especially for lubricants operating in the mixed-film regime.

ACKNOWLEDGMENT

This article was adapted from substantial content contained in Ref 1 and 2. Additional content was adapted from "Wire, Rod, and Tube Drawing," *Forming and Forging*, Volume 14, *ASM Handbook*, 1988, p 330–342.

REFERENCES

1. J.A. Schey, *Tribology in Metalworking, Friction, Lubrication, and Wear*, ASM International, 1983
2. Drawing of Rod and Wire, R.N. Wright, *Encyclopedia of Materials Science and Engineering*, Vol 2, M.B. Bever, Ed., Pergamon Press and the MIT Press, 1986, p 1227–1231
3. T.Z. Blazynski, *Metal Forming, Tool Profiles, and Flow*, Halstead Press, 1976
4. B. Avitzur, *Metal Forming: Process and Analysis*, McGraw-Hill, 1968, Krieger, revised 1979
5. B. Avitzur, Friction During Metal Forming, *Friction, Lubrication, and Wear Technology*, Vol 18, *ASM Handbook*, 1992, p 59–69
6. B.B. Hundy and A.R.E. Singer, *J. Inst. Met.*, Vol 91, 1962–1963, p 401–407
7. M. Mohamdein and T. Vinh, *CIRP*, Vol 25, 1976, p 169–172
8. B. Avitzur, *Wire J.*, Vol 7 (No. 11), 1974, p 77–86
9. R.W. Johnson and G.W. Rowe, *Proc. Inst. Mech. Eng.*, Vol 182, 1967–1968, p 521–526
10. R.W. Gottschlich and N.N. Breyer, *J. Met.*, Vol 15, 1963, p 364–367
11. G. Properzi, *Wire J.*, Vol 12 (No. 12), 1979, p 58–62
12. R. Sayer and R. Moller, *Wire J.*, Vol 11 (No. 2), 1978, p 68–72
13. N. Koichiro, M. Tawara, Y. Hayase, T. Katsube, M. Kyoda, and H. Nakagawa, Effective Production in the Erhardt Push Bench Pipe Manufacturing, *Tetsu-to-Hagane* (in Japanese); *J. Iron Steel Inst. Jpn.*, Vol 78 (No. 11), Nov 1992, p T209–T212
14. O. Pawelski and V. Rüdiger, *Arch. Eisenhüttenwes.*, Vol 47, 1976, p 483–487 (in German)
15. G. Dévényi, *Draht*, Vol 13, 1962, p 223–231 (in German)
16. A.B. Dove, *Steel Wire Handbook*, The Wire Association International, 1980
17. A.F. Gerds and F.W. Boulger, "Rod, Wire and Tube Drawing," DMIC Report 226, Metal Deformation Processing, Vol II, Defense Metals Information Center, Battelle Memorial Institute, 7 July 1966, p 78–124
18. *Recent Progress in Metalworking*, American Elsevier, 1964
19. B. Avitzur, *Handbook of Metalforming Processes*, John Wiley, 1983
20. M.J. Hillier, A Hydrodynamic Lubrication in Wire Drawing, *Int. J. Prod. Res.*, Vol 5, 1967, p 171
21. G.H. Tattersall, "Theory of Hydrodynamic Lubrication in Wire Drawing," Report MW/D/46/59, British Iron and Steel Research Association
22. A. Bobrowsky, "Pressure Box," U.S. patent 3,417,589, 24 Dec 1968
23. E. Felder, Elastohydrodynamics and Related Topics, *Proc. Fifth Leeds-Lyon Symposium on Tribology*, D. Dowson et al., Ed., Mechanical Engineering Publications, London, 1979, p 365–369
24. O. Pawelski, W. Rasp, and T. Hirouchi, in *Tribologie*, Springer, Berlin, 1981, p 479–506 (in German)
25. V.L. Kolmogorov, S.I. Orlov, and K.P. Selischev, *Drawing under Conditions of Hydrodynamic Lubrication*, National Lending Library for Science and Technology, Boston Spa, Yorkshire, England, 1968
26. T.W. Rees, *J. Appl. Metalwork.*, Vol 1 (No. 4), 1981, p 53–57
27. J. Bernot, *Wire J.*, Vol 12 (No. 12), 1979, p 8–81
28. W. Schmidt, *Stahl Eisen*, Vol 91, 1971, p 1374–1381 (in German)
29. H. Kudo, S. Tanaka, K. Imamura, and K. Suzuki, *CIRP*, Vol 25, 1976, p 179–184
30. O. Miki, S. Fuji, and M. Vemura, *Wire J.*, Vol 12 (No. 5), 1979, p 56–60
31. J.G. Wistreich, *Met. Rev.*, Vol 3, 1958, p 97–142
32. C.N. Rowe, *Wear Control Handbook*, M.B. Peterson and W.O. Winer, Ed., American Society of Mechanical Engineers, 1980, p 143–160
33. E. Rabinowicz, *Wear Control Handbook*, M.B. Peterson and W.O. Winer, Ed., American Society of Mechanical Engineers, 1980, p 475–506
34. R.G. Bayer, Design for Wear Resistance, *Materials Selection and Design*, Vol 20, *ASM Handbook*, 1997, p 603–614
35. S.R. Shatynski and R.N. Wright, *Wire Technol.*, Vol 7 (No. 4), 1979, p 59–62, 66–69
36. W. Papsdorf, *Stahl Eisen,* Vol 72, 1952, p 393–399 (in German)
37. S. Harper, A.R. Goreham, and A.A. Marks, *Met. Mater.*, Vol 4, 1970, p 335–339
38. The Selection of Materials for Tools for Drawing, Wire, Bar, and Tubing, *Properties and Selection of Metals*, Vol 1, 8th ed., *Metals Handbook,* 1961, p 752
39. The Selection of Materials for Tools for Drawing, Wire, Bar, and Tubing, *Properties and Selection: Stainless Steels, Tool Materials, and Special-Purpose Metals*, Vol 3, 9th ed., *Metals Handbook*, 1980, p 522

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 291-315
DOI:10.1361/hwpd2003p291

Chapter 20

Extrusion*

EXTRUSION is a deformation process used to produce long, straight, semifinished metal products such as a bar, solid and hollow sections, tubes, wires, and strips. Practically all metals can be extruded, although the extrudability of many alloys (including many extrudable alloys of aluminum and magnesium) falls within a narrow range of processing conditions. Extrudability depends on a variety of application-specific factors, such as:

- Mechanical workability (W), or resistance to cracking with plastic deformation, defined in broad terms as: $W \approx \varphi_f/\sigma_0$ (where φ_f is deformation to fracture, and σ_0 is flow stress)
- Extrusion temperature
- Permitted temperature range
- Load or specific pressure required (which depends on the work material and geometry of the extruded section)
- Exit speed at a constant load
- Maximum extrusion speed (to the onset of hot shortness)
- Maximum extrusion ratio (R_E), which is the ratio of the initial cross-sectional area (A_0) to the final cross-sectional area (A_1) after extrusion ($R_E = A_0/A_1$)

Special techniques to extrude different materials are, to a large extent, dependent on the extrusion temperature. Cold (room-temperature) extrusion is economically feasible for low-melting alloys and some aluminum alloys, while hot extrusion is required for harder alloys. Figure 1 (Ref 1) is a general classification of alloys in terms of extrusion temperature.

Metal flow from extrusion through a conical die is analogous to drawing (described in Chapter 19, "Drawing of Wire, Rod, and Tube," in this Handbook), although there are some important differences in these processes. One basic difference is that the workpiece during drawing is pulled through a die, while during extrusion, the workpiece is either pushed through a die (direct extrusion) or a die is pushed over a stationary workpiece (indirect extrusion). The favorable compressive state of stress during extrusion allows a high capacity for deformation, and so it is possible to extrude metals that can be only slightly deformed by other methods. This is shown schematically in Fig. 2, where the deformation to fracture (φ_f) is plotted against the ratio of mean hydrostatic pressure (σ_m) and flow stress (σ_0). This is also borne out from practice, because alloys such as tin bronze (CuSn8) and free-cutting aluminum (AlCuMgPb) are almost impossible to hot roll without cracking but can be extruded with a perfect finish (Ref 1). Nonetheless, successful production by extrusion frequently requires a narrow range of processing conditions that balance extrusion temperature with extrusion speed, as briefly described in this chapter.

This chapter also briefly introduces the methods and mechanics of extrusion, with an emphasis on the unique metalworking aspects of extruding. For example, one unique aspect of extrusion is the complex role of friction, which may be beneficial or detrimental, depending on the conditions of metal flow (see the section "Friction and Lubrication" in this chapter). Extrusion is, with few exceptions, a batch process that operates under discontinuous conditions. In extrusion, a second billet is not loaded until the first has been extruded. Non-steady-state conditions occur at the beginning and end of the extrusion process, and flow is nonsteady due to temperature variations during the extrusion cycle and variations of temperatures and friction over the length of the billet. Thus, nonuniform flow and property variations occur over the length of the part, as briefly described in the section "Non-Steady-State Effects and Defects" in this chapter.

Extrusion Methods

The process of extrusion can be done by several different methods, which can be generally classified as either conventional extrusion or hydrostatic extrusion. In conventional extrusion, a billet (typically in a container to support the workpiece under high extrusion pressures) is forced through the die by a ram acting directly on the workpiece or billet (Fig. 3a). In contrast, hydrostatic extrusion is performed by liquid pressure (Fig. 3b) rather than by direct application of the load with a ram. In hydrostatic extrusion, the work material is completely surrounded by a fluid, which is sealed off and is pressurized sufficiently to extrude the billet through the die. Hydrostatic extrusion can be used to extrude brittle materials that cannot be processed by conventional extrusion (as described in more detail in the section "Hydrostatic Extrusion" in this chapter). Hydrostatic extrusion also allows greater reductions in area (higher extrusion ratios) than conventional extrusion. Hydrostatic extrusion is done hot or cold, but hot hydrostatic extrusion is not very common—if done at all.

Conventional extrusion can be done hot, warm, or cold, depending on the work material and the size and shape of the desired product. For example, aluminum alloys are very suitable for economical production of parts by cold (i.e., room-temperature) extrusion, even when parts are large or complex in shape. Cold extrusion of steel is typically confined to relatively smaller parts (with starting plugs seldom more than 10 kg, or 25 lb), depending on the availability of equipment, tool materials, and the plasticity of the steel. Cold extrusion is always conducted with an effective lubricant. In contrast, hot extrusion may be done by either unlubricated or lubricated methods (Fig. 4), depending on the extrudability of the work material (see "Hot Extrusion" in this chapter). Hot extrusion is used to produce long, straight metal products of constant cross section, such as bars, solid and hollow sections, tubes, wires, and strips, from materials that cannot be economically formed by cold extrusion.

The mechanical process of extrusion is further classified as either direct or forward extrusion (Fig. 5a) or indirect or backward extrusion (Fig. 5b). In some cases, the extrusion process may be designed as a combined (direct and indirect) action. In forward (direct) extrusion, the die and ram are at opposite ends of the extrusion billet, and the product and ram travel in the same direction. Also, there is relative motion between the extrusion billet and the die. In backward (indirect) extrusion, the die is on a moving ram, and the product travels in the direction opposite that of the ram, either around the ram (as in the impact extrusion of cylinders) or up through the center of a hollow ram.

The shape of the part is usually the primary factor that determines the use of direct or indirect extrusion. For example, many cuplike parts are produced by backward extrusion, while shaftlike parts and long, hollow shapes can usually be produced more easily by forward extrusion. For many shapes, both forward and backward extrusion are used. Other factors that

*Portions adapted from publications listed in the Acknowledgments at the end of this chapter.

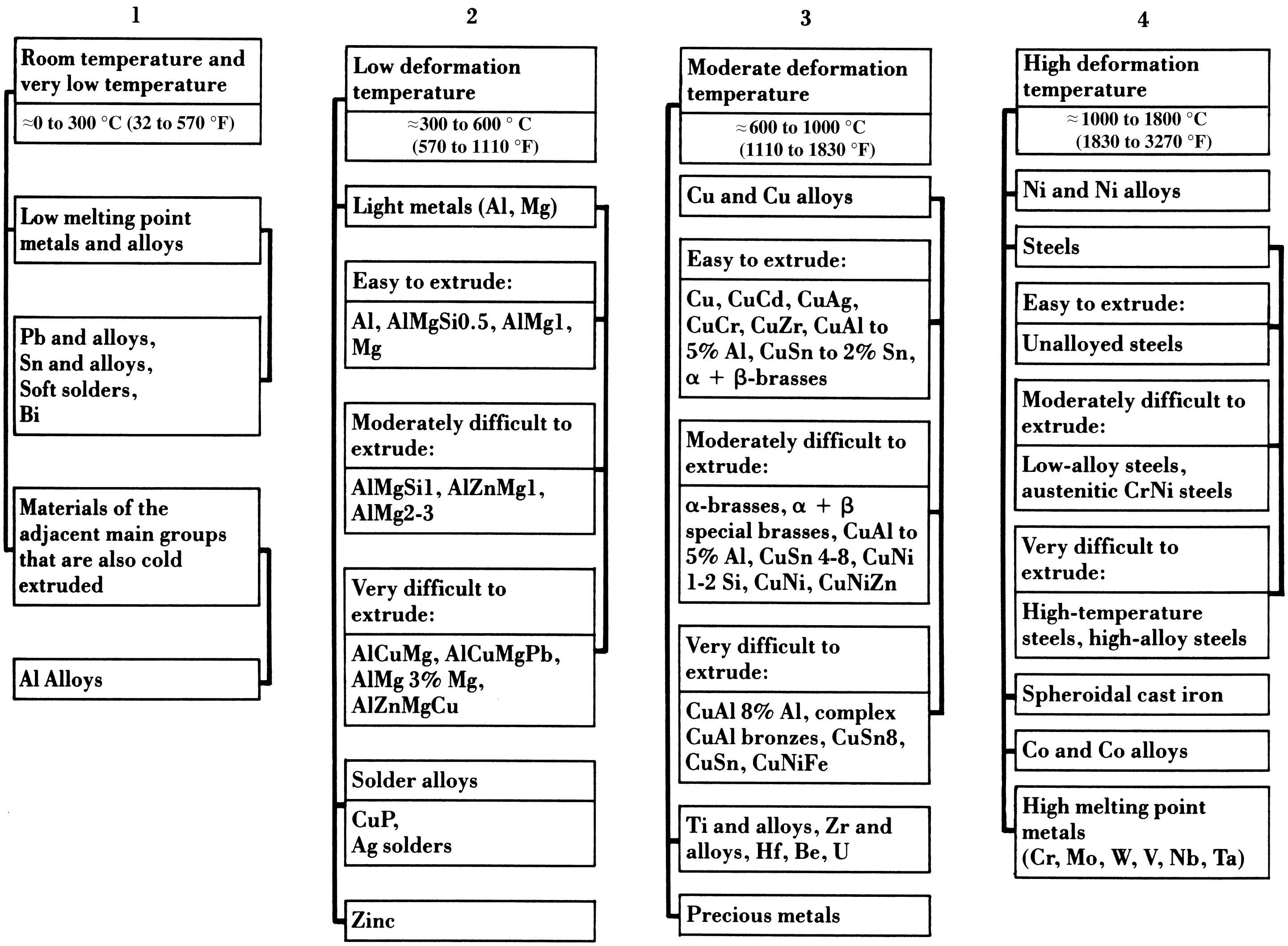


Fig. 1 Classification of extruded alloys on the basis of temperature. Source: Ref 1

influence procedure are the composition and condition of the work material, the required dimensional accuracy, quantity, and cost.

Dies

The tooling for extrusion consists of such components as containers, container liners, stems (rams), dummy blocks, mandrels, and dies. There is considerable variation in the tooling practice and design, depending on the processing temperature, lubrication, loading requirements, and whether the process is backward or forward extrusion. Thus, tooling is described in subsequent sections for the various types (direct/indirect) and processing conditions (hot/cold) of extrusion. However, one of the most critical aspects of tooling is die design. Die design is a crucial part of successful extrusion that depends very much on the judgment, skill, resourcefulness, intuition, and experience of the individual die designer and maker.

In general, the two most common types of extrusion dies are flat-face dies and shaped (conical) dies (Fig. 6). Spider or bridge dies also are used for producing hollow extrusions such as tubing. Flat-face dies (also termed square dies) have one or more openings (apertures) that are similar in cross section to that of the desired extruded product. Dies for lubricated extrusion (also called shaped, converging, or streamlined dies) often have a conical entry opening with a circular cross section that changes progressively to the final extruded shape required. Flat-face dies are easier to design and manufacture than shaped dies and are commonly used for the unlubricated hot extrusion of aluminum alloys. Shaped dies are more difficult and costly to design and manufacture, and they are generally used when lubrication is required; this includes cold extrusion and hot extrusion of harder alloys such as steels, titanium alloys, and nickel alloys.

The design of extrusion dies, whether of the flat or conical type, is still an art rather than a science. Nonetheless, rational design techniques and empirical guidelines are employed, often with the assistance of computer programs. Computer-aided design and manufacture has been developed to reduce the costs of designing and manufacturing extrusion dies, and a few examples are noted in Ref 2 to 4. Computer simulation programs also have been developed to model the metalworking process to more effectively select process variables, such as extruding speed and billet temperatures. However, finite-element modeling of deformation by extrusion can be complicated, because small variations in die shape, friction, or temperature distribution may result in a large change in the profile of the section. Moreover, deformation is complicated by the presence of quasi-shear surfaces close to the die exit (Ref 5). For these reasons, very few computer codes are available for three-dimensional simulation; most extrusion models are for axisymmetric problems.

Luckily, some internal equalization takes place in the product, and the process is fairly forgiving; in an example of computer-aided die design for extrusion of a T-section, material flow was much more uneven than anticipated, and yet, the extruded product was sound (Ref 6). The increased frictional resistance along the longer

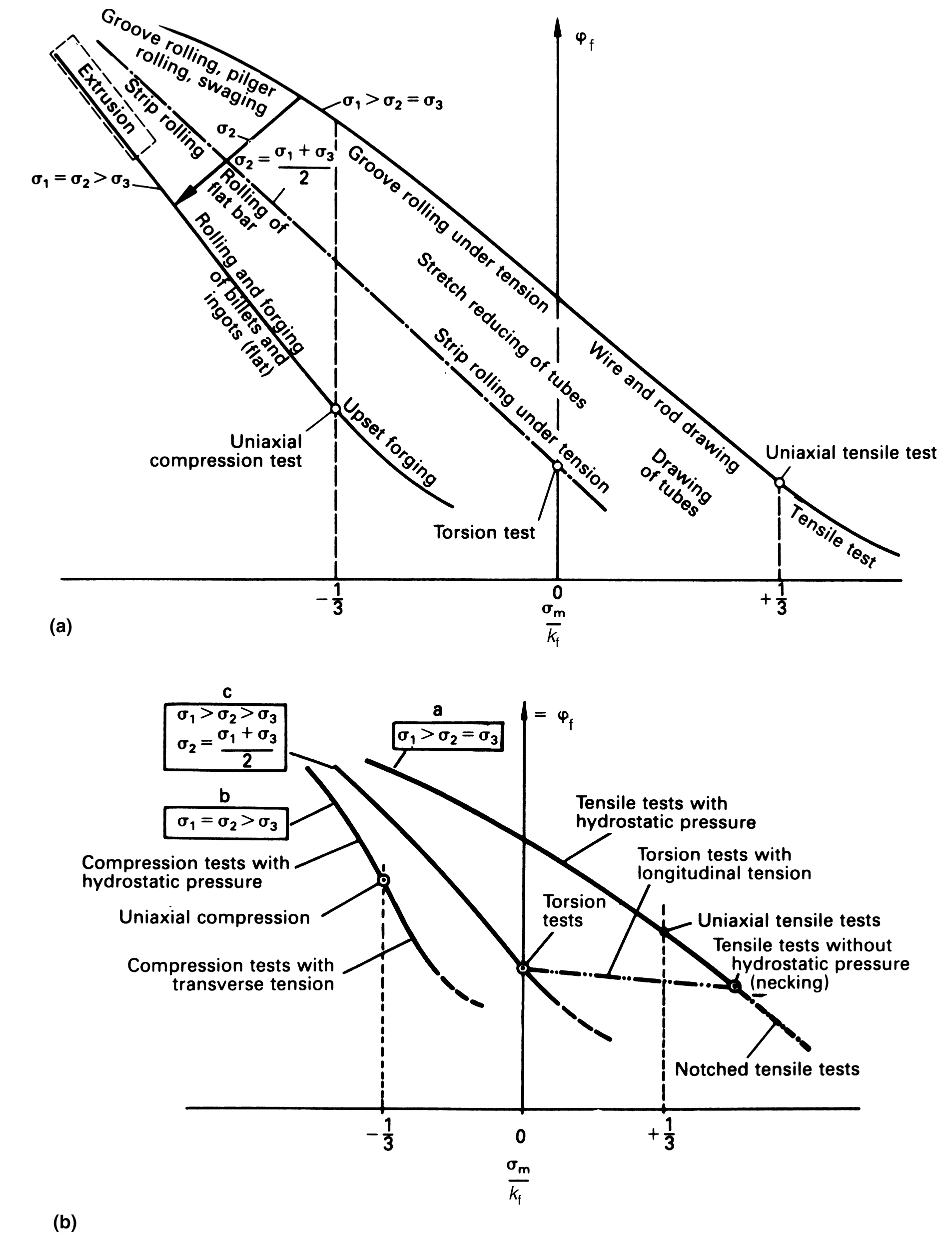


Fig. 2 Deformation to fracture, φ_f, versus the ratio of mean hydrostatic stress, $\sigma_m = \frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3)$, and flow stress, σ_0, in relation to the point of fracture for (a) various deformation processes and (b) in tension, compression, and torsion testing. $k_f = \sigma_1 - \sigma_3$ (see Fig. 8b). Source: Ref 1

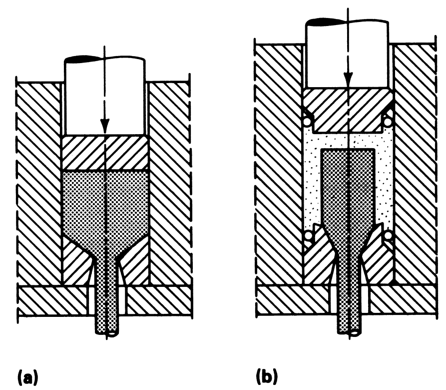


Fig. 3 Comparison of (a) conventional extrusion and (b) hydrostatic extrusion

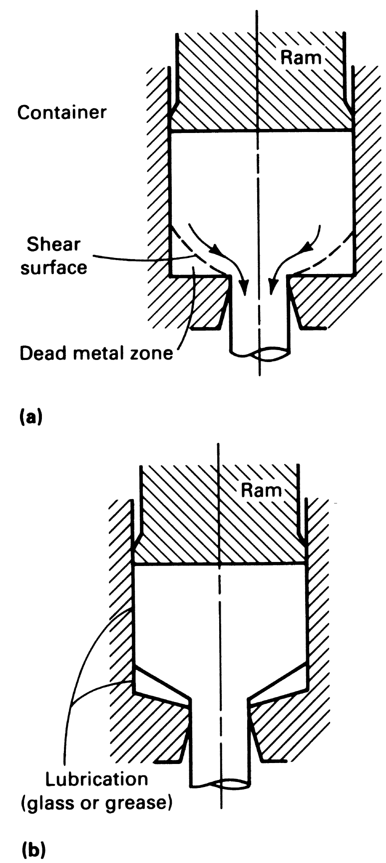


Fig. 4 Schematic of conventional hot extrusion (a) without lubrication and (b) with lubrication

die perimeter and the more complex flow increase the punch pressures relative to extrusion of a round bar of equal A_1 cross-sectional area.

General Die Design Considerations. The basic considerations of die design for direct extrusion are (Ref 7):

- The number of die openings based on the shape and size of the profile and the nature of exiting tooling
- Location of die openings relative to the billet axis
- Orientation of the openings around their centroids to match the handling system
- Determination of the final openings based on thermal shrinkage, stretching allowance, and die deflection (both dies and deep tongues)
- Optimization of bearing lengths to increase productivity

However, optimal design depends on a large number of factors, including the size of the shape to be produced, the maximum and minimum wall thickness, the press capacity, the length of the runout table, the stretcher capacity, the tool-stacking limitations, an understanding of the properties and characteristics of the metal to be extruded, and the press operating procedures and maintenance. Complete treatment of these factors is beyond the scope of this Handbook, so only some very basic factors in die design are addressed here. The key is to have a close working relationship between the designer, die maker, press operator, and die corrector.

Most design-for-manufacturing considerations for extrusions involve the difficulty of forcing the metal to flow uniformly from a large, round billet through small, complex die openings. This leads to some general rules that can help guide the design process. For example, even for the more easily extrudable aluminum parts, sections with both thick and thin sections are to be avoided. Metal flows faster through a larger opening than a smaller one, and so flow is faster where thicker

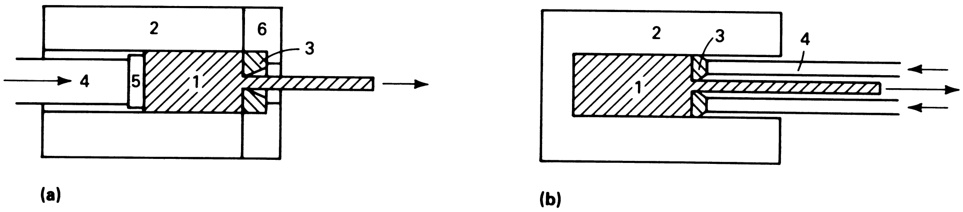


Fig. 5 Basic methods of conventional extrusion. (a) Forward (direct). (b) Backward (indirect). 1, billet; 2, container; 3, die; 4, stem; 5, dummy block; 6, die backer

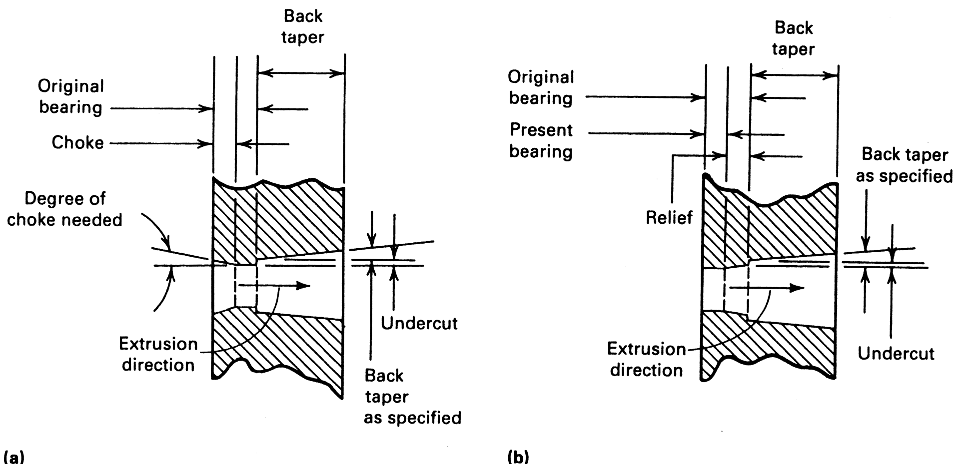


Fig. 6 Two basic types of extrusion dies. (a) Shaped (or conical-type) dies for lubricated extrusion. (b) Flat (or square-type) dies typical of unlubricated hot extrusion

sections occur, which thus results in distortions in the extruded shape. This must be compensated for in the design of dies. For example, when a section to be extruded has a thick wall and a thin wall, various means are employed to retard metal flow through the thick section and to increase the flow rate through the thin section of the die.

Long, thin wall sections should also be avoided, because such shapes are difficult to keep straight and flat. If such sections are absolutely necessary, then the addition of ribs to the walls help distribute the flow evenly. Hollow sections are quite feasible, and the designer can select various types of dies for hollow sections. These are discussed in the section "Tube Extrusion" in this chapter. It is best if hollow sections can have a longitudinal plane of symmetry. Semihollow features may also be considered but should be avoided, because a semihollow feature requires the die to contain a very thin—and hence, relatively weak—neck.

After die layout, tool strength analysis is also needed to determine correct openings with die cave and die deflection from the bending and shear stresses in the die, backer, and bolster due to the extrusion pressure. The need for conforming tools to support the die is also determined by the tool strength analysis. Die deflection analysis includes estimation of the bending of the die at various locations. This information is helpful in predicting the dimensional changes in the die orifices under load during extrusion. The dimensions of the die orifices can be modified to correct for these dimensional changes to obtain the desired tolerances in the extruded shape.

The next step in the design process is the determination of die-bearing lengths. The die bearing at any position is dependent on the section thickness at that position and on its distance from the die center. In the direct extrusion process, the frictional resistance at the billet-container interface slows down metal flow near the billet surface, and so the center of the billet moves faster than at the periphery. To balance flow, bearing length must be inversely proportional to its distance from the surface. Similarly, bearing length needs to be smaller to balance flow in a thinner section. The treatment of bearing surface at the front and back of the die aperture is termed the choke and relief, respectively (Fig. 7) (Ref 7). A choke can be provided on certain portions of the bearing surface if the die designer anticipates difficulty in filling sharp corners or completing thin sections of the extruded product. This slows the rate of metal flow and consequently fills the die aperture. For direct extrusion of hard aluminum alloys, the front of the bearing is generally choked at an angle up to 3° (Ref 7). Increasing the amount of back relief at the exit side of the bearing surface increases the rate of metal flow by decreasing the original bearing length.

The Extrusion Process (Ref 8)

In common with wiredrawing, the process of extrusion can be described by metal flow through a conical channel (Fig. 8). The starting material is usually a cylindrical, cast, or previously extruded or rolled billet that is placed into a container and pressurized, either by means of a punch attached to the ram of a hydraulic or mechanical press (direct extrusion) or by a die attached to a moving ram (indirect extrusion). Deformation in extrusion tends to be nonuniform (described more later), but it is still customary to calculate an average strain as either reduction, r, in area:

$$r = \frac{A_0 - A_1}{A_0} \qquad \text{(Eq 1)}$$

or extrusion ratio:

$$R_E = \frac{A_0}{A_1} \qquad \text{(Eq 2)}$$

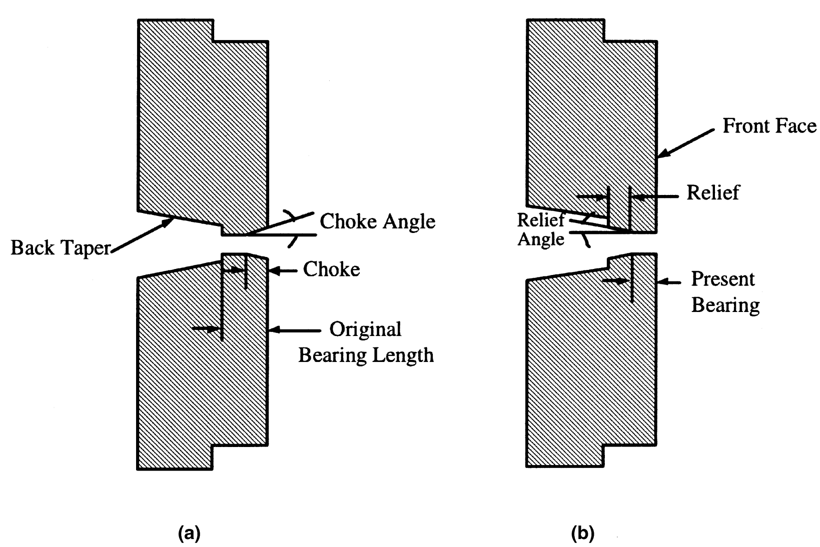


Fig. 7 Choke and relief in die bearing. (a) Choke at front of bearing. (b) Increased relief angle at the back or exit side of the bearing. Source: Ref 7

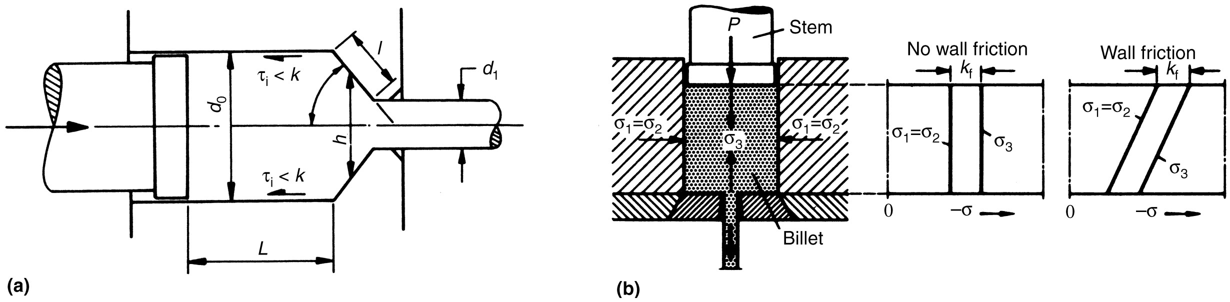


Fig. 8 Forward extrusion (a) with conical die entry (half arrows indicate interface shear stresses acting on the material. (b) Schematic variation of axial pressure, σ_3, and radial pressure, $\sigma_1 = \sigma_2$, over the length of the container for lubricated and unlubricated extrusion. $k_f = \sigma_1 - \sigma_3$

which conveniently leads to the true strain, ε:

$$\varepsilon = \ln R_E = \ln \frac{A_0}{A_1} \quad \text{(Eq 3)}$$

The extrusion ratio can reach very high values, in excess of 400 to 1.

Extrusion Pressure. Most extrusion theories lead to solutions in which the extrusion pressure p_e at the end of the stroke is proportional to strain:

$$p_e = \hat{\sigma}_0 (a + b\varepsilon) \quad \text{(Eq 4)}$$

where $\hat{\sigma}_0$ is the mean flow stress of the material, and the parenthetical term allows for the pressure required for uniform deformation, the inhomogeneity of deformation (redundant work), and die-face friction. The constants a and b may be calculated or determined from experiments. For very approximate calculations, it may be assumed that $a = 0.8$ and $b = 1.2$.

As in other processes, the effects of inhomogeneity of deformation and interface friction interact. Deformation becomes more homogeneous with a decreasing h/l ratio (Fig. 8), which implies, for a constant reduction, a smaller die half-angle, α. This results, however, in a longer surface over which friction must be overcome. Conversely, a larger die angle reduces the frictional components but leads to increased inhomogeneity; therefore, as in wiredrawing, an optimal die angle is found, the value of which increases with increasing extrusion ratio and friction (Fig. 9) (Ref 9).

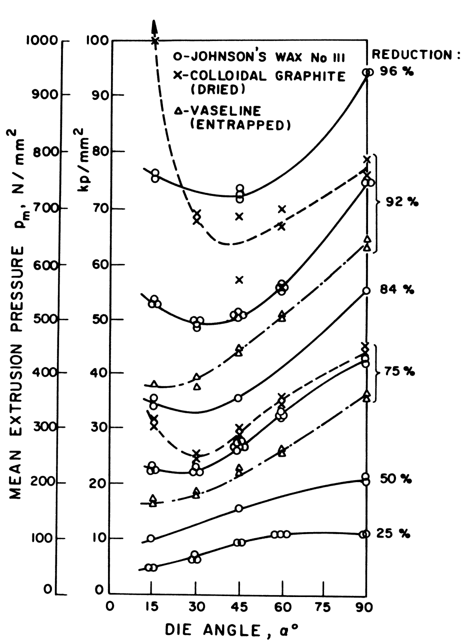


Fig. 9 Effect of lubrication on optimal die angle in forward extrusion of aluminum. Source: Ref 9

Flow becomes truly streamlined only when friction is very low. It can be shown that die shapes other than simple cones ensure even greater homogeneity of flow, especially with high friction. With low friction, the improvement does not, in general, justify an extra expense in diemaking, even though with numerically controlled machining this expense may be slight.

Stresses in the deformation zone are fully compressive when $h/l > 1$, and, for this reason, extrusion is particularly favorable for working of materials of limited ductility. If, however, the extrusion ratio is low and the die angle is large (and thus h/l is large, too), a hydrostatic tensile stress can develop in the center of the extrusion, and a centerburst defect similar to that found in wiredrawing occurs. Because increasing die friction increases the pressure required for extrusion and thus generates a hydrostatic pressure, higher friction is actually desirable in extending the range over which sound extrusions can be obtained (Fig. 10) (Ref 10).

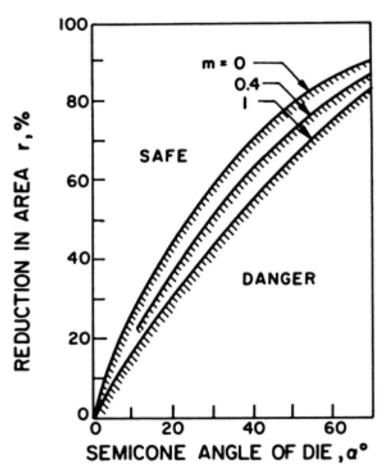


Fig. 10 Effect of friction factor (m) on danger of centerburst in extrusion. Source: Ref 10

On the basis of pressure considerations, a conical die entry would often be optimal. However, it has the disadvantage that the press stroke must be arrested before the punch can touch the die, thus increasing material lost in the remnant (butt). Furthermore, the extrusion must be pushed back out of the die if the remnant is to be cut off. Alternatively, the remnant may be pushed through with a deformable follower block (e.g., a graphite block), or another billet must follow the first billet without removal of the remnants. In the latter case, the interfaces between the billets must be clean to facilitate welding and prevent the development of internal defects, and thus the use of a lubricant is precluded. For practical reasons, extrusion is often conducted with a flat-face die (α = 90°), especially at heavier reductions where there is no danger of a centerburst defect.

Extrusion Force. The pressure calculated from Eq 4 applies to the container cross-sectional area A_0, and thus the extrusion force is:

$$P_e = p_e A_0 \quad \text{(Eq 5)}$$

For measurement of this net extrusion force, a billet may be pushed through the extrusion die without the support of a container (Fig. 11a). To avoid upsetting of the unsupported billet, the pressure p_e must be less than the flow stress (σ_0) of the workpiece material. Thus, the effects of friction and die angle can be explored but only for a rather limited range of reductions. This technique is, nevertheless, suitable for lubricant ranking. In order for the interface shear strength τ_i (or, if preferred, μ or friction factor, m) to be derived, the value of the average or mean flow stress ($\hat{\sigma}_0$) must be accurately known. Alternatively, the axial and normal (radial) forces can be simultaneously measured. Thus, the split die principle was employed for axial symmetry (Ref 11) and for plane strain (Ref 12).

If heavier reductions are to be explored, a conventional extrusion configuration is needed. It is possible, however, to measure separately the forces acting on the extrusion die by leaving a small, well-lubricated clearance between die and container, so that no fin is extruded, but free movement of the die is ensured even under load

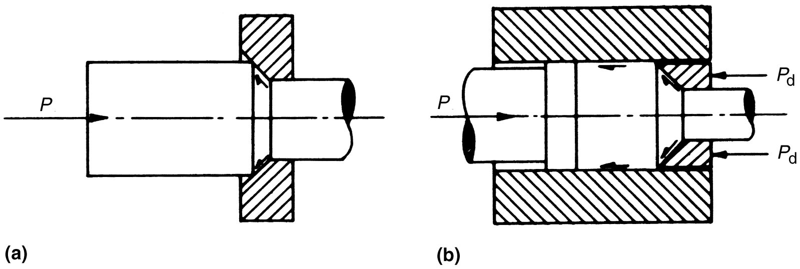


Fig. 11 Measurement of die friction by (a) containerless extrusion and (b) separation of die force

(Fig. 11b). Such measurements have been found to be reproducible (Ref 13, 14). Load cells for punch and die loads and strain gages on the container were used for measurement of internal pressure in cold forward and backward extrusion (Ref 15).

The extrusion force is limited by the potential failure of the ram or container. A short ram (or punch) can fail by upsetting, a long one by buckling. The container is an internally pressurized vessel that can be reinforced by a shrink-ring construction. Nevertheless, internal pressure is typically limited to 1.2 GPa (175 ksi) in hot extrusion and to 2.5 GPa (360 ksi) in cold extrusion. This means that for any given material, the extrusion ratio is limited according to Eq 4 (Fig. 12) (Ref 16).

Typical Flow in Extrusion (Ref 1). Metal flow varies considerably during extrusion, depending on the material, the material-tool interface friction, and the shape of the section. Fig. 13 shows the four types of flow patterns that have been observed.

Flow pattern S (Fig. 13a) is characterized by the maximum possible uniformity of flow in the container. Plastic flow takes place primarily in a deformation zone directly in front of the die. The major part of the nonextruded billet, pushed as a rigid body through the die, remains undeformed; therefore, the front of the billet moves evenly into the deformation zone.

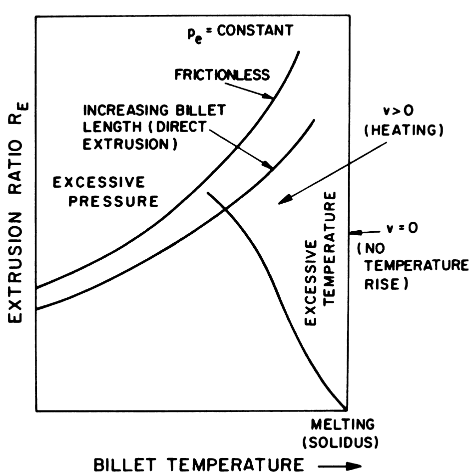


Fig. 12 Process limitations for hot extrusion arising from container friction and workpiece heating. Source: Ref 16

Flow pattern A (Fig. 13b) occurs in homogeneous materials when there is virtually no friction between the container and the billet but significant friction at the surface of the die and its holder. This retards the radial flow of the peripheral zones and increases the amount of shearing in this region. The result is a slightly larger dead-metal zone than that in flow type S, along with a correspondingly wider deformation zone. However, deformation in the center remains relatively uniform. Flow patterns of this type are seldom observed in nonlubricated extrusion; instead, they occur during the lubricated extrusion of soft metals and alloys, such as lead, tin, α-brasses, and tin bronzes, and during the extrusion of copper billets covered with oxide (which acts as a lubricant).

Flow pattern B (Fig. 13c) occurs in homogeneous materials if friction exists at both the container wall and at the surfaces of the die and die holder. The peripheral zones are retarded at the billet-container interface, while the lower resistance causes the material in the center to be accelerated toward the die. The shear zone between the retarded regions at the surface and the accelerated material in the center extends back into the billet to an extent that depends on the extrusion parameters and the alloy. Therefore, the dead-metal zone is large. At the start of extrusion, the shear deformation is concentrated in the peripheral regions, but as deformation continues, it extends toward the center. This increases the danger of material flowing from the billet surface—with impurities or lubricant—along the shear zone and finishing up under the surface of the extrusion. In addition, the dead-metal zone is not completely rigid and can influence, even if to a limited degree, the flow of the metal. Flow type B is found in single-phase (homogeneous) copper alloys that do not form a lubricating oxide skin and in most aluminum alloys.

Flow pattern C (Fig. 13d) occurs in the hot extrusion of materials having inhomogeneous properties when the friction is high (as in flow pattern B) and when the flow stress of the material in the cooler peripheral regions of the billet is considerably higher than that in the center. The billet surface forms a relatively stiff shell. Therefore, the conical dead-metal zone is much larger and extends from the front of the billet to the back. At the start of extrusion, only the material inside the funnel is plastic, and it is severely deformed, especially in the shear zone, as it flows toward the die. The stiff shell and the dead-metal zone are in axial compression as the billet length decreases; consequently, the displaced material of the outer regions follows the path of least resistance to the back of the billet, where it turns toward the center and flows into the funnel.

This type of flow is found in the (α + β) brasses, in which the cooling of the peripheral regions of the billet leads to an increase in flow stress, because the flow stress of the α-phase is much higher than that of the β-phase during hot working. As in the (α + β) brasses, flow type C occurs when there is a hard billet shell and, at the same time, the friction at the container wall is high. It can also occur without any phase change that leads to a higher flow stress if there is a large temperature difference between the billet and the container. This can take place in the extrusion of tin as well as of aluminum and its alloys.

Heat Generation (Ref 8). The work of plastic deformation and friction is converted into heat. Because strains can be very large and frictional energy is expended in a thin surface layer, heat generation in extrusion is substantial. With rapid extrusion, most of the heat is carried away by the extruded product, but heat is conducted back into the undeformed part of the billet when extrusion is relatively slow. In both cases, the butt end of the extrusion is heated more than the nose end. Preheating temperature in hot extrusion is usually high, partly to reduce mean flow stress ($\hat{\sigma}_0$) and thus allow a large extrusion ratio, and partly to ensure adequate ductility. It is possible that the temperature rise due to deformation and friction could heat the emerging product above the melting point (or, in alloys, above the solidus temperature). The parallel die land, while contributing only a few percent to the extrusion force, is a major source of heating. Thus, the temperature is highest on the surface, and

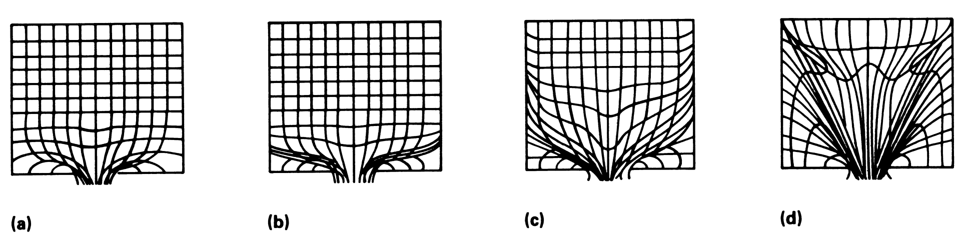


Fig. 13 Four types of flow patterns observed in the extrusion of metals. (a) Flow pattern S. (b) Flow pattern A. (c) Flow pattern B. (d) Flow pattern C. See text for details. Source: Adapted from Ref 1

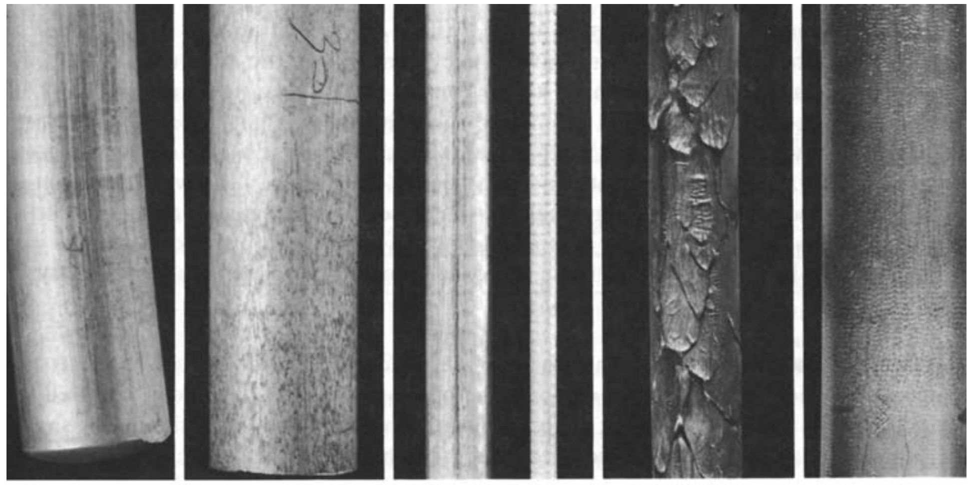
grain-boundary melting causes a disintegration of the product in the form of the characteristic "fir-tree" cracking (speed cracking). Grain growth can also occur on the surface due to the heat generation. Extrusion must be slowed down to minimize the rate of heat generation (Fig. 12). Alternatively, the billet can be unevenly preheated, with a temperature gradient dropping from front to back to compenstate for normal temperature rise at the butt end of the extrusion.

Non-Steady-State Effects and Defects (Ref 8). Non-steady-state conditions always exist at the beginning and end of the extrusion process. Non-steady-state conditions are more predominate in extrusion of short billets for the purpose of making parts, where end defects cannot be allowed to develop, because the unextruded portion of the workpiece forms part of the finished component. In contrast, steady-state conditions prevail during extrusion of long, semifinished products such as bars, sections, and tubes. In this case, end defects may be allowed to develop to a limited degree, because the residue of the remaining billet in the container is discarded.

Initiation of Extrusion. A pressure peak (breakthrough pressure) is observed just at the point where extrusion is initiated. In lubricated extrusion, it indicates that the lubrication mechanism typical of steady-state conditions has not yet been established. Thus, it is most marked when a square-ended billet is extruded through a tapered die; the sharp edge of the billet may actually scrape off whatever lubricant film is present on the die, and if this occurs, a new, low-friction layer can develop only after the die throat has been filled.

End of Extrusion. When extrusion proceeds to the point where the end face touches the deformation zone, flow becomes nonuniform even in lubricated flow. The flow rate from the sides to the center is insufficient, and a crater (pipe) forms. The depth of the pipe increases with increasing container and die friction but is reduced by friction on the back face of the billet. In extrusion of semifabricated products, the pipe increases the discard, especially if it fills up with oxides and lubricant residues; it can form a difficult-to-detect defect in the rear end of the product.

External Extrusion Defects. In addition to internal defects (such as centerburst) or metallurgical deficiencies (from variations in grain size and metallurgical structure, resulting from recrystallization after inhomogeneous deformation), there are also external, friction-related defects in both hot and cold extrusion that, when severe, make a product unsuitable for use. Surface cracking, whether circumferential (fir-tree cracking), or at roughly 45° to circumferential ("crow's feet" cracking), or even longitudinal ("splitting"), is a common problem in both extrusion and drawing (see Chapter 21, "Workability and Process Design in Extrusion and Wire Drawing," in this Handbook for illustrations). The fir-tree effect (speed effect) occurs in hot-short materials with friction on the die land as a contributory cause. A cold extrusion defect of similar appearance can also be attributed to the simultaneous occurrence of slip on the die face and sticking on the die land (Ref 17).



(a) (b) (c) (d) (e)

Fig. 14 Typical defects observed in cold extrusion of aluminum alloys. (a) Deep scoring. (b) Surface roughening from the presence of thick lubricant film. (c) Bambooing. (d) Shear cracks. (e) Machining marks stabilized by lubricant. Source: Ref 18

Other external defects include:

- Subsurface defects caused by intermittent stick-slip on the container wall may emerge at the surface in the form of laminations or flakes.
- Pickup on the die or on the die land leads to deep scoring of the extruded surface (Fig. 14a) (Ref 18). Pickup tends to be cumulative and may force a die change and cleanup, often after extrusion of each billet in hot extrusion.
- In lubricated extrusion, the surface roughens in the presence of a thick lubricant film (Fig. 14b), a phenomenon often described as "orange peel." An excessively thick, unstable film leads to bambooing (Fig. 14c) and, in some less-ductile materials, even to shear cracks (Fig. 14d). If the billet has been turned in preparation for extrusion, the machining marks are stabilized by the lubricant (Fig. 14e).

Friction and Lubrication (Ref 8)

Extrusion is one of the processes in which a thick lubricant film can be developed and maintained. The similarity to wiredrawing is evident; the difference is that lubricant pressure at the die inlet can be higher (even in nonhydrostatic extrusion) and thus, thick-film formation is promoted. Film thickness is governed by a number of factors. As would be expected from the plastohydrodynamic lubrication theory, film thickness increases with increasing viscosity and decreasing die half-angle. Film thicknesses may range from excessively thick full-fluid films to mixed films with predominantly boundary contact.

Another important aspect of extrusion is the complex role of friction. In wire drawing, the elimination of friction is desirable, because the introduction of friction promotes the formation of centerbursts during drawing (see Fig. 6a in Chapter 19, "Drawing of Wire, Rod, and Tube," in this Handbook). In extrusion, however, friction may deter centerbursts during extrusion. Nonetheless, there also are a number of friction-related surface defects, from either hot or cold extrusion, that can make a product unsuitable for use. Friction during extrusion can lead to the development of tensile stresses at the surface of heterogeneous alloys. Thus, lubrication is often required, although in some cases, hot extrusion is done without lubrication (especially for aluminum alloys).

Of all bulk deformation processes, extrusion is perhaps the most sensitive to lubrication, partly because the non-steady-state conditions at the beginning and end of extrusion are conducive to instabilities, and partly because any unwanted change in lubrication leads to defects that affect not only the appearance but also the properties and integrity of the product. In many respects, the problems of lubrication are similar to those of lubrication in forging.

From the standpoint of friction, it is of greatest importance whether the billet moves relative to the container (direct extrusion) or is at rest inside the container (indirect extrusion). In either case, the billet is first upset to fill out the container of cross section A_0. The pressure then rises until plastic flow begins, and a product of cross section A_1, which has a short, parallel land, emerges from the die. In principle, it is immaterial whether the extrusion is a round bar, a shaped section, or a tube. Products vary greatly in cross-sectional dimensions and in length, but they typically range from, say, 5 mm (0.2 in.) diameter wire to 500 mm (20 in.) diameter tubes or sections. Presses range from 2 to 150 MN (200 to 17,000 tonf) in capacity.

Table 1 Summary of the extrusion methods used for different materials

Method of extrusion: Without lubrication, without shell	Without lubrication, with shell	With lubrication, without shell
Type of material		
Materials with a lubricating oxide skin (e.g., Cu) and materials that flow according to type B and in which lubrication would lead to a poor-quality surface finish (e.g., Al alloys)	Materials that flow according to type C (e.g., brass) and materials that readily adhere to the container wall (e.g., complex Al bronzes) Materials that rapidly oxidize during billet heating, i.e., those in which any flowing in of the oxide skin in type B flow must be prevented at all costs (low-alloy Cu materials such as CuCr, CuZr, etc.)	Materials with container flow characteristics that can be made similar to flow type A by using a lubricant and usually conical dies Materials that are worked by cold extrusion, including hydrostatic extrusion, as well as alloys that are difficult to extrude, such as Ni alloys, steels, and high-melting-point metals that cannot be extruded—or only uneconomically—without lubrication

Source: Ref 1

The importance of lubrication during direct extrusion of a billet in a container is discussed in more detail in the section "Forward (Direct) Extrusion" in this chapter. It is important to note that lubrication can have a detrimental effect on the surface finish if the metal flow is inhomogeneous (flow type B or C in Fig. 13). If type B or C flow cannot be excluded, then container lubrication should be avoided. This is generally the case for aluminum alloys and for those copper alloys that are extruded without a shell and do not have a lubricating oxide skin (Ref 1, p 112). However, many difficult-to-extrude materials require lubrication to reduce extrusion load or to protect the die (Table 1).

Lubrication is advisable during cold extrusion or extrusion at low temperatures (less than 300 °C, or 570 °F). Lubrication may also be advisable in hot extrusion of some difficult alloys through conical dies. Lubricants can be classified into two groups by temperature (Ref 1):

- Greaselike lubrication below 1000 °C (1830 °F) with grease, graphite, MoS_2, mica, talc, soap, bentonite, asphalt, and plastic (e.g., high-temperature polymides)
- Glasslike lubrication above 1000 °C (1830 °F) with glass, basalt, and crystalline powders

Attributes of Extrusion Lubricants. An extrusion lubricant must have the usual attributes for metalworking, but some requirements assume special significance:

- An uninterrupted separating film between billet, die, and container and its low shear strength are important not only because they determine the allowable extrusion ratio and billet length in direct extrusion, but also because they control material flow and thus the properties of the product and the length of the discard.
- Heat insulation is imperative in nonisothermal extrusion when the extrusion temperature is high enough to cause softening of the die, mandrel, or container on prolonged exposure and thus when, in the absence of such insulation, the die would close in or open up (die wash), the mandrel would thin out or tear off, and the container would open up.
- Protection against oxidation is vital when the workpiece material forms an abrasive oxide, or when the lubricant does not wet as a result of oxidation. Because oxidation can be controlled not only by a preapplied lubricant film but also by the preheating atmosphere, the method of heating must be chosen with the entire tribological system in mind. Induction heating, resistance heating, and glass or salt baths can all be considered in addition to gas-fired or electric muffle furnaces, with or without a protective atmosphere. If scaling of steel is allowed, the scale may be removed with high-pressure water, but great care must be taken to avoid scale pockets that would break through any lubricant film.

Solid-Film Lubrication. Because of the high interface pressures, solids with pressure-insensitive shear strengths can be very useful as lubricants. With all forms of solid-film lubrication, pressures in direct extrusion of long billets may reach high enough values to satisfy the condition of sticking friction (Fig. 15). Sticking friction occurs when the frictional interface shear stress (τ_i) is higher than the condition for internal shearing of the material instead of sliding against the die surface. The condition of sliding friction (Fig. 15a) is not necessarily realistic for metalworking, and the coefficient of friction (μ) becomes meaningless with a high interface pressure (Fig. 15b). When sticking friction occurs with a long billet, then the billet length has to be reduced or the process changed to reverse extrusion. Solid lubricants may be applied to the billet when the pressure medium lacks sufficient lubricating properties. Continued development of a model for solid lubricant entrainment in hydrostatic extrusion has allowed for pressure-sensitive shear strength (Ref 19).

Oxides. Among the oxides grown on metals, only some (e.g., on copper) can serve as lubricants. Others (e.g., on titanium) serve merely as parting agents. Externally introduced oxides never repaid the effort put into investigating them, although researchers found that solids such as lime or slate gave adequate lubrication in the form of a friable pad trapped between die

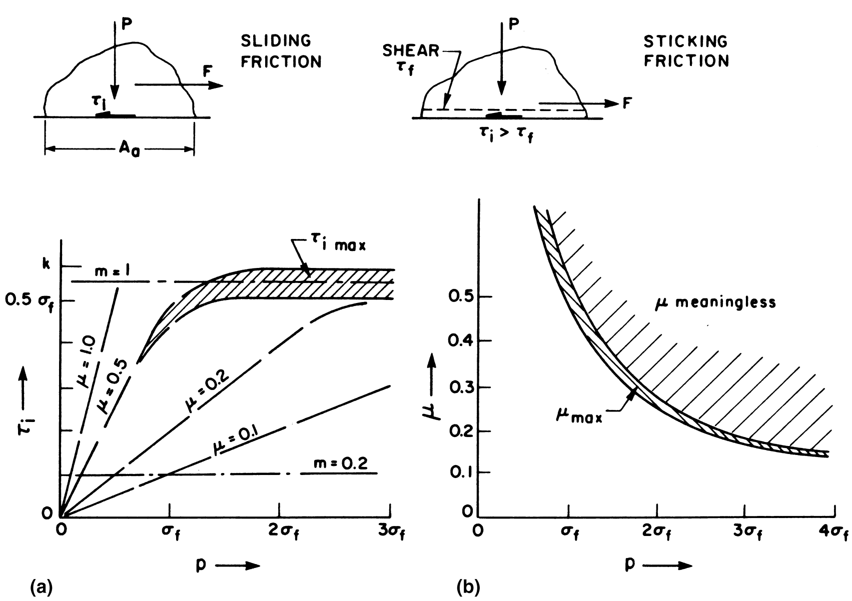


Fig. 15 Shear stress and coefficient of friction in (a) sliding friction and (b) sticking friction. (a) In sliding friction, the coefficient of friction, μ, is roughly constant, and the frictional (or interface) shear stress, τ_i, increases in proportion to the interface pressure, *p*. This condition is not necessarily realistic in metalworking when τ_i exceeds conditions for internal shear of the workpiece. (b) Sticking friction occurs when $\tau_i = \mu p > k$, where *k* is the value of μ_{max} where relative sliding does not occur at the interface. Calculated μ actually drops with increasing hydrostatic stress (or interface pressure, *p*). σ_f = flow stress

and billet (Ref 20). Lubrication must be attributed to a controlled form of wear (and perhaps melting) of the pad.

Canning and Lubricants. Simple canning techniques can be used with lubrication in the form of canning powders or glass lubricants for difficult-to-extrude materials in a steel can. Lubricants are applied as a coating or, if a thicker coating is required or a powder is to be extruded, a simple canning technique can be used with glass as the lubricant. To ensure smooth, defect-free material flow, extrusion is invariably done with lubrication, and then it is preferable to have a well-fitting billet with a conical nose. The can is often evacuated through a tube to prevent oxidation of the powder and puffing up of the can. Because the can material is removed after extrusion, preferably by mechanical means, a parting agent is often placed on the billet-can interface. In some instances, a thin sheet of diffusion barrier is also used. By the proper choice of can material, extrusion pressures can be reduced, as in extrusion of a high-strength Al-Cu-Mg alloy canned in pure aluminum (Ref 21). However, the can material must not be too soft or too thick, lest it be extruded ahead of the billet or behave similar to a fluid film with a typical unsteady bambooing flow. A can material of one-half to one-third the strength of the billet material gives satisfactory results. An example of a complex canning arrangement for extrusion of a tungsten tube is shown in Fig. 16. The can metal, because of its lower strength, allows roughening of the extrusion surface.

Glass Lubrication. The geometry of the extrusion process makes it particularly favorable for lubrication with glasses that melt gradually to form a lubricating film. As a lubricant, glass exhibits unique characteristics, such as its ability to soften selectively during contact with the hot billet and, simultaneously, to insulate the hot billet material from the tooling. For example, glass lubricants are used in the Séjournet process, which is a common method of hot extrusion for steels and titanium alloys (see the section "Hot Extrusion" in this chapter). Since about 1950, the technique of glass lubrication of the Ugine-Séjournet process and its further developments has spread to become dominant for extrusion of tubes, sections, and bars from steels and other materials when quantities are insufficient or shapes too complex to allow rolling. By far, the largest quantities of steel are extruded by phase-change lubrication of glass.

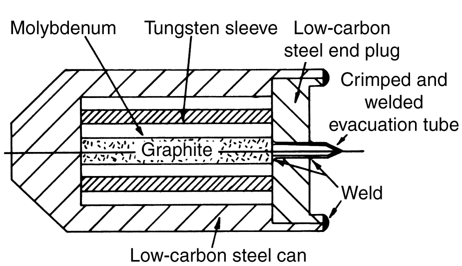


Fig. 16 Complex canning arrangement, including composite billet for extruding tungsten tubes at 1210 °C (2210 °F)

Glasses for lubrication are often proprietary developed for minimum cost compatible with performance requirements for uniform application of heat-insulating, low-friction glass. The desirable glass property is usually defined as a viscosity of 10^1 to 10^2 Pa · s at the billet temperature, although a viscosity-temperature curve of low slope would appear to be needed for optimum melt-away performance, and thermal properties such as heat conduction and specific heat should be just as important. Window glass or its equivalent, in powder or fiber form, mixed with a binder, is satisfactory for many steels and even for stainless steels. Some producers replace borosilicate glass with the cheaper basalt for Ni-Cr alloys and with basalt mixed with borosilicate glass for Ni-Cu alloys extruded at lower temperatures. Several improvements have been suggested, usually in the form of patents. For example, it has been suggested that a desirable viscosity gradient could be established by the use of glass layers of differing viscosities, either between the billet and container or between the front and back of the billet bore. The viscosity of glass used for making commercial foams is too high, and the graphite used as a foaming agent interferes with wetting of the metal; however, foam is successful as a carrier for powder.

Various methods of glass application have been used. For example, a pad of plate glass placed in front of the billet cracks in contact with the billet. Pads made of about 100-mesh powder have become standard. They are self-sintered or are made with binders such as sodium silicate or bentonite clay. Other forms of glass such as wool, cloth, and foam were also tried but with less success or economy. As extrusion begins, a thin film of glass melts off the pad. If melting characteristics are properly chosen and the press speed is adequate (typically, 0.15 to 0.5 m/s), the pad supplies glass to the end of the stroke. Another method is rolling a preheated billet down an inclined table covered with glass powder or fiber so that it becomes coated with a layer of molten lubricant. For tube extrusion, a woven glass sock is pulled over the mandrel, or glass powder is sprayed into the billet bore.

Other Solid Lubricants. Layer-lattice compounds are widely employed. When they are the only means of lubrication, uniform coverage of the billet is critical, because localized film failure would lead to die pickup and workpiece damage. The shear stress of the interface increases with temperature; lower friction is ensured with grease-based lubricants filled with copper and lead powder (Ref 22). To minimize breakthrough pressure, the billet should be nosed (Ref 23). Because of their limited spreading ability, these lubricants are more suitable for short billets.

Graphitic lubricants are often used for selective lubrication of die and mandrel, either alone or in conjunction with another lubricant. Polymers are used in cold extrusion as solid-film lubricants. In hot extrusion, they may be gradually melted to provide a form of phase-change lubrication.

Forward (Direct) Extrusion

A typical sequence of operations for the forward extrusion of a solid section is:

1. The heated billet and the dummy block are loaded into the container.
2. The billet is extruded by the force of the punch being pushed against it. This upsets the billet, then forces the metal to flow through the die. During extrusion, a thin shell of material (butt) may be left in the container.
3. The container is separated from the die, the extruded section with the butt, and the dummy block.
4. The discard (butt) is sheared off.
5. The shear die, the container, and the ram are returned to their initial (loading) positions.

In forward or direct extrusion, the billet is pushed bodily along the container wall, and the extrusion emerges in the same direction. However, the crucial aspect is that the billet moves relative to the container; therefore, in addition to friction on the die, friction on the container wall has to be overcome, too. This has a profound effect on all aspects of extrusion.

Process Characteristics

Container Friction. As previously noted, the large pressures of most practical extrusion processes require a billet supported by a container. At a first approximation, it can be assumed that pressure within the container propagates as in a hydraulic medium and that the extrusion pressure prevails at the billet-container interface (in reality, pressures are lower). Because even the minimum pressure (Eq 4) can reach high values at large extrusion ratios, friction at the container-to-billet interface assumes great importance. From the practical point of view, there are only two possibilities: either the interface is arrested (sticking friction), or sliding of the interface is ensured (lubricated flow). No intermediate condition can be allowed, because partial sticking and partial sliding invariably lead to unsteady material flow and a defective extruded product.

Material flow is affected by friction of the billet with the container wall, as previously noted (Fig. 13). These flow patterns extend to the product, as illustrated schematically in Fig. 17. In the total absence of friction, grid lines inscribed on the center plane of a split billet remain undistorted until they reach the deformation zone just ahead of the die (Fig. 17a). Even at the prevailing high interface pressures, this condition can be maintained only as long as $\tau_i << k$ (where k is the shear flow stress according to von Mises criterion).

With increasing friction, friction on the container wall causes the grid lines to curve. In

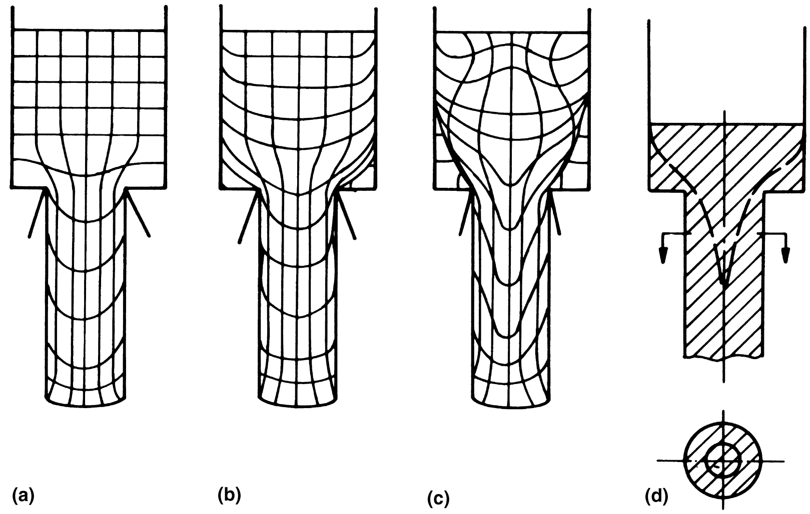


Fig. 17 Material flow in forward extrusion. (a) Without container friction. (b) With container friction. (c) With cooling. (d) Extrusion defect

seeking minimum energy flow, movement at the transition to the flat die is arrested, and a dead-metal zone is formed (Fig. 17b). Deformation now occurs in the body of the billet along the boundary of the dead-metal zone. The internal die angle thus formed increases with decreasing friction and increasing reduction. Dead-metal zone formation in lubricated extrusion may be avoided if the die is made to the appropriate angle. Sliding then takes place over the entire container-die interface but only if the lubricant thins out uniformly without suffering local breakdowns. Otherwise, an incipient dead-metal zone forms, and localized subsurface defects are generated. Extension of the lubricant is substantial; even at a modest extrusion ratio of 16 (extruding a round bar to one-quarter of the billet diameter), the surface increases fourfold.

If the previous conditions cannot be fulfilled, it is much better to extrude without any lubricant whatsoever. Then, sticking friction prevails over the entire container-die interface, and extrusion proceeds by shearing along the container wall (Fig. 18). A dead-metal zone of maximum angle forms, and this angle increases with reduction ratio. It is essential, though, that no trace of a lubricant or contaminant film be present on the billet surface, because otherwise, the sticking condition (A-B in Fig. 19a) is locally disturbed, and the lubricant is washed in below the product surface (at F-E in Fig. 19b) (Ref 24). This can lead to surface lamination and, if the product is subjected to subsequent heat treatment, to blistering.

When a hot workpiece is extruded in a colder container, chilling contributes to the inhomogeneity of deformation, and a very complex flow pattern develops (Fig. 17c). In later stages of extrusion, the dead-metal zone spreads throughout the entire length of the billet. Material near the punch face moves toward the center and carries with it surface oxide films. This leads to the development of the so-called extrusion defect (also called piping): a ring of oxide inclusions causes complete material separation in the form of a cone, the diameter of which increases as extrusion proceeds (Fig. 17d). Chilling in the container can further complicate material flow when the billet is allowed to rest directly on the container. Uneven temperature distribution also results in nonuniform metal flow, which in turn leads to internal stresses and curving of the extruded product.

Unlubricated (hot) extrusion with sticking friction has its advantages, too. The extruded product is smooth and shiny, because its surface is formed by shearing inside the body of the billet. Undesirable side effects can be neutralized by extruding with a punch (or follower block) of slightly smaller diameter than the container, thus leaving a skull that traps all oxides and other foreign material. The formation of the extrusion defect can be prevented by leaving a large butt, and yet, for reasons of economy, extrusion usually continues until a butt of 5 to 20% of the weight of the billet is left. The extruded product is then examined for soundness.

Effect of Friction on Forces. In direct extrusion of a billet, the pressure required to move the billet against the frictional stress, τ_i, has to be added to the extrusion pressure, p_e. The calculation is simple when τ_i is a constant, and also for sticking friction when $\tau_i = k \approx \sigma_0/\sqrt{3}$. The area to be sheared is of circumference $d_0\pi$ and length L, where L is the length of the unextruded billet measured to the die entry (Fig. 8) or to the beginning of the dead-metal zone (Fig. 18). Thus, the pressure at any point in the extrusion stroke is:

$$p_L = p_e + 4\tau_i\,(L / d_0) \qquad \text{(Eq 6)}$$

The punch force drops gradually (Fig. 20) to the minimum p_e measured at the point where the die or dead-metal zone is touched. Sometimes a higher initial breakthrough pressure is registered, which can be attributed to the initiation of flow in the workpiece material and/or to transients associated with the development of the lubricating film.

In extrusion with sliding friction, the magnitude of τ_i must be known. The shear strengths of some lubricating substances (for example, soft metals) are insensitive to interface pressure. In this case, then Eq 6 applies. If, on the other hand, the lubricant is pressure-sensitive and τ_i can be described by a constant μ (as is the case for many boundary lubricants and polymers), the punch pressure at a distance L can be obtained by integration in the following form (Ref 8, p.411):

$$p_L = p_e \exp(4\mu L/d_0) \qquad \text{(Eq 7)}$$

and the extrusion pressure rises exponentially. If, however, the condition $\tau_i = \mu p_L$ is satisfied anywhere along the container wall, sliding at that point is arrested, even in the presence of a lubricant (Fig. 20). The increased pressure (and particularly, breakthrough pressure) limits the

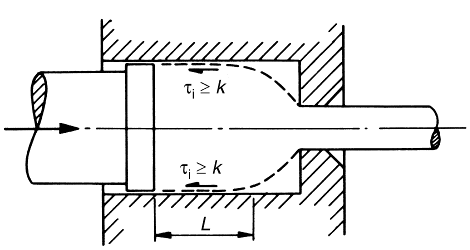


Fig. 18 Forward extrusion with sticking friction

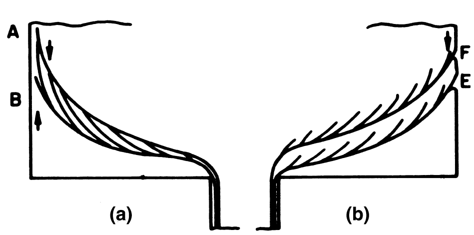


Fig. 19 Material flow in unlubricated extrusion (a) under normal conditions (left side) and (b) with intermittent lubrication (right side)

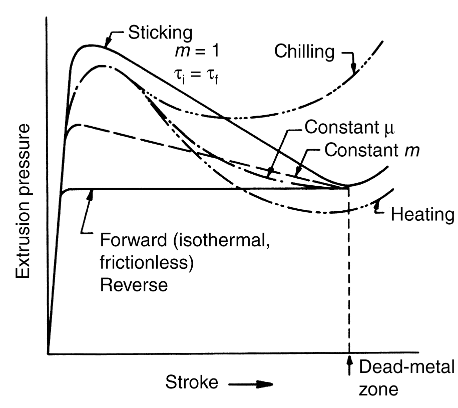


Fig. 20 Variation in extrusion pressure during direct extrusion under various container-friction and thermal conditions. m, friction factor; τ_f, shear flow stress

length of billet that can be directly extruded (Fig. 12).

Evaluation of Friction. In principle, measurement of the punch force at two points in the stroke offers a means of determining the interface shear strength from Eq 6 or the coefficient of friction from Eq 7. Of course, any such evaluation can be valid only if all other variables remain constant. This condition is seldom satisfied, because the mean or average flow stress ($\hat{\sigma}_0$) is bound to vary even in hot extrusion and is known to vary greatly in cold extrusion. If the billet heats up in the course of extrusion (as in cold extrusion and isothermal hot extrusion), the extrusion force drops steeply (Fig. 20), and a misleadingly high friction value is calculated. In contrast, chilling in nonisothermal hot extrusion raises $\hat{\sigma}_0$, and quite unrealistically low friction may be calculated. Sometimes, when the extrusion force actually rises during the stroke (Fig. 20), a negative friction value is obtained. Therefore, coefficients of friction quoted without further specification should be looked on with suspicion.

Nevertheless, the extrusion force and, even more so, its variation in the course of the stroke provide important clues regarding the efficiency of lubrication. They can be used for comparative lubricant evaluation, provided that other conditions are not allowed to change. More reliable quantitative data could be obtained by direct measurement of the frictional shear stress. However, because the high pressures and temperatures often encountered present serious experimental difficulties, few attempts have been recorded.

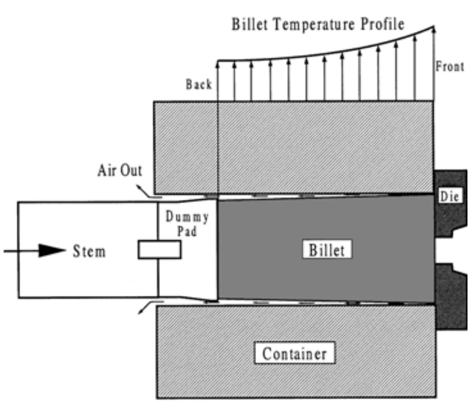


Fig. 21 Bleeding out air during upsetting

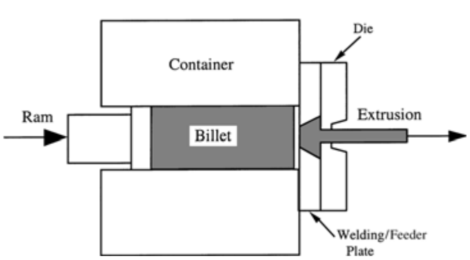


Fig. 22 Continuous-type extrusion using a welding plate in front of the die

Special Processes

Billet-on-billet extrusion is a special technique for alloys that are easily welded together at the extrusion temperature and pressure. Perfect welding of the billet in the container with the next billet must take place as the joint passes through the deformation zone, and the following requirements must be fulfilled (Ref 7):

- Good weldability at the deformation temperature
- Accurate temperature control
- Cleaned billet surface
- Sawn, cleaned billet end free from grease
- Bleeding of air from the container at the start of extrusion, to avoid blisters and other defects (Fig. 21)

Two methods of billet-on-billet extrusion have been developed. In one method, the discard is removed, and the following billet is welded to the one remaining in the welding or feeder plate (Fig. 22). In the other method, there is no discard. The subsequent billet is directly pressed onto the billet still in the container.

Extrusion of Shapes. One of the main advantages of extrusion relative to rolling of shapes is that the dimensions or shape of the extruded product can be changed relatively easily and inexpensively by fabricating a new die or die insert. If the extruded product is to be straight and free of internal stresses, the rate of extrusion must be the same in all parts of the extruded section. The rate of flow reflects the resistance that a given part of the section encounters in exiting from the die. Friction on the die face and over the die land hinders free flow and can be used to equalize flow rates. A thicker section emerges more rapidly than a thinner one and therefore curves the resulting extrusion. Two ways to counteract the tendency of the extrusion to curve is to increase the land length or reduce the die half-angle to increase frictional retardation. Products of small cross-sectional area usually are extruded through multiple holes. Flow rates can then be equalized by similar techniques. For example, another way to counteract this tendency for uneven flow is to incorporate what is called a bleeder hole in the die near the thin portion of an extrusion. The difference between the rates of flow through two holes of different diameters can be used for evaluating frictional conditions.

Tube Extrusion. Hollow products with a variety of cross-sectional shapes may be extruded, in principle, by two direct techniques: one method uses a fixed or floating mandrel, the other uses a bridge-type die (Fig. 23).

Tube Extrusion with Mandrel. A hole pierced or machined in a billet can be maintained in the course of extrusion by a fixed or floating mandrel long enough to reach into the die land (Fig. 23a). A solid billet can be used if a separately actuated piercing ram is available. Lubricated extrusion is possible only with a hollow billet; in working with a piercing ram, sticking friction must be encouraged in order to avoid subsurface defects on the internal tube surface. The effect of friction on material flow is similar to that found in extrusion of bars. Marker pins radially inserted into the billet show the development of a dead-metal zone in unlubricated extrusion (Fig. 24a), whereas pins are successively extruded in lubricated extrusion (Fig. 24b) (Ref 25). The residual skull clearly shows that the original skin remains in the container in the absence of a lubricant but is drawn in during lubricated extrusion.

Tube Extrusion with Bridge-Type Die. Alternatively, a solid billet is extruded through a bridge, spider, or porthole die (Fig. 23b) that divides and then reunites the flowing material to form a hollow product with several internal pressure welds. All traces of lubricant must be carefully kept out; otherwise, the weld quality is impaired. This technique is suitable only for materials that can be extruded at low enough temperatures to retain the strength of the die (in practice, lead and the more readily weldable aluminum alloys). Die design calls for great skill.

The additional frictional surface created on the mandrel or around the bridge adds considerably to the extrusion pressure. Toward the back of the billet, sticking may prevail between mandrel and billet, but toward the deformation zone, the inner surface of the tube slides on the mandrel, subjecting it to large frictional forces. Because the punch is exposed to high temperatures in hot extrusion, it may not be able to withstand these forces and may be torn off or thinned out by plastic flow. Therefore, lubrication of the mandrel surface is usually mandatory, at least above aluminum extrusion temperatures.

Friction-Assisted Extrusion. There are several ingenious schemes (Ref 26) in which friction is actually exploited to aid extrusion. In friction-assisted forward extrusion, the container is moved at a somewhat higher speed than, but in the same direction as, the punch (Fig. 25), and thus, friction on the container wall actually aids the movement of the billet, ensures more homogeneous material flow, and reduces the remnant.

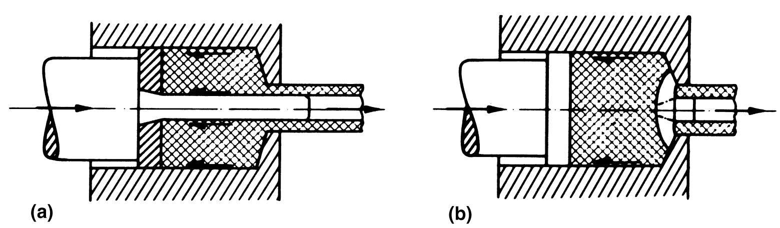


Fig. 23 Forward extrusion of a tube with (a) a fixed mandrel and (b) a bridge-type die

Backward (Indirect) Extrusion

Indirect extrusion is also called inverse, reverse, or back(ward) extrusion, because the billet is stationary in the container cavity, and a punch moving against the billet forces the extrusion to emerge in a direction opposite to the punch movement (Fig. 26a). The crucial point is that the billet remains at rest in the container, and thus, container friction does not come into play. Therefore, the extrusion load and the temperature generated by deformation and friction are reduced, effectively increasing the load capacity of the press. Indirect extrusion is often used for extrusion of harder alloys when shape is critical.

The sequence of operations for the backward extrusion of a solid section is:

1. The die is inserted into the press.
2. The billet is loaded into the container.
3. The billet is extruded, leaving a butt.
4. The die and the butt are separated from the section.

Backward extrusion offers a number of advantages:

- A 25 to 30% reduction in maximum load relative to direct extrusion
- Extrusion pressure is not a function of billet length, because there is no relative displacement between the billet and the container. Therefore, billet length is not limited by the load required for this displacement but only by the length and stability of the hollow stem needed for a given container length.
- Extrusion pressure is more uniform as a function of ram travel compared to direct extrusion (Fig. 27).
- No heat is produced by friction between the billet and the container; consequently, no temperature increase occurs at the billet surface toward the end of extrusion, as is typical in the direct extrusion of aluminum alloys. Therefore, in backward extrusion, there is a lesser tendency toward cracking of the surfaces and edges, and extrusion speeds can be significantly higher.
- The service life of the tooling is increased, especially that of the inner liner, because of reduced friction and temperatures.

The disadvantage of backward extrusion is that impurities or defects on the billet surface affect the surface of the extrusion and are not automatically retained as a shell or discard in the container. As a result, machined billets are used in many cases. In addition, the cross-sectional area of the extrusion is limited by the size of the hollow stem.

The components of a typical tool assembly used for the backward extrusion of steel parts are identified in Fig. 28. There is considerable variation in the tooling practice and design details of tool assembly components. A major problem in indirect extrusion is the construction of the hollow ram and the die; advances in the design of these components have contributed to the extended application of the process (Ref 1, 27). An entirely different form of back extrusion is obtained when a solid punch penetrates the billet.

Material Flow (Ref 8). In indirect extrusion, as in direct extrusion, material flow in the vicinity of the die depends on die geometry, reduction, and the presence of a lubricant. With a good lubricant and a tapered die entry, a homogeneity of flow comparable to that of hydrostatic extrusion can be obtained. In lubricated flow, the dead-metal zone is limited to a thin layer on the die face. The nonextruded part of the billet shows no effect of deformation (Fig. 29b). Because only the extrusion pressure, p_e, needs to be developed, heat generation is also reduced. As can be deduced from Fig. 12, for a given punch and container pressure, either a larger extrusion ratio can be taken or, more importantly for difficult-to-extrude materials, workpiece temperature can be lowered and extrusion speeds increased without running the danger of incipient melting. In direct extrusion, the temperature in the unextruded portion of the billet keeps rising, whereas in indirect extrusion, it remains essentially constant. This is of particular importance in extrusion with sticking friction.

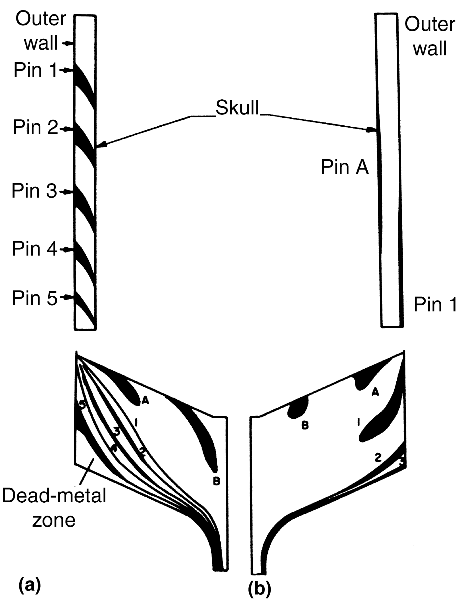


Fig. 24 Material flow in tube extrusion shown by marker pins. (a) Unlubricated. (b) Lubricated

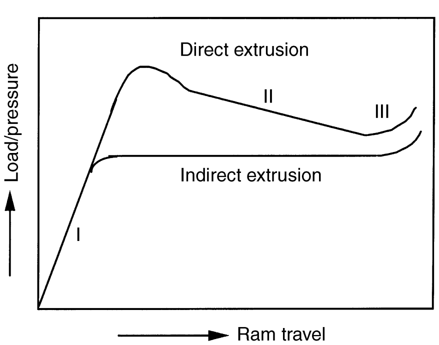


Fig. 27 Variation of load or pressure with ram travel for both direct and indirect extrusion. Region I: The billet is upset, and pressure rises rapidly to a peak value. Region II: So-called steady-state, where pressure drops with direct extrusion. Region III: Pressure reaches a minimum and then rises sharply as the discard is compacted.

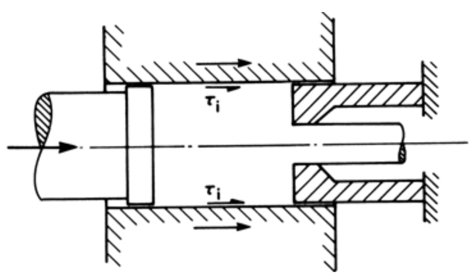


Fig. 25 Friction-assisted forward extrusion

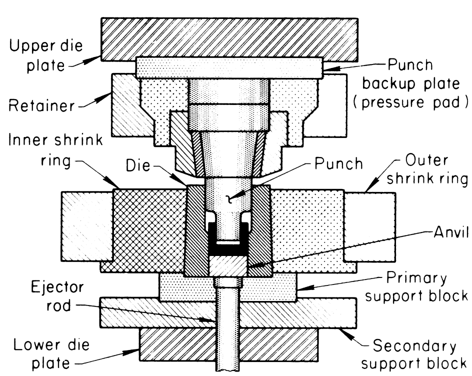


Fig. 28 Tools constituting a typical setup for the cold backward extrusion of steel parts

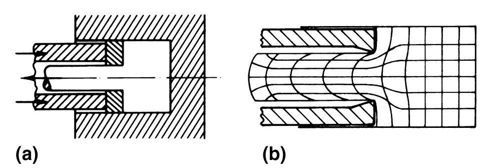


Fig. 26 (a) Backward extrusion and (b) resultant material flow

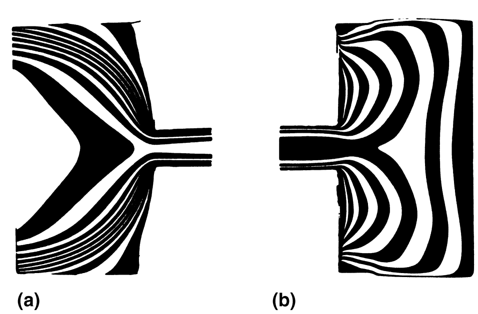


Fig. 29 Material flow in hot extrusion of a composite billet made up of disks of two aluminum alloys. (a) Forward extrusion. (b) Backward extrusion

Combination of Direct (Friction-Assisted) and Indirect Extrusion (Ref 8). A number of attempts have been made to take advantage of container friction. Because friction is strongly influenced by the movement of the billet relative to the container, it is recognized that many of these processes are combinations of direct and indirect extrusion. In one technique (Ref 28), the die is fixed in a hollow stem, which, at the beginning of extrusion, penetrates into the billet. Thus, extrusion begins without a high breakthrough pressure; the punch then advances, and the process continues as direct extrusion.

Several methods of making the process essentially continuous have been proposed. One method (Ref 29) suggests that the material be wedged into the groove of a roller with the aid of a stationary shoe. Friction on the groove walls moves the billet, while friction on the shoe hinders its movement, and the net extrusion force available is governed by the balance of the two. This technique has been used for extrusion of both aluminum and copper. In another technique (Ref 30), a closed pass is formed between a grooved roll and a roll provided with a flange, and the frictional forces developed on the roll surface and on the groove sides are used for extrusion through a die.

A longer contact length, and thus greater frictional force, can be obtained in a linear arrangement. In one process (Ref 31), articulated grips move a square bar from two sides, while lubricated stationary dies prevent its spread in the lateral direction. In another version (Ref 32), the frictional force between a tightly fitting, elastically deforming clamp and a precision billet is used to push the material through the extrusion die. It has also been proposed (Ref 33) to remove the need for close-fitting clamps by using segmented grips moving around all four sides of the billet. None of the previously mentioned processes has achieved production status, partly because of difficulties of construction and product quality. Nevertheless, they provide examples of using rather than combating friction.

Cold Extrusion

Cold extrusion is normally used to create near-net shape products, such as fasteners, automotive components, and so forth. Impact extrusion is a term employed for the cold extrusion of thin-walled products, such as toothpaste tubes. Cold extrusion is used when the process is economically attractive because of:

- Savings in material
- Reduction or elimination of machining and grinding operations because of the good surface finish and dimensional accuracy of cold-extruded parts
- Elimination of heat treating operations because of the increase in the mechanical properties of cold-extruded parts

Cold extrusion is sometimes used to produce only a few parts of a certain type, but it is more commonly used for mass production because of the high cost of tools and equipment. Cold extrusion is often combined with cold forging (e.g., cold heading) for mass production of near-net or net-shape parts, such as bolts, nuts, rivets, and many automotive and appliance components.

Cold extrusion and cold heading are often combined to produce hardware items and machinery parts that require two or more diameters. An intermediate-sized starting piece may be reduced to the smallest diameter by cold extrusion, followed by forming of the large-section portion by heading. Cold extrusion may also follow other operations that shape the starting piece. The use of symmetrical slugs as the starting material for extrusion is common practice, but other shapes are often used. For example, one or more machining operations sometimes precede extrusion in order to produce a shape that can be more easily extruded.

The use of hot upset forgings as the starting material is also common practice. Hot upsetting followed by cold extrusion is often more economical than alternative procedures for producing a specific shape. Axle shafts for cars and trucks are regularly produced by this practice; the advantages include improved grain flow as well as low cost. For example, the rear-axle drive shaft illustrated in Fig. 30 was fabricated by a three-step cold extrusion operation. This process improved surface finish (and thus fatigue resistance), maintained more uniform diameters and closer dimensional tolerances, increased strength and hardness, and simplified production.

The problems most commonly encountered in cold extrusion are:

- Tool breakage
- Galling or scoring of tools
- Workpieces sticking to dies
- Workpieces splitting on outside diameter or cupping in inside diameter
- Excessive buildup of lubricant in dies

Table 2 lists the most likely causes of these problems.

Cold Extrusion Tooling

Metals can be cold extruded by different tooling setups, depending mainly on the size and shape of the workpiece, the composition of the work metal, and the quantity requirements. The principal types of tooling employed are:

Table 2 Problems in cold extrusion and some potential causes

Problem	Potential cause
Tool breakage	Slug not properly located in die
	Slug material not completely annealed
	Slug not symmetrical or not properly shaped
	Improper selection or improper heat treatment of tool material
	Misalignment and/or excessive deflection of tools and equipment
	Incorrect preloading of dies
	Damage caused by double slugging or overweight slugs
Galling or scoring of tools	Improper lubrication of slugs
	Improper surface finish of tools
	Improper selection or improper heat treatment of tool material
	Improper edge or bend radii on punch or extrusion die
Workpieces sticking to dies	No back relief on punch or die
	Incorrect nose angle on punch and incorrect extrusion angle of die
	Galled or scored tools
Workpieces splitting on outside diameter or forming chevrons on inside diameter	Slug material not completely annealed
	Reduction of area either too great or too small
	Excessive surface seams or internal defects in work material
	Incorrect die angles
Excessive buildup of lubricant on dies	Inadequate vent holes in die
	Excessive amount of lubricant used
	Lack of a means of removal of lubricant, or failure to prevent lubricant buildup by spraying the die with an air-oil mist

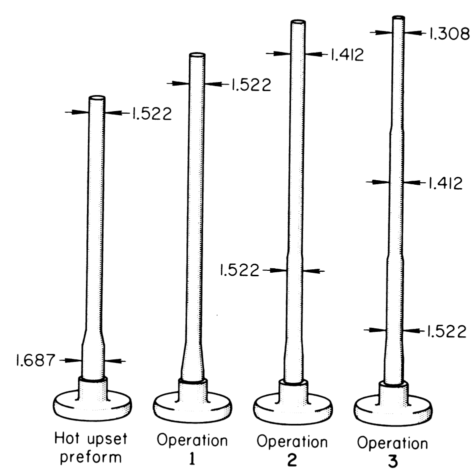


Fig. 30 Rear-axle drive shaft of 1039 steel produced by three-step cold extrusion of an upset forging (billet weight: 36 kg, or 79.5 lb). The drive shafts were hot upset forged to form the flange and to preform the shaft and then were cold extruded to lengthen the shaft. The flange could have been upset as a final operation after the shaft had been cold extruded to length, but this would have required more passes in the extrusion press than space allowed. Hot upsetting and cold extrusion replaced a hammer forging and machining sequence, after which the flange, a separate piece, had been attached. Dimensions given in inches

- *Single-station tooling* forms the part in one stroke of the press. Additional operations may be required for finishing. Closed-end containers, such as toothpaste tubes, are formed in this manner.
- *Multiple-station tooling* involves a series of separate dies arranged so that the rough blank is made into a preform, which then proceeds through successive operations until the required form is produced.
- *Transfer presses* are similar in concept to multiple-station tooling, because they can perform several operations in succession by mechanical transfer of the workpiece from one operation to the next.
- *Upsetters or headers* are used for continuous operation, frequently incorporating both backward and forward extrusion and cold heading.
- *Rotating dial or indexing* can be applied for manual or automatic production, where multiple dies are held in an index dial on the press table.

Punch Design. A major problem in punch design consists of assessing the nature and magnitude of the stresses to which the punch is subjected in service. Because the stresses are dynamic, fatigue effects arise, and these fatigue effects, in conjunction with the inherently brittle nature of hardened tool steels, necessitate care in avoiding design features likely to produce stress concentrations. The stability problems that may arise when slender punches are used are affected by the accuracy of alignment provided by the tool set or the press itself, or by factors in the extrusion operation, such as punch wander, initial centering, and use of distorted slugs. The ratio of punch length to punch diameter also affects stability; a ratio of approximately 3 to 1 is probably the maximum for cold extrusion of steel using tool steel punches.

The design of the punch nose has a significant effect on extrusion pressures and tool life. In backward extrusion, acceptable results are obtained with a nose profile consisting of a truncated cone having an included angle of 170 to 180°, with an edge radius of 0.51 to 2.54 mm (0.020 to 0.100 in.), and a land length of 1.27 to 1.9 mm (0.050 to 0.075 in.), with the shank relieved 0.1 to 0.2 mm (0.004 to 0.008 in.) on the diameter. Although they reduce initial punch stresses, small cone angles or large radii are undesirable because of rapid lubricant depletion and the risk of metal-to-metal contact. Design of the punch nose to distribute the lubricant properly during extrusion is essential for minimizing the pressures developed.

The area ratio between punch shank and head is also an important design factor. A large ratio has the effect of spreading the punch load over a large area of pressure pad. On the other hand, it requires a wider block of metal for its fabrication, with a resultant cost increase. Because pressure pads are less expensive than punches, it is generally advisable to favor the smaller ratios. The pressure pad, which transmits the load from the back of the punch to the die set, should be designed for economy, ease of replacement, and efficiency in reducing the number of punch failures.

Die Design. In forward extrusion, the die is under maximum pressure, and this pressure is not distributed uniformly. Therefore, the tool designer must calculate the hoop (tensile) stresses on the inner die wall and provide adequate reinforcement. Ordinarily, pressures of less than approximately half the yield strength of the die do not require reinforcement, while those in excess of this value do require reinforcement.

Extrusion dies are usually inserted in one or more shrink rings to provide reinforcement. These rings prestress the die in compression by providing interference fits between rings and die. This results in lower working stress and therefore longer fatigue life of extrusion tools. A similar technique is used to shrink radially segmented die inserts together to prevent the segments from separating under load. Permanent shrink-fit assemblies are sometimes made by heating the outer ring to facilitate assembly. Interchangeable die inserts are usually force fitted mechanically, using a tapered press fit and molybdenum disulfide as a lubricant. Of the two methods, shrinking-on by heating is generally preferred, because a cylindrical hole and shaft are easier to fabricate than a tapered hole and shaft. However, a taper fit has several advantages, such as:

- The hardness and yield strength of the various die components are not lowered (as they would be by heating) and can be measured with dependable accuracy.
- The prestress value is ensured by strict control of the input measurements.
- Release and exchange of the inner die bushings is quick, easy, and inexpensive.
- Die parts can be standardized.
- Hot working die steels are not required.

The most commonly used taper angle is $^1/_2$ to 1°. The conditions for obtaining the specified advantages of the taper force fit are careful preparation of the taper shell surfaces and exact agreement between taper angles of corresponding contact faces. If the shell surfaces do not provide uniform support over the entire die length, the prestresses will be unequal, and the reinforcement will not be fully effective.

In some setups, the first reinforcement is applied by taper force fit and the second (outer) reinforcement by shrinking-on. It is advisable to standardize on the size of reinforcing elements. In general, no further advantage is gained by making the outside diameter of a reinforcement more than four to five times the die diameter.

In forward extrusion, die angles are determined by the shape of the workpiece and by the operating sequence. In general, an angle of 2α = 24 to 70° is selected for the forward extrusion of solids, and an angle of 2α = 60 to 126° is preferred for extruding hollow parts, the angle varying inversely with wall thickness. Ejection pressure on the work increases with decreasing die angle, because greater friction must be overcome. This pressure also increases with an increase in the length of the part. Extrusion pressure causes elastic expansion of the die, which shrinks when the pressure is discontinued. Accordingly, very high wall pressures are developed, and these require correspondingly high ejection pressures.

Table 3 Typical tool steels used in extruding aluminum

Tool	AISI steel	Hardness, HRC
Die, solid	W1	65–67
Die sleeve(a)	D2	60–62
	L6	56–62
	H13	48–52
Die button(b)	H11	48–50
	H13	48–50
	L6	50–52
	H21	47–50
	T1	58–60
Ejector	D2	55–57
	S1	52–54
Punch	S1	54–56
	D2	58–60
	H13	50–52
Stripper	L6	56–58
Mandrel, forward	S1	52–54
	H13	50–52
Holder	H11	42–48
	H13	42–48
	4130	36–44
	4140	36–44

AISI, American Iron and Steel Institute. (a) Cemented carbide is sometimes used for die sleeves. (b) Maraging steel is sometimes used for die buttons.

Die Materials. Compressive strength of the punch and tensile strength of the die are among the most important factors influencing the selection of material for cold extrusion tools. Because the die is invariably prestressed in compression by the pressure of inner and outer shrink rings, the principal requirement for a satisfactory die is a combination of tensile yield strength and prestressing that prevents failure. Punches require sufficient compressive strength to resist upsetting without being hazardously brittle. Thus, almost without exception, and particularly for extruding steel, the primary tools in contact with the work must be made from steels that through harden in the section sizes involved.

Recommended materials for extrusion punches typically include M2 and M4 high-speed tool steels and tungsten carbide. Tool steel punches should be heat treated to a hardness of 62 to 66 HRC. Die inserts are usually fabricated from such alloy tool steels as D2, M2, and M4 and are heat treated to 58 to 64 HRC, depending on the steel. Typical tool steels and their working hardnesses for the cold extrusion of aluminum are given in Table 3.

Tungsten carbide is extensively used, because it provides good die life, high production rates, and good dimensional control. Tungsten carbide often finds application as a punch material in backward extrusion. Retainer rings or housings used for tungsten carbide dies should have sufficient strength and toughness to prevent splitting and failure of the working tools. Shrink rings should be fabricated from hot work die steels such as H11 or H13 heat treated to 46 to 48 HRC. Outer housings are often made from H13 die steel or from 4340 alloy steel.

Effects of Lubricant Viscosity. Even though most extrusion of semifabricated products takes place at elevated temperatures, the softer aluminum alloys are extruded at room temperature or at slightly elevated temperatures with the aid of lubrication. Aluminum was extruded at various temperatures around 150 °C (300 °F), using abietic acid so that viscosity could be varied over a wide range (Ref 34). At higher viscosities (lower temperatures) and a lower half-angle, pressure built up fairly gradually until extrusion began. The first part of the section to emerge was fairly bright, typical of boundary contact, but then the surface became duller, indicating the presence of a thicker film.

When conditions are not favorable enough to create a thick film, deformation begins in the die throat, and the surface is much brighter, with only occasional hydrostatic pockets, which is indicative of a predominantly boundary-type, mixed-film lubrication. With excessively high viscosity, the lubricant film becomes so thick that it cannot maintain stability, and periodic collapse leads to development of the bamboo defect. It has been noted that the wavelength of the bamboo defect was equal to the length of the die land, with the defect attributed to intermittent lubricant failure (Ref 35). Most information on this defect comes from research related to hydrostatic extrusion.

Cold Extrusion of Steel

The extrudability of steel decreases with increasing carbon or alloy content. The cold extrusion of steels containing up to 0.45% C is common practice, and steels with even higher carbon contents have been successfully extruded. However, it is advisable to use steels of the lowest carbon content that will meet service requirements. Most carbon and alloy steels that are extruded contain 0.10 to 0.25% C. However, in some applications, steels with more than 0.45% (especially alloy steels) are cold extruded.

For a given carbon content, most alloy steels are harder than plain carbon steels and are therefore more difficult to extrude. Most alloy steels also work harden more rapidly than their carbon steel counterparts; therefore, they sometimes require intermediate annealing. Steels that have been spheroidized by annealing are in their softest condition and are therefore preferred for extrusion. However, operations that precede or follow extrusion may make it impractical to have the steel in its softest condition. Extremely soft steels of low-to-medium carbon content have poor shearability and machinability; therefore, some extrudability is occasionally sacrificed. Annealing techniques that produce a partly pearlitic structure are ideal for many extrusion applications in which shearability or machinability is important.

Free-machining additives, such as sulfur or lead, are likely to impair extrudability. Free-machining steels, containing such additives as lead and sulfur, are not preferred for cold extrusion. Extrusions from these steels are more susceptible to defects than extrusions from their non-free-machining counterparts. In addition, because parts produced by cold extrusion generally require only minimal machining (this is often the primary reason for using cold extrusion), there is much less need for free-machining additives than when parts are produced entirely by machining.

The successful extrusion of free-machining steels depends on the amount of upset, the flow of metal during extrusion, and the quality requirements of the extruded part. Free-machining steels can generally withstand only the mildest upset without developing defects. If it is under compression at all times during flow, a free-machining steel will probably extrude without defects. However, rupture is likely if compressive force is suddenly changed to tensile force.

Nonmetallic inclusions, particularly the silicate type, are detrimental to extrudability. The fewer the inclusions, the more desirable the steel is for cold extrusion. Silicate inclusions have been found to be the most harmful. Therefore, some steels have been deoxidized with aluminum rather than silicon in an attempt to keep the number of silicate inclusions at a minimum. The aluminum-killed steels have better extrudability in severe applications.

Extrusion Quality. Carbon steel bars are available at additional cost in two classes of extrusion quality: cold extrusion quality A and cold extrusion quality B. The mill preparation for cold extrusion quality A is the same as that used for special-quality bars; cold extrusion quality B is a still higher quality. Higher quality refers primarily to fewer external and internal defects. Hot scarfing and more rigorous inspection of the billets are additional operations that are performed at the mill to prepare cold extrusion quality B material.

Alloy steel without a quality extra is used in applications similar to those of cold extrusion quality A for carbon steel. Alloy steels are also available as cold heading quality, which parallels cold extrusion quality B for carbon steel.

Table 4 Relative pressure requirements for the cold extrusion of annealed slugs of five aluminum alloys (alloy 1100 = 1.0)

Alloy	Relative extrusion pressure
1100	1.0
3003	1.2
6061	1.6
2014	1.8
7075	2.3

Boron-modified steels for heading and extrusion are also available. The advisability of paying the additional cost for cold extrusion quality B or cold heading quality steel depends on the severity of extrusion, the quality requirements of the extruded part, and the cost of rejected parts in comparison with the extra cost for these steels.

Cold Extrusion of Aluminum and Its Alloys

Aluminum alloys are well adapted to cold (impact) extrusion. The lower-strength, more ductile alloys, such as 1100 and 3003, are the easiest to extrude. When higher mechanical properties are required in the final product, heat treatable grades are used. Although nearly all aluminum alloys can be cold extruded, the five alloys listed in Table 4 are most commonly used. The alloys in Table 4 are listed in the order of decreasing extrudability based on pressure requirements. The easiest alloy to extrude (1100) has been assigned an arbitrary value of 1.0 in this comparison.

The cold extrusion process of aluminum parts is often considered, because high production rates can be achieved. Even when parts are large or of complex shape, lower production rates may still be economical. The impact-extruded part itself has a desirable structure. It is fully wrought, achieving maximum strength and toughness. It is a near-net shape. There is no parting line, and all that may be required is a trim to tubular sections. Surface finish is good. Impacts have zero draft angles, and tolerances are tight. Once impacted, sections can be treated in the same manner as any other piece of wrought aluminum.

From a design standpoint, aluminum impacts should be considered in the following situations:

- For hollow parts with one end partially or totally closed
- When multiple-part assemblies can be replaced with a one-piece design
- When a pressure-tight container is required
- When bottoms must be thicker than the walls, or the bottom design includes bosses, tubular extensions, projections, or recesses
- When a bottom flange is required

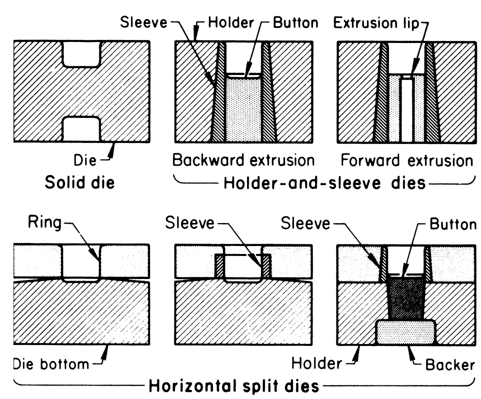


Fig. 31 Three types of dies used in the cold extrusion of aluminum alloy parts

- When bottoms, sidewalls, or heads have changes in section thickness

Equipment is readily available that can produce backward and forward extrusions up to 400 mm (16 in.) in diameter. Backward extrusions can be up to 1.5 m (60 in.) long. The length of forward extrusions is limited only by the cross section of the part and the capacity of the press. Hydraulic extrusion and forging presses, suitably modified, are used for making very large extrusions. Parts up to 840 mm (33 in.) in diameter have been produced by backward extrusion from high-strength aluminum alloys in a 125 MN (14,000 tonf) extrusion press. Similar extrusions up to 1 m (40 in.) in diameter have been produced in large forging presses. Because aluminum extrudes more easily, tools designed especially for extruding aluminum may be different from those used for steel. For example, a punch used for the backward extrusion of steel should not have a length-to-diameter ratio greater than approximately 3 to 1; however, this ratio, under favorable conditions, can be as high as 17 to 1 for aluminum (although a 10 to 1 ratio is usually the practical maximum).

Tooling. Three basic types of dies for extruding aluminum are shown in Fig. 31. Solid dies are usually the most economical to make. Generally, a cavity is provided in each end so that the die can be reversed when one end becomes cracked or worn. Compared to the die cavities used in the backward extrusion of steel, the die cavities for aluminum are notably shallow, reflecting a major difference in the extrusion characteristics of the two metals. Steel is more difficult to extrude, requiring higher pressures and continuous die support of the workpiece throughout the extrusion cycle. In contrast, aluminum extrudes readily, and when the punch strikes the slug in backward extrusion, the metal squirts up the sides of the punch, following the punch contours without the external restraint or support afforded by a surrounding die cavity.

Holder-and-sleeve dies are used when extrusion pressures are extremely high. This type of die consists of a shrink ring or rings (the holder), a sleeve, and an insert (button). The die sleeve is prestressed in compression in the shrink ring to match the tension stress expected during extrusion. Horizontal split dies are composed of as many as four parts: a shrink ring, a sleeve (insert), and a one-piece or two-piece base. Figure 31 identifies the one-piece base as a die bottom, and the components of the two-piece base as a holder and a backer.

Stock for Slugs. In general, the methods for preparing aluminum slugs are similar to those for preparing slugs from other metals and are therefore subject to the same advantages and limitations. Many extrusions are produced directly from slugs purchased in the O (annealed, recrystallized) temper. In other applications, especially when slugs are machined from bars, the slugs are annealed after machining and before surface preparation. When extruding alloys that will be heat treated, such as 6061, common practice is to extrude the slug in the O temper, solution treat the preform to the T4 temper, and then size or finish extrude. This procedure has two advantages. First, after solution treatment, the metal is reasonably soft and permits sizing or additional working, and, second, the distortion caused by solution treatment can be corrected in final sizing. After sizing, the part can be aged to the T6 temper, if required.

Slugs for extrusions are obtained by blanking from plate; by sawing, shearing, or machining from bars; or by casting. Rolled aluminum alloy plate is widely used as a source of cold extrusion stock. The high speed at which slugs can be prepared is the major advantage of blanking from rolled plate. When slug thickness is greater than approximately 50 mm (2 in.) or when the thickness-to-diameter ratio is greater than approximately 1 to 1, blanking from plate is uneconomical, if not impossible. Blanking is also excessively wasteful of metal, which negates a principal advantage of the cold extrusion process.

Sawing from bars is widely used as a method of obtaining slugs. More accurate slugs are produced by sawing than by blanking; however, as in blanking, a considerable amount of metal is lost. When "doughnut" slugs are required, they can be sawed from tubing, or they can be punched, drilled, or extruded. Machined slugs (such as those produced in an automatic bar machine) are generally more accurate but cost more than those produced by other methods. Cast slugs can also be used; the selection of a cast slug is made on the basis of adequate quality at lower fabricating cost. Compositions that are not readily available in plate or bar stock can sometimes be successfully cast and extruded. There is often a savings in metal when a preform can be cast to shape.

Surface Preparation and Lubricants. Slugs of the more extrudable aluminum alloys, such as 1100 and 3003, are often given no surface preparation before a lubricant is applied prior to extrusion. For slugs of the less extrudable aluminum alloys or for maximum extrusion severity or both, surface preparation may be necessary for retention of lubricant. One method is to etch the slugs in a heated caustic solution, followed by water rinsing, nitric acid desmutting, and a final rinse in water. For the most severe extrusion, slug surfaces are given a phosphate coating before the lubricant is applied.

Aluminum and aluminum alloys can be successfully extruded with such lubricants as high-viscosity oil, grease, wax, tallow, and sodium-tallow soap. Zinc stearate, applied by dry tumbling, is an excellent lubricant for extruding aluminum. In applications in which it is desirable to remove the lubricant, water-soluble lubricants are used to reduce the wash cycle. The lubricant should be applied to metal surfaces that are free from foreign oil, grease, and dirt. Preliminary etching of the surfaces (see previous information) increases the effectiveness of the lubricant. For the most difficult aluminum extrusions (less extrudable alloys or greater severity or both), the slugs should be given a phosphate treatment, followed by application of a soap that reacts with the surface to form a lubricating layer similar to that formed when extruding steel.

Other Nonferrous Alloys

Cold Extrusion of Copper and Copper Alloy Parts. Oxygen-free copper (C10200) is the most extrudable of the coppers and copper-base alloys. Other grades of copper and most of the copper-base alloys can be cold extruded, although there are wide differences in extrudability among the different compositions. For example, the harder copper alloys, such as aluminum-silicon bronze and nickel silver, are far more difficult to extrude than the softer, more ductile alloys, such as cartridge brass (alloy C26000), which can satisfactorily withstand cold reduction of up to 90% between anneals. Alloys containing as much as 1.25% Pb can be successfully extruded if the amount of upset is mild and the workpiece is in compression at all times during metal flow. Copper alloys containing more than 1.25% Pb are likely to fracture when cold extruded.

The pressure required for extruding a given area for one of the more extrudable coppers or copper alloys (such as C10200 or C26000) is less than that required for extruding low-carbon steel. However, the pressure required for extruding copper alloys is generally two to three times that required for extruding aluminum alloys (depending on the copper or aluminum alloy being compared).

The length of a backward-extruded section is limited by the length-to-diameter ratio of the punch and varies with unit pressure. This ratio should be a maximum of 5 to 1 for copper. A ratio of 10 to 1 is common for the extrusion of aluminum, and ratios as high as 17 to 1 have been used. The total reduction of area for copper or copper alloys, under the best conditions, should not exceed 93%. In applications involving minimum-to-moderate severity, copper slugs are often extruded with no special surface preparation before the lubricant is applied. However, for the extrusion of harder alloys (aluminum bronze, for example) or for maximum severity or both, best practice includes the following surface preparation before the lubricant is applied:

- Cleaning in an alkaline cleaner to remove oil, grease, and soil
- Rinsing in water
- Pickling in 10 vol% sulfuric acid at 20 to 65 °C (70 to 150 °F) to remove metal oxides
- Rinsing in cold water
- Rinsing in a well-buffered solution, such as carbonate or borate, to neutralize residual acid or acid salts
- Lubrication. Zinc stearate is an excellent lubricant for extruding copper alloys. Common practice is to etch the slugs as described previously and then to coat them by dry tumbling in zinc stearate.

Impact Extrusion of Magnesium Alloys. Impact extrusion is used to produce symmetrical tubular magnesium alloy workpieces, especially those with thin walls or irregular profiles for which other methods are not practical. As applied to magnesium alloys, the extrusion process cannot be referred to as cold, because both blanks and tooling must be preheated to not less than 175 °C (350 °F); workpiece temperatures of 260 °C (500 °F) are common. Pressures for the impact extrusion of magnesium alloys are approximately half those required for aluminum and depend mainly on alloy composition, amount of reduction, and operating temperature.

Length-to-diameter ratios for magnesium extrusions may be as high as 15 to 1. There is no lower limit, but parts with ratios of less than approximately 2 to 1 can usually be press drawn at lower cost. A typical ratio is 8 to 1, and parts with higher length-to-diameter ratios are more amenable to forward extrusion than to backward extrusion. At all ratios, the mechanical properties of magnesium extrusions normally exceed those of the blanks from which they are made, because of the beneficial effects of mechanical working.

Dies for the impact extrusion of magnesium alloys differ from those used for other metals, because magnesium alloys are extruded at moderately elevated temperature (usually 260 °C, or 500 °F). Common practice is to heat the die with tubular electric heaters. The die is insulated from the press, and an insulating shroud is built around the die. The top of the die is also covered, except for punch entry and the feeding and ejection devices. The punch is not heated, but it becomes hot during continuous operation; therefore, the punch should be insulated from the ram.

Punches and dies are usually made of a hot work tool steel, such as H12 or H13, heat treated to 48 to 52 HRC. In one application, tools made of heat treated H13 produced 200,000 extrusions. Carbide dies can be used and can extrude up to 10 million pieces. The sidewalls of the die cavity should have a draft of approximately 0.002 mm/mm of depth, which prevents the extrusion from sticking in the cavity. In normal operation, the part stays on the punch and is stripped from it on the upward stroke.

The tolerances for magnesium alloy extrusions are influenced by the size and shape of the part, the length-to-diameter ratio, and the press alignment. Magnesium has a relatively high coefficient of thermal expansion compared to steel. Therefore, in order to ensure that the magnesium extrusion, when cooled to room temperature, will be within dimensional tolerance, it is necessary to multiply the room-temperature dimensions of steel tools by a compensatory factor for the temperature at which the magnesium alloy is to be extruded.

Hot Extrusion

The techniques for extrusion of different materials are dependent, to a large extent, on the extrusion temperature. The essence of hot working is to reduce the yield stress and thus increase extrusion speed for a given press load. However, if the exit temperature is too high, surface liquation can result in hot shortness (that is, surface and surface cracks, such as in Fig. 32). The optimal extrusion temperature, $T_{optimal}$, is found at the intersection of the load limiting curve and the allowable extrusion speed for a given surface liquation temperature (Fig. 33). In terms of the right-side curve for hot shortness of an alloy, a lower overall exit temperature allows a higher extrusion speed, because more surface heating (from friction and less time for heat dissipation) can occur without surface melting.

When lubrication is used with extrusion, the right-side curve in Fig. 33 is also shifted to the right, because friction is reduced. Conventional hot extrusion may be either nonlubricated or lubricated (Fig. 4). As previously noted in the section "Friction and Lubrication" in this chapter, container lubrication should be avoided if non-homogenous flow (type B or C flow) cannot be excluded. Generally, hot extrusion of aluminum alloys is done without lubrication, while copper alloys, titanium alloys, alloy steels, stainless steels, and tool steels are extruded with a variety of grease/graphite or glass-based lubricants.

In nonlubricated hot extrusion, the material flows by internal shear, and a dead-metal zone is formed in front of the extrusion die. Deformation occurs with shearing the alloy at the die-to-billet interface. Only lower-melting alloys, including aluminum alloys, can be extruded hot without a lubricant. Two-dimensional shapes of great complexity and thin walls (e.g., architectural extrusions) can be made at a relatively low die cost. In the absence of a lubricant, divided material streams can be reunited and welded in a bridge-type die, allowing extrusions with one or more closed cavities, in a very wide size range (multihole tubes are extruded with walls as thin as 0.25 mm, or 0.010 in.). Nonlubricated hot extrusion is a relatively straightforward process once the conditions have been defined, as briefly described in the section "Unlubricated Hot Extrusion of Aluminum Alloys" in this chapter.

Hot extrusion of higher-melting alloys requires a high-temperature lubricant (usually glass or grease) between the extruded billet and the die. Shapes are more limited, and wall thickness is greater. Lubrication in extrusion reduces load and energy requirements, reduces tool wear, improves surface finish, and provides a product with nearly uniform properties. This technique is commonly used in the extrusion of shapes from steels, titanium alloys, and nickel alloys. Proper die design is critical in lubricated extrusion, especially when noncircular shapes are extruded. An effective die design must ensure smooth metal flow with consistent lubrication. It is desirable to use shaped dies, which provide a smooth transition for the billet from the round or rectangular container to the shaped-die exit.

Critical parameters for successful and economical hot extrusion include the method of billet preparation and heating, the amount of pres-

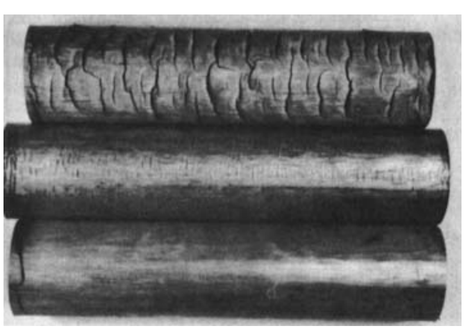


Fig. 32 Surface cracks (hot shortness) of different degrees of severity on an extruded CuSn8 tube

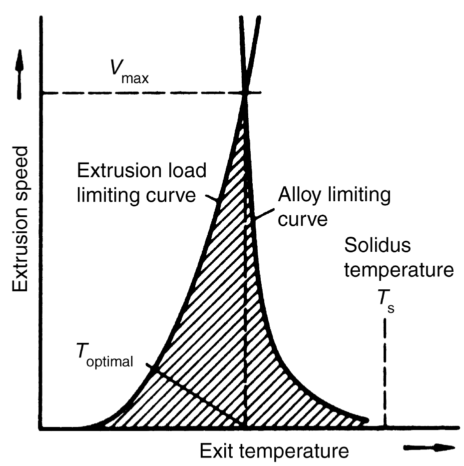


Fig. 33 Limit diagram of extrusion speed, V, versus temperature for a given extrusion load and the alloy limit for surface cracking (hot shortness). Note: This optimal temperature only refers to extrusion speed and not metallurgical development of properties.

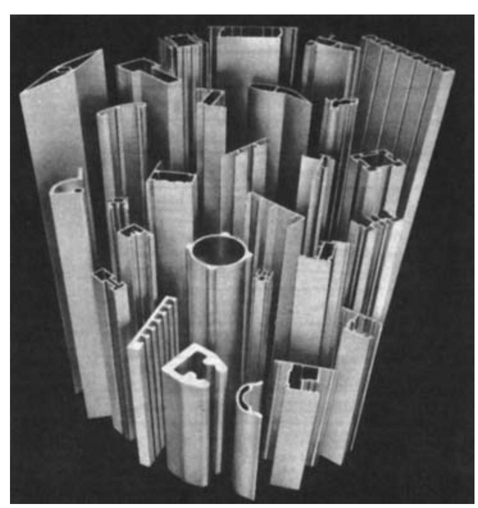


Fig. 34 Examples of extruded sections produced from easily extrudable aluminum alloys

sure and rate of speed used for extruding, and the type of lubricant employed. A billet temperature that is too high can cause blisters or other surface defects, including cracking. A temperature that is too low increases the pressure requirements for the extrusion and shortens tool life. All metals also shrink on cooling after hot extrusion; therefore, a shrinkage allowance must be provided in the design of the dies. Deformation of the die under high pressures and expansion resulting from the high temperatures must also be considered in die design.

Materials for Hot Extrusion

Figure 1 classifies various materials on the basis of typical extrusion temperatures. Commonly extruded metals include aluminum alloys, lead, tin, magnesium, zinc, and steel. Other metals that are hot extruded include titanium and titanium alloys, nickel and its alloys, superalloys, zirconium, beryllium, uranium, and molybdenum. Some titanium alloys are more difficult to extrude than steels. Nickel alloys also can be very difficult to extrude, and billet temperatures above 995 °C (1825 °F) are used. All of these metals are extruded into tubes, rods, and bars.

Lead and tin exhibit high ductility and are easy to extrude. The addition of alloying elements increases the force required, but extruding does not present a problem and is carried out with billets heated to a maximum temperature of approximately 300 °C (575 °F). Principal applications include pipes, wire, tubes, and sheathing for cable. Molten lead is used instead of billets for many applications. Vertical extrusion presses are sometimes used to produce protective sheathings of lead on electrical conductors.

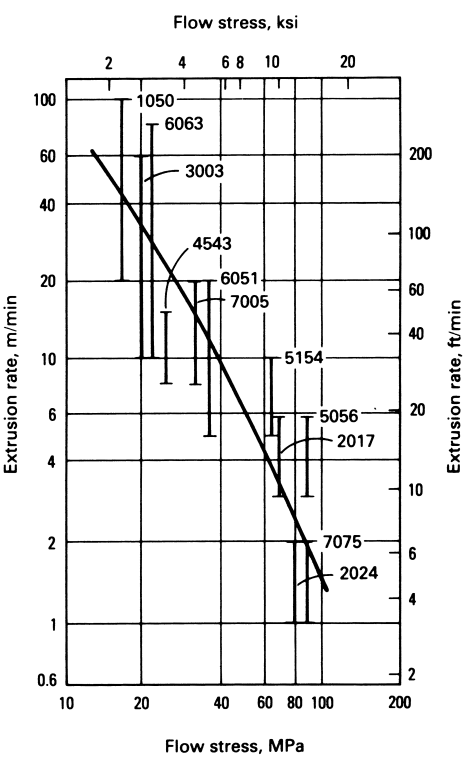


Fig. 35 Extrusion rate versus flow stress for various aluminum alloys. Source: Adapted from Ref 1

Aluminum and aluminum alloys are probably the ideal materials for extrusion. In the aluminum industry, rod, bar, wire tubular products, and shapes are termed mill products, as they are in the steel industry, even though they often are produced by extrusion rather than by rolling. Most commercially available aluminum alloys can be extruded, and any shape can be produced from easily extrudable alloys. Examples of various sections produced from easily extrudable aluminum alloys are shown in Fig. 34 (Ref 1). Most design-for-manufacturing considerations for aluminum extrusions depend on the difficulty of extruding through small, complex die openings. Hollow sections are quite feasible, although they cost approximately 10% more per pound produced. The added cost is often compensated for by the added torsional stiffness that the hollow shape provides. It is best if hollow sections can have a longitudinal plane of symmetry. Semihollow features should be avoided, because semihollow features require the die to contain a very thin—and hence relatively weak—neck. Sections with both thick and thin sections are to be avoided. Metal tends to flow faster where thicker sections occur, giving rise to distortions in the extruded shape.

The relative extrudability of aluminum alloys, as measured by extrusion rate, is given subsequently for several of the more important commercial extrusion alloys:

Alloy	Extrudability, % of rate for 6063
1350	160
1060	135
1100	135
3003	120
6063	100
6061	60
2011	35
5086	25
2014	20
5083	20
2024	15
7075	9
7178	8

In general, higher alloy content and strength increase the difficulty of extrusion and lower its extrusion rate. The easily extruded alloys can be economically extruded at speeds up to 100 m/min (330 ft/min) or faster, while speeds for moderately difficult or very difficult alloys are very low (Fig. 35). With a typical extrusion ratio of 40 to 1, exit speeds of the more difficult alloys are in the order of 0.6 to 1.2 m/min (2 to 4 ft/min).

Extrudability of the moderately difficult or very difficult alloys cannot be significantly increased by hot extrusion technology because of the narrow temperature interval between the extrusion-load limiting temperature and the temperature of hot shortness. In extrusion of high-magnesium alloys, a short cylinder of soft aluminum helps to initiate more homogeneous flow (Ref 36). Billet temperatures generally range from approximately 300 to 595 °C (575 to 1100 °F), depending on the alloy. Typical billet temperatures for some of the harder aluminum alloys are listed in Table 5. This does not include variations in exit temperatures.

Magnesium and Magnesium Alloys. Extruded magnesium and magnesium alloy products are used in the aircraft, aerospace, and nuclear power industries. With similar billet temperatures, the extrudability of these materials is approximately the same as that of aluminum, but longer heating periods are usually necessary to ensure uniform temperatures throughout the billets.

Zinc and Zinc Alloys. The extrusion of zinc and zinc alloys requires pressures that are higher than those necessary for lead, aluminum, and magnesium. Billet temperatures generally range from approximately 205 to 345 °C (400 to 650 °F). Applications include rods, bars, tubes, hardware components, fittings, and handrails.

Copper and copper alloy extrusions are widely used for wire, rods, bars, pipes, tubes, electrical conductors and connectors, and welding electrodes. Architectural shapes are extruded from brass but usually in limited quantities. Billet temperatures vary from approximately 595 to 995 °C (1100 to 1825 °F). Depending on the alloy, extrudability ranges from easy to difficult. High pressures (690 MPa, or 100 ksi, or more) are necessary for the extrusion of many copper alloys.

Steels. For the hot extrusion of steel, it is necessary to use glass or some other high-temperature lubricant to prevent the excessive tooling wear that can result from the high billet temperatures required (995 to 1300 °C, or 1825 to 2375 °F). In addition, high ram speeds are required in order to minimize contact time between the billets and the tooling. The products produced include structural sections (generally required in small quantities) and tubes with small bores. For economic reasons, steel structural shapes, espe-

Table 5 Typical values of billet temperature and extrusion speed of some harder Al alloys

		Billet temperature		Exit speed	
Alloy	Type	°C	°F	m/min	ft/min
2014–2024	Heat treatable	420–450	788–842	1.5–3.5	5–11
5083, 5086, 5456	Non-heat-treatable	440–450	824–842	2–6	7–20
7001	Heat treatable	370–415	700–780	0.5–1.5	2–5
7075, 7079	Heat treatable	300–460	572–860	0.8–2	3–7
7049, 7150, 7178	Heat treatable	300–440	572–824	0.8–1.8	2.5–6

Note: Temperatures and extrusion speeds are dependent on the final shape and the extrusion ratio, and it may be necessary to start with lower billet temperatures than mentioned in the table. Source: Ref 1, 7

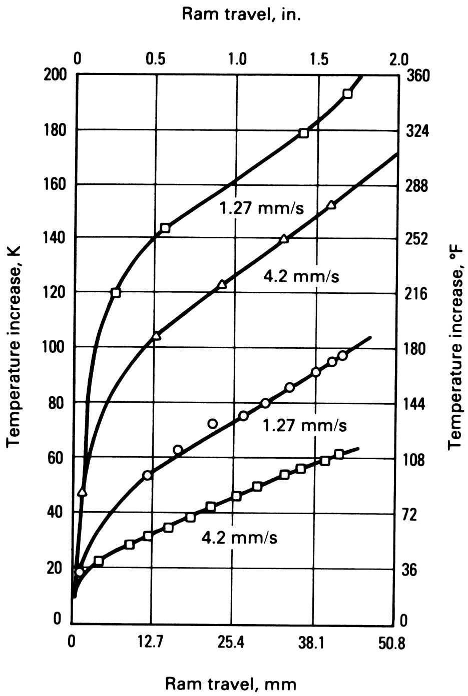


Fig. 36 Increase in emergent temperature versus ram travel in the extrusion of superpure aluminum. Ram speeds are indicated on the curves. Billet diameter: 38 mm (1.5 in.); billet length: 51 mm (2 in.); extrusion ratio: 16:1; starting temperature: 20 °C (70 °F). Source: A.R.E. Singer and J.W. Coakham, Temperature Changes Occurring During the Extrusion of Aluminum, Tin and Lead, *J. Inst. Met.,* Vol 89, 1961–1962, p 177

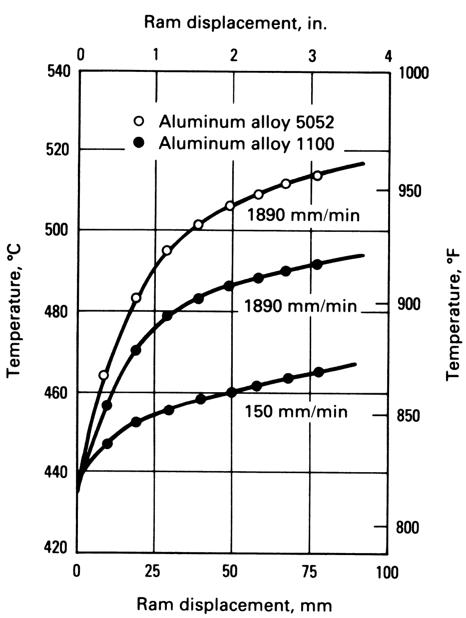


Fig. 37 Surface temperature of the extruded product versus ram displacement for two aluminum alloys. Ram velocities are indicated on the curves. Reduction ratio: 5:1; billet diameter: 71 mm (2.8 in.); billet length: 142 mm (5.6 in.); initial billet and tooling temperature: 440 °C (825 °F). Source: G.D. Lahoti and T. Altan, Prediction of Metal Flow and Temperatures in Asymmetric Deformation Processes, *Proceedings of the 21st Sagamore Army Materials Research Conference,* Aug 1974

cially those needed in large quantities, are better suited to the rolling process. Alloy and stainless steels are usually extruded in the form of either solid shapes or tubes.

Metal powders are extruded into long shapes by cold and hot processes, depending on the characteristics of the powders. Aluminum, copper, nickel, stainless steels, beryllium, and uranium are some of the powders that are extruded. The powders are often compressed into billets that are heated before being placed in the extrusion press. For many applications, the powders are encapsulated in protective metallic cans, heated, and extruded with the cans. Extrusion of metal powders is discussed in detail in the article "Extrusion of Metal Powders" in *Powder Metal Technologies and Applications,* Volume 7 of the *ASM Handbook.*

Process Factors

Exit Temperature. The temperature of the extruded product as it emerges from the die is one of the essential factors that influence product quality. The temperatures developed during extrusion significantly influence the speed at which the process can be carried out. This is especially true for alloys that are difficult to extrude, where the flow stress of the extruded material must be kept relatively low by increasing the billet preheating temperature.

In general, a higher rate of production can be obtained by increasing extrusion speed and/or the extrusion ratio (for a larger reduction in area). However, the flow stress of the extruded material must be kept relatively low to maintain extrusion pressure at an acceptable level. One way is to increase the billet preheating temperature, but this combination of high billet temperature with a large reduction in area and/or high extrusion speed causes a considerable rise in temperature in the extruded material, especially near the surface of the section.

Due to the mechanics and thermodynamics of the extrusion process, exit temperatures also vary with a constant ram speed. Data from this effect are plotted for aluminum and aluminum alloys in Fig. 36 and 37, respectively. Exit temperatures are influenced by a complex set of relationships among the material and process variables, such as billet material and temperature, friction, tool material and temperature, extrusion speed, shape of the extruded section, and reduction in area. This complex thermal situation exists as soon as a heated billet is loaded into the preheated container and extrusion begins.

Various methods have been investigated and used to maintain a more constant exit temperature during extrusion. These methods, referred to as isothermal extrusion, include:

- Reducing exit speed during extrusion with controls according to measurement of exit temperature
- Reducing extrusion speed according to preselected speed programs
- Nonuniform taper heating to give a lower temperature at the back of the billet

Taper heating (Fig. 21) is perhaps the most practical method of isothermal production in a production environment. Additional heaters can be used to transfer more heat to the front of the billet, or the back end can be cooled as the billet is transferred to the container. Several methods for isothermal extrusion of aluminum alloys are also described in Ref 7.

Billet Preparation. The more common metals that are to be extruded are generally cast in the form of cylindrical logs measuring 3.7 to 6 m (12 to 20 ft) or more in length. These logs are sawed or sheared into billets of varying length, depending on the cross-sectional area and the length of the product to be extruded. Additional billet preparation is sometimes necessary, depending on the material to be extruded. For example, it is necessary to machine the outer surfaces of some steel billets before they are heated; the outer surfaces must then be descaled after being heated to the extrusion temperature. Best results are attained in backward extrusion by scalping the billets before extrusion to remove oxides and other impurities from the billet skin. If this is not done, these impurities would find their way onto the surfaces of the extrusion because of the inherent nature of the metal flow in backward extrusion.

For the hot extrusion of such materials as brass, bronze, and other soft metals, the dummy block is made smaller in diameter than the billet. In extruding, no lubrication is provided between the bore of the container liner and the outer surface of the billet. Consequently, friction prevents the outer surface of the billet from sliding, and the undesirable skin of the billet is left in the container as the dummy block shears the metal during its forward stroke. An additional press stroke is required to remove this retained metal before the next billet can be charged into the container.

Pressure Requirements. The unit pressures needed for hot extrusion are significant considerations in press selection. The determination of pressure requirements is difficult for the extrusion of complicated shapes and sections—especially those with thin walls. Careful judgments based on past experience must be made for estimates. Formulas have been developed for estimating pressure requirements, using shape, friction, and other parameters. However, for less complicated shapes, such as round bars and tubes, a fair approximation of pressure requirements can be calculated by:

$$p = k \ln\left(\frac{A}{a}\right) \qquad \text{(Eq 8)}$$

where p is the extrusion pressure required; k is a numerical value representing the resistance to deformation, usually based on past experience in extruding a specific metal at a specific temperature; A is the cross-sectional area of the container liner or, in the case of tubes or other hollow shapes, the cross-sectional area of the liner minus the cross-sectional area of the mandrel (in square inches or square millimeters); and a is the total cross-sectional area of the extruded product (the shape area times the number of openings in the die) (in square inches or square millimeters).

The extrusion pressure requirements determined with Eq 8 are useful, but the values obtained are only approximations. The factor k varies with such factors as billet temperature, die design, type of metal extruded, amount of reduction (extrusion ratio), stem speed, and configuration of the extruded product. Billet length, nonhomogeneous metal flow, and friction also influence pressure requirements.

Unit pressures generally range from 450 to 760 MPa (65 to 110 ksi), with a maximum of approximately 1035 MPa (150 ksi). When practical, it is generally desirable to use a press with a capacity exceeding that actually required. This allows lower billet temperatures and faster stem speeds to be used and provides improved properties in the extruded products.

Ram Speeds. Optimal stem speeds are essential for hot extrusion. Excessive speed can cause overheating of the billet as well as tears and other surface defects. A speed that is too slow reduces productivity and increases the required extrusion pressure because of billet cooling. Slow speeds can also decrease tool life because of prolonged contact time between the tools and the hot billet. Typical stem speeds for various metals are:

	Stem speed	
Material	mm/s	in./s
Steel	152–203	6–8
Copper	51–76	2–3
Aluminum	12.7–25.4	–1
Brass	25–51	1–2

The use of variable-delivery pumps and adjustable valves facilitates control of stem speed. Automatic control is available for maintaining constant speed throughout the extruding cycle.

Lubrication is another important operating parameter. The types of lubricants used and the effects of lubrication are discussed in the section "Lubricated Hot Extrusion" in this chapter.

Dies and Tooling. The extrusion of long bars and tubes imposes some of the most severe demands on the various tooling elements. High pressures and sliding velocities combine with adhesive and abrasive action and, in hot extrusion, with sudden temperature fluctuations and prolonged exposure to high temperatures. It is, therefore, not surprising that some of the improvements in tool materials were first applied to extrusion dies.

Hot working die steels are generally adequate except in isothermal extrusion of titanium, for which superalloy dies are more satisfactory. Much work has gone into increasing wear resistance by selection of appropriate steels, by use of surface treatments such as nitriding and W_2C and W_3C coating, and by use of tungsten carbide inserts. Ceramic coatings, particularly ZrO_2 and Al_2O_3, have come into general use in extrusion of iron- and nickel-base alloys and have also found application as solid die inserts. The low thermal conductivity of many high-temperature materials is a disadvantage in that high tool temperatures can build up. The molybdenum-base alloy TZM has been shown to give a better compromise in properties.

Follower blocks suffer mostly from thermal fatigue and, if overloaded, from plastic deformation. The materials employed are similar to those used for containers. Mandrels are subject to high heat and frequently fail by plastic deformation. Therefore, they are often cooled internally or, between strokes, externally.

Lubricated Hot Extrusion. The Ugine-Sejournet process is the most commonly used for the extrusion of steels and titanium alloys. In this process, the heated billet is rolled over a bed of ground glass or is sprinkled with glass powder to provide a layer of low-melting glass on the billet surface. Before the billet is inserted into the hot extrusion container, a suitable lubricating system is positioned immediately ahead of the die. This lubricating system can be a compacted glass pad, glass wool, or both. The prelubricated billet is quickly inserted into the container, along with the appropriate followers or a dummy block. The extrusion cycle is then started.

As a lubricant, glass exhibits unique characteristics, such as its ability to soften selectively during contact with the hot billet and, simultaneously, to insulate the hot billet material from the tooling. The tooling is usually maintained at a temperature that is considerably lower than that of the billet. In the extrusion of titanium and steel, the billet temperature is usually 1000 to 1250 °C (1830 to 2280 °F), but the maximum temperature the tooling can withstand is 500 to 550 °C (930 to 1020 °F). Therefore, compatibility can be attained only by using the appropriate lubricants, insulative die coating, and ceramic die inserts and by designing dies to minimize tool wear. Glass lubricants have performed satisfactorily on a production basis in extruding long lengths.

The choice between grease and glass lubricants is based mainly on the extrusion temperature. At low temperatures, lubrication is used only to reduce friction. At moderate temperatures, there is also some insulation between the hot billet and the tooling from the use of partially molten lubricants and vapor formation in addition to the lubrication effect. At temperatures above 1000 °C (1830 °F), the thermal insulation of the tooling from overheating is of equal importance to the lubricating effect, particularly with difficult-to-extrude alloys. The lubrication film can also impede oxidation.

Lubricated Hot Extrusion of Aluminum Alloys. Most aluminum alloys are hot extruded without any lubricant whatsoever. Lubricated hot extrusion of aluminum alloys has found only limited application. It is not easy to identify an obviously best lubricant. There are few glasses of the right viscosity for phase-change lubrication, and the rough product surface typical of the plastohydrodynamic (PHD) mechanism is inferior to the accustomed surface. Graphite cannot maintain sliding friction over the entire course of extrusion of a long billet, although both graphite and MoS_2 have been used for mandrel lubrication in tube extrusion. Polytetrafluoroethylene has been found to be useful in cable extrusion. However, unlubricated hot extrusion of aluminum alloys remains dominant.

Unlubricated Hot Extrusion of Aluminum Alloys. By far, the largest quantity of aluminum alloys is hot extruded without any lubricant whatsoever. As long as the billet is of good, uniform quality, the die is free of major pickup, and no lubricant or contaminant finds it way into the container, the quality of the extrusion will be excellent (Ref 36). For best surface quality, a follower block of smaller diameter is used to leave a skull, which is removed by specially designed follower blocks or in a separate stroke. The skull entraps imperfections of the billet surface. When extrusion is done without a skull, the billet is scalped for highest quality. Shearing along the container wall contributes to heat generation; in direct extrusion it may reach 50% of the total force (Ref 27) and thus greatly adds to the temperature rise. Reverse extrusion imposes much smaller strains. The work of shearing along the container wall does not contribute to an increase in the temperature of the extruded product, and a threefold increase in speed is possible (Ref 36).

In nonlubricated hot extrusion of aluminum alloys, a flat-face (shear-face) die is often used. The main problem in extrusion with sticking friction is the die. The die is designed and made so that optimal material flow is ensured and an extrusion free of defects is produced. The die usually has a flat face, which is protected by the dead-metal zone. However, the high adhesion of aluminum to steel soon results in the buildup of a coating on the die and die land. Local lubrication of the die land or of the billet face with a graphitic lubricant is sometimes practiced, even though the lubricant is lost soon after the extrusion emerges (Ref 37) and extrusion proceeds over the bare die. No lubricant is allowed in some press shops, because it may encourage attempts at more extensive lubrication, with consequent material defects, as shown in Fig. 19. Die coatings can be helpful. Pickup is typical of bare tool steel.

Hydrostatic Extrusion

In hydrostatic extrusion, the billet is pushed through the die by the action of a liquid pressure medium rather than by direct application of the load with a ram. The billet is completely surrounded by a fluid that is sealed off and pressurized to extrude the billet through the die (Fig. 3b). For all practical purposes, hydrostatic extrusion is a form of lubricated extrusion, where pressure applied to the inlet zone aids in developing a hydrodynamic film. Hydrostatic extrusion has attracted a great amount of interest, because envelopment of

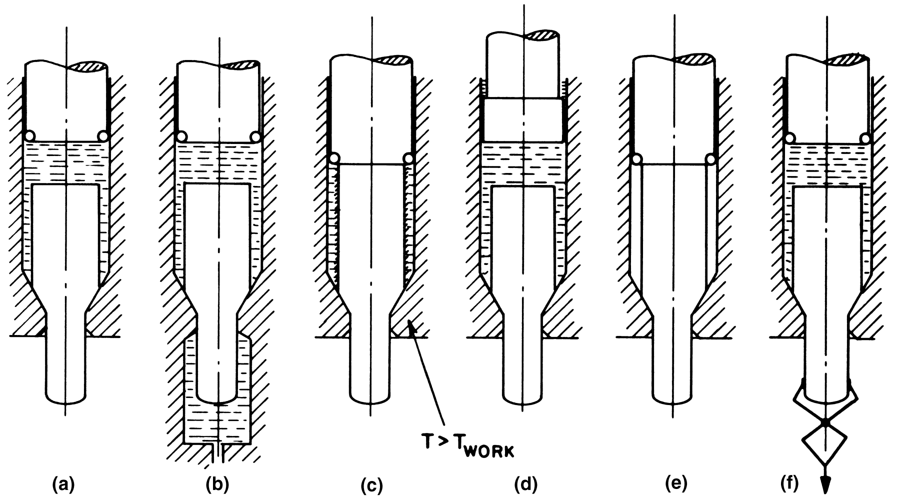


Fig. 38 Basic process of hydrostatic extrusion. (a) Simple hydrostatic extrusion. (b) Extrusion into a pressurized medium. (c) Thick film of semisolid lubricant to the surface of the billet. (d) Dynamic seal with a controlled leakage rate. (e) Punch with a high-pressure seal that bears on the billet. (f) Drawing augmented hydrostatic extrusion

the billet in a fluid medium reduces or practically eliminates major sources of friction such as container friction and, in extrusion of tubes, friction on the mandrel surface. These are major contributors to extrusion pressure and thus limit both billet length and extrusion ratio for given maximum punch and container pressures (Fig. 12).

The concept of hydrostatic extrusion can be traced back to 1893 (Ref 38), when an augmented hydrostatic extrusion process, known as Hydraw, was patented. In the Hydraw process, a liquid in a pressure chamber causes the wire to extrude through a die, while speed control is maintained by a moderate pull on the wire at the exit. This process is referred to as augmented hydrostatic extrusion process, because a slight drawing force augments the extrusion pressure. Although this particular method has been recognized as a potential for economical production of fine wire with the goal of replacing multiple-pass wire drawing with large reductions in a single pass or to produce wire of materials that cannot be drawn by conventional methods, the concept of augmented hydrostatic extrusion for the production of fine wire still has difficulties (Ref 39).

This section focuses on the method of simple hydrostatic extrusion (Fig. 38a) and some variants of it (Fig. 38b–f). The first commercial applications of simple hydrostatic extrusion have been in cases when conventional processes have shortcomings. Hydrostatic extrusion is primarily used when conventional lubrication is inadequate—for example, in the extrusion of special alloys, composites, clad materials, or brittle materials that cannot be processed by conventional extrusion. Hydrostatic extrusion allows greater reductions in area (higher extrusion ratios) than either cold or conventional hot extrusion. The large reduction capability of hydrostatic extrusion has prompted much work into wire production. The greater homogeneity of deformation also makes hydrostatic extrusion particularly favorable for cladding, and the much reduced mandrel friction makes it advantageous for extrusion of tubes. Applications of hydrostatic extrusion include the production of copper tubing, copper-clad aluminum wire, fine wires of precious metals, aluminum alloy tubes, and niobium-titanium superconducting wires (Ref 40).

Methods of Hydrostatic Extrusion

Because hydrostatic extrusion requires less power than conventional extrusion, it continues to attract attention. As noted, the primary advantages are:

- There is no friction between the billet and the container. Therefore, the pressure at the beginning of extrusion is much lower, and billets of any length can theoretically be extruded into semifinished products. Billet length is limited only by container construction, and, if desired, a spool of wire may be placed inside the container and long lengths of fine wire extruded.
- Friction at the die can be significantly reduced by a film of pressurized lubricant between the deforming metal and the die surface.
- The lower extrusion pressures and the reduced die friction of hydrostatic extrusion allow the use of either higher extrusion ratios or lower extrusion temperatures.
- The uniform hydrostatic pressure in the container means that billets do not have to be straight; coiled wire can also be extruded.

However, there are a number of practical difficulties that have led to different techniques. In the basic process of simple hydrostatic extrusion, the pressure is built up by the penetration of a punch that traps the fluid medium with a high-pressure seal (Fig. 38a). Contact between container and billet surfaces is eliminated, which can be particularly advantageous when long billets are to be extruded. If the pressure medium has lubricating qualities, friction is much reduced over the die face. This permits the use of lower die angles, which then ensures greater homogeneity of deformation.

The hydrostatic pressure acting on the workpiece can be increased by taking advantage of low die friction, by extruding with a low die half-angle (low *h*/*l* ratio), or by extruding into a pressurized medium (Fig. 38b). With the latter technique, even brittle materials can be deformed, although at the expense of high container pressures. These parameters are described in more detail in Ref 41, which describes an analytical model for optimization in terms of various process parameters, such as die cone angle, the reduction ratio, the friction coefficient, the extrusion and back-pressure ratio, and the die shape.

Nonetheless, a basic problem of all of the previously mentioned techniques is low production rate; the container has to be pressurized for each billet, and large quantities of fluid have to be supplied and drained. A thick film of semisolid lubricant applied to the surface of the billet allows the billet to be loaded as in usual cold extrusion but into a preheated container in which the lubricant melts to become the pressurized medium (Fig. 38c) (Ref 42). If the clearance around the punch is small enough to act as a dynamic seal with a controlled leakage rate (Fig. 38d), sufficient pressure for extrusion might be developed. Less practically, lubricant may also be trapped with a billet that has a flange at the punch end. Extrusion at high velocities is also helpful in minimizing leakage.

A common problem is that an excessively thick lubricant film not only roughens the product but may also generate periodic instability, leading to bambooing that manifests itself in periodic variations in the surface finish and diameter of the extruded product. To suppress this effect, a billet-augmented process has been considered. In this technique, the punch has a high-pressure seal that bears on the back face of the billet (Fig. 38e). Alternatively, the extruded product may be subjected to a drawing force (Fig. 38f). The combined extrusion-drawing process creates the advantage of lower die pressures and the possibility of heavier reductions. In extrusion of tubes, drawing also reduces the frictional stresses on the mandrel. In extrusion through two dies in tandem, the first reduction is equivalent to drawing under hydrostatic pressure with fluid-film lubrication, and the second reduction, where most of the work is done, corresponds to hydrostatic extrusion.

The general limitations of the hydrostatic extrusion process include:

- Containment of the fluid under high pressure (up to 2 GPa, or 290 ksi) requires reliable seals between the container bore surface and both the ram and die. The technology required to achieve dependable seals at these points is widely available, however. Also, sealing between the billet

nose and the die can easily be achieved by chamfering or tapering the billet nose to match the entry angle of the die.
- In addition to being tapered to match the die opening angle, the billet is also usually machined all over to remove surface defects that would otherwise reappear on the extruded product. This is especially true when cast billets are being used.

Other limitations of the process arise when a relatively large volume of fluid is used compared to the billet volume to be extruded. These include:

- Increased handling for injecting and removing the fluid for each extrusion cycle
- Reduced control of billet speed and stopping due to potential stick-slip and excessive stored energy in the compressed fluid
- Reduced process efficiency in terms of billet-to-container volume ratio
- Increased complications when extruding at elevated temperatures

The problems of billet speed and stopping control can be reduced by using viscous dampers and by improving lubrication at the billet-die interface. Another approach to minimizing this and the other problems cited previously is to keep the amount of pressurizing fluid to an absolute minimum, as in the Hydrafilm process (see the section "The Hydrafilm Process" in this chapter).

Equipment designs for hydrostatic extrusion presses are becoming more similar to those of conventional extrusion presses, as many of the problems have been resolved (Ref 43). The high pressures and the low-viscosity media necessitate the use of special containers and seals to prevent leakage of the pressure medium. Various methods are used to prevent the stick-slip phenomenon, which was once considered a serious problem in hydrostatic extrusion. Various ingenious schemes have been proposed to extrude bars continuously and semicontinuously, as reviewed in Ref 44 and 45.

The essential components of tooling for hydrostatic extrusion are the high-pressure container, the ram, the die, and, if tubing is being extruded, the mandrel. The container is subjected to very high pressures (up to 2 GPa, or 290 ksi), and a large hoop stress is imposed at the bore. The ram is under a uniaxial compressive load and generally has a much longer life than the container. The die and mandrel have considerably shorter service lives than the container, because they are usually subjected to larger stresses, higher temperatures, and more wear. The stability of extrusion increases also when an elastomer is used for the punch and a separate lubricant is applied to the billet.

Tool materials used for hydrostatic extrusion include alloy steels, higher-alloy tool steels, 18Ni maraging steels, nickel- and cobalt-base superalloys, and, in some applications, cemented carbides (usually 5 to 15% Co binder phase). Table 6 lists some tooling applications of these materials, and more detailed information on the suitability of materials for hydrostatic extrusion tooling is available in Ref 43.

Table 6 Tool materials for hydrostatic extrusion and some typical applications

Material	Typical applications
High-strength alloy steel	Outer container rings
Hot work tool steels (e.g., AISI H11, H12)	Inner container liners; rams for relatively low working pressure; dies
18Ni maraging steels	Dies; container liners and rams for high working pressures
Nickel- and cobalt-base superalloys	Dies for high-temperature applications
Cemented carbides	Die inserts; mandrel tips

AISI, American Iron and Steel Institute. Source: Ref 43

Hydraulic Medium and Lubrication (Adapted from Ref 8)

The hydraulic medium must be chosen to develop the required pressure without solidification. Thus, a medium with good lubricating properties, such as castor oil, may be used if pressures are not too high, as in model experiments with wax or in extrusion of soft metals. Otherwise, favorite substances are low-viscosity mineral oils, castor oil with methylated spirits (useful up to 1.4 GPa, or 200 ksi), glycerine with 25% ethylene glycol (2.0 to 2.8 GPa, or 290 to 400 ksi), and isopentane and gasoline (to 3.0 GPa, or 435 ksi).

Because of the limitation imposed by solidification, lubrication is frequently of the mixed-film type, as evidenced by the need for the pressure medium to be compounded with boundary and extreme-pressure additives (and even with layer-lattice compounds) and by the response of surface finish to die angle, reduction, and the effect of viscosity, η, and the velocity-pressure ratio, v/p, on film thickness (where film thickness is proportional to $\eta v/p$). Often, it is found to be more practical and effective to apply a separate lubricant film, such as a grease or, in the case of steel, a conventional phosphate/soap coating to the billet, and to use the hydrostatic feature only to reduce friction and improve the homogeneity of deformation. Oxidation of the surface prior to coating with soap has also been found to be effective. Of course, under such conditions, the benefits of hydrostatic operation become rather dubious, because pressures are hardly reduced relative to those encountered in well-lubricated conventional extrusion.

Because pressures in hydrostatic extrusion are sufficient to initiate yielding of the billet, the lubricant film may easily thicken to the point where it becomes unstable, and bambooing is often observed. In experiments on wax as a model material, it was noted that the defect was associated with fluctuations between PHD and boundary conditions, and that the calculated coefficient of friction dropped with increasing viscosity (Ref 25). Bambooing occurred only with the stiffer, less compressible castor oil and not with higher-viscosity silicone oils. It has been shown that the extrusion actually comes to a momentary rest but then accelerates sufficiently to establish, even though for only a brief time, a fluid film (Ref 46). Other research noted that the bamboo defect is suppressed with increasing extrusion speed (Ref 47). Stick-slip combined with poor die lubrication can even result in firtreelike cracking, and therefore, lubrication reduces the defect.

The origin of the bamboo defect lies in the lubrication mechanism, but its development, magnitude, and damping out depend on the entire system. Thus, billet augmentation suppresses fluctuations not only by physically restraining the billet but also by reducing the film thickness. This, incidentally, increases die friction. Because the billet is now pushed (Fig. 38e), it first upsets if the billet-to-container clearance is too large for the prevailing lubricant viscosity; in contrast, lubricant starvation sets in when the clearance is too small. Bambooing may still occur if $\eta v/p$ is excessive, α is small, or the die land is slightly tapered inward. If the augmenting pressure is too high, the billet is upset into the die throat, lubricant excess is reduced, and the entry geometry is impaired, resulting in a rise in friction and a change to boundary lubrication.

The same effects were noted in extrusion with a controlled follower block clearance (Fig. 38d) (Ref 48). The whole billet is extruded if loss of fluid around the follower block is limited. The leakage rate, Q, can be calculated from:

$$Q = \frac{\pi \alpha c^3}{6\eta} \frac{dp}{dl} \qquad \text{(Eq 9)}$$

where α is the radius of the container, c is the radial clearance, and dp/dl is the rate of pressure drop.

Mixed-film lubrication prevails also in hydrostatic extrusion/drawing (Fig. 38f), as shown by the roughening of the product surface with increasing speed (Ref 49). Deformation of the bar prior to entry into the die and a rougher emerging surface indicate a shift toward a more hydrodynamic contribution at higher extrusion/draw stress ratios, at least for a strain-hardening material such as copper.

Hot Hydrostatic Extrusion

Hydrostatic extrusion was originally used only for cold working; that is, the billet was not preheated. Recently, however, more emphasis has been placed on the development of hot hydrostatic extrusion processes. Hot hydrostatic extrusion is particularly applicable to difficult-to-work materials, such as high-strength aluminum alloys, titanium alloys, refractory metals and alloys, bimetallic products, and multifilament superconductors. The hot process has also been used for the production of copper tubing at extrusion ratios on the order of 500 to 1.

One of the main problems associated with hot hydrostatic extrusion has been that the pressure media used in cold or warm processes (usually castor or other vegetable oils) ignite and burn at high temperatures. Therefore, research in the area of hot hydrostatic extrusion has focused on the development of processes using pressure media that can withstand elevated temperatures. There is also the potential for hot extrusion with grease, polymer, or glass media, which melt at the extrusion temperature and thus transmit the extrusion pressure.

One team of investigators proposed the use of a viscoplastic pressure medium for hot hydrostatic extrusion (Ref 50). Because such materials are soft solids at room temperature, a piece of the pressure medium can be introduced into the container without the need for a charging pump. This simplifies machine design; the construction of a press used in this manner can be as simple as that of a conventional extrusion press.

Viscoplastic pressure media used for hot hydrostatic extrusion include:

- Waxes or fats, such as beeswax, carnauba wax, montan wax, lanolin, and complex waxes
- Soap-type greases composed of petroleum oil and such soaps as fatty acids or soaps of sodium, calcium, or lithium; such mixtures may also contain graphite.
- Mixtures of nonsoap greases and silica or other metal oxides; mixtures of petroleum oil and bentonite are heat resistant up to 1200 °C (2190 °F).
- High-molecular-weight polymers, such as polyethylene; the properties of these materials depend on their molecular weight and the additives used.

Some metal oxides, salts, and glass can also be used as pressure media for hot hydrostatic extrusion, but these materials may adhere to the extruded product and can be difficult to remove.

The Hydrafilm Process

The method known as the Hydrafilm, or thick-film, process was developed at Battelle Columbus Division to overcome many of the disadvantages of pure hydrostatic extrusion, in which a large volume of hydrostatic fluid is usually employed. The Hydrafilm process is characterized by the relatively low volume of pressurizing fluid used and by the relatively thick film of lubricant that is maintained around the billet (Fig. 39).

The small amount of pressurizing fluid used in the Hydrafilm process minimizes fluid-handling time, thus increasing billet cycle rates. It also minimizes billet chilling, which is important in hot hydrostatic extrusion. For cold or warm hydrostatic extrusion, the billet can be precoated with the pressurizing medium by dipping or spraying; lubricants that are different and separate from the pressurizing medium can also be used for some applications. Other potential advantages of the Hydrafilm process include (Ref 51):

- The process is simplified and is very similar in operation to conventional extrusion.
- Conversion of conventional presses to hydrostatic extrusion is facilitated.
- The Hydrafilm concept is applicable to high-speed mechanical presses; production rates of 30 to 40 billets/min may be possible.
- The process is adaptable to hot extrusion at billet temperatures up to and beyond 1200 °C (2190 °F).

The Hydrafilm process has been used for extruding tubes of various materials, including tubing of Ti-6Al-4V, American Iron and Steel Institute (AISI) 4140 steel, AISI type 316 stainless steel, and Zircaloy-2; aluminum, copper, and steel rod; and copper and steel shapes. More information on these applications is available in Ref 51.

Simple Hydrostatic Extrusion of Brittle Materials

Most brittle materials are subject to circumferential (transverse) and longitudinal surface cracking during hydrostatic extrusion. This cracking can be avoided through the use of either fluid-to-fluid extrusion or double-reduction dies. In fluid-to-fluid extrusion, the billet is hydrostatically extruded into a fluid at a lower pressure. This method has several disadvantages, including high tooling and operating costs, extrusion lengths that are limited to the length of the secondary chamber, and increased fluid pressure required for extrusion. For these reasons, the fluid-to-fluid process may not be suitable for many industrial applications.

The problem of extruding low-ductility metals was approached in a different way by researchers at Battelle Columbus Division (Ref 52), who established that the cracks or fracture first developed in the rear section of the die land, immediately before the exit plane, and that the surface cracking resulted from residual tensile stresses as the product left the die. The cracks observed were either longitudinal or transverse across the extruded product, depending on whether the predominating residual stresses were longitudinal or circumferential. This phenomenon was noted much earlier in rod and tube drawing (Ref 53). It was discovered that it is possible to reverse the residual stresses at the surface to compressive stresses by a subsequent draw with a low reduction in area (<2%).

The work of these investigators led to the development of the double-reduction die (Ref 53). Figure 40 compares a standard die with a double-reduction die. The double-reduction die used for the experiments was designed to give a 2% reduction in the second step. This method has been successfully applied to the extrusion of brittle materials, including beryllium and TZM molybdenum alloy, without any cracking. The lubricant used was polytetrafluoroethylene, and the pressurizing fluid was castor oil. The results may be applicable to conventional cold extrusion through a lubricated conical die (Ref 52).

It is believed that the small second reduction prevents cracking by imposing an annular counterpressure on the extrusion as it exits the first portion of the die. This counters the axial tensile stresses arising from residual stresses, elastic

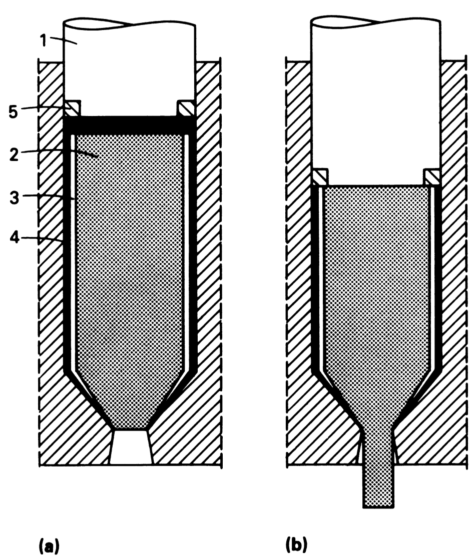


Fig. 39 Schematics of the Hydrafilm (thick-film) process. (a) Before extrusion. (b) During extrusion. 1, extrusion ram; 2, billet; 3, film of pressurizing liquid; 4, separate film of lubricant; 5, sealing ring. Source: Ref 51

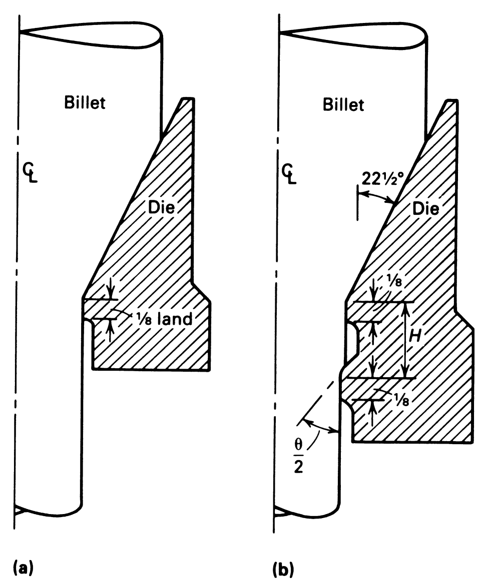


Fig. 40 (a) Standard die and (b) double-reduction die for the hydrostatic extrusion of brittle materials. H, distance between the start of each bearing; θ, included angle at second reduction. Dimensions given in inches. Source: Ref 51

bending, and friction. Prevention of circumferential cracks on exit from the second portion of the die is believed to be associated with the favorable permanent change in residual stresses in the workpiece caused by the small second reduction (Ref 51).

Summary (Ref 8)

- In extrusion of long, semifabricated products, friction is generally unnecessary and undesirable. Friction on the die increases extrusion pressures and, for a given die geometry, impairs the homogeneity of deformation. If, under the prevailing pressures, the shear strength of the interface exceeds the shear flow stress of the workpiece material, a dead-metal zone forms.
- In direct (forward) extrusion, friction on the container wall increases the extrusion pressure. In the limit, sticking friction develops, and extrusion proceeds by shearing through the billet, leaving behind a skull. For a given press capacity, the length of billet that can be extruded is limited by container friction.
- Friction on the container wall is immaterial in indirect (reverse) extrusion, because there is no relative movement.
- Friction contributes to heat generation and limits attainable reductions and speeds in hot extrusion. Friction on the die land promotes surface cracking (speed cracking, firtree effect).
- Friction is beneficial only when there is danger of internal (arrowhead) cracking or when the container is moved so that the frictional force contributes to the extrusion force. Friction on the die and die land may be used to equalize material flow in extrusion of complex shapes.
- Extrusion is a process that can only be conducted either totally unlubricated or fully lubricated. Intermediate conditions are undesirable, because changeover from unlubricated to lubricated flow during extrusion of a billet results in formation of subsurface defects.
- Unlubricated extrusion is essential for extruding hollow products with bridge, spider, or porthole dies, because use of a lubricant would prevent rewelding of the separated material streams. Unlubricated extrusion is desirable for hot extrusion of aluminum alloys with flat dies, because the surface of the extruded product is freshly formed by internal shear in the billet and is feasible also for nonisothermal hot extrusion of copper and its alloys.
- Lubricated extrusion is essential for hot extrusion of steel, nickel- and titanium-base alloys, refractory metal alloys, and high-strength copper alloys, as well as for all cold extrusion. The geometry of the process makes it particularly favorable for lubrication with glasses that melt gradually to form a lubricating film. The billet surface becomes the surface of the extrusion, modified by surface extension and by the presence of the lubricant.
- Friction is reduced, and the homogeneity of flow is improved, by lubrication. Lubricant film thickness increases with viscosity, extrusion speed, and decreasing die angle. Tapering of the billet eliminates breakthrough pressures. Film thickness increases also with the pressure applied to the lubricant, and thus, hydrostatic extrusion not only eliminates container friction but can also reduce die friction. It is particularly effective in reducing friction on the mandrel in tube extrusion. Excessive film thickness leads to rough surfaces and, possibly, to periodic collapse of the film (bamboo defect).
- Severe sliding and prolonged exposure to high temperatures and high pressures combine to make hot extrusion extremely demanding on die materials. Adhesive and abrasive wear, thermal fatigue, and plastic deformation may all occur. Surface treatments such as nitriding are beneficial even in unlubricated extrusion, and ceramic die coatings are often essential in high-temperature glass-lubricated extrusion.
- The most commonly used lubricants are given in Table 7. For purposes of preliminary calculation, typical μ values are also shown, although it must be understood that under hydrodynamic or mixed-film conditions, the coefficient of friction is greatly affected by process conditions. At the high pressures developed with high extrusion ratios, sticking may be attained even with relatively good lubricants.

Table 7 Commonly used extrusion lubricants and typical friction values

	Hot extrusion		Cold extrusion	
Material	Lubricant(a)(b)	μ(c)	Lubricant	μ
Steels	GR (for short pieces)	0.2	NA	. . .
	Glass (10–100 Pa · s) pad and coating	0.02		
Stainless steels and Ni alloys	Glass (20–200 Pa · s) pad and coating	0.03	NA	. . .
	Melting solids	0.05		
Al and Mg alloys	None	ST	Lanolin	0.07
	None (GR on die face)	ST		
Cu and Cu alloys	GR (for short pieces)	0.2	Castor oil (for hydrostatic extrusion)	0.03
	None (with or without skull)	ST		
	Glass (10–100 Pa · s) pad and coating	0.05		
Ti alloys	GR (for short pieces)	0.2	NA	. . .
	Glass (10–100 Pa · s) coating + GR on die	0.05		
	Glass (10–100 Pa · s) pad and coating	0.03		
Refractory metals	Glass (10–200 Pa · s) coating + GR on die	0.05	NA	. . .
	Glass (10–200 Pa · s) pad and coating	0.03		

(a) GR = graphite, often with a polymeric binder. (b) Viscosities of glasses are given for typical extrusion temperatures. (c) When $\mu p = k$ is satisfied, sticking friction (ST) sets in.

ACKNOWLEDGMENTS

Significant portions of this chapter were adapted from:

- Conventional Hot Extrusion, *Forming and Forging,* Volume 14, *ASM Handbook,* 1988
- Hydrostatic Extrusion, *Forming and Forging,* Volume 14, *ASM Handbook,* 1988
- K. Laue and H. Stenger, *Extrusion: Processes, Machinery, and Tooling,* American Society for Metals, 1981
- J.A. Schey, Chapter 8: Extrusion of Semi-Fabricated Products, *Tribology in Metalworking: Friction, Lubrication, and Wear,* American Society for Metals, 1983, p 403–441

REFERENCES

1. K. Laue and H. Stenger, *Extrusion: Processes, Machinery, and Tooling,* American Society for Metals, 1981
2. Y.-J. Huang, Integrated Computer Aided Extrusion Process Simulation, Die Design/Manufacturing, and Management Planning System, *Light Met. Age*, Vol 56 (No. 5–6), June 1998, p 76–81
3. C. Devadas, O.C. Celliers, D. Greliche, and X. Zhang, Mathematical Modelling of Extrusion Process to Enhance Die Design, *Aluminum 2000 Conference, Second International Congress on Aluminum,* Vol II, 29 March to 4 April 1993 (Florence, Italy), Interall Publications, 1993, p 353–365
4. C.O. Stockdale, The Application of CAD/CAM to Extrusion Die Design, *Fifth International Aluminum Extrusion Technology Seminar,* Vol I, 19–22 May 1992 (Chicago, IL), The Aluminum Association, 1992, p 391–394
5. R.H. Wagoner and J.-L. Chenot, *Metal Forming Analysis,* Cambridge University Press, 2001, p 224
6. V. Nagpal, C.F. Billhardt, and T. Altan, *J. Eng. Ind. (Trans. ASME),* Vol 101, 1979 p 319–325
7. P.K. Saha, *Aluminum Extrusion Technology,* ASM International, 2000
8. J.A. Schey, Chapter 8: Extrusion of Semi-Fabricated Products, *Tribology in Metalworking: Friction, Lubrication, and Wear,* American Society for Metals, 1983, p 420, 435
9. H. Kudo and H. Takahashi, *CIRP Ann.,* Vol 13, 1965, p 73–78
10. B. Avitzur, *Wire J.,* Vol 7 (No. 11), 1974, p 77–86

11. G.H. Townend and D.G. Broscomb, *Proc. Inst. Mech. Eng.,* Vol 169, 1955, p 671–676
12. N.H. Polakowski and L.B. Schmitt, *Trans. AIME,* Vol 218, 1960, p 409–416
13. I. Perlmutter, V. DePierre, and C.M. Pierce, in *Friction and Lubrication in Metal Processing,* ASME 1966, p 147–161
14. V. DePierre, *J. Lubr. Technol. (Trans. ASME),* Vol 92, 1970, p 398–405
15. B. Lengyel and D. Mutlu, in *Proc. 19th Int. MTDR Conf.,* Macmillan, 1979, p 265–269
16. T. Sheppard and A.F. Castle, in *Proc. 16th Int. MTDR Conf.,* Macmillan, 1976, p 535–545
17. R.J. Wilcox and P.W. Whitton, *J. Inst. Met.,* Vol 88, 1959–1960, p 145–149
18. R. Akeret and P.M. Stratman, *Light Met. Age,* Vol 31 (No. 3/4), 1973, p 6, 8–10; (No. 5/6), p 15–18
19. J.R. Johnson and W.R.D. Wilson, *ASLE Trans.,* Vol 24, 1981, p 307–316
20. J.A. Rogers and G.W. Rowe, *J. Inst. Met.,* Vol 95, 1967, p 257–263
21. H. Unckel, *Aluminium,* Vol 25, 1943, p 342–346
22. C.E.N. Sturgess and T.A. Dean, *J. Mech. Work. Technol.,* Vol 3, 1979, p 119–135
23. P. Loewenstein, J.G. Hung, and R.G. Jenkins, Refractory Metals and Alloys, Vol 111, The Metallurgical Society AIME, Gordon and Breach, 1965, p 151–167
24. R. Chadwick, *Met. Rev.,* Vol 4, 1959, p 189–255
25. A. Wuerscher and W.B. Rice, *J. Eng. Ind. (Trans. ASME),* Vol 94, 1972, p 795–799
26. B. Avitzur, *Wire J.,* Vol 11 (No. 3), 1978, p 78–84
27. R. Chadwick, *Int. Met. Rev.,* Vol 25, 1980, p 94–136
28. K.F. Ziehm, *Light Met. Age,* Vol 28 (No. 11/12), 1970, p 6–10
29. D. Green, *J. Inst. Met.,* Vol 100, 1972, p 295–300
30. B. Avitzur, Paper MF 75-140, SME, (Dearborn), 1975
31. J.T. Black and W.G. Voorhes, *J. Eng. Ind. (Trans. ASME),* Vol 100, 1978, p 37–42
32. F.J. Fuchs and G.L. Schmehl, *Wire J.,* Vol 6 (No. 11), 1973, p 53–57
33. B. Lengyel and S. Kamyab-Tehrani, in *Proc. 17th Int. MTDR Conf.,* Macmillan, 1977, p 423–428
34. P.W. Wallace, K.M. Kularni, and J.A. Schey, *J. Inst. Met.,* Vol 100, 1972, p 78–85
35. A.W. Duffill and P.B. Mellor, *Proc. Inst. Mech. Eng.,* Vol 180 (Pt. 3I), 1965–1966, p 260–269
36. R. Chadwick, *Met. Mater.,* Vol 4, 1970, p 201–207
37. G. Merk and S.E. Naess, *Z. Metallkd.,* Vol 68, 1977, p 683–687
38. H.L.D. Pugh, Hydrostatic Extrusion, *Cu '86—Copper Tomorrow; Technology—Products—Research, Conference Proceedings,* (Florence, Italy), Europa Metalli-LMI, 1989, p 67–83
39. B. Avitzur, Limitations on Wire Size During Drawing and Extrusion, *Wire J. Int.,* Vol 25 (No. 3), Mar 1992, p 51–56
40. N. Inoue, Introduction, *Hydrostatic Extrusion: Theory and Applications,* N. Inoue and M. Nishihara, Ed., Elsevier Applied Science Publishers, 1985, p 1–6
41. H. Elkholy, Parametric Optimization of Power in Hydrostatic Extrusion, *J. Mater. Process. Technol.,* Vol 70 (No. 1–3), Oct 1997, p 111–115
42. R.J. Fiorentino, G.E. Meyer, and T.G. Byrer, *Metall. Met. Form.,* Vol 39, 1972, p 200–203
43. M. Yamaguchi, Plant Equipment for Hydrostatic Extrusion, *Hydrostatic Extrusion: Theory and Applications,* N. Inoue and M. Nishihara, Ed., Elsevier Applied Science Publishers, 1985, p 164–193
44. J.M. Alexander, *Mater. Sci. Eng.,* Vol 10 (No. 2), 1972, p 70–74
45. E. Tuschy, *Metall.,* Vol 36, 1982, p 269–279 (in German)
46. D. Egerton and W.B. Rice, *J. Eng. Ind. (Trans. ASME),* Vol 98, 1976, p 795–799
47. P. Krhanek and J. Jakes, in *Engineering Solids Under Pressure,* H.L.I.D. Pugh, Ed., Inst. Mech. Eng., 1971, p 117–121
48. K.M. Kulkarni and J.A. Schey, *J. Lubr. Technol. (Trans. ASME),* Vol 97, 1975, p 25–31
49. P. Dunn and B. Lengyel, *J. Inst. Met.,* Vol 100, 1972, p 317–321
50. M. Nishihara, M. Noguchi, T. Matsushita, and Y. Yamaguchi, Hot Hydrostatic Extrusion of Nonferrous Metals, *Proceedings of the 18th International MTDR Conference,* 1977, p 91–96
51. R.J. Fiorentino, Selected Hydrostatic Extrusion Methods and Extruded Materials, *Hydrostatic Extrusion: Theory and Applications,* N. Inoue and M. Nishihara, Ed., Elsevier Applied Science Publishers, 1985, p 284–322
52. R.J. Fiorentino, B.D. Richardson, and A.M. Sabroff, Hydrostatic Extrusion of Brittle Materials: Role of Design and Residual Stress Formation, *Met. Form.,* 1969, p 107–110
53. H. Buhler, Austrian Patent 139,790, 1934; British Patent 423,868, 1935

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 316-322
DOI:10.1361/hwpd2003p316

Chapter 21

Workability and Process Design in Extrusion and Wire Drawing

Roger N. Wright, Rensselaer Polytechnic Institute

EXTRUSION AND WIRE DRAWING are very analogous processes with regard to the manner of metal flow through the die, and they have somewhat comparable workability problems. The major differences in workability behavior involve the difference in hydrostatic stress state (the flowing metal is generally subjected to much greater compression in extrusion than in wire drawing), differences in working temperature (extrusion is usually hot while drawing is usually cold), and the obvious differences in hardware and process layout.

Workability in these cases is the ability of the workpiece metal to undergo extrusion or drawing without fracture or defect development. As in any of the deformation processes, the problems of *metallurgical* workability are difficult to separate from problems of *mechanical* process optimization. In short, if process parameters are improperly chosen, even the best material may fracture. On the other hand, metal with outstanding intrinsic workability will in many cases withstand the abuse of improper processing without fracture or defect development. In any case, it rarely makes sense for a manufacturer to cope with inferior procedures by requiring metals of unusually good workability. The costs and/or supply problems for such superior materials provide considerable incentive for process optimization. Thus, mature, competitive manufacturing generally involves the use of processes flexible enough to accommodate metal of "standard" or "average" workability, which is readily available from competent, major suppliers.

This is not to say, however, that the measurement or estimation of workability is a minor consideration. Quite the contrary, it is essential to monitor day-to-day fluctuations in "standard" products. Not only will material of completely unacceptable quality be occasionally received, but fluctuations within acceptable limits can often be dealt with to advantage, with the superior material being directed to the more demanding process sequences. Beyond the case of the "standard" material, *new* alloy development and *new* process development require that the workpiece metal be carefully characterized for workability so that the most efficient processing schedule can be set forth.

Multipass Workability versus Single-Pass Workability

It is not quite enough to define workability in extrusion and wire drawing as the ability of the metal to be worked without fracture or fracture-type defects. It must be clear whether the ability to be worked concerns a single pass or a series of passes. The question mostly applies to wire drawing, because most extrusion is done as a single-pass operation. In a single wire-drawing pass, the material cannot be worked beyond a reduction that produces a drawing stress, σ_d, that is equal to the flow (yield) strength (σ_0) of the emerging wire. Otherwise it would simply break at the die exit (or at the point where the drawn wire contacts the capstan). The limiting *single-pass* reduction, or the reduction at which $\sigma_d = \sigma_0$ is a function of die design, friction, and the work-hardening character of the material. Thus, for a reasonably ductile metal one could argue that work-hardening behavior is an important factor in workability. Actually, the effect of work hardening on single-pass workability is rather small, and most commercial drawing passes involve drawing stresses well below the strength of the wire at the die exit (typically, $\sigma_d \simeq 0.5\sigma_0$).

The more practical wire-drawing consideration, then, is *multipass workability,* or the ability of the metal to withstand fracture or defect development in the face of progressive straining in multiple drawing passes without the benefit of intermediate annealing. Generally, multipass workability is tied to the resistance to ductile fracture and not directly to factors such as draw-stress limitation and work-hardening character. Of course, in some materials "multipass" workability may be so limited that fracture occurs at strains or reductions below those commonly undertaken in a single drawing pass (up to about 40%).

In the case of extrusion, the single-pass reduction of common interest may be limited by the allowable level of extrusion pressure (analogous to the drawing-stress limitation in wire drawing) or by the onset of some degree of ductile fracture. Of course, the extrusion-pressure limitation is not a workability limitation, per se, because workpiece defect development or fracture is not implicit.

General Observations on Fracture and Flaw Development

The limits of workability in extrusion and drawing may be associated with the development of localized surface cracks, centerbursts, or gross fracture. While the surface cracks and centerbursts are "subcritical flaws" not necessarily producing gross fracture, they are perhaps even more important than fracture itself. In particular, they may go unnoticed, leading to inferior physical properties in as-shipped material or to sudden failures under apparently safe subsequent process or service conditions.

Again, it is worth noting that such fracture and defect development can be produced in superior materials if processing is improperly undertaken. For the discussions in this chapter, one should assume for the most part that proper or orthodox procedures are being followed and that defects or fractures truly reflect practical metal workability limitations.

Surface cracking, whether circumferential ("fir-tree" cracking), or at roughly 45° to circumferential ("crow's-feet" cracking), or even longitudinal ("splitting"), is a common problem in both extrusion and drawing. A schematic representation of surface cracking is given in Fig. 1, and illustrative micrographs are shown in Fig. 2 to 5. There are several explanations commonly given for these flaws. In wire drawing, they are observed under conditions of boundary-lubrication breakdown and sometimes, under "thick-film" conditions, are developed with dry soaps (Ref 3). In the former case it is likely that local lubrication breakdown in the die causes simultaneous sticking and slipping of adjacent areas of the wire surface. This sticking and slipping of al-

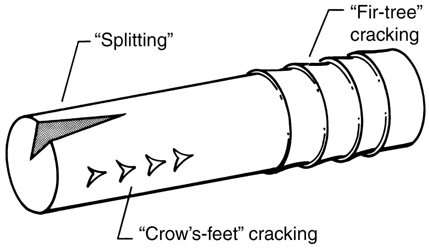


Fig. 1 Three types of surface cracks that occur in extruded and drawn products. Source: Ref 1, p 252

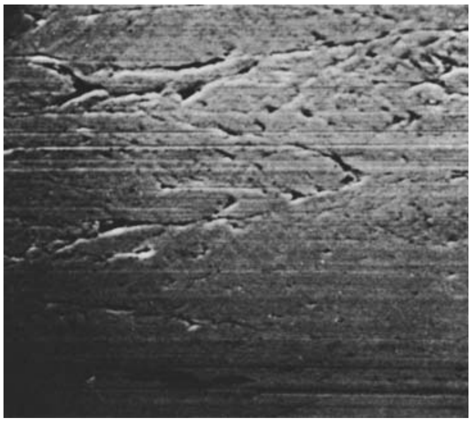


Fig. 2 "Crow's-feet" crack configurations on the surface of a drawn EC aluminum rod. Scanning electron micrograph. Magnification: 248×. Source: Ref 1, p 253

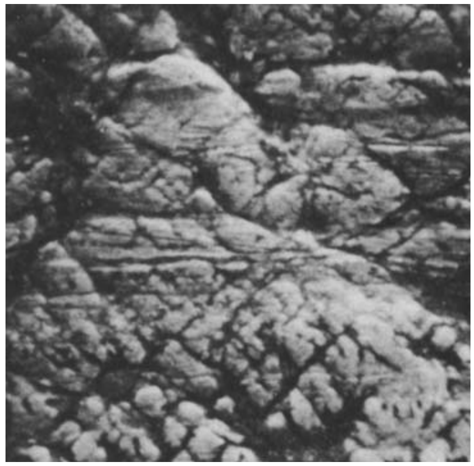


Fig. 3 Gross roughness and apparent cracking, resulting from thick-film lubrication, on the surface of EC aluminum rod drawn in calcium stearate. Scanning electron micrograph. Magnification: 320×. Source: Ref 1, p 254

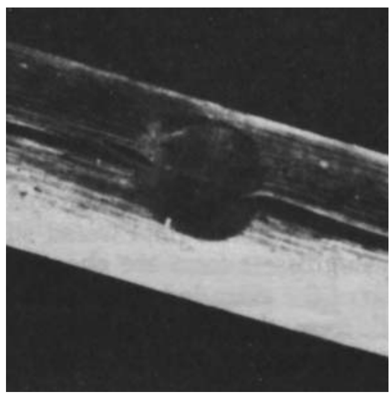


Fig. 4 A "split" in tungsten wire, induced by compression. Source: Ref 2

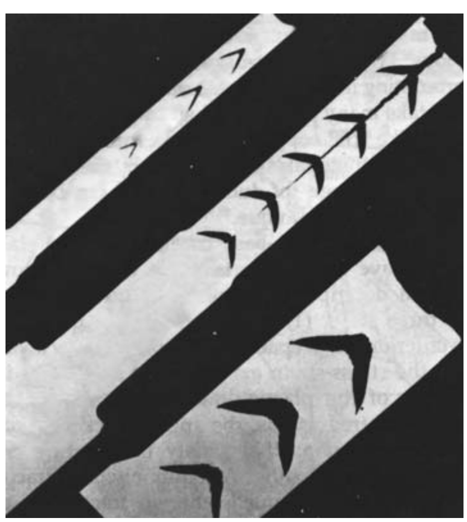


Fig. 5 Centerbursts in sectional, cold extruded steel bars. Courtesy of Bethlehem Steel Corporation

ternate areas presumably leads to intense shearing and tearing of local regions, which present the "crow's feet" appearance shown in Fig. 2 (Ref 4). In the case of "thick-film" dry-soap lubrication, it is suggested that the cracking comes from unconstrained, nonuniform surface flow permitted by the thick lubricant film and produced by irregularities in the soap-particle distribution and by differences in plastic flow from one grain to another (orange peel). In any case, the cross-hatched cracking shown in Fig. 3 is developed. In drawing of refractory metals, such as tungsten, deep longitudinal surface cracking or "splitting" is often observed. A "split" induced by indentation is shown in Fig. 4 (Ref 2). Apparently, very weak longitudinal interfaces are developed in the drawing of tungsten, perhaps as a result of "mechanical fibering" (Ref 5) and/or grain-boundary embrittlement. Restricted slip may also lead to such delamination (Ref 6). Moreover, it has been shown that this weakness is greatly increased by certain anneals (Ref 2). During drawing, these weak interfaces are conceivably opened through the action of tensile surface residual stresses (Ref 7).

Circumferential cracking and 45° shear cracking are seen in extrusion. The local sticking/slipping mechanism described previously for wire drawing is felt to cause intense shear zones and resulting surface cracks. On a somewhat grosser scale, transverse cracking is believed to be caused by momentary sticking in the die land followed by pressure buildup and breaking away (Ref 8). A particularly common cause of surface cracking in hot extrusion is surface heating. The extrusion temperature will often be specified so as to be below temperature ranges associated with poor ductility or "hot shortness." Unfortunately, surface frictional heating, intensified by excessive ram speed, may raise the surface temperature into the hot-shortness range anyhow, leading to surface fractures. Such fractures may be associated with low-melting-range second phases, or strain concentrating second phases, or may be intergranular in nature.

Yet another "surface" flaw observed in extrusion is the "extrusion defect," an intrusion of surface oxide fostered by high container wall friction. Although this results in a defect that is somewhat comparable to surface cracking, no fracture is involved, and the "extrusion defect" is not a true workability problem. On the other hand, surface fractures can provide paths that allow lubricant and oxide to intrude into the interior of the extruded product, producing some of the same problems presented by the "extrusion defect."

Centerbursts in their grossest form are chevronlike shear fractures that form in the center of the extruded or drawn product, as shown in Fig. 5. In more subtle form, they may be little more than aggregates of pores along the centerline of the product. They may develop almost spontaneously or may slowly build up with progressive working (as in multiple drawing passes). They may lead to outright fracture during the working process or may remain subcritical, reducing finished product ductility or posing the threat of catastrophic failure during service. The basic cause of centerbursting is nonuniform metal flow through the die and the associated development of tensile stresses in the center of the workpiece. Such nonuniform flow and centerline tensile-stress development may be projected from the *shape* of the deformation zone. A practical description of shape can be gotten from the Δ parameter, as described by Backofen (Ref 9). Basically, Δ is the ratio of the "height" of the deformation zone (perpendicular to the axis of extension or die axis) to the "length" of the deformation zone (parallel to the axis of extension or die axis). For round wire drawing and extrusion:

$$\Delta = \frac{\alpha}{r}\left[1+\sqrt{1-r}\right]^2 \qquad \text{(Eq 1)}$$

where α is the die semicone angle in radians and r is the reduction *per pass* in decimal form. Values of Δ less than one are associated with *uniform* metal flow. Centerbursting is associated with the nonuniform flow of higher Δ passes, usually where Δ is about 2.5 or more. The bursting may be slow to develop and may require considerable overall plastic strain before becoming obvious. Thus, light extrusion passes ($r < 0.20$) often will not result in centerbursting in spite of high Δ values (low r produces high Δ). Similarly, it may take many passes of a high-Δ

geometry before a failure-resistant material such as electrolytic tough pitch (ETP) copper displays centerbursts. This is not to say, however, that the material is undamaged by the early passes. Quite the contrary, the early stages of ductile fracture can be detected by an appropriate experimental technique (Ref 10, 11).

Beyond the semiempirical Δ-based criterion, several elaborate analytical approaches have been undertaken in order to provide a more detailed theoretical description of the centerburst process (Ref 12–15). These efforts involve three steps:

1. Establishment of a quantitative, generalized fracture criterion
2. Preparation of a plastic-flow analysis suitable for determining the stress-strain experience at each point in the workpiece
3. Use of the plastic-flow analysis to determine which processing conditions (reductions, die angles, friction conditions, etc.) produce stresses and strains that satisfy the fracture criterion and presumably lead to centerbursting

A commonly used fracture criterion is that of Cockcroft and Latham (Ref 16), wherein fracture is projected to occur when:

$$\int_0^{\bar{\varepsilon}_f} \bar{\sigma}\left(\frac{\sigma^*}{\bar{\sigma}}\right) d\bar{\varepsilon} = C \quad \text{(Eq 2)}$$

where $\bar{\varepsilon}$ is effective strain, $\bar{\varepsilon}_f$ is effective strain at fracture, $\bar{\sigma}$ is effective stress, σ^* is the highest tensile stress, and C is a material constant. Determining a proper value of C has been a major problem, in that upset testing and tensile testing can provide widely different values. This problem with variable results in using the Cockcroft-Latham criterion can be resolved using finite element analysis, but differences obtained by tension testing and compression testing should be recognized. In work on aluminum alloy 2024-T351, for example, C has been estimated at 0.3 times the average effective stress in tension and 0.08 times the average effective stress in compression (Ref 13, 14). Latham and Cockcroft (Ref 12) and Hoffmanner (Ref 13) have employed slip-line fields for plastic-flow analysis, whereas Chen, Oh, and Kobayashi have employed the matrix method (Ref 14, 17). The projections of Hoffmanner are typical of the results of these efforts and are shown in Fig. 6. The ordinate is simply the value of the lefthand side of Eq 2 divided by the average effective stress. When the value of this integral reaches or exceeds the product of C times the average effective stress, centerbursting is projected. Note that for reductions large enough for Δ to be less than about 2.5, no centerbursting is projected. No centerbursting is projected for small strains either.

With appropriate adjustment of the value of C, these analyses are in reasonable agreement with experimental findings. However, much refinement of these approaches remains to be undertaken. An application of such bulk workability testing to limited workability steels (T440C

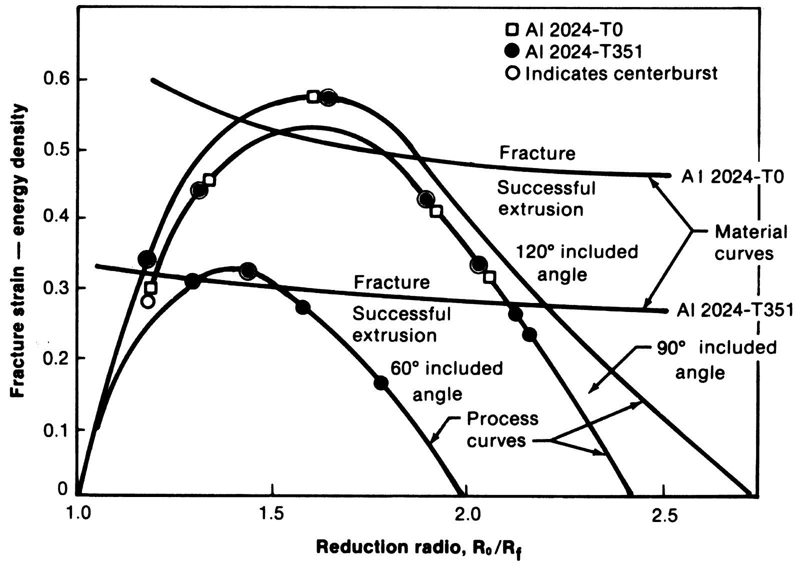


Fig. 6 Workability criteria for centerbursting in aluminum alloy 2024 based on a maximum tensile stress-strain energy criterion.Source: Ref 13

and M42) and to carbon steel in the warm working range has been set forth by Wright et al. (Ref 18).

A novel approach to predicting centerbursts has been employed by Avitzur (Ref 15). Using an upper-bound analysis, Avitzur has projected die angles and reductions for which the metal being worked will tend to flow apart at the centerline. As has been noted by Backofen (Ref 19), this approach seeks to provide a "necessary, but not sufficient" criterion for centerbursts. That is to say, the upper-bound analysis projects conditions whereby metal will tend to flow apart at the centerline, assuming that no energy is required to fracture the metal. A fully sufficient criterion would require incorporation of the fracture resistance of the metal. Moreover, Lee and McMeeking (Ref 20) have pointed out that the lower-bound theorem indicates that the workpiece will *always* tend to flow apart, regardless of die angle and reduction (the only thing that apparently keeps the metal from doing so is its fracture resistance, which is, of course, neglected in the upper-bound approach). They argue that the critical conditions for centerbursts in upper-bound analysis arise from the assumption of an unrealistic velocity field for this phenomenon.

In an approach of relative simplicity, Coffin and Rogers (Ref 10, 11) related centerline void development and centerbursting to the level of hydrostatic tension, based on a slip-line field analysis. It seems quite clear that such damage occurs with process geometries that produce hydrostatic tension (as opposed to compression) at the centerline. Normally, with low friction, such conditions are achieved with Δ values in excess of 2.5. In addition to the Hill and Tupper slip-line field analysis (Ref 21) used by Coffin and Rogers, the interested reader is directed to the hydrostatic tension projections of Chen, Oh, and Kobayashi (Ref 14) and of Lee et al. (Ref 22, 23). The effect of friction in enhancing centerline hydrostatic tension and centerbursting is well worth noting. This has been shown quite clearly by Coffin and Rogers (Ref 10). Moreover, Avitzur projects an increase in the occurrence of centerbursting with an increase in friction (Ref 15). Sticking friction can produce a centerline hydrostatic stress state that is more tensile than the frictionless case by as much as three-tenths of the flow stress.

For *short-term* problem solving, the following perspective seems best. First, it should be recognized that there is probably a *potential* for centerbursting anytime the Δ value for extrusion or drawing exceeds about 2.5. Moreover, metals of limited workability will develop centerbursts after small strains and certainly in single passes. On the other hand, metals of great workability may show little or no evidence of centerbursting until very high strains have accumulated in single or multipass reductions with Δ ≈ 2.5 (again, the r for Δ is the reduction in each pass). Such progressive centerbursts lead to failures in wire drawing called "cuppy-core" failures, with the "cup" presumably forming over many passes. Of course, many commercial extrusion and drawing operations involve Δ values above 2.5. Thus, to some extent, a potential for centerbursting is the rule and not the exception. In this regard it cannot be overemphasized that even where no gross centerbursting has occurred, damage in the form of incipient ductile fracture may be present. Such damage (centerline porosity, etc.) can manifest itself in reduced tensile properties in the as-extruded or drawn product. Other, more insidious problems may arise. For example, caps cut from bar stock will occasionally be sources of vacuum leaks, due to centerline porosity.

Fractures. Occasionally a surface crack or centerburst will form to such an extent and with such spontaneity as to cause outright fracture. Normally, however, the compressive stress state in the die in extrusion and drawing limits such crack growth, and truly spontaneous fracture is avoided.

The important exception is fracture of the wire in drawing due to excessive drawing load. The drawing stress is composed of the work required for uniform deformation, "redundant" deformation, and friction. Using Wistreich's redundant-work data (Ref 24), Wright has formulated the following relationship for practical drawing-stress calculations (Ref 25):

$$\frac{\sigma_d}{\hat{\sigma}_0} \simeq \left(\frac{3.2}{\Delta} + 0.9\right)(\alpha + \mu) \qquad \text{(Eq 3)}$$

where $\hat{\sigma}_0$ is average flow stress and μ is coefficient of friction. Rarely is a commercial drawing pass designed so that the extent of uniform and redundant deformation involves excessive drawing stress. Rather, the occurrence of drawing breaks is usually attributable to excessive friction, to inertial loadings, and, most often, to flaws in the wire, which reduce load-bearing capacity. These flaws may be gross inclusions, cut or abraded surfaces, "rolled-in dirt," or enlarged surface cracks or centerbursts developed in the manner discussed previously. Again, the well-known case of "cuppy-core" fracture in wire drawing is little more than a drawing break that occurs at a point weakened through the growth of a centerburst.

In the case of hot drawing, fracture of the wire often occurs at a point beyond the die where the wire has cooled into a brittle range. This takes place most often on contact with the capstan. The capstan, unless heated, chills the wire and simultaneously imposes tension as the wire assumes the capstan radius.

Shapes and Tubes

Most of the remarks in the previous section address flaw development as it occurs in round stock. In the case of tubular products, the remarks regarding surface cracking and fracture are still largely applicable. Although there may be nothing quite like centerbursting in tubular products, cracks may in some cases develop on internal surfaces for some of the same reasons that govern external surface behavior. The consideration of shapes is quite complicated. For rather regular shapes (hexagonal cross sections, ovals, etc.), the remarks concerning round stock are directly applicable. For shapes with thin sections (I-beams, L-shapes, etc.), the surface-flaw problems persist, but centerbursts are less likely to occur. In effect, Δ values are low in the thin sections. On the other hand, problems involving surface friction are apt to be more severe in tubes and thin section shapes.

Some Comments on Fracture Mechanisms

Now that the general observations on cracking in extrusion and drawing have been set forth, it is important to consider the relevant mechanisms of fracture from a viewpoint basic enough to allow design of sound workability test procedures or selection of pertinent technical data from published sources.

First of all, the driving forces for fracture, stress, and strain have nothing directly to do with the inherent resistance of the metal to fracture (or its workability). There will merely be some limit of stress and/or strain beyond which cracking will occur. The more workable the metal, the higher the limit. For example, as mentioned previously, Cockcroft and Latham (Ref 16) have proposed that fracture occurs, at a given temperature and strain rate, when the integral in Eq 2 reaches a critical value of *C*. More workable metals simply have higher values of *C*.

Thus, in the discussion of general cracking observations, various causative factors involving stress and strain have been alleged (frictional effects, centerline tension, excessive drawing stress, etc.). Therefore, it is helpful to have some idea of the levels and natures of these stresses and strains in the process and to have comparable states of stress and strain in any test that is to be used as a basis for forecasting workability.

In most instances, fractures in extrusion and drawing are *ductile* in nature. The mechanisms of ductile fracture typically involve the growth of pores with strain under a general tensile stress state (hydrostatic tension). Eventually the pores become gross enough so that tearing occurs in the intervening material, and a fracture results. The fracture path may follow bands of intense shear strain or may be largely perpendicular to the direction of the largest tensile stress.

The fundamental question is which factors inherent to the metal affect its ability to withstand given stresses and strains without cracking. (A comprehensive discussion of this question is beyond the scope of this chapter.) Such factors include aspects of basic crystal structure; crystal-defect structure; grain size, shape, and orientation; inclusion and second-phase content; defects inherited from primary processing (porosity, exogenous inclusions); and handling damage (cuts, scrapes, rolled-in particles, etc.).

Crystal structure, substructure, and microstructure are factors related to the ability of a metal to deform readily without the need for large stresses. Hot working, of course, allows deformation to be undertaken at relatively low stress levels. On the other hand, in some hot working ranges the development of fracture is nonductile, due to the melting (hot shortness) or considerable softening of certain second phases, or to excessive grain-boundary deformation. In any case, it is important to recognize that the ability of a metal to deform extensively without fracturing is very sensitive to temperature. Thus, processing should reflect careful attention to temperature optimization and control, and workability evaluations must be undertaken in the temperature ranges to be encountered during processing.

Somewhat similarly, the stresses and fracture tendencies developed by a deforming metal are sensitive to strain rate. Generally, higher strain rates are consistent with increased stress requirements and a greater likelihood of fracture. (An important exception is carbon steel in the 200–400 °C, or 390–750 °F, range, wherein dynamic aging effects reverse the stress-strain rate relationship.) In extrusion and drawing, on the other hand, these effects are complicated by frictional heating. Moreover, the rather high strain rates involved in commercial drawing and extrusion are not easy to develop in simple, well-defined test formats (short of setting up a pilot model of the process itself). At a minimum, however, the investigator should be aware that flow stress and fracture behavior are fundamentally strain-rate sensitive.

Lastly, and of most practical interest, second-phase content, inclusions and internal cavities, and external damage are of great importance to workability. In the extreme case, they reduce workpiece cross section or cause extraordinary strain concentrations in their vicinities, to the extent that fracture follows almost immediately. More commonly, they serve as sites ("fracture centers") for the development of pores or cavities in the early stages of ductile fracture. Thus, the larger or more numerous these defects are, the sooner or more rapidly ductile fracture will develop. Even very fine scale precipitation can produce a marked acceleration in the development of ductile fracture. An interesting case of this can be seen in the work of Coffin and Rogers, where centerline porosity (the initial stages of centerbursting) is shown to be more pronounced in drawn ETP copper than in drawn oxygen-free high-conductivity (OFHC) copper (Ref 10, 11). Interesting observations on the evolution of centerburst fractures in extrusion have been set forth by Pepe (Ref 26).

Thus, for a given stress state, temperature, and strain rate, the workability variations demonstrated by an array of materials will strongly reflect any variations in inclusions, second phases, internal cavities, and external damage. This fact has led to extensive use of approaches such as metallographic characterization, nondestructive testing, and dye checking to predict workability. These techniques can provide much useful information. However, they are no substitute for mechanical testing.

Projecting Workability from Mechanical Tests

The ultimate workability test is the process itself. Indeed, in some instances manufacturers may wish to set up a pilot line for the purpose of qualifying and evaluating new materials. Beyond this, careful observation of manufacturing operations can yield much information on workability, even when the metal being processed is of "acceptable" quality. Nonetheless, it is usually desirable to have abstract mechanical test procedures at hand for the purpose of projecting workability.

Unfortunately, wire drawing and extrusion have not had the benefit of the extensive developmental work that has gone into the projection of workability limits in, say, sheet metal deformation. Thus, it is not often possible to project *absolute* levels of workability in wire drawing and extrusion from simple mechanical test formats (an important exception is that of the maximum possible single-pass wire drawing reduction, where rather good quantitative estimates are readily made). Rather, this discussion focuses on testing that can reveal *relative* degrees of workability from one sampling of feedstock to the next. Of course, pertinent results from these tests can always be correlated with observed processing limitations to allow empirical projections of absolute workability.

The tensile test is probably the most important of all mechanical tests, and it can be useful in projecting drawing and extrusion behavior. The concept of the tensile test, as used here, should be extended to include aspects of various "stretch tests" as commonly employed in the wire industry and short-time hot tensile tests such as can be run on a hot tensile (Gleeble) testing apparatus (Ref 27) (see chapter 7). The first consideration is that the test temperature be relevant. Such a requirement of course demands knowledge of the process itself, and, particularly at the workpiece surface, processing temperature can be difficult to assess. Beyond this, the *cumulative* heating effects of thermomechanical working must be considered, particularly in wire drawing, where high-speed "cold" processing can generate wire surfaces that are actually at several hundred degrees Celsius. In any case, the test temperature should be carefully considered. The selection of a relevant strain rate in a tensile test (or any simple mechanical test) can be very difficult in view of the wide range of strain rates, and the very high strain rates, encountered in the processes themselves. In general, it is recommended that the tests be run at the highest practical strain rate.

The tensile-test results that are germane to workability in extrusion and drawing are the *flow-stress level* and the *area reduction at fracture.* Flow-stress data are of course directly useful in estimating limitations due to extrusion-pressure capacity. The relevance to wire drawing is less clear-cut since the draw stress and capacity for draw stress are both based on flow stress. Beyond this, the flow-stress level has many uses in assessing frictional effects, die and tooling pressures, thermomechanical effects, and other factors somewhat outside the scope of workability. However, the really important piece of data is the area reduction at fracture. This is an expression of the intrinsic ductility at fracture, in the presence of hydrostatic tension, and should be very much related to the capacity to resist centerbursts and ductile process fractures in general.

From a somewhat more sophisticated basis, the flow stress and the area reduction at fracture can be used to estimate the value of C in the analysis of Cockcroft and Latham (Ref 16). For the case of a simple tensile test:

$$C = \int_0^{\bar{\varepsilon}_f} \bar{\sigma}\left(\frac{\sigma^*}{\bar{\sigma}}\right) d\,\bar{\varepsilon} = \int_0^{\bar{\varepsilon}_f} \sigma d\varepsilon = \hat{\sigma}_0 \varepsilon_f \qquad \text{(Eq 4)}$$

where $\hat{\sigma}_0$ is again the average flow stress and ε_f is the true strain at fracture or ln (A_i/A_f), where A_i and A_f are the initial and final cross-sectional areas, respectively. The use of the Bridgman correction for stress in the *necked* region (Ref 28) is recommended for accurate determination of the average flow $\hat{\sigma}_0$ stress. Even if the Cockcroft and Latham analysis cannot be applied to the prediction of absolute workability, the value of C, or even the area reduction itself, is an excellent index of relative workability in wire drawing.

In the case of very ductile metals, the area at fracture may be very small and difficult to measure. In this regard, the scanning electron microscope (SEM) can be very useful. (As a word of caution, the magnifications set forth on the SEM may not be as precise as desired for such measurement. This problem can be surmounted to some extent by placing objects of known dimensions in the field of view.)

Most of the other tensile test data are of little use in workability assessment. In particular, the elongation (uniform or total) has limited meaning because it is largely dictated by the onset of necking. Necking is a plastic instability that does not occur in normal wire drawing and extrusion, and so the elongation has little bearing on relative workability. It is unfortunate that such wide use is made of elongation or stretching data in an effort to forecast workability.

Compression testing (see Chapters 5, "Cold Upset Testing" and 6, "Hot Compression Testing") can also be useful in projecting workability in drawing and extrusion. There are some major limitations, however. First, the tensile hydrostatic stress state germane to centerbursts is not developed. Second, the flow stress may be obscured by uncertain "friction-hill" effects at the sample ends. Third, the fractures are generally at the surface, and the workability of material in the interior of the workpiece may not be reflected. On the other hand, knowledge of the resistance to surface cracking is important.

With proper analysis of the stress and strain at the specimen surface, the compression test may be used to determine a value of C for the Cockcroft and Latham analysis. Work along this line has been undertaken by Hoffmanner (Ref 13), who observed strain to fracture on barreled cylindrical surfaces during upsetting. The strains could be monitored through the deformation of a grid printed on the specimen surface. Stress state could be related to strain through the relation:

$$\frac{d\varepsilon_2}{d\varepsilon_3} = \frac{(2\sigma_2 - \sigma_3)}{(2\sigma_3 - \sigma_2)} \qquad \text{(Eq 5)}$$

where ε_2 and σ_2 are the circumferential strain and stress, respectively, and where ε_3 and σ_3 are the strain and stress, respectively, in the axial direction.

The work of several authors (Ref 29–31) has resulted in the following criterion for surface fracture in the compression of cylindrical specimens:

$$\varepsilon_1 + \frac{1}{2}\varepsilon_2 = a \qquad \text{(Eq 6)}$$

where ε_1 is the circumferential strain, ε_2 is the axial strain at the equatorial free surface, and a is a material constant. Thus the value of a may be used to index workability at least as far as *free-surface* fracture is concerned. For example, a values of 0.32 and 0.18 have been measured, respectively, for type 1020 steel and type 303 stainless steel (Ref 31). Unfortunately, the surfaces undergoing fracture in extrusion and drawing generally are not *free* surfaces. Even so, exploration of the use of Eq 6 in assessing surface workability in extrusion and drawing may be worthwhile.

In summary, however, the compression test is less suited to assessing workability in extrusion and drawing than might be imagined.

The torsion, or twist, test is widely used to assess workability in cylindrical stock (see Chapter 8, "Torsion Testing to Assess Bulk Workability"). Even where the load is not measured, the number of turns to failure is felt to be a useful measure of *surface-region* fracture resistance. It should always be emphasized, however, that the strain is highest at the surface in this test. That is:

$$\gamma = \frac{R\Delta\phi}{L} \qquad \text{(Eq 7)}$$

where γ is the torsional shear strain, R is the radial position, L is the gage length, and $\Delta\phi$ is the angle, in radians, through which one end of the gage length has been twisted relative to the other (one revolution = 2π radians). Moreover, the surface is a free surface. Thus, the fracture resistance manifested in this test is only that of the surface material, and of the unconstrained surface, at that. Thus, as in the compression test, the torsion test has little relevance to centerburst problems and may not be an accurate indicator of surface fracture behavior under constraint of the extrusion or drawing die.

Even so, the torsion test can be recommended for extrusion and drawing workability testing on a qualified basis. It has the advantage of allowing very large strains to be generated. Thus, growth of surface flaws can be rather carefully evaluated in very ductile materials. In general, fracture (or failure via observable surface cracks, etc.) will occur when the maximum torsional shear strain, γ_{max}, reaches a critical level. The value of γ_{max} is a measure of workability and may be calculated from:

$$\gamma_{max} = \frac{R_{max} \cdot N_{max}}{L} \qquad \text{(Eq 8)}$$

where R_{max} is the actual radius of the cylindrical stock, and where N_{max} equals $2\pi n$, with n being the number of turns to fracture or failure. Of course, for any fracture strain or level of workability, a maximum number of turns exists and n becomes an index of workability.

Bend testing is also widely used to assess workability of rod and wire. Basically, the rod or wire is bent or wrapped around a mandrel. In bending, the strain is maximum along a surface locus, with fracture (or failure via observable surface cracks, etc.) occurring when the maximum tensile strain (at the outer bend radius) reaches a critical level. This maximum tensile strain (ε_{max}) can be calculated from:

$$\varepsilon_{\max} = \frac{1}{\left(1+\dfrac{D}{d}\right)} \qquad \text{(Eq 9)}$$

where d is the rod diameter and D is the mandrel diameter. For any given fracture strain (ε_f) a minimum value of D/d (or D) exists, such that:

$$(D/d)_{\min} = \left[\frac{1}{\varepsilon_f}\right] - 1 \qquad \text{(Eq 10)}$$

If the rod or wire is itself used as a mandrel, the test is often called a wrap test. For a wrap test D/d is 1.0, and ε_{max} is 0.5. Thus, if a rod or wire passes a wrap test, ε_f exceeds 0.5.

As in the case of torsion, bend testing is most relevant to workability limitations associated with surface cracking. It has no relevance to workability limitations related to structure peculiar to the centerline.

More Elaborate Test Formats. A major problem with the previously described tests is the absence of a die and the associated metal flow constraints. It is desirable to measure workability with a test that comes closer to process mechanisms. The most practical approach to this is to use a laboratory-scale drawing system. Even a tensile-testing machine, outfitted with a die holder, can be used to draw short lengths of stock. Of course, deformation at hot working temperatures and at high strain rates will require more elaborate equipment.

In any event, some practical approaches for assessing workability on this basis are:

- Draw a series of specimens once each through respective dies with different die angles, using the same reductions for each specimen. The value of Δ increases with increasing die angle, of course, and the die angle at which centerbursting is first noticed can be used as an index of centerburst resistance for the material; the higher the value of α, the better the resistance. A 10% reduction with α values of 4 to 16° will provide a range of Δ values from 2.7 to 10.6.
- For materials with too much workability for failure in a single pass, set up a sequence of dies allowing successive reductions to be taken at relatively high values of Δ (Δ ≈ 5). The reductions and die angles can be the same in each pass. The *total* reduction at which centerbursting is first observed can then be used as an index of centerburst resistance; the greater the total reduction, the higher the resistance. A series of 15% reductions at an α value of 12° (Δ = 5.2) is recommended.
- As an alternative to waiting for gross centerbursting to occur, one can assess "ductile damage" or incipient ductile fracture at the centerline. This can, of course, be done metallographically or by *careful* density measurements (Ref 10, 11). Pores on the order of 1 μm in diameter are fairly easy to resolve with conventional metallographic techniques, and Coffin and Rogers detected ductile damage-related density changes of a few parts in ten thousand.
- Some indication of resistance to surface cracking can be obtained by drawing the stock under conditions of little or no lubrication or under conditions of thick-film, dry-soap lubrication. Scanning electron microscopy will reveal the extent of surface cracking, and comparative evaluations can be made at a given cumulative reduction (Ref 3). Three or four passes of 20% each may be necessary to establish a consistent pattern of surface damage. With the inconsistent boundary lubrication conditions that will develop with little lubricant, stick-slip behavior should promote surface cracking. On the other hand, the rough, somewhat unconstrained surface that develops with dry soaps will also provide local strain concentrations predisposed to cracking. The results of such testing must be viewed with caution, however, owing to problems in maintaining consistent lubricant performance.

Process Design

Extrusion-Process Design Implications. In many instances, extrusion-process design maximizes productivity and thereby maximizes exit speed. Limitations to extrusion speed involve press capacity and surface temperature, as shown in the workability or extrusion-limit diagram of Fig. 7. This is a widely used format for extrusion-limit diagrams (Ref 33 and 34). The maximum extrusion speed, V_{max}, is to be found at the intersection of two curves. The scored area under the two curves is the practical process regime, with maximum productivity associated, in principle, with the highest speed, at the top of the scored area.

The left-side curve in Fig. 7 expresses extrusion speed, for a given available press load, with the speed increasing as flow stress diminishes with increasing temperature. The right-side curve expresses allowable extrusion speed, for a given surface liquation temperature, with speed increasing with lower overall exit temperature, since there can be more surface heating (from friction and lack of time for dissipation) without surface melting. When lubrication is used with extrusion, the right-side curve in Fig. 7 should be shifted to the right, as friction

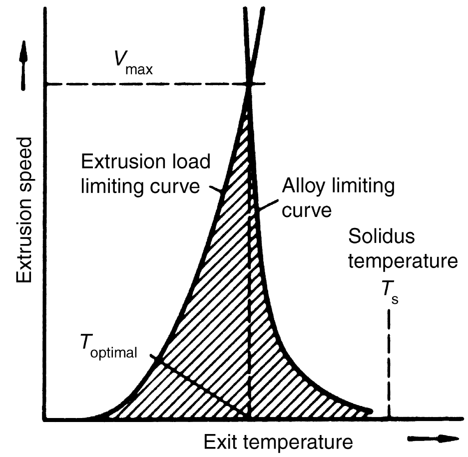


Fig. 7 Generic extrusion limit diagram. Source: Ref 32

is reduced. Thus, V_{max} would be increased. However, many extrusion operations are undertaken without lubrication, and the surface traction is fixed at a large fraction of the extrudate shear strength (at the prevailing temperature and strain rate). Except in the cases of certain light sizing or shaping passes, or certain complex shapes, extrusion reductions are generally so large that centerline stresses remain generally compressive, and centerbursting is not a problem.

Drawing-Process Design Implications. From a single-pass perspective, there is a limiting reduction that may be taken, associated approximately with the condition where $\sigma_d/\hat{\sigma}_0$ is unity. (The precise limit would be for the case where the drawing stress equaled the flow stress at the die exit, which may differ from the average flow stress.) This limiting reduction can be maximized if the drawing pass is designed to minimize the draw stress (see Chapter 19, "Drawing of Wire, Rod, and Tube"). However, only rarely is this single-pass reduction limit approached or desired in practice. Moreover, it is not a function of the intrinsic workability of the drawing stock.

Many multipass-drawing-process design concepts involve reductions consistent with standard gages, and the remaining design issues can be explained in terms of the Δ parameter. There are a wide variety of drawing technologies that benefit from low Δ process designs. With lower values of Δ (~1.0 to 2.0), one reduces redundant work and general strain accumulation, one reduces centerline tension and the propensity for centerbursts, and one lowers die pressure. The minimization of centerline tension and general strain accumulation should allow maximum multipass reduction to be achieved without general ductile fracture or centerline fracture. Intermediate Δ (~2.0 to 3.5) drawing-process designs have been common, historically, and may offer the advantages of reduced "fine" development (a "fine" is a small particle of the wire surface that becomes liberated during drawing) (Ref 35). Higher Δ drawing-process designs may be employed when high friction is unavoidable or when a high die pressure is desired to iron the wire surface and brighten the wire.

When drawing-temperature control is a priority, it may be useful to use "tapered" drawing schedules that reduce the drawing reduction as the wire work hardens, thus achieving a relatively constant draw stress, and relatively constant degree of thermomechanical heating, from pass to pass. This concept is largely independent of the workability of the wire.

The achievement of maximum drawing productivity through maximum drawing speed may be highly problematical where lubrication is marginal, owing to very high surface temperatures and related surface failure. The increased speed not only limits the dissipation of frictional heating, but it limits the time available for cooling between drawing passes. However, where good lubrication is readily achieved, increased speed may abet lubrication.

REFERENCES

1. G.E. Dieter, Ed., *Workability Testing Techniques,* American Society for Metals, 1984
2. A.W. Funkenbusch, F. Bacon, and D. Lee, Technical Information Series Report No. 78CRDO17, General Electric Corporate Research and Development, Feb 1978
3. R.N. Wright and A.T. Male, *Trans. ASME, Lubr. Eng.,* Vol 97, Series F (No. 4), 1975, p 134
4. R.N. Wright, *Wire J. Int.,* Vol 35 (No. 8), 2002, p 86
5. W.A. Backofen, *Deformation Processing,* Addison-Wesley, 1972, p 303
6. D. Lee, *Metall. Trans.,* Vol 6A, 1975, p 2083
7. R.N. Wright, *Wire Technol.,* Vol 6 (No. 3), 1978, p 131
8. R.J. Wilcox and P.W. Whitton, *J. Inst. Met.,* Vol 88, 1959–1960, p 145
9. W.A. Backofen, *Deformation Processing,* Addison-Wesley, 1972, p 88
10. L.F. Coffin, Jr. and H.C. Rogers, *Trans. ASM,* Vol 60, 1967, p 672
11. H.C. Rogers and L.F. Coffin, Jr., *Int. J. Mech. Sci.,* Vol 13, 1971, p 141
12. D.J. Latham and M.G. Cockcroft, "The Effect of Stress System on the Workability of Metals," Report No.216, National Engineering Laboratory, Feb 1966
13. A.L. Hoffmanner, Technical Report AFML-TR-69-174, Air Force Materials Laboratory, June 1969
14. S.-I. Oh, C.C. Chen, and S. Kobayashi, *Trans. ASME, J. Eng. Ind.,* Vol 101, Series B, 1979, p 36
15. B. Avitzur, *Trans. ASME, J. Eng. Ind.,* Vol 90, Series B, 1968, p 79
16. M.G. Cockcroft and D.J. Latham, *J. Inst. Met.,* Vol 96, 1968, p 33
17. C.H. Lee and S. Kobayashi, *Trans. ASME, J. Eng. Ind.,* Vol 95, Series B, 1973, p 865
18. R. N. Wright, T. A. Kircher, and J. R. Vervlied, *J. Met.,* Vol 39 (No. 10), 1987, p 26
19. W.A. Backofen, *Deformation Processing,* Addison-Wesley, 1972, p 150
20. E.H. Lee and R.M. McMeeking, *Trans. ASME, J. Eng. Ind.,* Vol 100, Series B, 1978, p 386
21. R. Hill and S.J. Tupper, *J. Iron Steel Inst.,* Vol 159, Part 4, 1948, p 353
22. E.H. Lee, R.L. Mallett, and R.M. McMeeking, *Numerical Modeling for Manufacturing Processes,* Jones, Armen, and Fong, Ed., ASME, 1977, p 19
23. E.H. Lee, R.L. Mallett, and W.H. Yang, *Comput. Meth. Appl. Mech. Eng.,* Vol 10, 1977, p 339
24. J.G. Wistreich, *Proc. Inst. Mech. Eng.,* Vol 169, 1955, p 654
25. R.N. Wright, *Wire Technol.,* Vol 4 (No. 5), 1976, p 57
26. J.J. Pepe, *Met. Eng. Q.,* Vol 16 (No. 1), 1976, p 46
27. W.F. Savage, *High Speed Testing,* Vol III, Interscience Publishers, 1962, p 55
28. P.W. Bridgman, *Trans. ASM,* Vol 32, 1944, p 553
29. H. Kudo and K. Aoi, *J. Jpn. Soc. Technol. Plast.,* Vol 8, 1967, p 17
30. S. Kobayashi, *Trans. ASME, J. Eng. Ind.,* Vol 92, Series B, 1970, p 391
31. P.W. Lee and H.A. Kuhn, *Metall. Trans.,* Vol 4, 1973, p 969
32. K. Laue and H. Stenger, *Extrusion: Processes, Machinery, Tooling,* American Society for Metals, 1981
33. W.Z. Misiolek, "Engineering Concepts in Extrusion," Aluminum Processing Program Workshop, Rensselaer Polytechnic Institute, 18 April 1994
34. H. Stenger, *Drahtwelt,* Vol. 59, 1973, p. 235 and 371
35. G. Baker, "Workpiece Wear Mechanisms in the Drawing of Copper Wire," Ph.D. dissertation, Rensselaer Polytechnic Institute, 1994

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 325-336
DOI:10.1361/hwpd2003p325

Chapter 22

Multidisciplinary Process Design and Optimization: An Overview

H.L. Gegel, Air Force Research Laboratory, Wright-Patterson Air Force Base (Retired)
R. Grandhi and J.S. Gunasekera, Wright State University
Charles Gure, Consultant

ENGINEERING DESIGN is the reverse of engineering analysis. It is the process of product realization that involves the synthesis of knowledge and techniques from engineering, science, aesthetics, economics, and psychology in establishing specifications for products and their associated production processes. Multidisciplinary process design and optimization (MPDO) is an engineering environment that supports process design and optimization. Its purpose is to automate, optimize, and accelerate the design process. The MPDO design environment is the process by which engineering descriptions and specifications are integrated with material process analysis techniques to ensure that a product possesses the desired behavior, performance, quality, and cost, not to mention satisfying the requirement of delivering the product when the customer needs it.

The MPDO concept represents an integrated infrastructure of software, standards, and databases that allows companies to solve their toughest design problems by making more effective use of their design software and human resources. The MPDO environment is a result of a series of technical developments that have occurred over a period of four decades of materials and manufacturing research. Some of the technical and business developments include the following:

- Analytic and numerical mesh-based and material models for understanding and evaluating materials processing
- Computer-aided design/computer-aided manufacturing (CAD/CAM) and computer-aided engineering (CAE) tools for creating part and die geometries and automatically meshing the geometries for analysis
- Workflow process modeling, which involves defining and implementing a collection of activities (structured or ad hoc) that must be performed with different types of human activities and processes performed by tools
- Evolutionary computation, expert systems, and numerical optimization to achieve more effective solutions than solutions created by human engineers. (Evolutionary computation is a term for biologically inspired artificial-intelligence technology such as genetic algorithms and neural networks. The idea of a genetic algorithm is based on the evolutionary concept of mutation that "breeds" solutions, which may be better than those devised by human engineers. Neural-network methods refer to simulation models based on massively parallel processing similar in concept to the structure of neurons in a brain. These programs learn by looking at examples as images or collections of processing data. They presumably may spot relationships that may escape human notice.)

These developments are part of a collection of advanced engineering and business design techniques that provide design excellence and high-quality products for a given product line or company. Additionally, these tools are used when a need exists to shorten lead times, get products to market faster, enhance value-stream analysis, and eliminate scrap or activities that do not add value for the customer. They are capable of reducing design time and giving engineers the means to focus on design decisions that maintain upstream and downstream quality and value for the customer.

This chapter briefly introduces MPDO as a type of engineering productivity software. The commercial needs (Ref 1) for MPDO are categorized as:

- Modeling to reduce iteration of trials (speed and cost)
- Modeling to optimize an existing process to reduce cost
- Modeling to improve a customer-specified quality (reduced process and product variability)
- Modeling to reduce product delivery time.

A multistage design process is described to show the benefits of the MPDO concept in upstream process design of a forging. However, MPDO can be utilized to create optimal designs for all metalworking processes as forging, extrusion, rolling, shape rolling, and cog milling. The design techniques are mostly the same for various bulk forming processes, only the design rules, material databases, and the equipment characteristics will vary.

Additional examples of process design and optimization are also provided in the chapters that follow. "Optimal Design of Thermomechanical Processes," Chapter 23, demonstrates the application of dynamical modeling and process design techniques using simplified analysis models and trajectory optimization for solving a class of problems where shape is not a controlling factor. "Computer-Aided Optimization for Improved Process Engineering Productivity of Complex Forgings," Chapter 25, emphasizes the use of a multiobjective system for the integrated design of forging process conditions, preform shapes, and stages. Nonlinear viscoplastic finite element methods and mathematical optimization tools are utilized for simulating metal forming and shape optimization for preform design. "Process Design of Gas Turbine Engine Components Using iSIGHT for Process Optimization," Chapter 24, demonstrates the benefits of using a commercially available computer-aided optimization (CAO) application (iSIGHT) to obtain an optimal forging product and process design.

Why Use MPDO?

Competitive forging suppliers are streamlining their manufacturing processes to compete more effectively in the global market. By shifting the focus from manufacturing hours to manufacturing parts, a new performance indicator for how well a manufacturing job has been done has evolved (Ref 2). Competitive companies now measure delivery performance based on the date that the customer wants the order, not on the date someone at the company feels comfortable about pledging.

Improving the practice of engineering design and manufacturing is essential for achieving industrial excellence (Ref 3 to 5). In the 1990s, best practices in the industry, which focused the vision on what could be done to convert losses into gains for the industry, were identified. Studies were done on different types of companies to shed some light on the variables that influence productivity growth at the local level (Ref 6). The types of companies were:

- Process companies, where manufacturing is characterized as being highly connected and automated
- Batch processing, where manufacturing is characterized as a disconnected line-flow organization of work
- Several different batch processes, where manufacturing is characterized by a very rapid change in both product and processes

In none of these companies did the traditional profit and loss statements provide adequate up-to-date information about performance.

A new measuring system called *total factor productivity* (TFP) was devised to measure the overall efficiency of a manufacturing facility (Ref 6). It identified the following practices, which, when managed correctly, really make a difference:

- Investing in new equipment
- Reducing waste
- Cutting work-in-process

The study showed unequivocally that capital investment in new equipment and design tools are essential for sustaining growth in TFP over a long time; that is, a decade or more. A negative correlation between waste rates; that is, scrap and rework, was expected, but the magnitude of the effect of waste on TFP was not anticipated. A one-tenth reduction in waste increased the TFP by 3%. The strength of this relationship was amplified when a decision to boost the production throughput rate was made. This decision was made to theoretically raise TFP because of large-fixed components in labor and capital costs. Unfortunately, the waste ratio increases, and this increase indicates the powerful impact of waste reduction on productivity.

Cutting the amount of work-in-process (WIP) for a given level of output is much greater than can be explained by reductions in working capital. A reduction in WIP by one-tenth produced a 9% rise in TFP. Empirical evidence from studying both Japanese and U. S. companies supports the observed WIP relationship and TFP. Cutting WIP generally leads to faster, more reliable delivery times and lower rejection rates. Reducing WIP clearly cuts overhead costs, and WIP is invariably a symptom of poor product and material process design, lack of process understanding, erratic process yields, unreliable equipment, unfocused business processes incapable of producing outputs that satisfy customer requirements, poor responsiveness to changing customer expectations, and insufficient critical measures of performance—cost, quality, speed, and service.

Dramatic increases in efficiency are thus possible by solving the problems that create them in the first place. Creating order and promoting understanding of business processes make for a powerful combination—and a powerful lever for competitiveness.

Lean Manufacturing. In the mid-1990s, high-technology suppliers began moving toward lean manufacturing by exercising some common-sense principles such as (Ref 7):

- Precisely specifying *value* for each job activity and product
- Identifying the *value stream* for each product
- Making *value flow* without interruptions
- Letting the customer *pull* value from them
- Always pursuing *perfection*

Lean manufacturing always starts with *value as defined by the customer* (Ref 8–10). A lean-manufacturing business strategy includes close contact with the client. To please the client, an all out effort is made to find out all the necessary information needed to satisfy the client's expectations and to reach an optimal solution. Competitiveness is achieved by applying both workflow and process modeling to the part manufacturing process.

To compete effectively, some forging companies also have adopted a step-by-step strategy to reach a highly efficient mode of lean manufacturing (Ref 11, 12). The lean manufacturing of forged products, when compared with the traditional manufacturing hour business strategy, is characterized by less human effort, less rework and scrap, less time, less space, and smaller economic lot size. Process-activity diagrams describe the workflow in terms of specific activities in engineering offices, purchasing offices, sales departments, and activities in the forge shop. Causes for bottlenecks in a targeted work area were identified to reduce the time parts are processed out of sequence and to maximize value flow within the work area.

Process Design Methods

The traditional build-and-test methods of developing a complete manufacturing process were heavily experienced based until the end of the 1980s (Ref 13). Major barriers existed to increasing the use of models in industry with how to implement existing modeling technology. One of the most important aspects was deciding whether to use analytical modeling or mesh-based models. Since then, many software applications have been developed for analyzing the process. By the end of the 1980s, mesh-based modeling became the dominant method for most industrial manufacturing processes, and powerful mesh-generating and remeshing software became commercially available. The finite element method (FEM) was selected for the following reasons: (a) engineering applications tend to involve complex shapes; (b) the process physics of the bulk metalworking processes is highly nonlinear, including geometry deformation and material and boundary condition behavior; and (c) the stability of the process is governed by coupled phase transformations, plastic flow, and heat flow. Additionally, by relaxing some conditions on the FEM, a one-step simplified finite element method (SFEM) that has computational speeds corresponding to analytic models can be used to speed up the design process. Most industrial design for manufacturing is done by teams of specialists, not by a single individual.

Manufacturing processes inherently involve the processing of materials, and understanding these processes at some level of detail is necessary if predictive models are to be used to optimize the processes. The MPDO techniques require a menu of analysis tools, some complex and others simple, which are capable of predicting the time evolution of internal variables and input-output relationships for the system. The design focus places emphasis on solving processing problems and using appropriate modeling techniques to optimize the material, process, and equipment variables within delivery time and quality constraints set by the customer. The type of model is related to design objectives such as shape complexity, temperature, and allowable error. A summary of various design objectives is given in Table 1 for several analytical and numerical methods used in the modeling of deformation processes. These analytical and numerical methods are described in Chapter

Table 1 Methods of deformation analysis and typical design capability ratings

Method(a)	Die load	Die pressure distribution	Shape complexity	Sculptured surface	Strain rate	Temperature	Error, %	CPU time	Rating
SLAB	G	G	…	F	…	…	25	1–5 s	M
UB(b)	M	…	…	G	M	…(c)	15	1 min	M
UBET	M	…	M	…	F	…(c)	10	3–5 min	F
SLF	M	P	P	…	M	…	10	5–7 min	M
SFEM	M	M	G	VG	VG	G	10	1–5 s	G
FEM	G	G	G	VG	VG	G	5	1–4 h	G

See text for discussion. (a) See also Chapter 1 for references and an introduction to modeling methods. (b) More suitable for metalworking operations that are steady state such as extrusion, drawing, and rolling. (c) A temperature module must be added separately. VG, very good; G, good; F, fair; M, moderate; P, poor.

1, "Workability and Process Design—An Introduction," while the following provides a brief description of design capabilities of these methods.

Table 1 summarizes the results from the development of different analysis methods that are important in the context of industrial process modeling to simulate metalworking processes. The Sachs (Slab) method (SM), upper bound method (UBM), and the slip-line field (SLF) theory, which are relatively simple methods compared to the FEM, use a rigid plastic approximation for the constitutive equation of the workpiece. The approach for the practical use of the Slab and UB methods is to assume a velocity field and for this velocity field to estimate the average strains, strain rates, and temperatures within each distinct zone of the velocity field. Knowing the effective flow stress and friction, one applies the necessary equations for predicting the stress distribution and forming load in the Slab method and the average forming pressure in the UBM.

These simple models require simple geometries and uniform boundary conditions even though they can cope with some complexity in the constitutive behavior of a material (Ref 14). The slip-line field (SLF) is an older method that can be used for plane strain and plane stress problems, as it is not applicable for three-dimensional or axisymmetric problems. The material deforms plastically along trajectories of maximum shear stress called slip lines. Plastic equilibrium equations are made up of slip lines, and the characteristics for stress and velocity are parallel, resulting in the strain-rate for the plane strain and plane stress geometries being characterized by a single slip line field (Ref 15). These simple models still have interest for industrial process modeling for specific cases where they are known to be reliable, because they provide a basis for comparison. These specific cases include those that can be modeled by steady-state approximations as extrusion, plane-strain forging, rolling, and drawing.

Mesh-based models are needed, if the process design objective is to improve quality by reducing variability in properties, microstructure, and scrap. The upper-bound elemental technique and the finite element method (FEM) are useful for this class of industrial problem.

Upper Bound Elemental Technique Analysis

The upper bound elemental technique (UBET) is useful for parts that are mainly two-dimensional or axisymmetrical. The part is automatically divided into rectangular or triangular zones where the velocity field has a prescribed form. At the interelement boundaries, the normal velocity is continuous, and the elements are designed to ensure appropriate boundary conditions on the tools. As Table 1 shows, the computing time is low for computing the forging force, material flow, and the local normal stress using an approximate method.

Analysis by UBET provides a value that is equal to the actual force or higher, and the predicted total power consumption inside the die is higher than the power actually consumed. In general, the approach to using this analysis method begins by assuming a pattern of deformation, which is usually based on the modeler's conception of the actual flow. The velocity components satisfy the volume continuity of the material at entry and exit of the die and boundary conditions. The strain-rate components that satisfy the incompressibility condition are calculated, making the velocity field "kinematically admissible;" (that is, the analytical method postulates a velocity field that obeys continuity requirements and the prescribed boundary velocity requirements). On obtaining the velocity field, the total power consumption inside the die is computed numerically for different conditions of friction, strain-rate, and flow stress. The average pressure can be calculated since the total power consumption is obtained. In its present form, the UBET method appears to be close to the FEM with the advantage that the velocity field is truly compressible.

Finite Element Method (FEM)

The FEM discretizes the continuum body, such as the billet, into many small units called elements just like the UBET method. The difference is that the FEM has nodes, which, on the element boundary, connect neighboring elements. The unknown parameters are located at the nodes as nodal velocity, temperature, and so forth. The velocity field or temperature field is assumed within the element using shape functions and the nodal unknowns. The material properties, which can be a function of temperature, strain-rate, and strain, are located at the numerical integration points inside the elements. The local unknowns are then assembled into a global linear equation system. After solving this normally large set of linear equations, the nodal unknowns are obtained such as velocity, temperature, or displacements. Iterations are required because of the nonlinearity of the material and the geometry deformation. The UBET method does not appear to have as much flexibility as the FEM when the components being modeled have complicated flow patterns, and three-dimensional analysis is presently difficult (Ref 16).

The FEM is widely used by industrial process modeling teams, because the major advantage for this method is its ability to generalize (Ref 15–17). The FEM has applicability to a wide class of boundary value problems with few restrictions on geometry. To illustrate the advantage of the FEM, if the objective of an industrial design team is to improve the quality of a part with a complex shape by reducing variability in the microstructure, mechanical properties, and waste, complex models of the workpiece, forge press, lubricant, and dies that consider second-order effects may be required. Only the FEM has the flexibility to handle this class of industrial metalworking problem.

In practical metalworking processes, a number of preforming operations are required to transform a simple initial geometry into a complex shape while achieving desired tolerances, residual stresses, microstructures, and properties. In a multistage forging process, the various workpiece shapes are achieved by using dies of various shapes. Therefore, the advantages of the FEM are needed to fully treat the nonlinearity of the boundary conditions of the dies and the constitutive behavior of the workpiece material in the process analysis. The major disadvantage of the conventional FEM (as Table 1 shows) is the central processing unit (CPU) time.

In order to make the FEM useful in the early design stage to correspond to the simplified method time scale for analysis, a simplified FEM can be used to speed up the design process. As the following discussions briefly note, the simplified finite element method (SFEM) has the following characteristics:

- Only one-step computation is required.
- History variables are used as initial conditions only.
- Geometry-based FEM models can be used.
- Parametric geometry input can be used, specifically for die and preform shape design.

The one-step FEM approach dramatically reduces the computation time, because forming simulations usually require hundreds of steps. This use of the FEM is in correspondence with the simplified methods listed in Table 1 where these simplified methods are applied to analyze a typical step of the forming process during the design stage.

In many steady-state forming processes such as extrusion, profile rolling, ring rolling, and drawing, a one-step calculation is adequate in the early design stages to obtain the forming load and material flow. History variables such as temperature and strain, which are important for the transient problem, are not considered in the SFEM. However, they can be used as initial conditions to obtain the correct material properties for the analysis. Since SFEM is not intended for detail design, the effect of history variables can be ignored.

Geometry-based FEM models allow the user to easily set up the FEM model. This type of FEM model is set up by inputting the geometry and assigning boundary and initial conditions, which are normally known to the designer regardless of what computation method he or she chooses to use. This is a good approach to use in industrial process modeling, because no FEM knowledge is required at this stage. The FEM model is created internally based on very few control parameters such as the element size or time step. The geometry-based system is now possible because of the new technology of generating finite element meshes. The Delaunay and Advancing Front triangulation methods allow the user to generate unstructured finite element meshes automatically on any arbitrary two-dimensional and three-dimensional geometry. The only input required from the user apart from the part geometry is the element size or mesh density.

In industrial process modeling, it is important in the design to allow the design team to change or modify their previous design easily and quickly without having to set up the FEM model again. The parametric geometry attributes make rapid design changes possible. When one attribute changes the whole geometry, the FEM model associated with it will change accordingly, without user intervention. This flexibility and automation of FEM model creation make the SFEM a useful analysis tool in conjunction with computer-aided optimization to automate material process design and optimization.

Several SFEM simulations are subsequently illustrated to demonstrate the usefulness of the SFEM for process design and optimization.

Compute Preform Size. Computing the preform size is the first thing that a designer will do. The SFEM allows the designer to compute the preform size based on the final part geometry at the click of a mouse. The examples in Fig. 1 and 2 calculate the sizes of the different parts. An advantage that this method has over other simplified analysis methods is that an inexperienced analyst does not have to be concerned about assuming an admissible velocity field or deriving any analytical equations for any particular forming process. The SFEM approach to simplified modeling, like the full FEM, is less restrictive, and, with parametric geometric attributes, previous design alternatives can be investigated with less human effort.

Comparison of One-Step FEM with the Full FEM. A series of simplified finite element method (SFEM) simulations can show how the FEM can be made to be in correspondence with the upper bound (UB) methods given in Table 1. The Figure 3 geometry represents the initial billet geometry, which is a simple cylinder. Figures 4 and 5 represent intermediate shapes between the initial billet geometry and the final geometry shown in Fig. 6. The final geometry in Fig. 6 represents an engine disk typically made from a high-temperature titanium alloy. These one-step simulations correspond to what a forging designer often describes as a "short-shot" when studying die filling by making partial forgings in a trial-and-error design process. The reason for doing these one-step simulations is to provide a means for the process designer to evaluate any preform shape and die geometry at any forming stage of the forging process. The SFEM method corresponds to the one-step UB analytical methods in terms of simulation time. Figure 7 shows that the CPU time for the SFEM method is about 70 times less than the time needed for the full FEM simulation.

The SFEM simulations and CPU times shown in Fig. 3 to 6 are for a full elastoviscoplastic material with thermal coupling and unloading, which is required for analyzing residual stress patterns in the forged part. The one-step SFEM allows the forging designer to compute in a single step the forming loads, material flow, and die stresses, and, in general, the material flow simulations are comparable to the results of the full simulation. The SFEM has the capability for rapidly analyzing any type of 2D and 3D forming process while having many advantages not offered by other simplified analysis tools summarized in Table 1. One important advantage of this method is that the process designer does not have to be an expert in the FEM to create the simulation model, as the SFEM model is created automatically by the analysis system. The forging designer does not have to be concerned about assuming an admissible velocity field or deriving any analytical equations for any particular forming process. This is not offered by the other simplified analysis tools summarized in Table 1.

Table 1 thus summarizes some of the important analytical and mesh-based analysis tools that are available for industrial process modeling. The advantages of the FEM are mainly in its

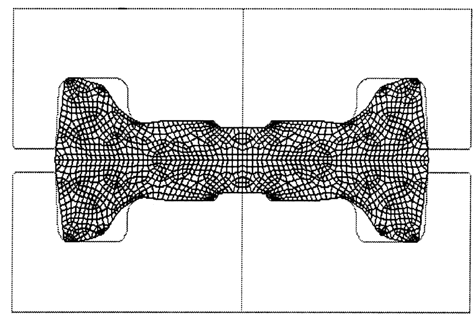


Fig. 5 Partial die fill. One-step Analysis at 1.0 s. CPU time 5.53 s

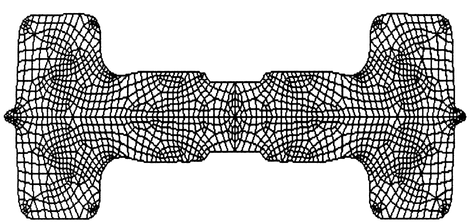


Fig. 1 Engine disk. Central processing unit (CPU) time less than five s using a Pentium personal computer. Model was an elasto-plastic material with thermal analysis.

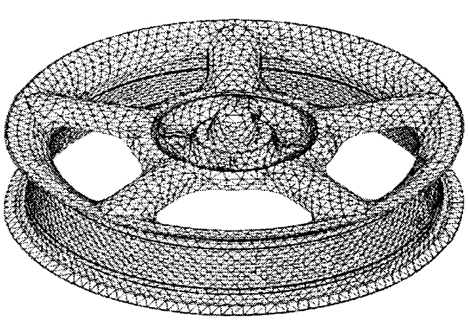


Fig. 2 Three-dimensional (3D) wheel. CPU time was less than five minutes for a one-step simulation of an automobile tire rim.

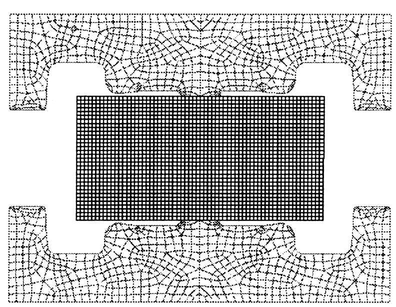


Fig. 3 Initial billet geometry

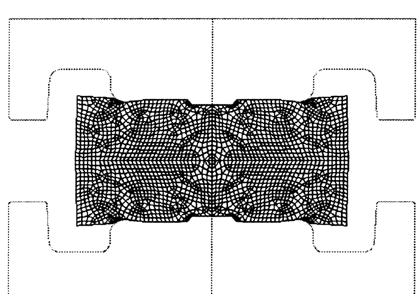


Fig. 4 Partial die fill. One-step FEM analysis at 0.25 S. CPU time 7.21 s

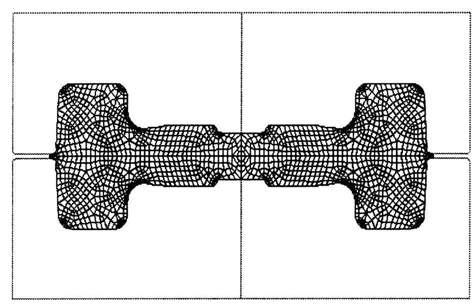


Fig. 6 Die fill with flash. One-Step Analysis at 1.2 s. CPU time 7.20 s

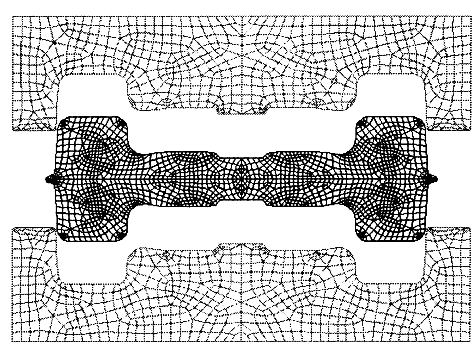


Fig. 7 Die fill with flash. Simulation time 2 S. CPU time 488 s

flexibility for the treatment of complex, 3D shapes and complex constitutive laws. The older analytical methods still have value because they provide a basis for the comparison of specific cases for which they give reliable results. The barriers to a widespread use of meshed models no longer exist as commercial software for automatically generating an unstructured FEM mesh is available. The FEM models with parametric geometry attributes make it possible for the user to modify the previous design easily and quickly without having to set up the FEM model again, as the FEM model associated with an attribute modification will change accordingly without user intervention. Complex shapes are nearly always found in secondary processing such as forging, whereas primary processing, such as in billet conditioning, nearly always produces simple shapes. Thus, mesh-based methods are the dominant methods for modeling secondary processes and nonmesh, lumped parameters or a state variable model are dominant methods in the primary processes.

Computer Aided Optimization

The competitive metalworking companies of today are striving to meet the objectives of lean manufacturing and value-stream analysis to give customers exactly what they want on time. Conventional design techniques rely heavily on handbooks and the experience of the design engineer. These design techniques provide adequate designs, which are validated through extensive physical modeling, a process that entails considerable expense and long lead times. The conventional design techniques may not satisfy the speed requirements of a lean manufacturing business strategy. Human involvement is great with the conventional design techniques. Too much time is spent trying to squeeze more out of local performance, leaving insufficient time for engineers to search for the correct combinations of downstream process parameters needed for arriving at a global-optimum solution.

New approaches to process design are now being started in some supplier companies to improve the efficiency of upstream process design by carrying out computer-aided optimization (CAO) to achieve improvement and ever-higher standards. Today, industrial process design teams have a choice of commercially available process modeling tools for simulating metalworking processes. Cost-effective design optimization has been demonstrated using generic CAO software, which can run multivendor simulation applications, acting as a "software robot" to drive these programs without human intervention to an optimal solution (Ref 18). Activity-based models (Ref 13) of a modern forging operation can also be connected to design activities (Fig. 8) by using computer-aided optimization (CAO) techniques. Figure 8 is an activity-based model that was one of the first attempts to modernize the forging design process as part of an air force (AF) Manufacturing Science project in the 1980s. Followed by an air force Man-Tech program to modernize the forging industry and to learn how to design an integral blade and rotor (IBR) forging. In the 1990s, DARPA funded a similar program with the U.S. engine industry to reduce the cost of manufacturing precision cast turbine engine components. The activity-based model in Fig. 8 is still appropriate today, although many more new software products are commercially available compared to the analysis tools (like ALPID, NIKE and AFD) available when Fig. 8 was developed.

The global approach to automation is straightforward. Process design automation tools are needed especially when the design team is challenged by customer-imposed quality standards, delivery time, and manufacturing cost and the team encounters a general problem of a new material and new product shape, which they have never previously manufactured. Because the potentially viable processing routes are numerous, many process simulations are necessary to identify optimal methods, unless some means of restricting the solution space are carried out. The problem of large design solution spaces is not restricted to the design of the deformation process. It also applies to the selection of correct combinations of processing temperatures and strain-rates and to the identification of optimal heat treatment techniques, material removal techniques, customized nondestructive evaluation (NDE) and product testing.

Multidisciplinary process design and optimization (MPDO) provides the framework to conform metalworking processes to changing customer and market requirements and to do it fast. In this framework, a system is viewed as an interconnection of various interacting components, each being described by a set of relations describing the time evolution of internal variables and input-output relationships (Ref 19). The MPDO process consists of formulating appropriate material, process, and equipment models together with carefully chosen design objectives and constraints. The selection of models, objectives, and constraints must be considered concurrently to be consistent.

The MPDO process offers a means for automating and satisfying the process design activities. It allows a process design team to combine analytic, numeric, evolutionary computing, and knowledge-directed heuristic search techniques to narrow the design space across engineering disciplines. From a lean manufacturing perspective, using MPDO software to do iterative, costly design tasks address the following basic issues:

- Integrating uniformly diverse engineering and business disciplines
- Using engineering tools correctly
- Freeing the engineer from doing costly design analysis iterations
- Freeing the engineer to do other cross-functional manufacturing or business work

An MPDO program automates existing design environments and allows engineers to analyze a complex problem by predicting the results of a particular decision. The MPDO process integrates and couples process design simulation codes and runs them automatically to produce an optimal process design. An MPDO environment enables process designers to explore many more alternatives than an engineer can study manually and reach optimal designs faster, resulting in reduced product cycle time, reduced scrap and rework, lower manufacturing cost, and increased process productivity.

The MPDO process permits engineers to focus on parts-manufacturing decision making, and it eliminates an iterative, costly trial-and-error manipulation of simulation codes by the engineer to obtain an optimal process design. Figure 9 shows different types of commercially available CAO software that can be packaged with appropriate process modeling software, business models, and standards to form an MPDO infrastructure.

Manual Design Process

The manual design process is costly and complex. The complexity of the forging process is illustrated in Fig. 10 by the physics of a nonsteady-state forging process. The complex physics can be broken down into seven design tasks (Ref 13). These tasks must be performed in proper sequence, and the information generated in each task must be correctly transferred between tasks, making certain that any redundancy in work is eliminated. The manual process can be prone to error, if a uniform user interface is not maintained. An overview of the tasks that must be performed in developing a forging manufacturing plan is presented in Fig. 11. This representation shows inputs, redesign when analysis invalidates a design, and database information used by a process designer or planner. The tasks that lead to the creation of a manufacturing plan are:

- *Task 1*: Integration of geometry representation
- *Task 2*: Initial-guess finished forging
- *Task 3*: Forging sequence design
- *Task 4*: Processing parameter selection
- *Task 5*: Forging sequence analysis
- *Task 6*: Die design and stress analysis
- *Task 7*: Generate key results

This section provides a conceptual description of these tasks.

Design Task 1: Integration of Geometry Representation. Forging process design is conducted backwards. It starts with the finished-part drawing and ends with a forging design of the part, including the preform shape, flash, and die design for the finisher die. The finished part representation is usually in the form of a geometry database plus the structure and property specifications. The latter is typically not archived in the database at this point in the design process.

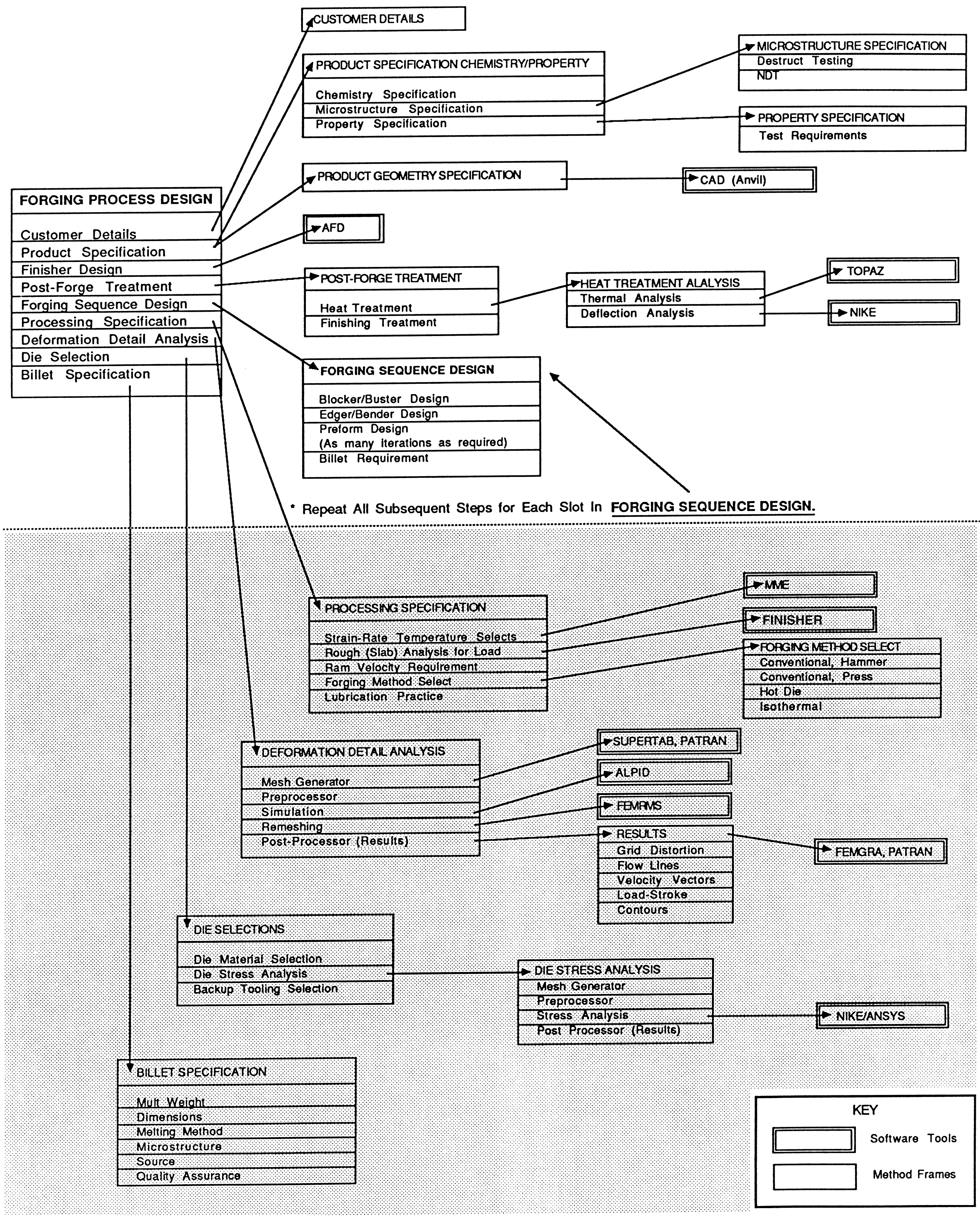


Fig. 8 Activity-based model for planning and specification of forging process. Analysis tools (like ALPID, NIKE and AFD) in this diagram were those available through government sources in the 1980s. ANVIL was a commercial CAD/CAM graphics package commonly used by forging designers. Today, many more and better software products are commercially available for use.

This phase includes the collection of information about the customer's specifications. With a lean manufacturing strategy, close contact with the client is sought to discover if any client expectations for the product exist that are not deducible from the purchase order description of the deliverable product, including the required delivery time. The effort is to find out all of the necessary information needed to satisfy client requirements and to reach an optimal solution that embodies the necessary constraints and requirements. Fixed solution ideas often have an adverse effect on the outcome, so at this stage of the design process the following information must be obtained from the client:

- The objectives the client expects to have satisfied
- The properties the finished forging must have
- Any general requirements not completely specified by the customer

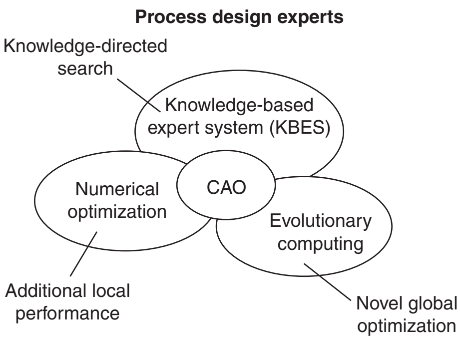


Fig. 9 Multidiscipline process design and optimization (MPDO)

Design Task 2: Initial-Guess Finished Forging. The initial-guess finished forging geometry is generally experientially based. A design expert will take into account material and equipment variables such as the workpiece material, recommended process, and available equipment characteristics. The interdependence of forging process parameters is shown in Fig. 12. A design approach that fully utilizes expert knowledge and supplements it with numerical optimization is warranted. Concept variants must now be evaluated by analysis to eliminate any variants that do not satisfy the demands of the customer's specifications. During this task, the chief criteria are technical in nature, and economic criteria are beginning to play a part. It may be that several preform variants look equally promising, but the final decision must be based on a more concrete level of knowledge. An overall objective is to find a preform shape that will fill the die cavity with the least amount of energy and pressure. Microstructure-property relationships, equipment characteristics, and any downstream requirements for heat treatment and material removal must be factored into the design. Using MPDO software as a software robot to run process simulation analyses can eliminate most of the engineering oversight needed to create an optimized solution by designing the preform shape and die geometry concurrently.

One way to softly automate this task is by using artificial intelligence techniques such as genetic algorithms and neural network analysis that automatically provide a list of possible process design alternatives and are dependent on the functional requirements (FRs) defined during design clarification. This approach allows the design team to deal with real-world problems, where the best is only a theoretical ideal that is often unattainable or not cost effective. Through the use of soft optimization techniques, a modest goal of being just good enough can be

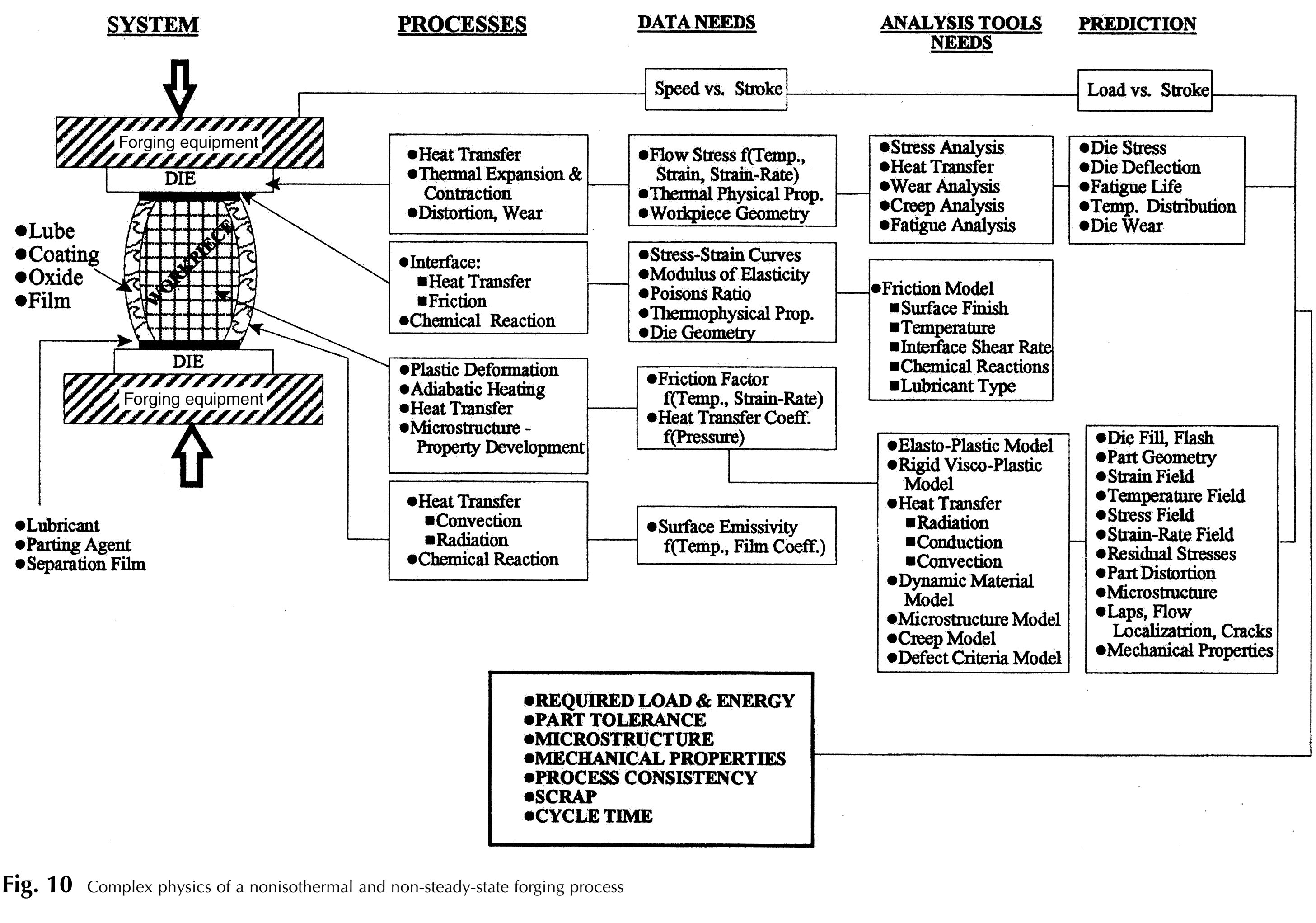


Fig. 10 Complex physics of a nonisothermal and non-steady-state forging process

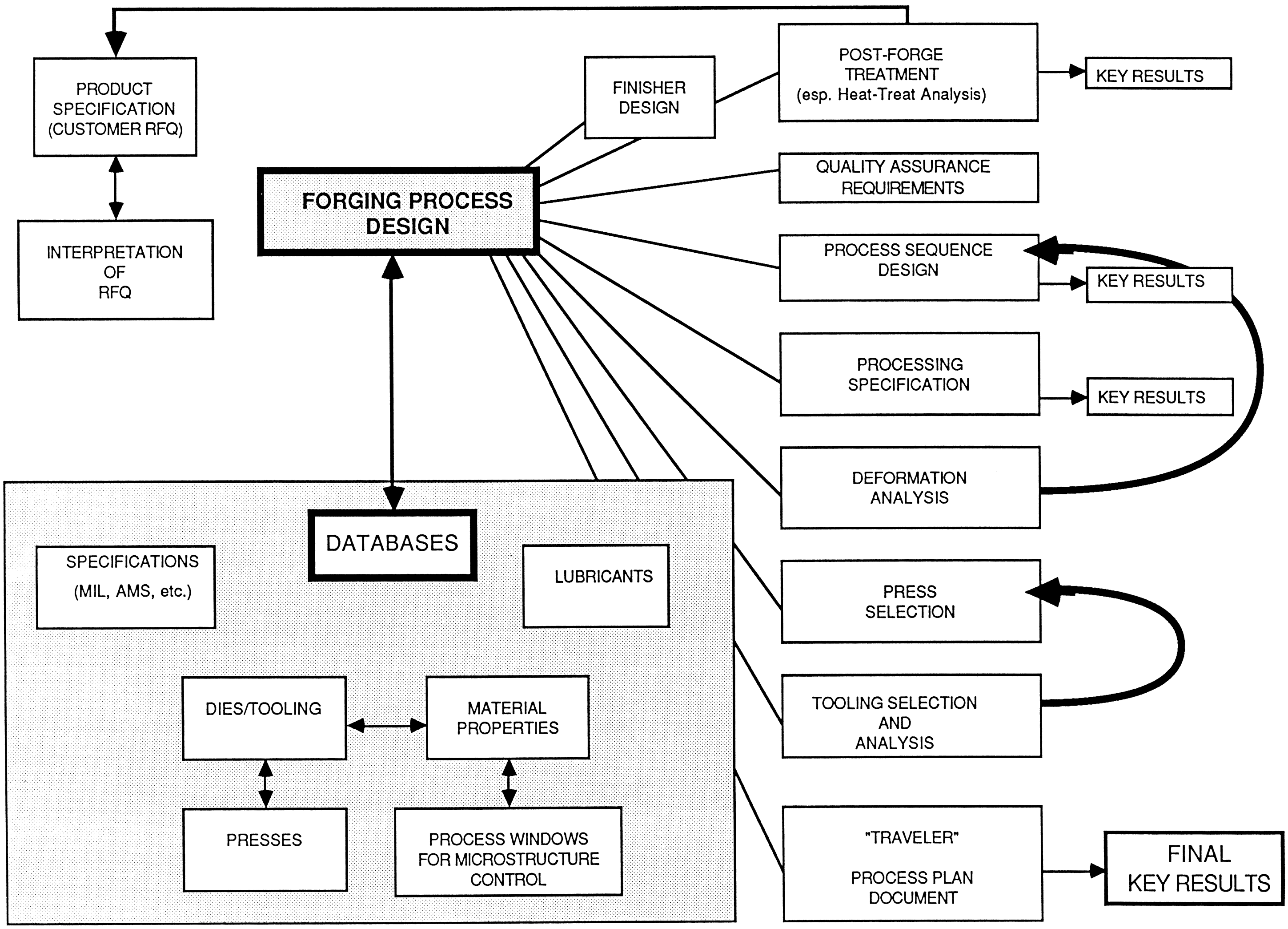


Fig. 11 Precision forging design process. RFQ, request for quote

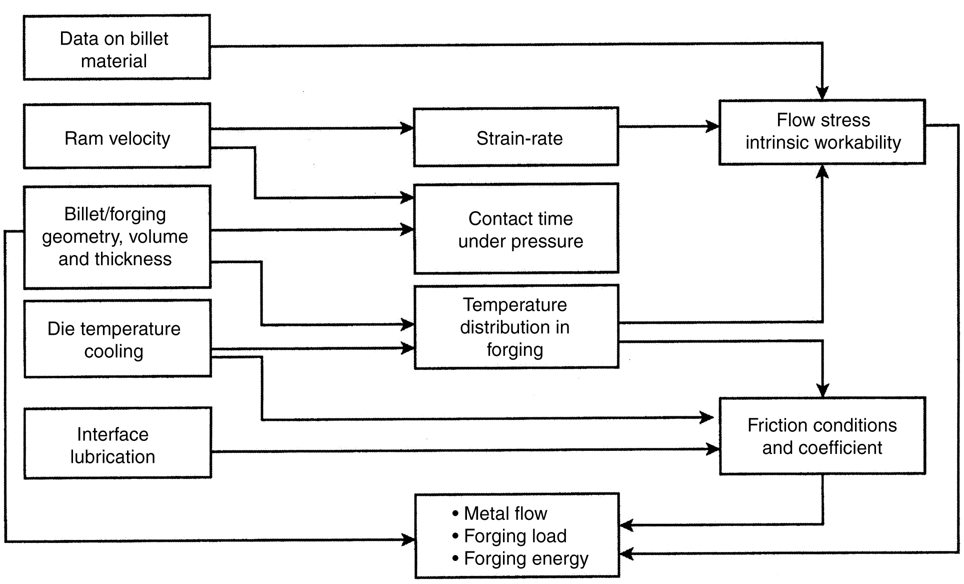


Fig. 12 Interdependence of forging process parameters

achieved even for problems in manufacturing currently considered to be beyond reach by calculus-based methods. This approach to conceptual design is expected to reduce the time for arriving at a set of design parameters (DPs). It is highly unlikely that efficient algorithms for the solution of these problems of arbitrary size will be found. Figure 8 shows an integrated computer-aided value-flow diagram. The overall objective is to find a solution for the finished forging that will have an acceptable level of residual stress and a uniform distribution of the residual stress after quenching hot from the finisher die. The objectives included getting the correct service properties while minimizing any distortion during subsequent machining (Fig. 13).

Design Task 3: Forging Sequence Design. Forging sequence design requires tracing backwards to acceptable starting billet geometry. This design task is required when several intermediate shapes may be required to reach the desired finished shape. An important aspect of sequence design is distributing uniformly the energy dissipated between the different sequence stages. Normally, the process designer knows the principal design parameters that affect the design objective or violated constraint but does not know the direction or amount to change these parameters. Frequently, evaluation of individual variations may lead to the selection of one that is particularly promising but may be further improved by considering other appropriate combinations of process parameters and the elimination of weak links. A knowledge-based expert system can support this scenario by incorporating other appropriate ideas and solutions, allowing the designer to specify the parameters to change and then call a numerical optimization with just these design parameters to determine their best set of values. If no design knowledge is available, a hybrid optimization approach may be combined to get a better global optimum.

Design Task 4: Processing Parameter Selection. A strong relationship exists between the process variables and the machine variables, and this design task is concerned with establishing optimum process and machine variables. Process parameters include the workpiece deformation rate, which is expressed as the slide velocity and the effective strain rate, and the temperatures of the workpiece before forging and the temperature during forging. These process variables are evaluated during this stage of design, with the objective of determining optimum values in terms of overall workability. This task can encompass the meaning of workability in its widest sense; that is, workability criteria defined by defect prevention, effective equipment utilization, and proper development of microstructure.

Evolution of microstructure during hot working is a major factor in process design, and many methods have been investigated and developed (e.g., see Chapter 1, "Workability and Process Design—An Introduction," for a general introduction). One method is the so-called dynamic material model (DMM) (Ref 20), which is a fundamental methodology for selecting temperature and strain rate for stable deformation and microstructure evolution. The DMM is a criterion for mechanical and microstructure stability, which is based on continuous field variables that have an immediate and direct influence on the process. Microstructure-based models, in contrast, may be more difficult to use, as they are generally based on the secondary effect of the field variables on some microstructural condition such as grain growth above a certain size. This can require iterative coupling with the rest of the process model. Microstructural models also may be tough to use, because of the complex physical metallurgy of hot working processes.

The designer is usually concerned with defect avoidance, for example, preform buckling; underfill; flow related defects such as shear bands, lap or fold, suck-in, shearing, or flow through defects; free surface cracking; grain boundary cavities; and triple junction cracks. Typically, temperature and strain-rate ranges are entered as *nonholonomic constraints* to avoid overconstraining the system. (A nonholonomic constraint is an inequality that avoids overconstraining a set of equations, making it possible to get a solution for the unknown parameters. This approach is used as a means for incorporating material behavior information into the FEM model to define the stability ranges of temperature, strain rate, and strain for the workpiece material that is akin to the concept of material workability. It is a consistent approach for applying process optimization methods to achieve a stable process, preform and die shapes, and an optimal set of forging microstructures and properties.)

Besides stable material flow, the database information correlating processing-microstructure relationships can be used to identify the process window within which both stable material flow and appropriate microstructure evolution is achieved. The optimal combination of temperature and strain-rate values will produce the least variation in service properties such as grain size, tensile strength, fracture toughness, and fatigue strength (Ref 19, 21, 22). When the workpiece material is processed within the range of temperature and strain rate given by the material stability analysis, the process is not sensitive to perturbations of either process parameter, and the material stability is path independent. The DMM analysis also helps the design team to define process variables such as die temperature, the workpiece temperature before entering the die, and the workpiece temperature range during deformation.

A strong relationship exists between the process variables and the machine variables. Figure 12 shows the interdependence of process variables. Forging companies have access to a variety of types of forging equipment (Ref 23). However; the choice of equipment to produce a particular part is often limited to the equipment that is available in the forge shop. Accordingly, it is not always possible to select the optimum type of equipment for a particular job. It is for this reason that it is important to understand the principal differences between the various classes of equipment and how these differences affect the metal forming process and die design decisions. This particular knowledge is important for the manufacturing of high quality precision forgings. The important characteristics that must be understood when selecting the forging equipment are listed as follows:

- *Workpiece deformation rate:* This parameter is expressed as the slide velocity equipment parameter and the effective strain-rate of the workpiece material, a process variable. The workpiece deformation rate should be strictly referred to as the relative velocity between the dies when referring to the equipment. The choice of relative die velocity is determined by the intrinsic workability of the workpiece material, which is determined by the prior thermomechanical history of the forging billet.
- *Dimensional control:* This process variable reflects the capability of the equipment to consistently achieve part tolerance. Dimensional control must be achieved along with achieving a controlled set of microstructures and properties in the finished product on a repeatable basis. Consistency is achieved only when all components in the forging cell are individually in control.
- *Production rate:* Production rate equates to the number of parts per unit time. The production rate is dependent on the type of press and the ability of the workpiece mate-

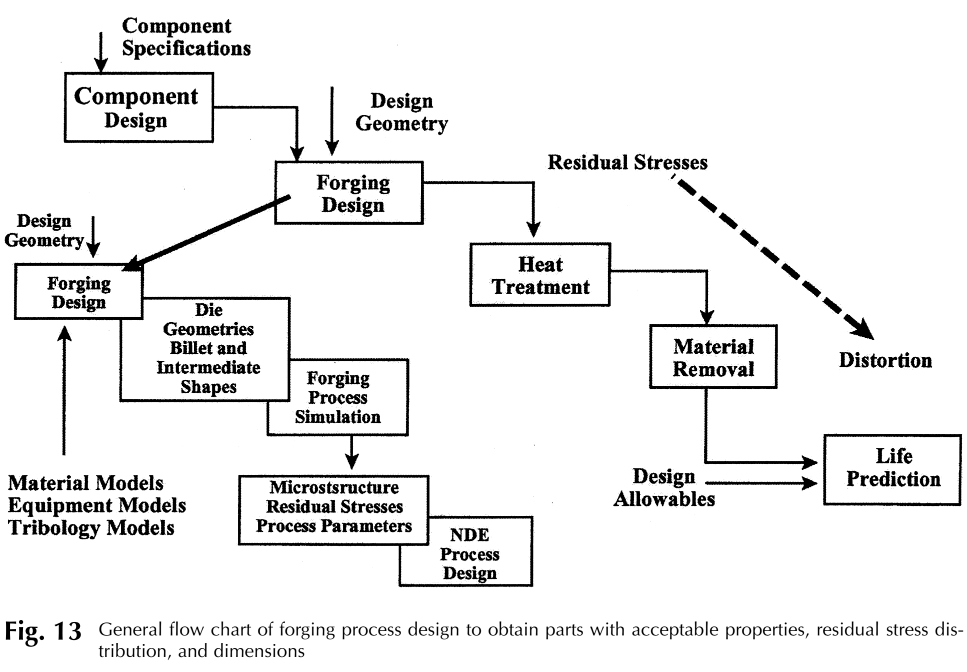


Fig. 13 General flow chart of forging process design to obtain parts with acceptable properties, residual stress distribution, and dimensions

rial to respond to the demands of the process. The forging press and the workpiece material must be properly matched to achieve an optimal production rate. Materials such as titanium and nickel-base alloys that have a high strain-rate sensitivity must be forged at lower die velocities as in hydraulic presses to avoid a rapid rise in flow stress and die stress values and to avoid the formation of defects and plastic instabilities. The equipment parameters that control the production rate are the stroking rates, ram velocity, machine energy, and the degree of automation and computer control.

Design Task 5: Forging Sequence Analysis. The detail forging sequence design is the phase of the design process in which the arrangement, form, dimensions, surface properties of all the individual components, for example, the equipment, workpiece material, lubricants, separating agents, dies and tooling, are finally laid down. All of the materials used in the process are specified, and the technical and economic feasibility are rechecked. All the drawings and other production documents are produced. It completes the substantiation of the forging process with final instructions about the layout, shapes, dimensions, microstructures and properties, and a final scrutiny of the manufacturing methods and costs.

Unlike the finisher die and workpiece preform design that goes on in the backward direction from the final requirements, the forging sequence analysis task progresses in the forward direction. This reversal of methodology is done because the initial conditions are not generally known for intermediate steps in a nonisothermal forging sequence without reheats. This design stage is quite crucial, as very often corrections must be made during this work and preceding steps repeated. The crucial activities are optimization of the principle and optimization of the layout and intermediate workpiece shapes. These activities influence each other and overlap considerably. In general, optimization of the intermediate workpiece shapes, geometries of the dies, and, hence, optimizations of the manufacturing process begin to assume growing importance as substantiation proceeds.

Sequence design involves studying a combination of circumstances that, when considered together, will result in lowering the overall cost of the forging and satisfying customer specifications. Attention must be given to optimize such factors as throughput, material and energy consumption, and tool and die life extension, the number of tool changes, and tooling cost. Many intermediate results are obtained for the forging sequence and process parameter designs. These results, in the manual approach, often require an experienced engineer to do trial-and-error simulations over several weeks, which involves costly, iterative manipulation of simulation program parameters. The tools of MPDO potentially can help the engineer reduce the design space, and a combination of optimization techniques may be required to provide the best solution. Optimization studies show that no single technique works for all design problems. Because an experienced designer uses experiential rules to decide sequencing, it is likely that a knowledge-directed approach supplemented by another optimization method, for example, numerical optimization, may be needed to get the best solution.

Design task 6: die design and stress analysis provides the detailed loading distribution on the die surface. Equilibrium elasticity analysis of the dies is done to make certain enough die material exists to distribute the load satisfactorily and to determine elastic die deflections. An experienced designer does die block sizing generally, making it feasible to fully utilize expert knowledge and to supplement it with other optimization techniques. At this stage of process design, the designer typically compares the elastic deflections of alternative die materials. A complete design interaction approach, where there are varying sizes, strain levels, and forces to specify the die diameter and thickness compatible with the die support system, is generally done.

The designer does a separate analysis of the strain concentration factors. This analysis is done within the die impression and on the die outer shapes such as ejector holes, and threaded surfaces. The local shape change produces effects that are confined to the specified zone, and

1. Billet material specification

Alloy:
Specifications:
Melting specs:
Material source:
Heat no.
Conversion:
Surface condition:

2. Billet material quality assurance

(Mill data or acceptance tests)

Composition:
Ultrasonic:
Property data:
Macro etch:
Microstructure:

3. Billet preparation

Conversion:
Billet shape:
Cut length:
Conditioning:
Identification:
Coating:

4. Forging practice

(This will be repeated for each step in forging sequence)

Preform conditioning:
Lubrication practice:
Fritting:
Furnace:
Heating practice:
Temperature:
Maximum time:
Atmosphere control:
Press:
Tooling assembly ID:
Press controls:
Velocity:
Stroke:
Load:
Identification (serial no.):
Trimming:
Hot inspection:

5. Post forge inspection

Dimensional check:
Defects:

6. Heat treatment

Furnace preheat:
Furnace setpoint:
Time at temperature:
Cooling from furnace:

Repeat these blocks as appropriate

7. Finishing operations

Machine to sonic shape:
Cut test samples:

8. Final quality assurance

Test type:
Test locations:
Identification:
Ultrasonic:
Mag particles:
Die penetrant:
Microstructure:
Macro slice:
Fracture toughness:
Tensile:
Fatigue:
Creep rupture:

Fig. 14 A completed schematic process and inspection traveler these local zones are the ones where failure is likely to occur. If the die loading during the forging process is cyclical, the designer should do a stress factor calculation to include the effect of fatigue on die life. A wear analysis can be done to show areas that will be prone to wear by abrasion, and candidate surface regions can be identified for surface hardening by ion implantation or chemical vapor deposition. Die drawings are prepared as required based on the detailed results of simulation-based design. This activity includes any open draw dies, preform dies and finish dies or inserts, including impressions, bearing and cleared surfaces, locking and attaching methods, ejection designs, and thermocouple holes.

Design Task 7: Generate Key Results. The end product of process design and analysis is a manufacturing plan. A complete process specification called a process and inspection traveler is created. The traveler generally accompanies the forging for critical parts, which, if failure occurred in service, would threaten the lives of human beings. A gas turbine aircraft engine disk is an example of a forged critical component that requires close process and inspection controls. A schematic of a process and inspection traveler is shown in Fig. 14.

A part traveler should be automatically generated by an integrated software system that is capable of fusing all of the important process and material information into a unified document. All key results should be made available, including key intermediate results from process sequence design and process parameter design. Some key design results for these two critical design tasks are shown in Fig. 15.

Traveler documents are an authenticated, extended, and complex substantiation of the total forging process, which is made up of complicated and related parts. The previous discussion of the interconnected design tasks provides an integrated design approach that is based on the use of commercially available computer-aided optimization (CAO) tools for integrating, automating, optimizing, and accelerating the entire design process and using a knowledge-based system for process design. The implementation of such a system requires a detailed specification of the methodology that is involved. A representation of an activity-based method description was shown in Fig. 8.

PROCESS SEQUENCE DESIGN RESULTS	PROCESS PARAMETERS DESIGN RESULTS
FINISH FORGING: Volume:_______ Weight:_______ Drawing No./CAD File:_______	FINISH FORGING: Strain-Rate:______ Temperature:______ Lubrication:______ Load Estimate:______
BLOCKER: Volume:_______ Weight:_______ Drawing No./CAD File_______	BLOCKER: Strain-Rate:_______ Temperature:_______ Lubrication:_______
BENDER: Volume:_______ Weight:_______ Drawing No./CAD File:_______ (Repeat for as many preforms as required.)	BENDER: Strain-Rate:_______ Temperature:_______ Lubrication:_______ Load Estimate:_______
FINAL PREFORM: Number of PREFORMING Operations:_______ Volume:_______ Weight:_______ Drawing No./CAD File:_______ (Repeat for as many preforms as required.)	FINAL PREFORM: Number of PREFORMING Operations:_______ Strain-Rate:_______ Temperature:_______ Lubrication:_______ Load Estimate:_______ (Repeat for as many preforms as required.)
BILLET: Volume:_______ Weight:_______ Drawing No./CAD File:_______	

Fig. 15 Examples of Key Intermediate Results from the critical steps of process-sequence design and process-parameter design

REFERENCES

1. P. Sargent and H. Shercliff, *Modelling Materials Processing* (Version 2). [CD-ROM]. CUED/C-MATS/TR.206, Cambridge, England: Department of Engineering, Cambridge University, Cambridge, CB2 1PZ England, November 1993
2. J.R. Wright, "Forging a Culture of Competitiveness," *Forg.*, Fall 1995, p 23
3. P.F. Drucker, The New Productivity Challenge, *Harvard Bus. Rev.*, Vol 69 (No. 6), 1991, p 69–79
4. J.S. Brown, Research that Reinvents the Organization, *Harvard Bus. Rev.*, Vol 69 (No. 1), 1991, p 102–111
5. M. Hammer, Reengineering Work: Don't Automate, Obliterate, *Harvard Bus. Rev.*, Vol 68 (No. 4), 1990, p 104–112
6. R.H. Hayes and K.B. Clark, Why Some Factories Are More Productive Than Others, *Harvard Bus. Rev.*, Vol 69 (No. 1), 1991, p 66–73
7. M. Imai, *Kaizen*, McGraw-Hill Publishing Company, 1986
8. R. Charan, How Networks Reshape Organizations, *Harvard Bus. Rev.*, Vol 69 (No. 5), 1991, p 104–115
9. E.A. Haas, Breakthrough Manufacturing, *Harvard Bus. Rev.*, Vol 65 (No. 2), 1987, p 75–81
10. "Knowledge-Based Integrated Design System," *Computer-Aided Materials Selection During Structural Design*, Committee on Application of Expert Systems to Materials Selection during Structural Design, NMAB-467, National Academy Press, Washington, D.C., 1995, p 57–62
11. J.P. Womack and D.T. Jones, *Lean Thinking: Banish Waste and Create Wealth In Your Corporation*, Simon and Schuster, 1995, p 9–90
12. D.K. Sobek, II, J.K. Liker, and A.C. Ward, Another Look at How Toyota Integrates Product Development, *Harvard Bus. Rev.*, July–August 1998, p 36
13. H.L. Gegel, J.C. Malas, S.M. Doraivelu, and V.A. Shende, Forging Process Design, *Forming and Forging*, Vol 14, *ASM Handbook* (formerly *Metals Handbook*, 9th ed.) ASM International, 1988, p 409
14. T. Altan, S. Oh, and H. Gegel, *Metal Forming—Fundamentals and Applications, ASM Series in Metal Processing*, H.L.

Gegel, Ed., American Society for Metals, 1983, p 131–142
15. R.H. Wagoner and J.-L. Chenot, *Fundamentals of Metal Forming*, Wiley, 1997
16. R.H. Wagoner and J.-L. Chenot, *Metal Forming Analysis*, Cambridge University Press, 2001
17. T. Altan, S. Oh, and H. Gegel, *Metal Forming—Fundamentals and Applications, ASM Series in Metal Processing*, H.L. Gegel, Ed., American Society for Metals, 1983, p 329–339
18. "iSight Technical Overview," Engineous Software, Inc., Morrisville, NC 27560, 1997
19. J.C. Malas and W.G. Frazier, Optimal Design of Thermomechanical Processes Using Ideal Forming Concepts, *The Integration of Material, Process and Product Design,* Zabaras et al., Ed., A.A. Balkema, Rotterdam, 1999, p 229–236
20. H.L. Gegel, J.C. Malas, S.M. Doraivelu, J.M. Alexander, and J.S. Gunasekera, Material Modeling and Intrinsic Workability for Simulation of Bulk Deformation, *Adv. Technol. Plast.*, Vol 1, Springer-Verlag, 1987, p 1243
21. J.C. Malas III, "Methodology for Design and Control of Thermomechanical Processes," Ph.D. dissertation, College of Engineering and Technology, Ohio University, November 1991
22. V. Srinivasan, E.A. Medina, J.C. Malas, III, S. Medeiros. W.G. Frazier, W.M. Mullins, and R. Srinivasan, Optimization of Microstructure During Deformation Processing Using Control Theory Principles, *Scr. Mater.*, Vol 36 (No. 3), 1997, p 347–353
23. T. Altan, S. Oh, and H. Gegel, *Metal Forming—Fundamentals and Applications, ASM Series in Metal Processing*, H.L. Gegel, Ed., American Society for Metals, 1983, p 5–36

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 337-345
DOI:10.1361/hwpd2003p337

Chapter 23

Optimal Design of Thermomechanical Processes

J.C. Malas, W.G. Frazier and H.L. Gegel (Retired), Air Force Research Laboratory, Wright-Patterson Air Force Base
V. Srinivasan, Indira Gandhi Centre, Kalpakam, India

OPTIMAL DESIGN of a multiple-step processing system consists of formulating appropriate models of materials, process, and equipment behavior together with carefully chosen design objectives and constraints. In order to be consistent, the selection of materials and process models, objectives, and constraints must be considered as an integrated whole, and not independently. The goal is to bring all of the process elements together to address the commercial need to reduce cost and to save time by reducing iteration of trials and improving quality by reducing product and process variability, while avoiding over-optimization of any particular element of the system. Parts suppliers are continually pushed to reduce prices, perhaps at the expense of quality or other considerations. To deal with quality issues, companies need to become more effective in using their existing design software and adopt upstream design processes that enable them to deliver customized services and products at relatively low cost.

As noted in the preceding chapter, "Multidisciplinary Process Design and Optimization: An Overview," multidisciplinary process design and optimization (MPDO) is a software infrastructure that links tools from different environments, while providing a software framework for automating, optimizing, and integrating the engineering design processes (Ref 1–5). In this framework, a connection exists between various interacting components of the metalworking process, which includes models for the workpiece, lubrication, tooling, metalworking equipment, inspection tools, analysis, and so forth. The benefits of MPDO are measured in terms of the ways it helps process design teams do their jobs. Multidisciplinary process design and optimization helps the design team:

- Create innovative design alternatives
- Manage the dynamical and material stability effects in the design of thermomechanical materials processes
- Handle explicitly design objectives and constraints on workpiece materials, processing equipment, and tooling
- Improve product design
- Reduce cost and time-to-market
- Use simplified models and concepts cost effectively to identify optimal workpiece material trajectories

Substantial improvements in effectiveness and efficiency can be realized through integrated approaches that optimize the whole system performance and not just individual subsystems such as workpiece materials behavior, material flow in dies, and equipment responses. Vendors who find creative ways to reach out to their customers to use products and services more effectively become more indispensable to their customers.

Design engineers work under considerable time pressure, and they need to have easy-to-use tools to speed the design of manufacturing processes, such as hot metal forging. Metalworking industries seeking to be responsive to the competitive needs of their customers must become more effective in using their existing design software and human resources. It is increasingly important to adopt upstream technical and business processes that enable them to deliver customized services and products at relatively low costs and deliver them with the expected quality when the customer expects the parts to be delivered. Basic design tasks, such as selecting the number of forming steps and specifying the processing conditions for thermomechanical operations, may also result in feasible solutions that are not necessarily optimal. Therefore, new process design and control methods are still sought for optimization of cost, quality, and productivity of thermomechanical processes.

This chapter describes a strategy for systematically calculating near optimal control parameters for hot deformation processes. The focus is on a design concept of "ideal forming" that is based on state-space control-theory concepts, materials stability analysis, and geometry mapping between the initial and final states to predict important process parameters. The application of dynamical modeling and process design techniques using ideal forming concepts and trajectory optimization are discussed with examples to illustrate potential benefits. The objective is to show a system design approach that holds promise for substantially reducing manufacturing cost, risk, and delivery time to a customer, while resulting in an improved product that satisfies customer requirements and expectations.

Critical issues such as efficient material flow and the sensitivity of the variation of finished parts properties to small changes in key process parameters are addressed using lumped-parameter models of microstructure evolution and optimization techniques for controlling microstructure during thermomechanical processing. The use of the state-space (lumped parameter) method provides considerable insight into the controllability of metalworking processes. It is a constrained optimization method, where the overall design is specified as constraints and objectives to accommodate multiple physical and economical requirements. This approach identifies optimal workpiece material trajectories that are generally needed for achieving customer product specifications.

Optimization is just one method for performing a process design. Whether the resulting design is the best design depends entirely on the criteria specified by the product designer. As illustrated in the section "Material Trajectory Optimization" in this chapter, optimization techniques require the specification of two types of criteria: objectives (wants) and constraints (needs). To achieve the desired goals, the designer must specify all relevant criteria and carefully determine whether the criteria are either objectives or constraints. For example, it may be desirable to minimize the production costs (objective) while maintaining specified product quality standards (constraint). On the other hand, the opposite scenario may be desired, i.e., maximize the product quality (objective) while not exceeding a specified cost (constraint). Effective optimization strategies consider the entire processing design problem, not just some parts, thereby avoiding over-optimization of parts of the process at the expense of the whole manufacturing enterprise.

Concepts of Dynamic Modeling in Optimal Design

A metalworking manufacturing enterprise for the purpose of this discussion consists of a sequence of thermomechanical processing steps to produce a component. Each unit process, such as forging, heat treatment, material removal, and nondestructive evaluation (NDE), can be decomposed into subsubsystems, i.e., the workpiece material, tooling, equipment, control system, and so forth, which can be decomposed further. A typical block diagram for a hot forging equipment system is shown in Fig. 1. Each component of the manufacturing system is represented by a set of relations, known as *state equations*, that describe the time evolution of internal variables and input-output relationships. When these state equations and input-output relationships are viewed as a whole, they provide a complete system description (Ref 1).

The systems approach to process design, when linked to computer-aided optimization models for design and analysis, allows simplified forming models to be used. Manufacturing processes can be mathematically modeled as nonlinear dynamical systems using a state-variable formulation, i.e., a system of coupled, first-order nonlinear differential equations. Symbolically, this is often written in the form:

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}) \qquad \text{(Eq 1)}$$

where $\mathbf{x}$ and $\mathbf{u}$ are vectors of the system's state and control variables, respectively. The function f defines the relationship among the current state of the system, the current control variables, and the rate of change of the state. The important point concerning dynamical models is that they are valid for a broad range of control signals, unlike algebraic models, which are only valid for a particular class of process controls, such as constant temperature and strain rate. The use of dynamical models provides greater predictive capability over algebraic models and provides more degrees of freedom in the time domain for optimal process design.

A material trajectory is defined here as the time evolution of materials state with respect to strain, strain rate, temperature, and temperature rate. The development of materials processing models invariably involves quantitative descriptions of the underlying process physics. Materials process models always involve descriptions in terms of sets of partial differential equations that are dependent on time, spatial dimensions, field variables, and internal states. Because several different processing histories (trajectories) can lead to nearly the same end result, it is important to be able to find the best trajectory to realize the desired objectives. Consequently, models and design techniques for controlling microstructure during thermomechanical processing have been developed to address critical issues such as stability, transient and steady state response, and robustness of material trajectories.

If the required shape does not involve complex geometries (as in primary processing operations), constraints are simplified, and the process model can be simplified to solve a lumped-parameter type problem. Lumped-parameter problems can be represented by a set of ordinary differential equations, and they can be solved using simpler means than the complex apparatus of numerical analysis that is required for solving partial differential equations. Whether the physical shape is sufficiently complex to justify developing meshed models is a difficult decision to make. One does not usually know a priori if some geometrical feature is significant. The idea of shape determining a materials process model type may not be a physical shape, but it could be the shape of a heat field. In this case, the material determines shape simplification. Materials such as titanium-, nickel-, and iron-base alloys have smaller heat capacities and thermal conductivities than aluminum, beryllium, and copper alloys have. The materials with a sufficiently low heat capacity and thermal conductivity frequently can be modeled as semi-infinite and highly symmetric, making an analytic model often entirely adequate. Time constants for different mechanisms such as heat diffusion and bulk or grain-boundary diffusion of mobile species should also be considered as factors in simplifying the process model. Time constants for mechanisms that are appreciably longer than another mechanism can be considered to be operating in a continuum and amenable to an analytic model.

The use of the state-space (lumped parameter) method provides considerable insight into the controllability of metalworking processes. It is a constrained optimization method in which the overall design is specified as constraints and objectives to accommodate multiple physical and economical requirements. This approach identifies optimal workpiece material trajectories that are generally needed for achieving customer product specifications. Critical issues such as efficient material flow and the sensitivity of the variation of finished parts properties to small changes in key process parameters can be addressed by using lumped-parameter models of microstructure evolution and optimization techniques for controlling microstructure during thermomechanical processing. Dynamic models of equipment systems are also useful and essential in determining the desired adjustable parameter settings for coincident tracking of the equipment response with optimized commands. In general, furnaces and forming machinery possess a range of time-varying performance capabilities that can be tuned to satisfy the needs of a given process. In one example, a 1000 ton hydraulic forge press is evaluated using a dynamical model (Ref 6).

Materials Modeling and Stability. Dynamic models of materials behavior are especially important because of the time-varying behavior of quantities such as microstructure, flow stress, and defect formation. Industrial process design and optimization depends on the type of materials model created and the existence of a cost-effective and efficient method for obtaining materials parameters that must be measured (Ref 1, 4, 7–12). The type of materials model created for industrial process simulation depends on two issues:

- The existence of an industrially sensible method for obtaining the needed materials parameters

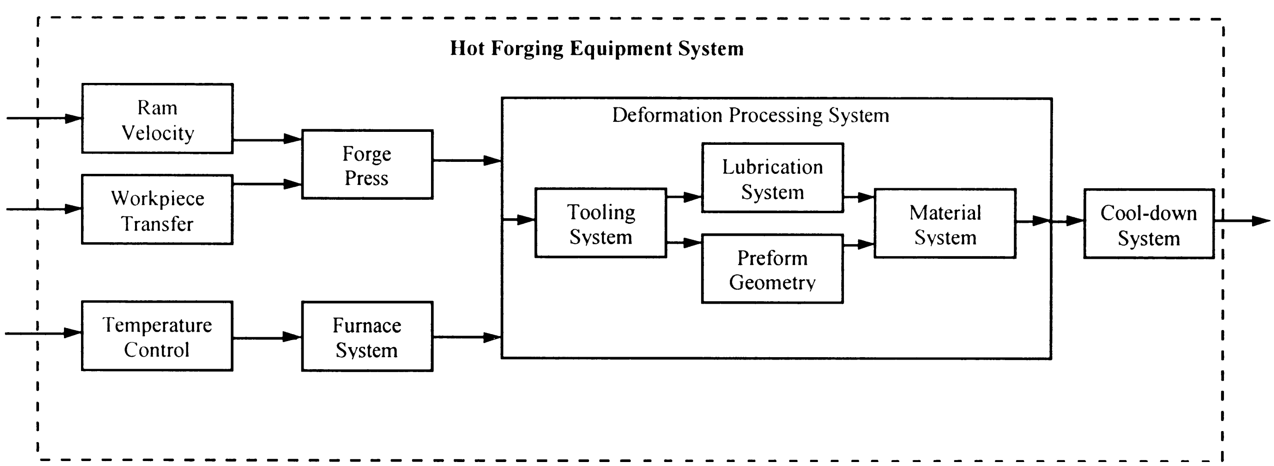


Fig. 1 Block diagram of metal forging system

- The danger that measured parameters may have little value or relevance for another material with a slightly different prior thermomechanical history

The latter item may be particularly troublesome because there are several reasons why materials properties may be variable or erratic. For example, even engineering alloys of similar composition can have lot-to-lot variations in properties, making it difficult to describe precisely the initial conditions needed for a materials model. Engineering alloys may also have multiple mechanisms that provide different degrees of freedom for dissipating energy while undergoing forced dissipative flow during deformation processing. Some of the dissipation mechanisms include lamellar kinking, adiabatic shearing, grain-boundary cracking, dynamic grain growth, grain-boundary diffusion, dynamic recrystallization, and dynamic recovery. These mechanisms may operate either in series or in parallel and at unpredictable times. Thus, the models of alloy deformation and workability can produce erratic results, especially when materials stability is path dependent.

Processing predominantly in a region of materials stability can reduce the sensitivity of erratic results due to materials variability. This is a key point in the concept of process design optimization. A stable processing space is a region defined by temperature, effective strain rate, and time, where the desired microstructures and associated properties evolve in a stable fashion regardless of the state of stress. In contrast, instability is defined as any set of conditions during processing that cause plastic instabilities such as flow localization, fracture, grain-boundary cavitations, and microstructural instabilities such as dynamic grain growth to occur. The occurrences of these defects are strongly material path- or trajectory- (Ref 1) dependent when processing occurs outside the limits for stable material flow.

Materials Stability Criteria and Intrinsic Workability. For lumped-parameter modeling, considerable attention is paid to materials stability criteria, constitutive laws, and physical insight about the materials processes. The concept of materials stability is akin to the idea of workability (Ref 14); it can be expressed in terms of key process parameters, temperature, and effective strain-rate ranges as a function of time, where microstructures and mechanical properties evolve without forming deleterious structures regardless of the state of stress.

Materials being formed (forged) undergo large, nonlinear, irreversible deformation that is made possible by linking together several atomic processes. This string of processes, in addition to providing the degree of freedom needed for producing large deformation, is also responsible for the evolution of structure. Each of these processes can be thought of as a channel for dispersing energy supplied by the forging process, and each can give rise to entropy production. Forging, as a rule, is a shape-changing process in which the deformation is generally inhomogeneous and transient over a large volume of the deforming continuum. Steady-state phenomena often represent a limiting condition that can be achieved only under unique combinations of temperature T and effective strain rate, $\dot{\bar{\varepsilon}}$. Therefore, it is highly desirable to be able to define stable regions in terms of the kinematic variables T, effective strain ($\bar{\varepsilon}$), and $\dot{\bar{\varepsilon}}$.

A solution to this problem requires identification of the limiting conditions for the loci of bifurcation points where two atomic mechanisms are operating simultaneously to produce a maximum in the energy dispersal rate at a unique combination of temperature T and effective strain rate $\dot{\bar{\varepsilon}}$. Under these unique conditions, the deformation process would be steady state. One branch of the deformation process would be stable, and the other would be unstable. Processes such as grain-boundary cavitation, wedge cracking, and cleavage are considered to cause the deforming continuum to become unstable because the free surfaces formed increase, rather than decrease, the free energy of the system. In this sense, workability can be viewed as an intrinsic characteristic of the material; the ability of the material to dissipate power at any state of stress by favorable metallurgical mechanisms generates entropy at a faster rate than it is applied and maintains the total energy of the system at the lowest level possible.

Identification of the limiting conditions for stable plastic deformation is an important concern, as discussed briefly in the next section, "Materials Stability during Forced Dissipative Flow." Materials stability criteria also can be easily incorporated in a simulation model as a so-called nonholonomic constraint (that is, a constraint based on an inequality condition, which does not excessively constrain a set of simulation equations) (Ref 13, 14). These stability limits are the constraints needed for designing preform and other intermediate product shapes, which are required for optimizing metalworking processes used to produce components with complex shapes.

Materials Stability during Forced Dissipative Flow. Successful industrial process design and optimization demand serious consideration of materials data needs. In addition to fundamental knowledge about materials properties such as the moduli of elasticity, heat capacity, thermal expansion, Poisson's ratio, density, compressibility, and the transport properties, it is essential that process designers understand how the workpiece material responds to the demands of the process. Whether material flow dissipates in nonhomogeneous flow and fracture or leads to large-scale order depends on the nature of instabilities in the system (Ref 15).

Decisions about forging process parameters depend on important materials properties, which are based on fundamental thermodynamic stability parameters such as mechanical, thermal, and diffusion stability. Following are some important process parameters (Ref 16) that are controlled by the workpiece material:

- Workpiece temperature at the time it is removed from the heating equipment
- Workpiece temperature after transferring it to the die system
- Workpiece temperature at the start of deformation
- Workpiece temperature at finish forging to control shrinkage
- Optimal die temperature and die material
- Proper die velocity and material strain-rate range
- Press size needed for required force and energy
- Forge press type compatible with workpiece strain-rate sensitivity
- Prefinish shapes for proper metal flow
- Service properties of forged components
- Lubricant type
- Maximum die life

Industrial simulation of deformation processes requires mathematical relationships that predict how the workpiece material will respond to the demands of the process. Overall, this response is expressed as a set of constitutive equations that relate the effective stress (effort variable), $\bar{\sigma}$, and the effective strain rate (flow variable), $\dot{\bar{\varepsilon}}$, in the current state of the material system. The scalar product of these two variables is the instantaneous power. The state of the system is determined by the prior thermomechanical history of the workpiece material and its current temperature (Ref 9, 17).

It is not practical to develop a general constitutive equation for all materials. When large plastic deformations are involved, it is reasonable to neglect the elastic strains. However, when the elastic strains are of the same size as the plastic strains, such as when predicting residual stresses, constitutive equations should include elastic and plastic behavior. Two general approaches are used to develop constitutive equations for material flow analysis. One approach is micromechanical and the other is empirical. A lack of basic quantitative understanding of the fundamental material energy dispersal mechanisms, their high variability from material to material, and different modeling needs for the various processes all work against the usefulness of micromechanical and general models, making empirical materials modeling the most useful for industrial process modeling.

Empirical descriptions of materials behavior that are based on a series of well-lubricated isothermal, constant strain-rate compression tests is effective and affordable because it is a simple test that is easy to understand and analyze. When conducted over the process parameter range of interest, all materials information needed for calculating the force and energy of forging and deciding the important process parameters needed for achieving mechanical and microstructural stability is generated (e.g., Ref 18, 19). This approach is based on careful measurements of the material flow stress over a range of temperature and effective strain rates and strain. The experimental results can be described analytically to produce constitutive equations, or results can be tabulated in spreadsheet form as a function of temperature, effective strain, and effective strain rate. The empirical approach is useful for industrial process modeling also, because it requires the least amount of information needed for determining the intrinsic workability of the workpiece material. The workability of the workpiece material is derived from the constitutive equation data.

A constitutive equation for describing the flow stress behavior based on compression testing data that is corrected for adiabatic heating is expressed as:

$$\bar{\sigma} = \bar{\sigma}(\dot{\bar{\varepsilon}}, T, t) \qquad \text{(Eq 2)}$$

This equation can be represented in differential form such that:

$$d \ln \bar{\sigma} = md \ln \dot{\bar{\varepsilon}} + \varphi dT + \psi\, dt \qquad \text{(Eq 3)}$$

where m is the well-known strain-rate sensitivity parameter, φ is the temperature dependence of the flow stress, and ψ is the rate of change of the flow stress. The coefficients are given by their respective partial derivatives:

$$m = \left[\frac{\partial \ln \bar{\sigma}}{\partial \ln \dot{\bar{\varepsilon}}}\right]_{\bar{\varepsilon},T} = \left[\frac{\dot{\bar{\varepsilon}}}{\bar{\sigma}} \frac{\partial \bar{\sigma}}{\partial \dot{\bar{\varepsilon}}}\right]_{\bar{\varepsilon},T} \qquad \text{(Eq 4a)}$$

$$\phi = \left[\frac{\partial \ln \bar{\sigma}}{\partial T}\right]_{\dot{\bar{\varepsilon}},\bar{\varepsilon}} \qquad \text{(Eq 4b)}$$

$$\psi = \left[\frac{\partial \ln \bar{\sigma}}{\partial t}\right]_{T,\dot{\bar{\varepsilon}}} \qquad \text{(Eq 4c)}$$

In plasticity analysis, these derivatives represent materials properties. They are as important as the flow stress for characterizing materials behavior during processing. These derivatives have been used to define two parameters for axiomatically defining the mechanical and structural stability constraints as a function of T, $\dot{\bar{\varepsilon}}$, and $\bar{\varepsilon}$ for use during materials process design and optimization. The coefficients m and φ are used to define arguments for two Liapunov materials stability criteria. The argument for mechanical stability is the strain-rate sensitivity parameter m and the temperature coefficient of the flow stress φ leads to another coefficient s that is an entropy production rate parameter, which serves as another argument for deciding structural stability. It is important to realize that all of the necessary information for designing a robust forging process is contained in the constitutive equation for the workpiece material.

A stability map can be automatically created for any effective strain level using an appropriate software application to aid the process design team in deciding the forging process parameters previously listed. Because the materials stability map is independent of process equipment and tooling parameters, it is representative of the so-called *intrinsic workability* of the workpiece material. The rate of change of the effective flow stress (ψ at constant temperature and effective strain rate) approaches zero asymptotically as steady state is moved toward. When steady state is reached, the processing map no longer changes with time (strain) and the microstructure and properties are no longer evolving with deformation. This materials behavior at large-scale, stable deformations helps to explain why simple compression test data can be used to simulate different classes of bulk hot-working processes, although the effective strain levels can be as large as 2.0% using constitutive equation data that were generated as a function of effective strain levels that were no larger than 0.4 and 0.6.

Observations such as the latter imply that steady state is a long-term attractor, and it is a fundamental materials behavior characteristic, which can be approximated with a Liapunov function. This has led to the development of a method referred to as *dynamic material modeling* (DMM), which is briefly introduced in Ref 14 and Chapters 1, "Workability and Process Design—An Introduction," and 2, "Bulk Workability of Metals," of this book. More detailed discussions of the DMM methodology are described in Ref 8 to 10, and Ref 8 is a compendium of compression data for 160 alloys with processing maps in support of this modeling method.

The DMM method identifies process design windows for materials stability based on concepts from continuum mechanics and irreversible thermodynamics. The stability of the workpiece material and the evolution of microstructure and properties is described using a classical Liapunov stability criterion that depicts how fast the deforming workpiece material is approaching steady state conditions. The theoretically derived Liapunov stability function cannot be solved directly because the forcing function cannot be specified a priori. The reason for this dilemma is that the initial state for the workpiece material cannot be precisely specified, and the plastic deformation process involves instabilities that are stochastic. Therefore, the Liapunov function must be solved axiomatically by finding a sufficient number of dimensionless system variables that describe its evolution by trial and error. This is the concept of the DMM technique, and it is possible to determine whether the deforming system, which is controlled by selected process mechanisms, is approaching steady state. The system parameters that control mechanical stability and microstructural stability during forced dissipative flow are the strain-rate sensitivity parameter m and an entropy production rate parameter s. Both of these parameters are derived from the constitutive equations for the workpiece material that are used for process simulation.

The mechanical and structural stability criteria used for creating the intrinsic workability maps are summarized subsequently. The Liapunov Function for mechanical stability is $L_m = m(\ln \dot{\bar{\varepsilon}}) > 0$, and the mechanical stability criterion is $\dot{L}_m = \frac{\partial m}{\partial \ln \dot{\bar{\varepsilon}}} < 0$. This condition ensures that the system is approaching steady state and the thermodynamic stability criterion for mechanical stability is satisfied. Similarly, another system parameter s that can be used in the Liapunov sense to evaluate microstructural stability has been defined. The Liapunov function for structural stability is $L_s = s(\ln \dot{\bar{\varepsilon}}) > 0$, and the condition for structural stability in the Liapunov sense is given by $\dot{L}_s = \frac{\partial s}{\partial \ln \dot{\bar{\varepsilon}}} < 0$. This parameter ensures that the thermodynamic diffusion stability criterion is satisfied and that all of the phases given by the material phase diagram can form. After the stability criteria are calculated, the negative values are grouped to decide the stable regions of the material in processing space. If the workpiece material is processed under stable conditions, the necessary and sufficient conditions for stable material flow and microstructure evolution are satisfied regardless of the state of stress. Conversely, if the workpiece material is deformed in an unstable region, it has the necessary but not sufficient conditions to form a structural defect. In this case, the state of stress, $\sigma_m/\bar{\sigma}$, may prevent a fracture-related defect from forming if this ratio is less than zero.

Thus, the DMM methodology is a simple technique to use in an industrial environment. It is the type of materials model for which the field variables such as temperature, strain rate, and strain have an immediate, direct influence, and they appear in the stability criteria. It describes the dynamic path a material element takes in response to an instantaneous change in $\dot{\bar{\varepsilon}}$ at a given temperature T and $\bar{\varepsilon}$. As such, it is a map that graphically describes power dissipation by the workpiece in stable and unstable ways. In contrast, models involving real variables that can be measured, such as grain size, grain growth rate, and fraction transformed, are very difficult to devise and calibrate because microstructure parameters tend to be affected by nearly every physical effect in the process. Thus, it is much easier to optimize a process to avoid a defect than to predict its occurrence or the numeric level of some microstructurally related feature. Physically based material models tend to make difficult demands in terms of model parameters and the experimental data necessary for initializing the model, and the results are often no better than an empirical model.

Multidisciplinary Process Design and Optimization. An understanding of the interaction between materials properties and processing variables is important in controlling the microstructures of forged parts. Multidisciplinary Process Design and Optimization is a tool for optimally controlling the deformation temperature and deformation rate to produce a desirable set of microstructures and properties in a finished forging. The workpiece stability maps of the DMM method can be used to define the critical temperature and strain-rate range for evolving the desired set of microstructures, which subsequently respond to heat treatment, to achieve the objective service properties. Multidisciplinary process design and optimization can also establish the other key process temperatures parameters discussed previously. Strain in the critical regions of the part is optimized through die and preform design, as discussed in more detail in subsequent chapters.

The goal of process design optimization is to create a robust process that meets product requirements and satisfies customer expectations. Meeting these objectives is enhanced through optimization of important process interactions (e.g., see Chapter 14, "Process Design in Impression Die Forging," Fig. 3, for interactions of the forging process). Important process variables include the type of forge press, the flow stress, the interface friction, and the part geometry. The size and geometry of the part determine the forming load and energy, and the flow stress increases with increasing deformation rate and decreasing temperature. The magnitudes of these variations depend on the specific material under consideration, and the frictional conditions deteriorate with die chilling.

Temperature variations during deformation processing are largely influenced by a number of variables for a given workpiece initial temperature. The key process variables include:

- Surface area of contact between the dies and the workpiece
- Workpiece thickness and volume
- Die temperature
- Amount of heat generated by deformation and friction
- Contact time under pressure

The strain rate and temperatures, which are both extremely critical when processing high-temperature alloys, are influenced by the characteristics of the forging press used for processing.

Material Trajectory Optimization

Optimization concepts are based on the notion that ideal deformation conditions can be defined for a material and/or a process. In the context of stable dissipation under forced conditions (described by DMM charts), dynamic state-space models of microstructure evolution can be developed from control-theory concepts to identify precise material trajectories for achieving the design objectives of finished shapes. Analytic models of microstructure evolution are applied in the stable regions, and the variable nature of materials behavior is managed by designing the materials process to operate primarily in the regions of materials stability for a process-parameter space. The deleterious effects from erratic behavior are dampened in the regions of materials stability, where the apparent activation energies that correspond to beneficial energy dissipation mechanisms are relatively low compared with materials failure mechanisms. In this approach, the analytic microstructure evolution models are only reliable when applied in the stable regions, where microstructure and mechanical property parameter variance is not sensitive to material path trajectory.

The use of optimization concepts to obtain valid and realistic design solutions applies to a variety of process analysis methods that are available to process design teams for almost any materials process design problem. When shape or volume is a relevant factor in optimization, mesh-based numerical techniques can be used to analyze geometrical relationships between the initial and final shapes. The simplified finite element method (SFEM), for example, corresponds to analysis methods such as the Sachs (Slab) and upper bound (UB) methods with respect to computation time and the analysis of a typical step of the forming process. Optimization of forging shape and properties may also involve factors associated with subsequent operations such as machining and heat treatment. The geometry of a near-net-shape forging geometry is influenced by several practical considerations. For example, one optimization factor may be the uniformity of residual stress distributions. Another type of optimization factor may be die life, which is controlled by the stresses within the die impression and on the die outer shapes. For reasons such as these, designers do not usually know a priori if some geometric feature is sufficiently significant to rule out the use of simplified process models, such as state-space (lumped-parameter) models.

Nonetheless, modeling of dynamical systems in state-space form provides a natural framework for describing the time evolution of the states and outputs of a system under the influence of multiple, time-varying inputs. This provides an effective framework for diverse optimization strategies over the entire materials processing design problem, not just some parts of the process at the expense of the whole manufacturing enterprise. This approach to optimized process design is a stark contrast to the traditional philosophy that a more predictive capability for materials process models leads to better design solutions. Some designers run detailed finite element simulations looking for trends in the behaviors of certain key process field parameters such as the material flow pattern and how they are affected by different values of die velocity, die temperature, and workpiece temperature. This ad hoc approach to materials process design is frequently but mistakenly called *process optimization.* It is very labor intense and costly, and the ad hoc approach may not be effective in finding an optimal solution.

Optimization of microstructures is one key application of state-space control theory. The following section describes this concept in more detail, and two examples are given. One example describes a state-space model for dynamic recrystallization of steel, and the second example describes trajectory optimization techniques of the extrusion process (Ref 20). The extrusion process is an example of a steady state process that can be effectively modeled using this method of analysis. For extrusion, material flow occurs predominantly along an axial direction, but high interface friction can affect the stresses and can cause nonhomogeneous deformation in the radial direction. In this class of problem, microstructure evolution is influenced by the flow pattern, the rate of energy input, and the ability of the workpiece material to dissipate energy under stable conditions.

Optimization of Microstructure Development Using Control Theory (Ref 21)

As shown in Figure 2, the design methodology is separated into a microstructure optimization stage (or problem) and a process optimization stage (or problem). Goals of the first stage are to achieve enhanced workability and to obtain prescribed microstructural parameters. In this investigation, an attempt has been made to control microstructural parameters such as final grain size (average) and volume fraction recrystallized during deformation. In the second stage, a primary goal is to achieve the thermomechanical conditions required by the first stage for predetermined regions of the deforming workpiece. In Figure 2, desired microstructural features, such as volume fraction recrystallized and grain size, serve as inputs to the microstructure optimization problem. Trajectories of strain, strain rate, and temperature are the outputs, which depend on the details of the microstructure development model and the criterion used for optimization. Next, these trajectories become the inputs to the process optimization problem. The outputs of the process optimization are the die shape, ram velocity, and billet temperature for the case of extrusion.

The design approach requires three basic components for defining and setting up the optimization problem: a dynamical system model, physical constraints, and an optimality criterion.

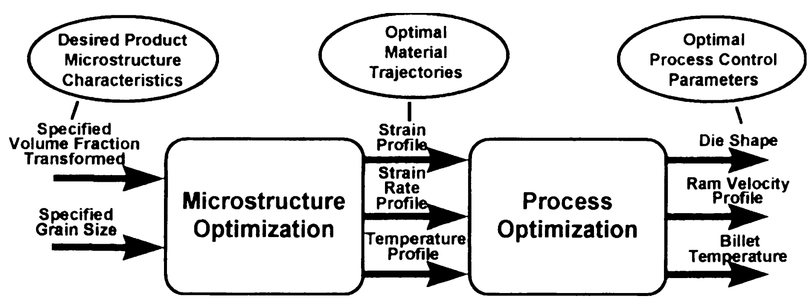


Fig. 2 Schematic representation of the two-stage approach

In metal forming, the system models of interest are material behavior and deformation process models; constraints include the hot workability of the work piece and the limitations of the forming equipment. Optimality criteria can be chosen to achieve a particular set of final microstructural features, to regulate temperature, and to maximize deformation speeds.

Materials Behavior and Process Modeling. Materials behavior models that describe the kinetics of primary metallurgical mechanisms such as dynamic recovery, dynamic recrystallization, and grain growth during hot working are required for analysis and optimization of materials system dynamics. As an illustration, consider a case of dynamic recrystallization for which a possible state-space model is:

$$\begin{bmatrix} \dot{D} \\ \dot{\chi} \\ \dot{\varepsilon} \\ \dot{T} \end{bmatrix} = \begin{bmatrix} f_1(T,\dot{\varepsilon},D) \\ f_2(T,\dot{\varepsilon},D,\chi) \\ u \\ \eta\sigma\dot{\varepsilon}/\rho C_p \end{bmatrix} \quad \text{(Eq 5)}$$

The current state at any point in a deforming body is described by the state variables grain size D; volume fraction recrystallized χ; accumulated strain ε; and current temperature T. The time rates of change of these variables, i.e., $\dot{D}$, $\dot{\chi}$, $\dot{\varepsilon}$, and $\dot{T}$, are functions of the current state variables and input variable, which is strain rate. In Eq 5, f_1 and f_2 are known functions; u is the system input, which is the strain rate in the present case; η is a coefficient that determines how much of the mechanical work is converted into heat and contributes to the increase in temperature of the billet; σ is flow stress; and the product ρC_p is the heat capacity (ρ is density and C_p is the specific heat) of the material.

In addition to dynamic system models, the formulation of an optimal control problem requires a statement of physical constraints and specification of an optimality criterion for producing the desired hot-worked microstructural characteristics. The limiting process conditions for acceptable hot workability are important materials behavior constraints in the first stage of the control strategy. The ranges of temperature and strain rate over which the material exhibits a table processing window have to be determined. This defines a region for any particular thermomechanical trajectory that may be determined using the prescribed optimality criterion.

Formulation of the Optimal Control Problem. The design problem is formulated into an open-loop optimal control problem (Ref 22) that can be formally stated as follows. Find input variables u to minimize the optimality criterion:

$$J = h[x(t_f)] + \int_{t=0}^{t_f} g[x(t),u(t)]dt \quad \text{(Eq 6)}$$

with the constraint that the system state obeys the state-space model:

$$\dot{x}(t) = f[\mathbf{x}(t),\mathbf{u}(t)],\ x(0) = x_o \quad \text{(Eq 7)}$$

In Eq 6 and 7, t is time; $x(t)$ is a vector of state variables; u is the system control input; t_f is the duration of the process; $h(x(t_f))$ is the penalty associated with violating the desired final state; $g(x(t),u(t))$ is the integrand of the penalty associated with deviating from desired state and control input trajectories; $f[\mathbf{x}(t), \mathbf{u}(t)]$ is a vector function that describes the process dynamics; and x_o is the initial state condition. A graphical description of the optimal control problem represented by Eq 6 and 7 is given in Fig. 3 for a case of a system with one input and one state. Suppose that an optimality criterion of the type given in Eq 6 has been defined and that several possible input trajectories have been evaluated according to that optimality criterion. Figure 3(a) shows several of the infinite number of trajectories that the system input can follow. The corresponding trajectories of the state variable (called *trial trajectories*) are given in Fig. 3(b). Figure 3(c) gives the value of the optimality criterion that corresponds to each of these trial trajectories as a function of trial index, i. The objective is to find the input trajectory that, together with the corresponding trajectory of the state variable, will give a minimum value of J, labeled as "Optimal, J^*" in Fig. 3(c). It is important to note that minimization of an optimality criterion implies that the system has been optimized with respect to that particular criterion, and that whether the original design objectives underlying the formulation of the optimality criterion have been met is a different issue.

Optimality Criteria for Microstructure Development. Careful selection of optimality criteria is critical for finding the most appropriate design solutions. In the control of microstructure development during hot metal deformation, common design criteria include producing specified microstructural features and gradients of microstructure within a specified variance on a repeatable basis. These optimality criteria and others can usually be formulated as functions to be minimized, and are often lumped together into a single scalar optimality criterion J in the form:

$$J = J_1^F + J_2^F + \ldots + J_{N_F}^F + J_1^T + J_2^T + \ldots + J_{N_T}^T \quad \text{(Eq 8)}$$

where the superscripts F and T refer to requirements on desired final states and trajectories, respectively. In the case where it is desired that microstructure feature x achieve a value x_{des} at the termination of the deformation process, the corresponding term in J often has the form:

$$J_1^F = \beta_1(x(t_f) - x_{des})^2 \quad \text{(Eq 9)}$$

where β_1 is a weight factor. This type of function can also be used to include certain fixed process parameters and other final values for nonmicrostructural quantities such as strain and temperature in optimization calculations. The terms J_j^T in the optimality criterion define requirements on the desired state and control input trajectories during the forming process and have integral forms.

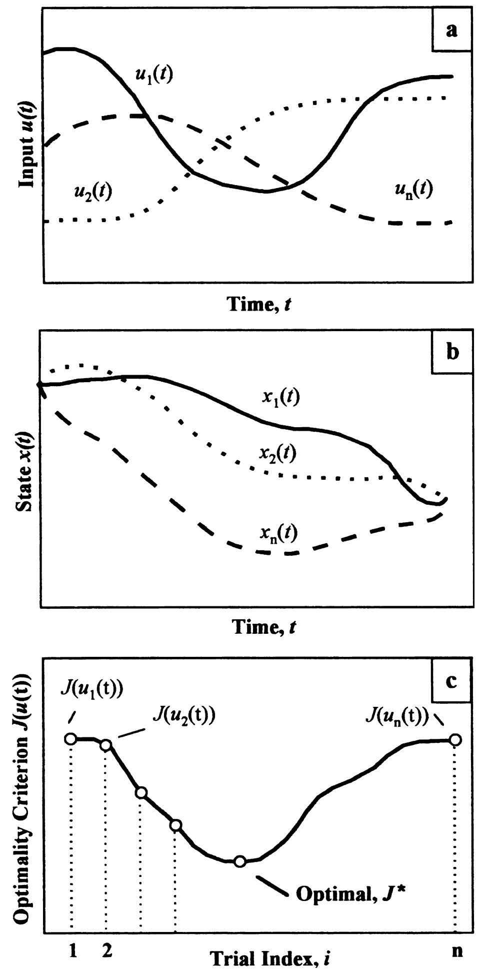


Fig. 3 A one-input, one-state example of the optimal control problem. (a) Several possible input trajectories. (b) Corresponding state trajectories. (c) Corresponding values of the optimality criterion

Table 1 shows some examples of typical optimality criteria for microstructure development during hot metal deformation. Both final value and trajectory specifications are shown. The general formulation of this approach allows new terms to be defined according to the specific needs of each design problem. The quantities $f_x(x,a)$ and $f_x(x,a,b)$ in Table 1 are penalty functions that can be used to constrain optimized design solutions within acceptable process parameter ranges imposed by material workability or equipment limitations. These functions evaluate to virtually zero for values of x in the acceptable range and attain very high values when x is outside that range. Scalars a and b define the acceptable ranges for process parameters, such as temperature or strain rate. An example of a penalty function is shown in Fig. 4.

The weight factors β, serve three purposes. First, they are used to scale the terms in J so that they have comparable influence in the overall optimality criterion. Second, they are increased for certain terms according to their relative importance to achieve overall design requirements. Third, they may be adjusted in order to avoid possible conflicts in design requirements and obtain a satisfactory compromise solution.

The optimality criterion J, which is to be minimized in order to determine, ε, $\dot{\varepsilon}$, and T, can incorporate a number of physically realistic requirements. For the specific problem of hot metal deformation, one possible optimality criterion is:

$$J = \beta_1 (D(t_f) - D_{des})^2 + \beta_2 (\chi(t_f) - \chi_{des})^2 + \int_0^{t_f} \{ f_2^w [\dot{\varepsilon}(t), \dot{\varepsilon}_{min}, \dot{\varepsilon}_{max}] + f_1^w [T(t), T_{min}, T_{max}] \} dt \qquad \text{(Eq 10)}$$

where D is the average recrystallized grain size; D_{des} is the desired final grain size; χ_{des} is the desired final volume fraction recrystallized; $\dot{\varepsilon}_{min}$ and $\dot{\varepsilon}_{max}$ are the minimum and maximum limits of strain rate, respectively; T_{min} and T_{max} are the minimum and maximum limits of temperature for acceptable workability, respectively; $\dot{\varepsilon}(t)$ is the nominal strain rate; and $T(t)$ is the nominal temperature. The functions f_1 and f_2 ensure that the nominal strain rate and temperature during deformation will be kept within the prescribed limits of the safe processing window.

Finding an analytical solution to the problem given in Eq 10 is highly unlikely due to the complexity of the resulting functional form. However, it is possible to formulate a numerical algorithm that can yield a practical solution (Ref 22).

Examples

The formulation of state-space equations is described in the previous section for microstructural evolution during hot working. The following two examples use trajectory optimization with mathematical relationships for dynamic recrystallization of plain carbon steel. These equations are valid within the stable processing window determined by the DMM method.

Example 1: State-Space Models for Dynamic Crystallization of Steel. An empirical model developed by Yada (Ref 23) and critically reviewed and assessed by Devadas et al., (Ref 24) shown in Table 2, was used to describe the relationship between microstructural parameters

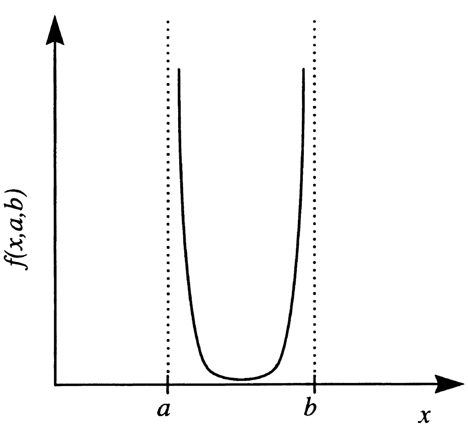


Fig. 4 A typical penalty function, *f*(*x*,*a*,*b*), for constraining process parameters between *a* and *b*

D and χ and process parameters $\dot{\varepsilon}$, T, and ε. The volume fraction recrystallized χ is zero until a critical amount of strain ε_c has been imposed. Beyond this critical strain, the kinetics of recrystallization at any temperature T is characterized by $\varepsilon_{0.5}$, the strain required to reach 50% recrystallization. The time derivative of the fraction recrystallized can be approximated by applying the chain rule of differentiation to the equation for χ in Table 2 with some simplifying assumptions. The result is shown in Table 3.

If the desired objective is to achieve a specified grain size, it is possible to obtain the desired $\dot{\varepsilon}$ from the fourth equation in Table 2, assuming T is fixed. Because of the dependence of $\dot{T}$ on $\dot{\varepsilon}$ (Table 3), this assumption is not valid, although it may be adequate in some applications. The purpose of this example, however, is to demonstrate a systematic approach applicable to any problem, not to find a solution for plain carbon steel only. It should be clear that, for many materials, the dependence of the size of the forming grains on the temperature, strain rate, and strain may be considerably more complex. In fact, the coupling among these parameters combined with multiple constraints on temperature, strain rate, and strain can easily eliminate straightforward solution approaches, while the state-space trajectory optimization approach can reliably solve the problem.

The equation for the rate of change of temperature due to deformation given in Table 3 states that a fraction η of the mechanical work is converted to heat and increases the temperature. The microstructural state of the material is given by the state vector $\mathbf{x} = [\chi, T, \varepsilon]^T$, which evolves

Table 1 Examples of typical terms for the optimality criterion for microstructure development during hot metal deformation

Design objective	Term in the optimality criterion
Achieve final average grain size D_{des}	$J_i^F = \beta_i (D(t_f) - D_{des})^2$
Achieve final strain of ε_{des}	$J_i^F = \beta_i (\varepsilon(t_f) - \varepsilon_{des})^2$
Maintain strain rate between $\dot{\varepsilon}_{min}$ and $\dot{\varepsilon}_{max}$ because of workability considerations	$J_j^T = \int_0^{t_f} \beta_j(t)\, f(\dot{\varepsilon}, \dot{\varepsilon}_{min}, \dot{\varepsilon}_{max})\, dt$
Limit deformation heating; initial temperature is T_0	$J_j^T = \int_0^{t_f} \beta_j(t)\, (T - T_0)^2\, dt$
Keep strain rate under $\dot{\varepsilon}_{max}$ because of equipment limitations	$J_j^T = \int_0^{t_f} \beta_j(t) f(\dot{\varepsilon}, \dot{\varepsilon}_{max})\, dt$
Maintain temperature between T_{min} and T_{max} because of workability considerations	$J_i^T = \int_0^{t_f} \beta_j(t)\, f(T, T_{min}, T_{max})\, dt$
Limit energy consumption; $\sigma(t)\,\dot{\varepsilon}(t)$ is a measure of power, and σ is flow stress	$J_j^T = \int_0^{t_f} \beta_j(t)\, \sigma(t)\, \dot{\varepsilon}(t)\, dt$

Table 2 Yada equations for the dynamic recrystallization of steel

Parameter	Equation
Volume fraction recrystallized	$\chi = 1 - \exp(\ln(2)((\varepsilon - \varepsilon_c)/\varepsilon_{0.5})^2)$
Critical strain	$\varepsilon_c = 4.76 \times 10^{-4}\, e^{8000/T}$
Plastic strain for 50% volume fraction recrystallization	$\varepsilon_{0.5} = 1.144 \times 10^{-3}\, D_0^{0.28}\, \dot{\varepsilon}^{0.05}\, e^{6420/T}$
Average recrystallized grain size	$D = 22{,}600\, \dot{\varepsilon}^{-0.27}\, e^{-0.27(Q/RT)}$
Activation enery and gas constant	$Q = 267$ kJ/mol, $R = 8.314 \times 10^{-3}$ kJ/mol K

Table 3 Equations used in the state-space model of microstructural evolution

Parameter	Equation
Time derivative of volume fraction recrystallized	$\dot{\chi} \approx \frac{\partial \chi}{\partial \varepsilon}\frac{d\varepsilon}{dt} = \frac{2 \ln 2}{(\varepsilon_{0.5})^2}(\varepsilon - \varepsilon_c)(1-\chi)\,\dot{\varepsilon}$
Time derivative of temperature	$\dot{T} = \frac{\eta}{\rho C_p}\, \sigma(\varepsilon, \dot{\varepsilon}, T)\, \dot{\varepsilon}$
Flow stress (kPa)	$\sigma = \sinh^{-1}[(\dot{\varepsilon}/A)^{1/n} e^{Q/nRT}] / 0.0115 \times 10^{-3}$ $\ln A(\varepsilon) = 13.92 + 9.023 / \varepsilon^{0.502}$ $n(\varepsilon) = -0.97 + 3.787 / \varepsilon^{0.368}$
Activation energy and gas constant	$Q(\varepsilon) = 125 + 133.3 / \varepsilon^{0.393}$, $R = 8.314 \times 10^{-3}$

according to the equations given in Table 3 for time derivatives of χ and T and the obvious relationship between the input strain rate u and the evolution of strain. Since the grain size does not influence the other state variables χ, ε, and T, it is treated as an output of the dynamical system and not included as one of the state variables.

Example 2: Application of a Two-Stage Approach for Optimization of Grain Size in Extruded Steel. Trajectory optimization has been applied to the design of extrusion processes (Ref 20). Optimized material trajectories in strain, strain-rate, and temperature space were determined for extruding a plain carbon steel into a 26 mm diameter bar shape with 26 μm grain size. It was a two-stage approach. In the first stage, the optimized strain, strain-rate, and temperature solutions were obtained for the grain size objective. In the second stage, the optimized material trajectories were used to directly calculate the corresponding extrusion die throat geometry for achieving the optimized strain and strain-rate trajectories.

Stage 1: Optimizing the Microstructural Trajectories. For the case studied here, the optimality criterion was chosen so as to attain a given final strain of 2, while ensuring that the recrystallized grain size was kept at a desired value of 26 μm. The average grain size of the raw stock prior to extrusion was 120 μm. The optimality criterion chosen was:

$$J = 10(\varepsilon(t_f) - 2.0)^2 + \int_0^{t_f} (D(t) - 26)^2 \, dt \qquad \text{(Eq 11)}$$

where a desired final strain of 2, with a weighing factor of 10, and a desired grain size of 26 μm have been specified. The results of the first stage of the microstructural optimization problem are shown in Fig. 5(a). Starting at an initial temperature of 1273 K, the temperature of the material increases approximately to 1295 K for deformation to a strain of 2.0. The strain rate is initially slightly below 1.0 s^{-1} and increases gradually to a little above 1.0 s^{-1}. The recrystallized grain size, which is initially 120 μm, decreases to 26 μm beyond the critical strain of approximately 0.25. Subsequently, since the recrystallized grain size depends both on T and $\dot{\varepsilon}$, the simultaneous increase in both of these variables maintains the grain size constant. The results of two additional optimization runs to achieve grain sizes of 30 and 15 μm are shown in Fig. 5(b) and (c), respectively. The initial billet temperature was 1273 K for the second case and 1223 K for the third case.

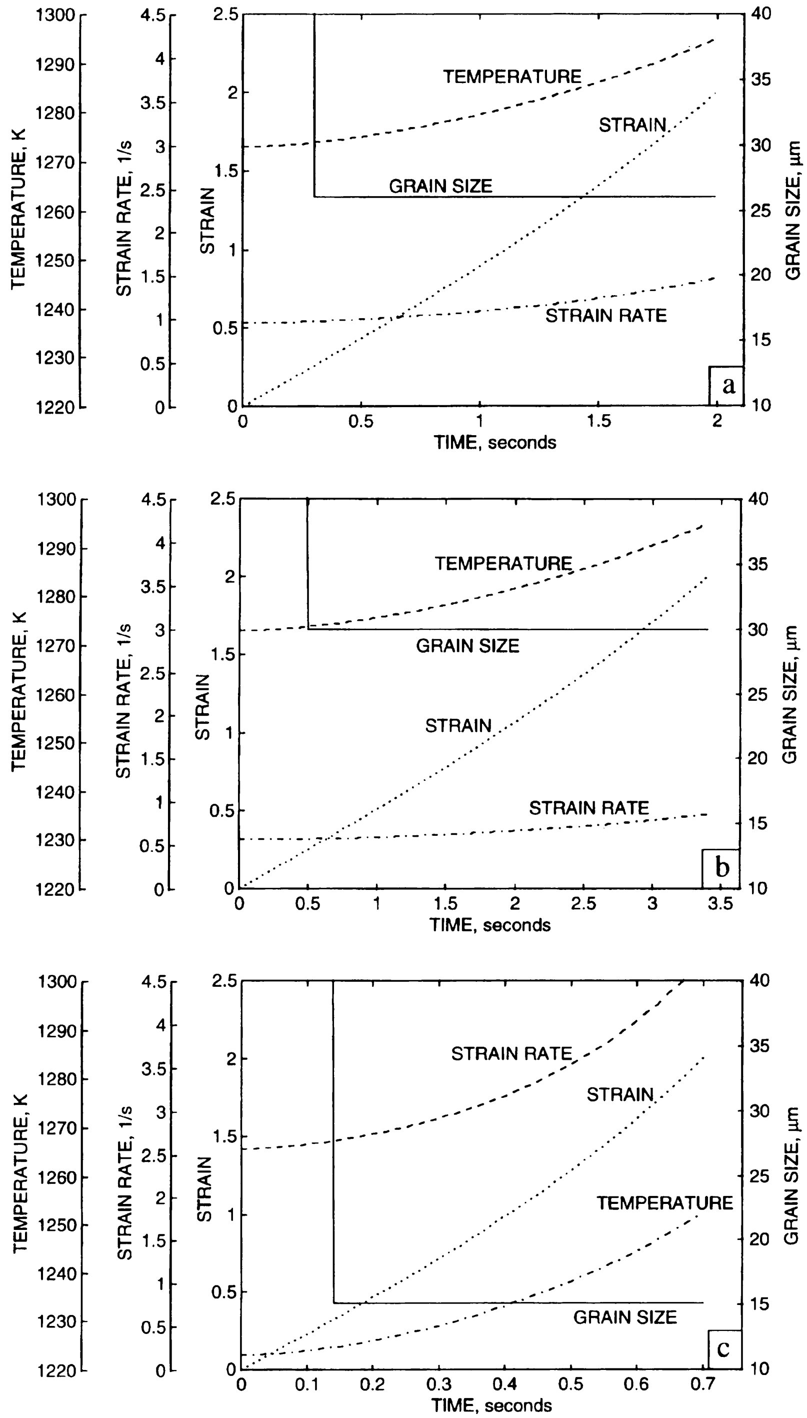


Fig. 5 Trajectories of strain, strain rate, temperature, and grain size for achieving the desired final grain size of (a) 26 μm, (b) 30 μm, and (c) 15 μm

Stage 2: Optimizing the Process Parameters. It is not physically possible to ensure that all the points in the deforming piece will undergo the strain, strain rate, and temperature trajectories obtained in stage 1. However, process parameters such as die geometry, ram velocity, and billet temperature can be designed to ensure that selected regions of the material will experience trajectories that approximate those designed. It is feasible to formulate a second optimization problem that determines values for process parameters that will attempt to achieve the desired trajectories at predetermined points in the material piece. In such an approach, each evaluation of the objective function for the optimization process usually requires a detailed analysis of the deformation process by the finite element method or some other technique.

In the case of round-to-round extrusion, it is possible to analytically calculate the die profile and ram velocity necessary for achieving the desired strain and strain rate profiles at the centerline of the piece. Given that r_0 is the die entrance radius (equal to the billet radius), L is the die length, and $\varepsilon(t)$ is the required strain trajectory, the ram velocity is:

$$V_{ram} = \frac{L}{\int_{t=0}^{t_f} e^{\varepsilon(t)} dt} \qquad \text{(Eq 12)}$$

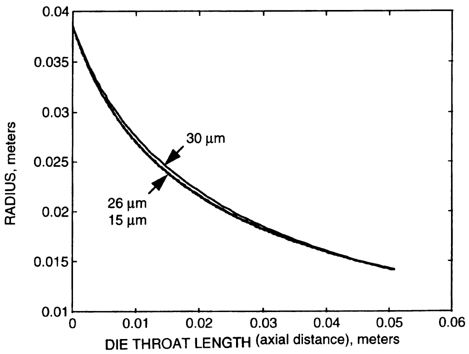


Fig. 6 Optimal die profile for achieving the final grain size of 26, 30, and 15 µm

The die shape can be described by the radius r and axial distance (die throat length) y, where $r(t) = r_0 e^{-\varepsilon(t)/2}$ and

$$y(t) = V_{\text{ram}} \int_0^t e^{\varepsilon(\tau)} d\tau \qquad \text{(Eq 13)}$$

Figure 6 gives the optimal die profile for achieving final grain sizes of 26, 30, and 15 µm obtained using this approach. The optimum ram velocities for achieving these grain sizes were 8.43, 5, and 25.1 mm/s, respectively. Note that the die shape is almost the same for the three optimization cases. This means that the same die can be used to achieve the different recrystallized grain sizes by changing only the velocity of the extrusion press.

REFERENCES

1. J.C. Malas and W.G. Frazier, *Optimal Design of Thermomechanical Processes Using Ideal Forming Concepts, The Integration of Material, Processes and Product Design,* Zabaras et al., Ed., A.A. Balkema, 1999, p 229–236
2. H. Cheng, R.V. Grandhi, and J.C. Malas, Design of Optimal Process Parameters for Non-Isothermal Forging, *Int. J. Numer. Methods Eng.,* Vol 37, 1994, p 155–177
3. R.E. Skelton, *Dynamic Systems Control,* John Wiley and Sons, Inc., 1988
4. K. Chung and O. Richmond, A Deformation Theory of Plasticity Based Upon Minimum Work Paths, *Int. J. Plast.,* Vol 9, 1993, p 907–920
5. K. Chung and O. Richmond, The Mechanics of Ideal Forming, *Trans. ASME,* Vol 61, 1994, p 176–181
6. W.G. Frazier et al., Modeling and Simulation of Metal Forming Equipment, *J. Mater. Eng. Perform.,* Vol 6 (No. 2), 1997, p 153–156
7. R.H. Wagoner and J.-L. Chenot, *Metal Forming Analysis,* Cambridge University Press, Cambridge, UK, 2001, p 206–255
8. Y.V.R.K. Prasad and S. Sasidhara, *Hot Working Guide: A Compendium of Processing Maps,* ASM International, 1997
9. Y.V.R.K. Prasad, et al., Modeling of Dynamic Material Behavior in Hot Deformation: Forging of Ti-6242, *Metall. Trans. A,* Vol 15A, 1984, p 1883–1891
10. H.L. Gegel, *Synthesis of Atomistics and Continuum Mechanics, Computer Simulation in Materials Science,* R.J. Arsenault et al., Ed., ASM International, 1988, p 291–344
11. Y.V.R.K. Prasad, et al., Titanium Alloy Processing, *Adv. Mater. Process.,* Vol 157 (No. 6), 2000, p 85–89
12. D. Cebon and M.F. Ashby, Information Systems for Material and Process Selection, *Adv. Mater. Process.,* Vol 157 (No. 6), 2000, p 44–48
13. J.C. Malas, "A Thermodynamic and Continuum Approach to the Design and Control of Precision Forging Processes," master's thesis, Wright State University, Dayton, OH, 1985
14. J.C. Malas and V. Seetharama, Use of Material Behavior Models in the Development of Process and Control, *J. Met.,* Vol 44, 1992, p 8–13
15. I. Prigogine, *From Being to Becoming: Time Complexity in the Physical Sciences,* W.H. Freeman and Company, New York, 1980
16. T. Altan, S. Oh, and H. Gegel, *Metal Forming: Fundamentals and Applications,* ASM International, 1983
17. H.L. Gegel, Synthesis of Atomistics and Continuum Mechanics, *Computer Simulations in Materials Science,* R.J. Arsenault, J.R. Beeler, Jr., and D.M. Esterling, Ed., ASM International, 1988, p 291–344
18. P. Dadras and J.F. Thomas, Jr., Metallurgical Transactions, Characterization and Modeling for Forging Deformation of Ti-6242, *Metall. Trans. A,* Vol 12A, 1981, p 1867
19. J.F. Thomas, Jr. and R. Srinivasan, Constitutive Equations for High-Temperature Deformation, *Computer Simulations in Materials Science,* R.J. Arsenault, J.R. Beeler, Jr. and D.M. Esterling, Ed., ASM International, 1988, p 269–290
20. W.G. Frazier et al., Application of Control Theory Principles to the Optimization of Grain Size During Hot Extrusion, *J. Mater. Sci. Technol.,* Vol 14, 1998, p 25–31
21. J.C. Malas et al., Optimization of Microstructure Development during Hot Working using Control Theory, *Metall. Mater. Trans. A,* Vol 28A, Sept 1997, p 1921–1930
22. D.E. Kirk, *Optimal Control Theory: An Introduction,* Prentice-Hall, 1970, p 29–46, 184–309
23. H. Yada, *Proc. Int. Symp. Accelerated Cooling of Rolled Steel, Conf of Metallurgists,* CIM, Pergamon Press, 1987, p 105–120
24. C. Devadas et al., *Metall. Trans A,* Vol 22A, 1991, p 335–349

Handbook of Workability and Process Design
G.E. Dieter, H.A. Kuhn, and S.L. Semiatin, editors, p 346-367
DOI:10.1361/hwpd2003p346

Chapter 24

Application of Multidisciplinary Optimization (MDO) Techniques to the Manufacturing of Aircraft Engine Components

Shesh K. Srivatsa, GE Aircraft Engines

MANUFACTURING INDUSTRIES face great challenges in the current dynamic and globally competitive market environment to develop new products and processes to satisfy market demands in a timely, efficient, and cost-effective manner. Both original equipment manufacturers (OEMs) and their supplier organizations face several design and manufacturing challenges that cannot be met by traditional methods. There is a strong need to dramatically reduce product development cycles and product design and manufacturing costs. New methodologies such as multidisciplinary optimization (MDO) are needed to overcome major obstacles to producing good product and process designs quickly.

This chapter discusses how MDO technologies can help reduce cost, cycle time, and delivery time. The focus in this chapter is on process optimization of aircraft engine components, specifically those manufactured by bulk forming processes. However, the MDO methods are equally applicable to other components in other industries. Multidisciplinary optimization technologies are becoming key design techniques for the new-economy businesses to ensure delivery speed and reliability. This has resulted in productivity enhancements by increasing design speed, minimizing human error, and improving designs, a critical requirement.

Turbine Disk Design and Manufacturing

The design and manufacturing of gas turbine engine components is a complicated, highly coupled, and iterative multidisciplinary process. For life-critical rotating parts and disks, the focus of this chapter, it involves both thermomechanical design for life prediction and manufacturing process simulation including forging, heat treatment, and machining. The individual steps of the disk design process, broken down into the mechanical design and the process/manufacturing aspect, are shown in Fig. 1. The thermomechanical design, for example, starts with a simple analysis to obtain a rough thickness distribution of the disk. As knowledge about the design increases, more complex analysis models are created up to a full three-dimensional (3D) finite-element analysis with tens of thousands of elements. The objective during the mechanical design phase is the determination of the final disk shape, as early as possible in the design timeline in order to be able to release the forgings, which tend to require a long lead time. A shape is to be determined that meets mission requirements at minimum weight and/or minimum cost subject to other technical and business constraints.

Historically, a sequential and a largely empirical process had been employed where the shape of the disk as created by the engine manufacturer was handed to the forging supplier. The engine manufacturer determined the final shape of the disk, added a protective layer around it (resulting in the so-called "ship shape"), and passed it to the forging supplier. The protective layer is added to account for variation and uncertainty inherent in any manufacturing process and to ensure that the geometry will clean up to the final required shape. The supplier then designed a forging die based on empirical and test results and prior experience to produce a disk to fulfill the engine manufacturer's requirements. Starting in the 1980s, the largely empirical process of die design was gradually replaced by material modeling methods and finite-element modeling (FEM) techniques. Two decades later, at least in the aerospace and automotive industries, these analysis techniques for metalworking processes are accepted, routine practices. Die designs are usually verified with finite-element analysis and reviewed upfront with the engine manufacturer prior to committing to production. A detailed simulation of the various steps in the manufacturing process is conducted to determine both residual stresses and final distortions of the finished part after machining operations. These residual stresses, in turn, are used in the subsequent life prediction of the part.

After the initial die design is completed, the part is forged, heat treated, inspected, and machined. If test or analysis results show that the disk does not meet its life requirement, or machining distortion exceeds the allowable limit, or forging and heat treatment do not produce the desired material microstructure and mechanical properties, or the press load required is beyond the capacity of the forging press, or the material does not fill the dies completely and without defects, or any other requirement is not fulfilled, then part of or all of the process has to be repeated. If that is not possible, the finished disk shape has to be changed and the mechanical design—at least in part—has to be revised. The same applies if the design does not meet life requirements. This process of iterating between design and manufacturing to ensure that the parts meet all their design requirements is a very expensive and time-consuming process.

A basic requirement of any forging shape and process design methodology is an accurate description of the deformation mechanics involved in the process. The deformation analysis technology has developed over the last couple of decades to a mature level and meets the basic design requirements. Process analysis for cost effectiveness and rapid product development continues to advance in sophistication through the use of modern design methods and tools. The role of computers in process development continues to expand through applications of data acquisition and management, modeling and sim-

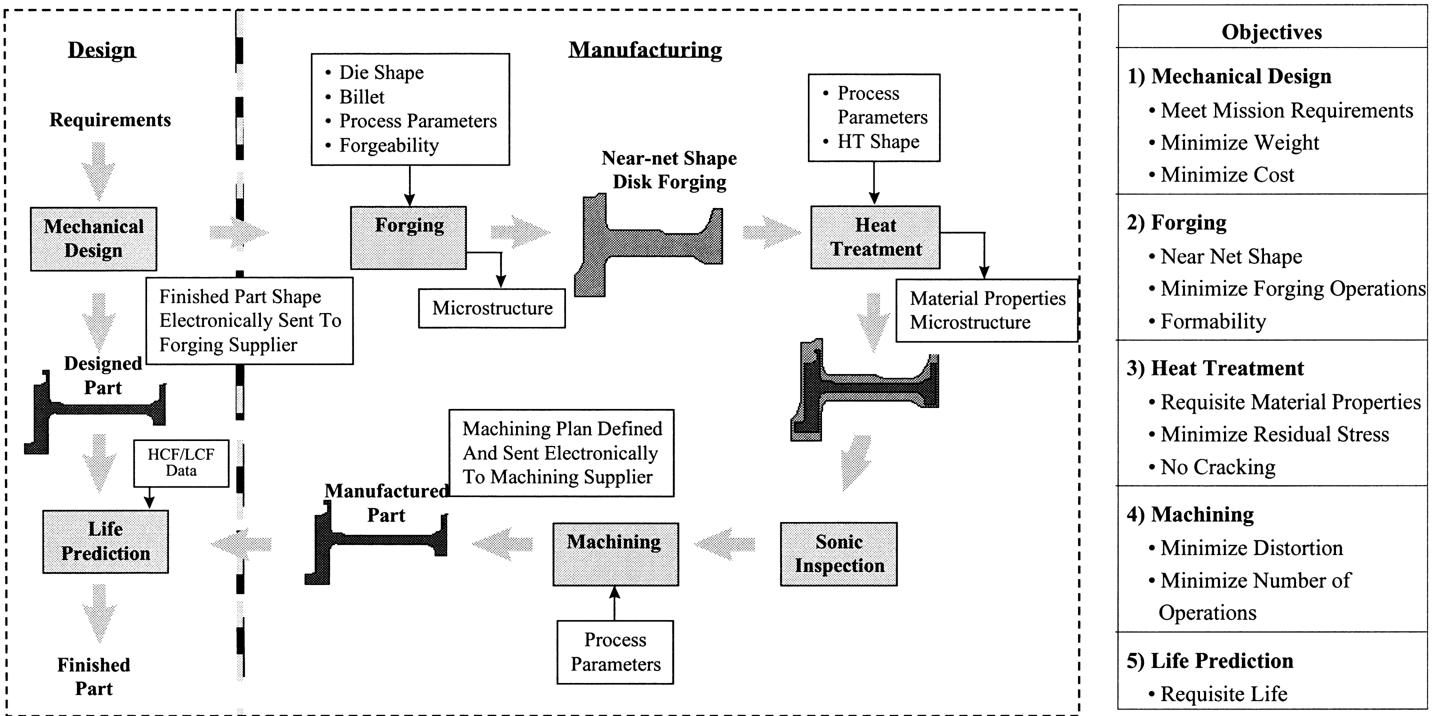


Fig. 1 Key activities in producing a forged component demonstrating the application of geometry and finite-element-based multidisciplinary optimization for complex product and process development

ulation, and (model-based) process control. The material modeling methods and finite-element modeling techniques developed in the 1980s and 1990s for metalworking processes are accepted, routine practices in well-managed part supplier companies. However, in spite of the increased use of modeling technology, actual process design variables (number of dies, shapes of dies, etc.) still require a significant amount of expert knowledge. Design rules are secrets closely guarded by the forgers and based to a considerable extent on previous experience. Simulation methods have just moved some of the trial-and-error design methodologies from the shop floor into the office. To overcome this dependency on a small group of industry experts and to reach designs that are more robust and ultimately less costly, optimization-based design techniques need to be employed.

In addition, the process just described results in merely an acceptable design, that is, one that meets the specifications, but not an optimal one with respect to user-defined criteria such as weight, cost, and performance. In today's cost-competitive environment, it is imperative that optimal solutions and not just acceptable solutions be obtained and be obtained rapidly. Manual optimization based on experience is time consuming, expensive, and tedious due to the multidisciplinary nature of the task and because of the complexity of computer-aided design (CAD) and computer-aided engineering (CAE) tools. Accurate simulation tools for individual stages of the disk design and manufacturing processes are widely available and routinely used. However, opportunities for optimization of the individual steps in this process are not fully utilized. Engineering analysis and process simulation tools have to be integrated into an optimization environment. An integrated procedure that addresses both mechanical design and manufacturing processes is almost mandatory to reduce design cycle time and eliminate wasteful iterations because of the multidisciplinary nature of the process and the prohibitive costs involved if changes become necessary once actual parts are being produced. By automating the design and analysis tools involved and also incorporating optimization techniques, it is possible to significantly reduce the lead time and engineering cost as well as to achieve an optimal design.

Despite the emergence of new materials and processes, bulk forming process continue to be critical to the production of parts for high-performance equipment. Many of the new alloys for critical applications are hard to work and require sophisticated workability and design methods for successful processing. Forged components are the backbone of the modern gas turbine engine; the entire rotor structure is composed of disks, spools, shafts, blisks, and seals that are machined from cast and wrought or powder materials. These components experience high centrifugal and torsional loads at elevated temperatures during service and constitute the most demanding structural applications for today's advanced aerospace materials. These components represent more than 40% of the weight and cost of a gas turbine engine and undergo a sophisticated series of manufacturing operations. Of these manufacturing processes, forging, heat treatment, and machining are critical operations that have an interacting effect on the finished engine component. Many metal alloys are currently employed in the manufacture of automotive, aerospace, and other hardware components. The high cost of manufacturing critical components can be significantly reduced with the development of mathematically and physically sound computational methodologies for process design, optimization, and control.

Forging is one of the most common processes for manufacturing rotating disks in gas turbine engines that can sustain the high stresses in an aircraft engine or power turbine. As a result, it is one of the most critical procedures in the turbine disk manufacturing process. During the forging step, a block of metal, which may be cylindrical or another shape, is transformed to a shape close to that of a disk in one or several operations. The objective during the simulation of the forging process is the determination of the die shape on the one hand and of an optimal forging process on the other that ensures proper die filling without compromising mechanical properties of the disk through the violation of stress, strain, strain rate, or temperature limits. The subsequent heat treatment process is designed to generate acceptable mechanical properties in the forged disk. Following heat treatment, the disk is machined to its final shape for assembly. A simulation of the machining process results in the final disk shape with accurate residual stresses and distortions. Each step has special process requirements involving different physical phenomena. All the conditions must be satisfied simultaneously for each step. Therefore, the disk manufacturing process is highly complicated and multidisciplinary.

Die designers in today's cost-competitive environment have to produce a design as near net shape as possible, that is, a design that requires the least amount of forging material and yet simultaneously meets all of the design specifications. For aerospace components made from high-cost superalloys, material cost is the most significant cost driver and is approximately 50% of the forging cost. Therefore, weight reduction is the most effective cost-reduction strategy.

Significant amounts of extra material are included in a forged part, which is subsequently machined to get the final disk shape. This is the preferred way to ensure good materials properties and avoid having to scrap the part due to process variability. However, decreasing the material consumption by even a modest percentage can drastically reduce the manufacturing cost of turbine disks. Previous attempts to manually achieve a balance between these two conflicting objectives have had limited success due to the complexity of the problem. With the advance of optimization, numerical simulation, and geometry modeling, it is now possible to attack this difficult problem. Optimization of the forging process results in a minimum-weight, forgeable disk that meets all constraints in terms of process parameters within specified processing windows and rules while maintaining properties. The optimization of the heat treatment process reduces residual stresses while maintaining required cooling rates through the control of surface heat transfer coefficients. Current industry approaches involve at least a few (~2–5) iterations of the process model to yield an acceptable solution. It is a challenge for a manual design process to generate a good forging design, let alone an optimal one.

Optimization Techniques

Numerous optimization techniques have been developed and demonstrated successfully on relatively simple problems with few parameters and simple analytical models. However, design optimization is not used as widely today because of many deployment difficulties, including the need to educate designers on optimization technologies and the need to modify simulation codes for optimization. Furthermore, optimization of an individual discipline does not necessarily produce an overall optimal design. To achieve the goal of evaluating many more complete design alternatives in less time, one must address the problem of how to perform optimization across disciplinary, organization, and complexity boundaries.

Circumstances frequently prevail in the design process for a team of engineers from diversified disciplines to combine their knowledge to produce an overall optimal solution. Additionally, the requirement exists to combine an understanding of process design with the ability to use simulation codes correctly and to overcome the tedious process of manual iteration to obtain an optimal design. Overcoming these obstacles and freeing engineers to do other engineering duties is now a priority task for the parts supplier industries. Proven software optimization tools, such as iSIGHT (Engineous Software, Inc.) described in this chapter, can drive simulation codes from multiple disciplines toward achieving superior designs in less time. The iSIGHT system is a generic software shell that can be applied to the design of many different products and processes. Also, iSIGHT is a comprehensive application with a user-friendly graphical user interface (GUI) that allows end users to easily use it.

This chapter addresses one of the important aspects of the turbine disk forging process—the design of die geometry to achieve near-net-shape forging. Using numerical optimization tools, the capability to achieve significant material savings on top of manually optimized forging designs for jet engine disks has been demonstrated. The design process is formulated as a parametric geometry and high-fidelity analysis-based shape optimization problem. The objective is to minimize the forging weight by changing the die geometry parameters. The forging constraints include producibility, materials properties, geometry, and forging time conditions. The forging weight and cost are minimized within prescribed processing windows and forging rules including bounds on strain, temperature, strain rate, press capacity, dwell time, sonic coverage, fillet radius, draft angles, die fill, and so forth. A generic and noninvasive procedure was developed to integrate CAD and CAE tools for forging design optimization. A fully automated analysis and optimization system that works in a heterogeneous and networked computing environment was built on the top of three commercial software packages: DEFORM for simulating metalforming process, a commercial CAD tool Unigraphics for defining and manipulating geometry, and iSIGHT for software integration and optimization. The analysis involves thermal-mechanical coupling, nonlinear material behavior, friction modeling, very large deformation, time-dependency, and multiple operations. The iSIGHT system replaces the manual iterative portion of the traditional design process with an automated, computer-controlled procedure. The software helps link all of the relevant simulation tools, then automatically changes the design parameters, runs the analysis codes, assesses the output against the target objectives, and changes the design variables based on instructions from an optimization algorithm chosen for the specific problem.

Through a series of iterations among the different software tools, a user-defined optimal scenario is established. The whole process iterates without human intervention until an optimal solution is achieved. The optimization system has been applied to selected disk-design problems and achieved significant weight reductions. The total forging weight was reduced by 5 to 10% over manual design optimization. If the saving in weight is expressed in terms of the excess material added over the smallest geometrically feasible shape to ensure forgeability and satisfaction of other constraints, the weight saving is 50 to 80%. Clearly, this is a significant material savings over the manually optimized forging designs for jet engine disks. Die designers are now able to explore, with the help of a fully automated software system, a bigger design space and find a feasible design of the least forging weight. In addition, a shape optimized with iSIGHT can have better properties than the manually designed shape, thus allowing for additional weight reductions for equivalent mechanical performance as the manual design.

By utilizing this technology, the quality of turbine disks is improved, the design process variation is minimized, and the manufacturing cost is reduced. Several case histories and different ways of using the optimization tools show the benefits of using MDO for solving product and process design problems by exploring more design options and parameter trade-offs in a given time than is possible by a human designer.

This chapter is not intended to serve as a review of optimization methods. There are several such review articles, for example, Ref 1. The focus here is on the application of MDO techniques to solve real-life manufacturing problems. Previous work related to the topic can be found in the literature. Kodiyalam, Kumar, and Finnigan (Ref 2) proposed a constructive solid geometry approach to three-dimensional structural shape optimization. Hardee and coworkers (Ref 3) demonstrated how to use ProEngineer as the CAD/CAE system to perform structural design optimization based on continuous sensitivity analysis. Fourment and Chenot (Ref 4, 5) presented their work on a metalforming problem, where they optimized the initial shape of the part and the shape of the preform tool in a two-step forging operation. Wright and Grandhi (Ref 6) as well as Gao and Grandhi (Ref 7) published their studies on forging preform shape optimization. All the aforementioned research related to forging optimization has utilized internally developed CAD/CAE systems. Thus it is possible to develop analytical design sensitivities and solve the optimization problem efficiently. This chapter presents a method of integrating commercial CAD/CAE systems and addressing the interprocess communication problem. The approach is generic and noninvasive and readily lends itself to more complicated applications. The method is demonstrated with several examples.

Metalforming Process Optimization. Optimization is the process of searching for the best solution, which includes finding the best value for the objective function while satisfying the design constraints. In metalforming, the objectives of process optimization may include one or more of the following:

- Forging cost
- Forging weight
- Press load
- Total number of operations
- Die cost
- Lead time

The primary objective here is to use optimization methods to design the necessary sequences of forming processes, select appropriate dies and preforms for each process, and control/design process parameters for each process such that, for a given raw material with a given initial geometry, one can obtain a final product with the desired microstructure and shape. Desired objectives may include one or more of the following criteria and may represent, for example, the deviation of the resulting material state of the final product from a desired state for the process defined by design parameters:

- Uniform deformation in the final product
- Desired mechanical properties in the final product
- Desired microstructure in the final product
- Required shape of the final product
- Residual-stress distribution to minimize distortion during manufacturing and/or during service
- Minimum deformation and wear of the die
- Minimum energy or load spent in deforming the material
- Minimum number of forging steps
- No defects such as forging laps or die nonfill

These objectives can be satisfied by controlling design variables that can represent design of the preform: design of the material state (microstructure) in the initial billet, appropriate selection of the process parameters (ram speed and pressure history, operating temperature, etc.), the initial shape of the workpiece, shape of the dies, thermal boundary conditions, and so forth. In most industrial forming applications, the desired objectives indicated previously are seldom simple enough to be achieved in a single forming operation. As a result, intermediate deformation or performing steps are used to efficiently transform the initial geometry into a final shape and/or with desired materials properties. Thermal processing is also used in between deformation stages to control the microstructure or product quality. Various examples of optimization formulations are described in this chapter.

Challenges to Forging Optimization. Optimization of both the forging and the heat treatment process individually has been successful, but the complete MDO scenario still faces a number of obstacles. Parametric CAD tools are not as robust for complicated geometries, especially 3D, as would be necessary in an automatic optimization environment. The same applies to the interface between and integration of CAD and CAE tools. Automatic 3D meshing (hexahedrons) capability is still somewhat limited. Computational resources constitute another bottleneck; for industrial applications, fast turnaround (generally less than 8 h) is a requirement. For a complicated structural optimization problem, the effort of getting the objective function, constraints, and corresponding sensitivities is computationally expensive. Therefore, a variety of approximation methodologies have been developed for reducing the computational cost. At the time of publication, the industry had been moving toward adopting off-the-shelf commercial software because of the high cost associated with software development and maintenance. The drawback of this approach is that the software capability becomes one of the constraints. For instance, analytical sensitivities are generally not available with commercial software, and thus other methods such as finite difference for gradient evaluations have to be employed for the sensitivity analysis. Software licensing agreements (multiple versus single user/computer licensing costs) can prevent parallelism for gradient evaluations. Even when it is possible to use analytical sensitivities, it is complicated and engineering labor intensive to develop these for general shape optimization problems. Barriers between various engineering and manufacturing departments are also a factor and are often difficult to break down.

Despite all these obstacles, though, progress toward a comprehensive optimization is apparent.

Collaborative Optimization Environment

The objective is to develop a flexible design environment that can be used to design complex products and processes to achieve affordable globally optimized designs. This is achieved through the creation of a collaborative optimization environment (COE). The COE provides a generic framework for design automation, system integration, and design optimization. There is collaboration between optimization algorithms, between disciplines, between computing systems, and between organizations. The key concepts are:

- Address multidisciplinary and multilevel optimization
- Interdigitation—innovative approach to leveraging and combining multiple optimization techniques (Ref 8–11) such that the user has a suite of optimization tools available, including gradient based and heuristic search techniques, genetic (adaptive) algorithms, and simulated annealing, which can be used in any combination during the optimization process
- Noninvasive coupling of commercially available CAD and CAE tools
- Multidisciplinary design optimization language (MDOL) to enable formulation and solution of complex problems

System Components

The forging and heat treatment optimization tools discussed here in detail involve the integration of commercially available software to reduce development and maintenance cost:

- iSIGHT: optimization and program control
- DEFORM: CAE code for simulating manufacturing process
- Unigraphics: CAD code for geometry manipulation

The three software tools are briefly described in the sections that follow.

iSIGHT for Optimization and Program Control. The commercial software package iSIGHT (Ref 9) is used as the integration and optimization framework. The system provides a holistic approach to optimization by embedding the concept of interdigitation (Ref 8–11), which was originally developed at GE Corporate Research and Development Center as part of the Engineous system (Ref 10). Users have access to a suite of optimization techniques including gradient-based, exploratory, and AI/expert system based algorithms within iSIGHT. These techniques may be used alone or collectively to efficiently solve complicated optimization problems. Another characteristic of iSIGHT is its ability to nonintrusively integrate external programs. iSIGHT provides excellent utilities to manage data flow between simulation codes, automate program execution, and allow external codes to reside on their respective platforms irrespective of where iSIGHT is installed—an important point concerning leasing and maintenance cost for those software tools that may be licensed only on one specific workstation.

DEFORM: CAE Code for Simulating Manufacturing Process. Advanced-process simulation tools are becoming more and more available for all stages of the disk design and manufacturing process. Simulation tools such as DEFORM (Ref 12), ABAQUS (Ref 13), or MARC (Ref 14) can accurately predict the mechanical behavior and properties during the manufacturing process. Therefore, these tools have become the state-of-the-art and are widely used. In combination with numerical optimization techniques, these tools offer the opportunity to improve individual steps in the overall process (Ref 15). DEFORM was chosen as the tool to be applied in forging and heat treatment optimization procedures described in this chapter. The finite-element-based package DEFORM has been demonstrated to be able to capture the physics of forging processes and predict the macrobehavior of the material. It is, therefore, widely accepted and used in the forging industry. It provides all the forging data, such as press load, strain, strain rate, stress, temperature, and so forth, which may constitute constraints and objectives for the optimization.

Unigraphics: CAD Code for Geometry Manipulation. General-purpose commercial CAD systems have evolved considerably and are becoming mature and widely used in the manufacturing industry. One of the main attractions of these systems is that they house a rich collection of user functions or application interfaces (APIs) for geometry queries, manipulation, display, checking, and so forth. These utilities make the difficult task of geometry handling easier. For the present application, unigraphics (UG) (Ref 16) has been chosen as the CAD system, but the method of integration outlined here is generic and applicable to other CAD systems.

A number of modules were developed to support design optimization on top of the UG Open API library (Ref 16), including:

- Graphically select feature and sketch parameters as design variables
- Update geometry model from an external process
- Monitor shape change
- Evaluate geometry properties such as volume, which is the objective function for disk forging optimization
- Automatically convert UG geometry to the format accepted by the CAE code

There are some geometry manipulation features built into DEFORM. However, CAD systems come with a number of standard features such as checking for geometry errors, invalid geometries, use of parametric modeling, feature-based modeling, creating offset geometries automatically (forging covers over ship or heat treat shapes). Building these features into DEFORM involve much more effort than using an off-the-shelf CAD package. Therefore, the decision here was to demonstrate the optimization system using Unigraphics.

System Integration and Operation

System Integration. A general procedure is introduced to generically and noninvasively integrate commercial CAD and CAE systems to support geometry and detailed analysis-based optimization. The integration and optimization environment iSIGHT has been adopted as the framework of the system. A client-server architecture is established to allow iSIGHT to drive the CAD system for geometry manipulation. The CAE system is wrapped and called upon within iSIGHT to perform detailed simulation. Data passing between the optimization environment and the CAD and CAE tools are handled through interprocess communication. The integrated system involving three software tools is shown in Fig. 2

iSIGHT uses DEFORM as its simulation code that supplies the objective and constraint information. iSIGHT controls a number of DEFORM parameters, mainly the "design variables." Because DEFORM does not directly calculate constraint and objective information for the optimization procedure, a specific DEFORM postprocessing function has been developed for that purpose. DEFORM is executed by iSIGHT fully automatically each time iSIGHT needs the solution for a new data point. The procedure is implemented in iSIGHT in the form of three separate "tasks." The first task, preprocessing, executes the DEFORM preprocessor, which reads the respective DEFORM keyword file and creates the DEFORM analysis database. The preprocessor also starts the DEFORM simulation run, which is executed in batch mode. Once this has been accomplished, the preprocessing task exits and passes the process ID of the DEFORM batch job back to the main task. The second task, run control, then checks on the execution of this batch job in 30 s intervals. Once the batch process has terminated, the run-control task exits, and the third task, postprocessing, is started. It executes the modified DEFORM postprocessor, which determines the global response quantities that can be used as constraints and objectives for the optimization. The DEFORM postprocessor extracts the relevant outputs from the DEFORM database.

System Application for Geometry-Based Optimization. This section describes the use of the iSIGHT system for geometry-based optimization, followed by a description of its GUI. To use the system for geometry-based optimization, the user first opens a UG interactive session, loads relevant part files, graphically selects geometry parameters that are design variables, and chooses objects whose volumes and other geometric properties are desired. The geometry is defined in parametric form to enable easy manipulation and updating. After completing the problem specification, the user may start the iSIGHT run. The routines within iSIGHT will automatically launch a noninteractive UG session and establish the interprocess connection, which is based on Expect (Ref 17), a Tcl extension. iSIGHT changes the design variable values and writes them into the file storing the selected geometry parameters. The noninteractive UG process reads the files and updates the geometry models. It is also possible to monitor the geometry change in real time through an interactive UG session.

The system has been applied to the forging shape optimization of turbine disks, where the weight of the disks is minimized under maximum press load and geometric constraints. In the present study, the forging process is modeled as a time-dependent, plastic-deforming, and either isothermal or nonisothermal process. Since the forging simulation is conducted in an optimization environment, some of the process and geometry parameters are modified in each DEFORM run. Therefore, it is necessary to regenerate the mesh and redefine the boundary conditions. Furthermore, the user needs to postprocess the analysis results and extract information on the optimization objective and constraint functions. Several modules based on Expect have been developed that drive DEFORM to accomplish the following tasks:

- Import geometry and then regenerate die and billet meshes
- Create appropriate boundary conditions
- Start DEFORM simulation in batch mode
- Monitor DEFORM runs
- Postprocess simulation results to extract maximum press load, strain, temperature and so forth

Each of these modules acts like a separate executable, or "simcode" in iSIGHT terminology. iSIGHT executes these "simcodes" in a predefined sequence, including potential looping and branching.

A summary of the tasks to perform forging optimization and the typical engineering hours and skill sets needed are summarized, providing a checklist for implementing MDO techniques in an industrial environment:

1. Parameterize die geometry and select design variables.
   - 8 to 16 h
   - Experience with UG parametric modeling
   - Forging expertise to identify design parameters
2. Create a standard DEFORM keyword file as template.
   - 4 to 8 h
   - Familiarity with DEFORM
3. Formulate optimization problem, and define forging operation sequence and control parameters.
   - 8 to 16 h
   - Knowledge of iSIGHT

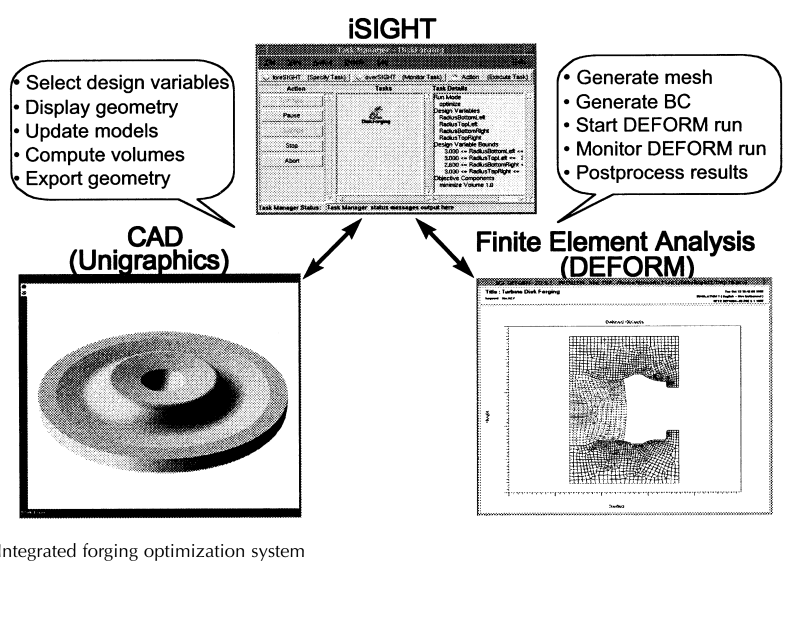


Fig. 2 Integrated forging optimization system

4. Start forging optimization run.
   - 25 to 50 h (computer elapse/cycle time) on typical HP or SGI workstations. The computer times refer to year 2000 state-of-the-art and are continually decreasing with improvements in computer technology
   - Knowledge of iSIGHT
5. Postprocess optimization results and review.
   - 8 to 16 h
   - Knowledge of iSIGHT and DEFORM

It is thus possible to complete one optimization problem with about 2 to 5 days of engineering effort and within an elapse time of one week. As computer hardware and software technology continually improve, this time is expected to drop to less than one day within a few years.

Graphical User Interface (GUI). A GUI has been developed that helps the user set up the specific optimization case as shown in Fig. 3. Via the iSIGHT GUI, the user has complete control over the simulation of the forging or heat treatment process and the selection of the design variables. In principle, any number of design variables can be chosen, but both in terms of run time for the optimization procedure and in terms of actual process and geometry control, about 15 can be considered the upper limit.

All the parameters describing the process and optimization sequence are set in the description file The GUI allows the creation of a fresh description file, or as is more often the case, the user may want to modify an existing file to save time. When the file is saved, some error checking is performed. Once the forging or heat treatment application—which resides in its description file—is loaded into iSIGHT, the procedure is ready to run. Inside the foreSIGHT module of iSIGHT, the user has the option to change the control of the problem, switching between "single run" and "optimization," to change any of the input variables (especially the design variables), to activate or deactivate design variables, to change objectives and constraints, or to modify the optimization plan. The default for the forging and heat treatment optimization is "ADS—Modified Method of Feasible Directions." This has mostly produced satisfactory results so far, and it is not recommended to replace it with another technique unless the user is familiar with the various optimization techniques iSIGHT offers. The technique "ADS—Sequential Linear Programming" has also been successfully applied. The user has the option to access the control parameters for each optimization technique, which are explained in detail in the respective manuals of the individual optimization packages inside iSIGHT, for example, ADS (Ref 18). Inside the overSIGHT module of iSIGHT the user has the option to attach tables and/or graphs to the iSIGHT run to permit interactive tracking of the changes in the design variables, constraints, and objectives on the screen. The use of the corresponding buttons in the task manager window is pretty straightforward.

The steps performed during one iteration of forging optimization are:

1. The dimensions of the top and bottom dies are sent to the UG CAD software from iSIGHT, and the CAD geometry is updated. The size of the billet is automatically updated to correspond to the necessary volume to fill the die.
2. The 3D axisymmetric CAD geometry is cut to produce the two-dimensional (2D) cross section that is used by DEFORM 2D. This 2D geometry is then incorporated into the DEFORM master file.
3. iSIGHT then executes DEFORM 2D to perform the forging simulation on this design.
4. Postprocessing information is extracted from the DEFORM 2D database, that assesses the quality of the forged part in terms of user-defined constraints. This data is retrieved by iSIGHT.
5. The output of the forging process is compared against requirements that have been specified by the engineer with iSIGHT. This gives a numerical assessment of the current design.
6. Now, iSIGHT again changes the dimensions of the top and bottom dies. The iSIGHT software has numerous exploration methods that can be used to intelligently make the decisions on which die dimensions to change, and by how much, to improve the forging process.

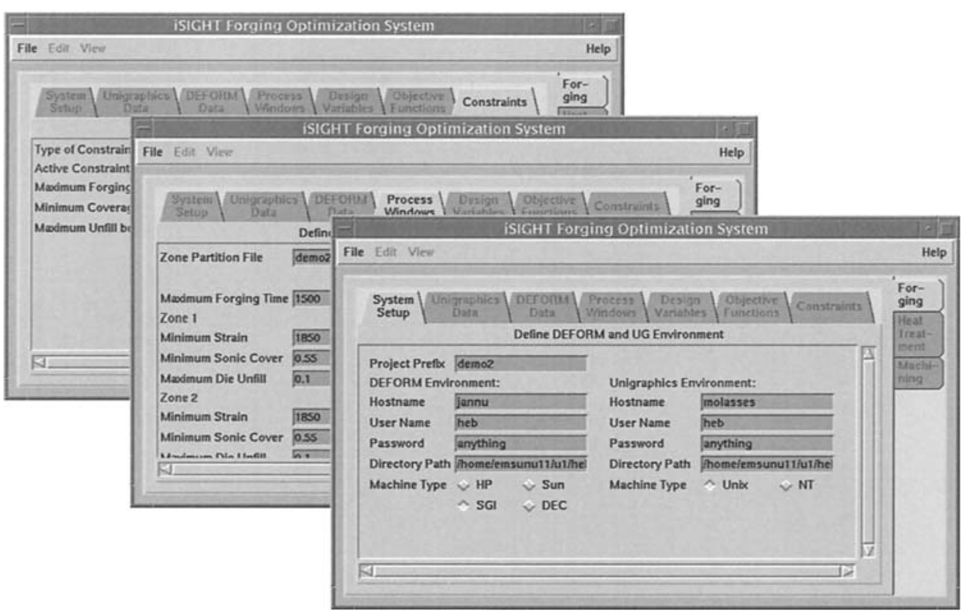


Fig. 3 Graphical user interface of iSIGHT optimization system. The tabs on the right are used to select the process (forging or heat treatment), and the tabs at the top are used to define the process and solution-control parameters.

Formulation of Forging Shape Optimization

This section describes some methods and examples for shape optimization of forgings. Real disk shapes have been optimized, but due to the sensitive nature of that information, only simplified examples that demonstrate the principles and techniques can be presented here. The various optimization features of the methods using iSIGHT include:

- Utilized feature-based modeling in Unigraphics
- Customized UG utilities including geometry error checking
- Developed modules on the DEFORM side to help with data processing
- Started from most aggressive shape (geometrically minimum weight) and added material to satisfy constraints
- Used modified feasible direction method with finite-difference gradients
- Conducted computations in a heterogeneous environment over the network
- Modeled forging process as time-dependent, plastic-deforming, and isothermal/noisothermal, rigid dies, and 2D axisymmetric
- Simulated all stages of forging process including the operations of:
  a. Transfer from furnace to dies
  b. Resting on dies prior to deformation
  c. Forge under strain-rate control to maximum tonnage
  d. Forge under load control
  e. Dwell at maximum tonnage
- Incorporated:
  a. Equipment constraints: press capacity, transfer, and dwell times
  b. Material-related constraints: strain, strain rate, and temperature-processing windows
  c. Geometrical constraints: forging rules defining the sonic and final machine part

coverage, minimum draft angle as well as corner and fillet radii

d. Invalid geometry and topological checks
e. Die-nonfill check and lap-formation checks

In summary, the optimization procedure ties the complex, multidisciplinary aspects of good process design together. This shows the link between the metalworking equipment, the material system (tooling, lubricant, and workpiece), and the equipment controls. MDO techniques are shown to optimize the process and the product performance. The overall objectives and benefits for the example of turbine-disk forgings include:

- Forging cost reduction is key to reducing cost of aircraft engines
- Use of optimization tools is a key enabler in achieving this goal
- Developed an integrated system for forging die-shape optimization
- Formulated forging shape optimization problem for turbine disk forging
- Developed a generic procedure to noninvasively integrate external CAD and CAE systems for geometry and detailed analysis-based optimization
- Formulated ways to handle the optimization and constraints for a highly nonlinear problem with a very irregular response surface
- Applied the system to both isothermal and nonisothermal forging shape optimization of turbine disks
- Demonstrated feasibility and success of forging and heat treatment process optimization

Formulation of Objectives and Constraints

Objective and Design Variables. The optimization problem is formulated to minimize the weight of the forging subject to several process constraints that are aligned with standard forging practices. The shape of the top and bottom dies determine the weight of the forging and are defined by a set of geometrical parameters. The design variable set $\{\pi_i, i = 1, \cdots, N\}$ is a subset of geometry parameters that defines the die shape, where N is the number of design variables. The forging volume, V, which is a smooth function of the shape design variables, is the objective function.

Optimization Constraints. There are five types of constraints.

Equipment Constraints. The forging press must have the capacity to forge the part. This is a producibility condition. The load P required to forge the part should be less than the maximum tonnage available on the press (P_a) on which it is to be forged:

$$P \leq P_a \quad \text{(Eq 1)}$$

Alternatively, this can be formulated as an energy constraint when hammer or screw presses are used. Additional equipment constraints can be imposed on transfer and dwell times.

Material Constraints. The forged part must possess the right materials properties. Close correlation has been observed between materials properties and forging and heat treat processing conditions. Materials properties are dependent on the thermomechanical processing path characterized by equivalent strain rate $\dot{\varepsilon}$, equivalent strain ε, and temperature T. When working with difficult-to-process alloys, as is common with aircraft engine materials, the thermomechanical processing path must be tightly controlled in order to achieve acceptable microstructures and mechanical properties in the final product. The forging can be partitioned into various zones with each zone having its own requirement on materials properties. For example, the bore of a disk may be strength limited whereas the rim may be creep limited, requiring different zones of the forging to have the appropriate thermomechanical processing path to obtain the needed mechanical properties. These thermomechanical rules or processing windows are material dependent. They are represented in the optimization framework as max/min limits on strain, strain-rate, temperature, adiabatic heating, and postforge cooling rate and define the processing window for the material to meet microstructural and property requirements.

$$\dot{\varepsilon} \leq \dot{\varepsilon}_a$$

$$\varepsilon \geq \varepsilon_a \quad \text{(Eq 2)}$$

$$T_a^{lb} \leq T \leq T_a^{ub} \quad \text{(Eq 3)}$$

Geometrical Constraints. Based on experience, die designers have developed several geometry-related forging rules. Geometric constraints are defined to ensure adequate cover over the finished shape to ensure the forging will clean up in spite of process variability (process parameters, die stack concentricity, die wear/deflection, etc.) and to account for all subsequent manufacturing operations such as heat treatment, machining, inspection, and joining. These are represented as bounds on design variables, protective cover h over the finished part shape for sonic inspection, the extent of die-nonfill δ, any lap formation, minimum fillet and corner radii, minimum forging thickness, and minimum draft angles. These geometrical constraints are dependent on the material being forged and ensure manufacturability and inspectability of the forging. Proper bounds on geometrical design variables avoid the creation of invalid geometries and prevent the occurrence of topological change. The utilities in the UG CAD system are used for invalid geometry and topological checks.

$$(\pi_i)_a^{lb} \leq \pi_i \leq (\pi_i)_a^{ub} \qquad (i = 1, \cdots, N) \quad \text{(Eq 4)}$$

$$h \geq h_a$$

$$\delta \leq \delta_a \quad \text{(Eq 5)}$$

Die stress cannot exceed a critical threshold value. Otherwise, a die can be damaged due to excessive stress, which leads to costly repair and even complete replacement of the die. The requirement on die life and wear is represented as:

$$\sigma \leq \sigma_a \quad \text{(Eq 6)}$$

The time spent in forging a part must be minimal—that is, as quick as possible—which is an important productivity consideration. The forging time limitation is represented as:

$$t \leq t_a \quad \text{(Eq 7)}$$

In Eq 1 through 7 superscripts lb and ub stand for lower and upper bounds, respectively, and the subscript a represents a prescribed allowable value.

Many of these design variables are dependent on each other. These dependencies are impossible to formulate in a closed form. This considerably increases the complexity of the optimization process.

Implementation of Constraints. Here are some additional comments on the implementation of the various constraints within the optimization framework:

- Press load and strain-rate conditions are implemented in the forging software, that is, DEFORM. The strain rate is controlled by adjusting the forging die speed as the calculation proceeds so that Eq 2 is implicitly satisfied. The maximum strain-rate constraint is enforced throughout the forging process.
- The strain and temperature constraints are enforced within the sonic geometry, which is the disk shape for sonic inspection, and not within the entire forging volume. There is no need to enforce the constraints in the material that is machined away.
- The die-nonfill check is also used to detect lap formation and ensure a successful forging process. The objective is to completely fill the finishing dies during the last operation. A numerical measure of this criterion is obtained by comparing the outline of the desired part and that of the actual forged part.
- Some of the constraints may not be needed for certain forging processes and/or materials. For instance, there is no temperature constraint for isothermal forging. In such cases, the appropriate constraints are rendered inactive.

The strain, temperature, sonic coverage, die-nonfill, and stress in Eq 2 to 6 may be functions of space and time. To simplify the problem, these conditions are enforced only at the end of the forging operation and, therefore, are functions of space and not of time. Despite this simplification, Eq 2 to 6 still need to be satisfied at every node or element of a finite-element model. A further complication is that some type of automatic unstructured mesh generation has to be employed in the analysis, due to the large deformation involved in the forging process. The introduction of unstructured meshes causes the number of nodes and elements as well as node and element locations to be different from one design to another. Therefore, it is impractical and also impossible to impose constraints of a field function form.

One solution is to introduce minimum and maximum types of constraint functions. They are global functions defined for the entire model and thus can be uniquely computed even when an unstructured mesh is used. The global strain, temperature, sonic coverage, and die-nonfill are defined as:

$$
\begin{aligned}
\bar{\varepsilon} &= \min_{\mathbf{x}\in\Omega} \varepsilon \\
\bar{T}_{\min} &= \min_{\mathbf{x}\in\Omega} T \\
\bar{T}_{\max} &= \max_{\mathbf{x}\in\Omega} T \\
\bar{h} &= \min_{\mathbf{x}\in\Gamma} h \\
\bar{\delta} &= \max_{\mathbf{x}\in\Gamma} \delta \\
\bar{\sigma} &= \max_{\mathbf{x}\in\Omega'} \sigma
\end{aligned}
\qquad \text{(Eq 8)}
$$

where Ω and Γ are the workpiece and its boundary, respectively, Ω' represents the dies, and $\mathbf{x}$ is the spatial coordinates. Thus, the strain, temperature, sonic coverage, die-nonfill, and stress constraints can be rewritten as:

$$
\begin{aligned}
&\bar{\varepsilon} \geq \varepsilon_a \\
&\bar{T}_{\min} \geq T_a^{lb} \\
&\bar{T}_{\max} \leq T_a^{ub} \\
&\bar{h} \geq h_a \\
&\bar{\delta} \leq \delta_a \\
&\bar{\sigma} \leq \sigma_a
\end{aligned}
\qquad \text{(Eq 9)}
$$

One problem with this approach is that the constraint functions may not be smooth functions of the design variables, and this can cause difficulties with the gradient-based algorithms. Another problem is that this formulation depends on point (nodal or elemental) information that is dependent on the finite-element mesh used and is not always a unique quantity. To resolve this issue, the following averaging schemes are proposed to compute the global strain, temperature, sonic coverage, and die-nonfill:

$$\bar{\varepsilon} = \lim_{\lambda\to 0}\left[\int_{\Omega} \varepsilon\, g\left(\frac{\varepsilon}{\varepsilon_a}\right)dV + \lambda\varepsilon_a\right] \Big/ \left[\int_{\Omega} g\left(\frac{\varepsilon}{\varepsilon_a}\right)dV + \lambda\right] \qquad \text{(Eq 10)}$$

$$\bar{T}_{\min} = \lim_{\lambda\to 0}\left[\int_{\Omega} T\, g\left(\frac{T}{T_a^{lb}}\right)dV + \lambda T_a^{lb}\right] \Big/ \left[\int_{\Omega} g\left(\frac{T}{T_a^{lb}}\right)dV + \lambda\right] \qquad \text{(Eq 11)}$$

$$\bar{T}_{\max} = \lim_{\lambda\to 0}\left[\int_{\Omega} T\, g\left(\frac{T_a^{ub}}{T}\right)dV + \lambda T_a^{ub}\right] \Big/ \left[\int_{\Omega} g\left(\frac{T_a^{ub}}{T}\right)dV + \lambda\right] \qquad \text{(Eq 12)}$$

$$\bar{h} = \lim_{\lambda\to 0}\left[\int_{\Gamma} h\, g\left(\frac{h}{h_a}\right)d\Gamma + \lambda h_a\right] \Big/ \left[\int_{\Gamma} g\left(\frac{h}{h_a}\right)d\Gamma + \lambda\right] \qquad \text{(Eq 13)}$$

$$\bar{\delta} = \lim_{\lambda\to 0}\left[\int_{\Gamma} \delta\, g\left(\frac{\delta_a}{\delta}\right)d\Gamma + \lambda\delta_a\right] \Big/ \left[\int_{\Gamma} g\left(\frac{\delta_a}{\delta}\right)d\Gamma + \lambda\right] \qquad \text{(Eq 14)}$$

where $g(x)$ may be defined in different ways and λ is a nonnegative parameter. For example:

$$g(x) = \begin{cases} 1 & \text{if } x \leq 1 \\ 0 & \text{if } x > 1 \end{cases} \qquad \text{(Eq 15)}$$

and

$$g(x) = \begin{cases} (1-x)^2 & \text{if } x \leq 1 \\ 0 & \text{if } x > 1 \end{cases} \qquad \text{(Eq 16)}$$

Equation 15 is a step function that is discontinuous and tries to weight strain, temperature, sonic cover, and die-nonfill violation equally. The averaged quantities computed using this function are not going to be smooth, which can again be problematic for the gradient-based optimization algorithms. Equation 16 is able to overcome these difficulties. It is a smooth function and tends to weight more on the more severely violated portion of the strain, temperature, sonic cover, and die-nonfill.

By carefully examining the definition of these global functions, one may conclude that the global functions have the following properties: Equations 3 and 5 are equivalent to Eq 9, where the global functions $\bar{\varepsilon}$, $\bar{T}_{\min}$, $\bar{T}_{\max}$, $\bar{h}$, and $\bar{\delta}$ may be computed using min/max formulation (Eq 8) or the averaged formulation (Eq 10–14) with either step function (Eq 15) or smooth function (Eq 16).

The following relations hold when constraints are violated:

$$
\begin{aligned}
&(\bar{\varepsilon})_{\text{min/max}} \leq (\bar{\varepsilon})_{\text{average1}} \leq (\bar{\varepsilon})_{\text{average2}} \\
&(\bar{T}_{\min})_{\text{min/max}} \leq (\bar{T}_{\min})_{\text{average1}} \leq (\bar{T}_{\min})_{\text{average2}} \\
&(\bar{T}_{\max})_{\text{min/max}} \geq (\bar{T}_{\max})_{\text{average1}} \geq (\bar{T}_{\max})_{\text{average2}} \\
&(\bar{h})_{\text{min/max}} \leq (\bar{h})_{\text{average1}} \leq (\bar{h})_{\text{average2}} \\
&(\bar{\delta})_{\text{min/max}} \geq (\bar{\delta})_{\text{average1}} \geq (\bar{\delta})_{\text{average2}}
\end{aligned}
\qquad \text{(Eq 17)}
$$

where $(\ )_{\text{min/max}}$ represents min/max formulation (Eq 6), while $(\ )_{\text{average1}}$ and $(\ )_{\text{average2}}$ are averaged formulation (Eq 10–14) computed using step function (Eq 15) and smooth function (Eq 16), respectively.

Numerical Examples

Consider the forging shape optimization of a generic turbine disk. A cylindrical billet is forged into a disk of the shape shown in Fig. 4 (top). The die geometry is captured in a CAD Unigraphics parametric model (Fig. 4, middle), and the problem is to design the die shape so that the forging weight is minimized. It is required that the maximum press load not exceed a prescribed value and that the forged part satisfy the condition of minimum cover over the sonic shape. Because the disk is axisymmetric, only its radial-axial cross section needs to be considered (Fig. 4, bottom). Several fillet radii R_1 to R_6 have been chosen as design variables. Both invalid geometry and intrusion into the minimum coverage over the sonic shape can be prevented by putting simple bounds on the design variables. It should be noted that simple bounds may not be sufficient to guarantee geometry validity in a more general situation. They work here because there is no coupling among the selected design variables.

Thus, the optimization problem is formulated as:

Minimize: V

Subject to:

$$R_{i\text{lb}} \leq R_i \leq R_{i\text{ub}} \ (i = 1, \ldots, 6)$$

$$P \leq P_{\text{ub}}$$

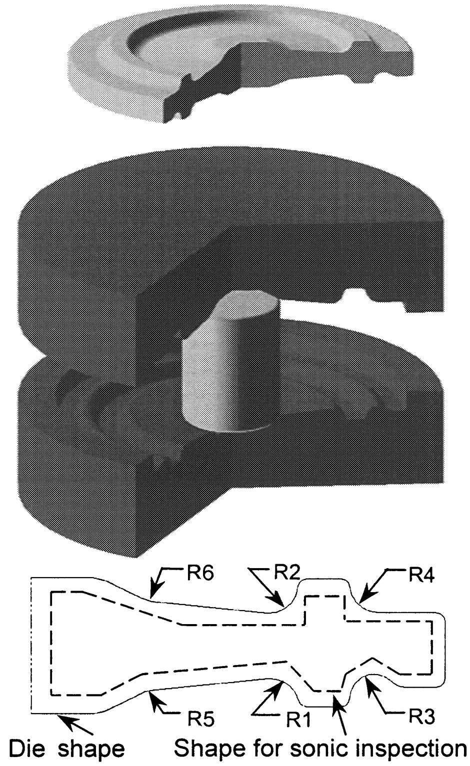


Fig. 4 Shape of generic turbine disk (top) and a CAD parametric model of a cylindrical billet and die shape (middle). Several fillet radii R_1 to R_6 are chosen as design variables. Zones can also be defined for regions with different constraints and property requirements (e.g., see Fig. 12).

where V is the volume of the workpiece, P and P_{ub} are the maximum press load and its upper bound, respectively, and R_{ilb} and R_{iub} are the given lower and upper bounds of the fillet radii, respectively. The most aggressive shape, which corresponds to the lowest volume V, has been chosen as the initial design. It is relatively easy to get this shape from the specified disk design by adding a minimum cover. This shape will have small corner radii (all the design variables at their lower bounds) making it difficult for the metal to flow around such sharp corners. As a result, the press load constraint may be violated for this design, and thus the fillet radii R_i, $i = 1, \ldots, 6$, have to be increased, which results in a larger volume. Subject to the press load constraint P_{ub}, the optimizer should choose the optimal values of design variables R_i. As is pointed out in the previous section, iSIGHT provides a suite of optimization algorithms from ADS (Ref 18) and other packages. The modified method of feasible directions is employed in this study. Since analytical design sensitivities are not available, the gradient information has to be obtained through finite differences.

The numerical computation involved three different workstations: a SUN Ultra-1 running iSIGHT v3.0, a HP9000 running Unigraphics v11.1, and a 16-processor Silicon Graphics Origin 2000 running DEFORM. This demonstrates that the integrated system can run in a heterogeneous environment over the network. Both isothermal and nonisothermal forging optimization are presented for this shape. More details of the optimization results are presented in Ref 19 and 20.

Isothermal Forging Optimization. In one example, a time-dependent, plastic-deforming, isothermal, closed-die forging process is considered. The top and bottom dies are assumed to be rigid. The maximum load P normally occurs at the end of the forging stroke as the dies fill out and the material starts to move into the flash region. The load changes rapidly with the stroke at this stage of the process. Therefore, it is difficult to accurately compare the loads at the end of the stroke from different die designs due to the inherent noise in the load predictions. For this reason, the load P is arbitrarily set to be the stroke-averaged load in between 98 and 99% of the final stroke. This enables a truer comparison of the load requirements for the various designs rather than considering the load at a unique value of the forging stroke. A good estimate on the real maximum press load may be obtained by multiplying P with a correction factor. Automatic mesh regeneration is enabled to accommodate the large deformation that is inevitable in the forging process. Four design variables R_1 to R_4 are used in this application. Due to repeated remeshing during the forging simulation, nonsmoothness is introduced in the finite-element solution. Therefore, a 10% perturbation on the design variables has to be used during sensitivity analysis using finite differences to smooth out the design space. Although the design sensitivities so calculated may not be very accurate locally, they provide the optimizer with the right search directions in a global perspective.

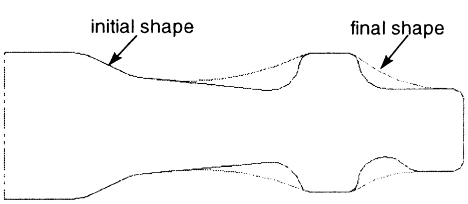


Fig. 5 Initial and final shapes for isothermal forging optimization

The initial and final shapes of the disk are shown in Fig. 5. The objective (volume) and constraint (press-load) function values versus simulation runs are shown in Fig. 6. All the data except the number of forging simulations have been normalized. The results suggest that the optimization is close to convergence after 15 simulation runs. The smaller fluctuations are the result of finite-difference perturbation, while the larger ones are due to line search of the optimizer. Since the abscissa shows the number of simulation runs as opposed to the number of optimization iterations, the results of both finite differencing and line search have been included. The upper bound of the press load $P_{ub} = 77.8$ is shown as a dashed line in Fig. 6 (right). It is apparent that the forging press load far exceeds this limit initially. As a result of the optimization, the normalized press load drops from 96.9 for the initial design to 77.8, which is within the available press capacity, a 19.7% reduction. The normalized volume, however, has increased by 12.4% from 82.4 of the initial minimum-weight shape to 92.6 of the final optimized shape. Intuitively, the result makes sense: the press load is reduced by designing a set of dies with smoother boundary shapes to facilitate streamlined metal flow, but the cost is a larger volume of the forged part, that is, more material is used. The optimization and simulation took about 25 h on a single-processor HP3600 workstation.

Nonisothermal Forging Optimization. In the next application, the single-step isothermal forging process of the previous example is replaced with a multistep, in particular, three-step nonisothermal one. Many of the forging processes are nonisothermal. The temperature of the workpiece may change significantly due to heat exchange with the environment and the dies. The temperature change will, in turn, alter materials properties of the workpiece and elevate the press load requirement. Multistep forging whereby both the workpiece and dies are reheated in between forging steps are introduced to compensate for the heat loss and to obtain workpiece and die temperatures close to the desired values. Here it is assumed that 75% of the total stroke is reached in the first forging step, while 90% and 100% of the stroke are achieved in the second and third strokes, respectively. Six design variables R_1 to R_6 are used.

Optimization results are shown in Fig. 7 and 8. The comparison between the initial and final shapes are shown in Fig. 7, while the disk volume and press load histories against simulation runs are shown in the left and right of Fig. 8. The optimization has converged after about 30 simulations. The press load, which violates the constraint severely at the beginning, has been reduced from 99.4 to 79.6, a 19.9% decrease. The value 79.6 is the load capacity of the forging press and is shown in Fig. 8 (right) as a dashed line. The volume is increased from 82.4 of the initial minimum-weight design to 97.6 of the final optimized shape, a 18.4% change. These results are qualitatively similar in trend to the isothermal simulations described earlier. The simulation and optimization took about 35 h on a single-processor HP3600 workstation.

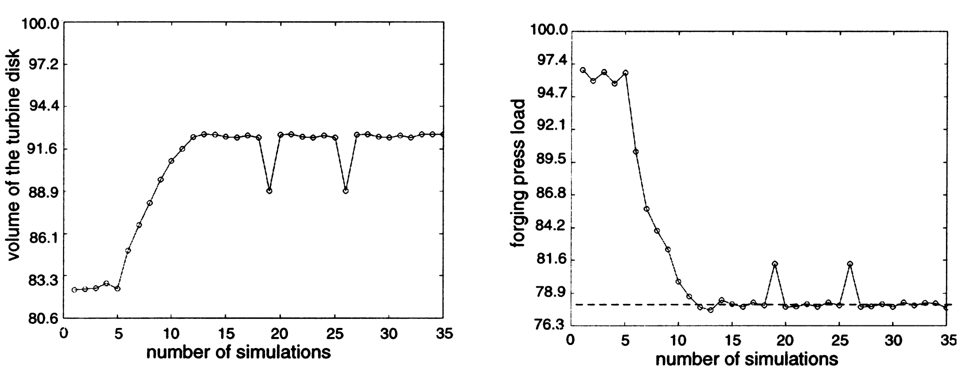


Fig. 6 Disk volume and press load versus the number of forging simulations for isothermal forging optimization

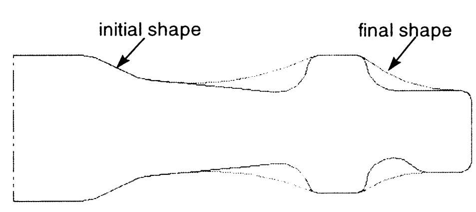


Fig. 7 Initial and final shapes for nonisothermal forging optimization

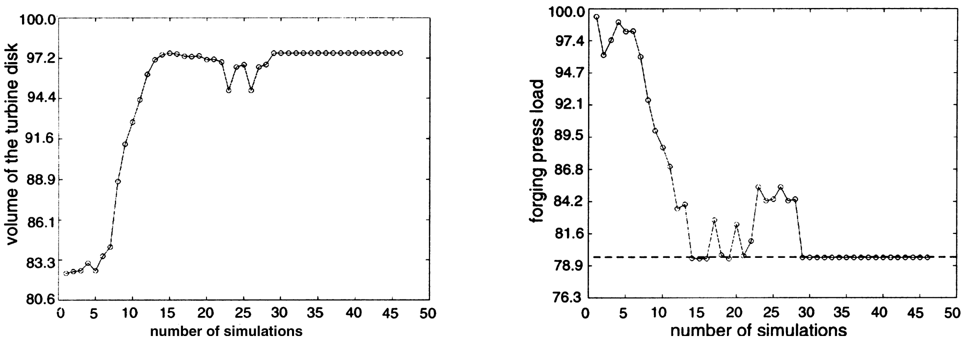


Fig. 8 Disk volume and press load versus the number of forging simulations for nonisothermal forging optimization

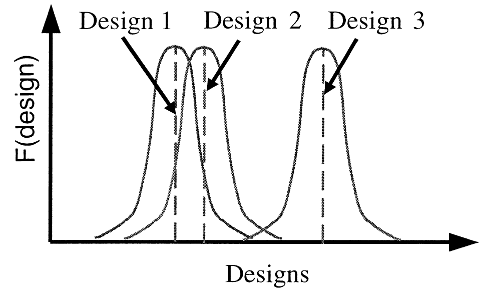


Fig. 9 Influence of model resolution: the noise makes it difficult to differentiate Designs 1 and 2. Designs 1 and 3, which are farther apart from each other, can be easily distinguished

Discretization, Approximation, and Searching (DASA) Optimization

In the methodology described previously, the shape optimization problem for die design was formulated in terms of global field constraint functions, and commercial CAD and CAE tools for forging design were integrated into a generic and noninvasive procedure for shape optimization. However, there are several issues with the application of this methodology to forging optimization. One way to overcome these difficulties is a new method known as discretization, approximation, and searching (DASA) (Ref 21). It has three essential ingredients: discretized design space, approximation in the neighborhood of a base design, and searching in the discretized neighborhood for a better design in each of the optimization iterations.

Issues with applying the conventional numerical optimization methods, which are designed for a continuous design space as well as smooth objective and constraint functions, to this application include:

- There are a few important elements in the forging process simulation that may cause nonsmoothness (noise) and even discontinuity in the objective and/or some of the constraint functions. Examples include rigid plastic material assumption, friction model, repeated remeshing due to large deformation, and so forth.
- Data mapping following each mesh regeneration introduces a differing amount of error in each iteration of the optimization process. Optimization works by comparing different designs and identifying the better one. If designs under evaluation are beyond the resolution of the analysis tool, it will be impossible to choose the better design since differences between various designs are lost in the noise of the calculations (Fig. 9).
- Commercial software packages are employed in this work and the finite-difference method, which is expensive and prone to truncation and discretization errors, is the only resort to obtain design-sensitivity information. As a side note, the industry is moving toward using off-the-shelf code. Although this greatly reduces the software maintenance and development cost, it creates some limitation for the application of first-order optimization methods since the analytical and semianalytical sensitivity information is often unavailable. Difficulties with field constraint functions and enforcing constraints at every node or element is impractical.
- An averaging technique is needed to define global or domain-based constraint functions.
- A practical discrete approximation searching method is needed to differentiate designs and reduce/eliminate the influence of the noise in the finite-element numerical results.

Gradient-based methods are probably the most efficient ones for many types of optimization problems. However, there are two big issues with this approach. First, there is no reliable and efficient way to obtain the sensitivity information since generally one does not have access to the source code of the commercial software used to conduct the simulations. Second, and more important, the response surface for this complex process is inherently nonsmooth, which creates difficulties for algorithms based on the assumption of smooth objective and constraint functions.

Another possible approach to this nonsmooth forging optimization problem is to build several global response surfaces for the objective and constraint functions using design of experiment (DOE) type of methods. Then gradient-based optimization algorithms can be applied to the global response surfaces. However, a detailed study showed that these response surfaces are of poor quality and computationally very time-consuming to obtain. Therefore, building local response surfaces along the optimization path is the more reliable and efficient approach.

DASA takes advantage of the response surface methodology, incorporates the notion of neighborhood searching common to most numerical optimization algorithms, and controls the influence of inaccuracies in modeling through discretizations of the design space. It proves to be effective for noisy and nonsmooth optimization problems. No rigorous convergence study similar to those in gradient-based algorithms has been completed, and DASA is not guaranteed to yield a solution that is optimal in the stringent mathematical sense, but it secures an improved solution, which is often sufficient in industrial applications. DASA has been incorporated into the iSIGHT Forging Optimization System.

Features of DASA. The basic ideas of DASA are:

- Discretize the design space to the extent that the analysis tool is able to resolve differences between two designs. Then represent the design space with these discrete design points only. This is consistent with the actual design practice where die designers treat a geometry parameter as a series of discrete values with certain spacing, say a quarter of an inch, rather than a single continuous variable.
- Conduct a sequence of local searches in the neighborhood of a base design in each design iteration or cycle. This is to avoid the expensive exhausted search, but there is a danger of being trapped at the local optimum similar to other numerical optimization methods. One strategy for reducing the chance of a local optimum is to start from different initial designs. It is desirable in forging optimization to start optimization at the most aggressive design—a design of least forging weight after satisfying geometry conditions including sonic cover, draft angle, as well as fillet and corner radii. Such a starting design is closest to the global optimum and thus the possibility of getting the best design is improved.
- Build a local approximation or response surface in the neighborhood of each base design. Because each forging simulation may take hours to complete depending on the size of a model, some type of response surface has to be introduced to reduce analysis time. A local approximation strikes the balance between speed and accuracy. The local search will exhaust all the design points surrounding a base design where the function values at these points are computed using the local response surface as opposed to the actual finite-element analysis. However, the design selected in each design iteration will be validated with the more detailed model.

The DASA algorithm involves the following steps and is summarized in a flow chart (Fig. 10):

1. Discretize the design space, choose initial design X_b, and read algorithm data.

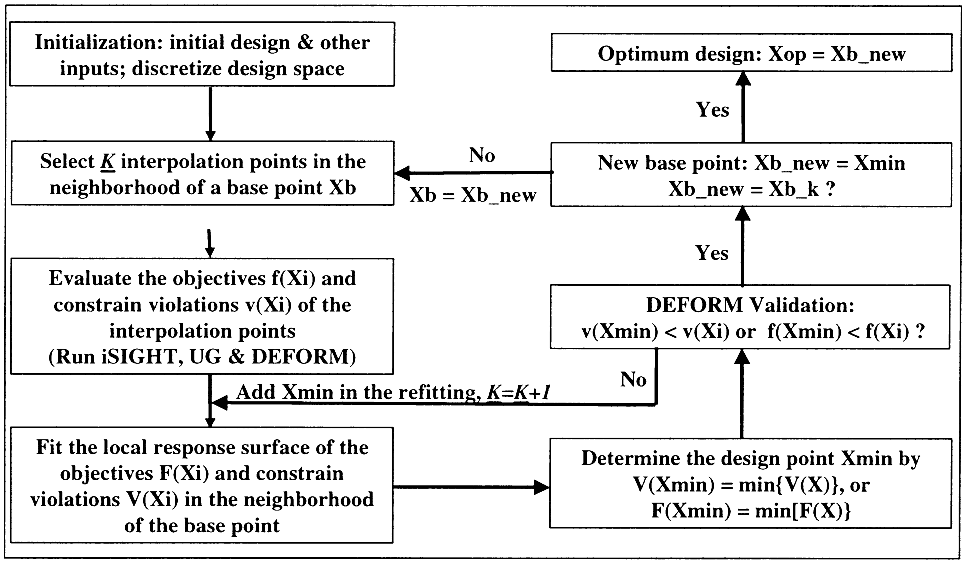


Fig. 10 Flow chart of discretization, approximation, and searching (DASA) optimization

2. Select the initial interpolation point set $\{X_i, i = 1, 2, \cdots, K\}$ in the neighborhood of the base design X_b.
3. Evaluate the objective $F(X)$ and constraint functions $V_j(X), j = 1, 2, \cdots, M$ at the interpolation points in the neighborhood of X_b.
4. Fit local response surfaces for the objective function $F(X)$ and each of the constraint $V_j(X), j = 1, 2, \cdots, M$ in the neighborhood of X_b using function values obtained at the interpolation points.
5. Select X_{min} among all points in the neighborhood of X_b such that it represents the best design. Here, all unknown objective and constraint functions are evaluated using the local response surfaces.
6. Conduct exact objective and constraint function evaluations to validate that X_{min} is indeed the best design among points in the neighborhood of X_b whose objective and constraint functions have been obtained through the exact analysis.
7. If step 6 confirms that X_{min} is the best design in neighborhood of X_b then go to step 8. Otherwise, add this point to the interpolation point set and go to step 4 to reconstruct the local response surfaces.
8. If $X_{min} = X_b$, it is the optimal solution and stop. Otherwise, set $X_b = X_{min}$ and go to step 2.

DASA Implementation. The following are a few additional details of the implementation of DASA. The design space is discretized using a uniform rectangular grid in step 1. Thus, the neighborhood of a base design is a hypercube. There are several strategies to select the initial interpolation point set $\{X_i, i = 1, 2, \cdots, K\}$ in step 2. Figure 11 shows three examples in the 2D case. The first method is used in the numerical example given in the next section. Different fitting methods and type of response surfaces may be used in step 4 to create the local approximation of the objective and constraint functions. The least-square fitting of second-order polynomials has been used here. The criteria for choosing the best design X_{min} in the neighborhood of X_b in step 5 are:

- If there is no constraint violation at all, then X_{min} is the point where $F(X)$ takes the smallest value.
- If there is a constraint violation, then X_{min} is the point where the sum of all constraint violations $\sqrt{\sum_l [V_l(X_{min})]^2}$ is the smallest.

Similar conditions are used in step 6 to check if X_{min} is indeed the best design in the neighborhood of X_b. Note that in this step X_{min} is only checked against all points whose objective and constraint functions have been calculated through the exact analysis.

Example Application of DASA. Consider the forging design of a generic gas turbine disk. Because of axisymmetry of the disk, a 2D analysis is needed. There are 10 design variables, and the section is divided into four zones (Fig. 12). Design constraints on strain, temperature, sonic cover, die nonfill, corner and fillet radii, and draft angles are imposed within each zone. In addition, there is a limitation on the forging time. Therefore, there are a total of 21 optimization constraints.

Some additional details of the problem include:

- Limitation of the forging load is maintained within DEFORM.
- Top and bottom dies are assumed to be rigid.
- The forging simulation is time dependent and nonisothermal.
- Global constraints are computed using the step-function formulation (Eq 16).
- Initial design starts at the most aggressive geometry—the minimum weight design after satisfying sonic cover, corner and fillet radii, and draft angle constraints.
- The numerical computation involves two different workstations: a SUN Ultra-1 running iSIGHT v4.0 as well as UG v15, and a 16-processor Silicon Graphics Origin 2000 running DEFORM.
- A typical optimization run takes about 50 h on a single-processor HP3600 workstation.

Figure 13 shows the initial and final cross sections of the forging design as well as the sonic shape. Figure 14 depicts the optimization history

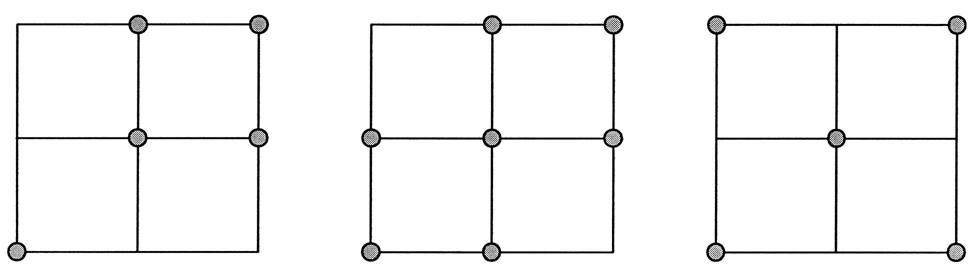


Fig. 11 Different strategies for selecting the initial interpolation point set in the two-dimensional space ($N = 2$): left ($N + 3$) points, middle ($2N + 3$) points, and right ($2^N + 1$) points

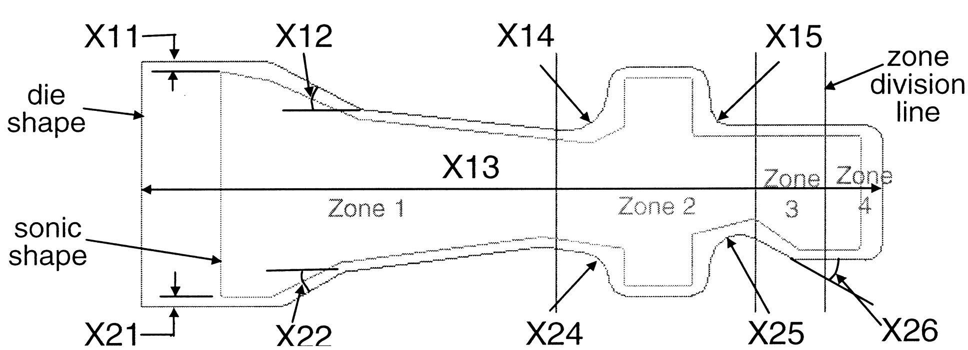


Fig. 12 Example of 10 design variables and four zones for design constraints

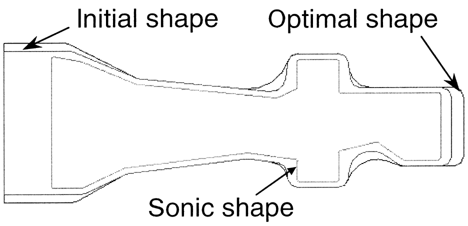


Fig. 13 Discretization, approximation, and searching (DASA) results: initial and optimal designs surrounding the sonic shape; the initial shape is inside the optimal design shape

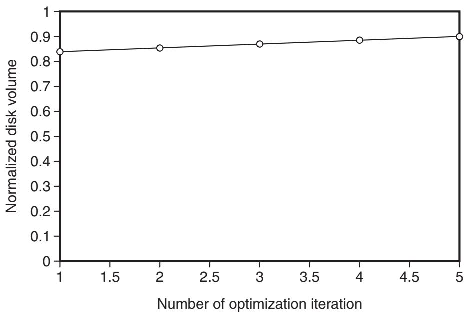


Fig. 14 Objective function (forging volume) versus design iterations with the discretization, approximation, and searching (DASA) optimization method

of the forging volume. Since the optimizer is started from the most aggressive geometry, it is expected the optimizer will add material to meet the design requirements since the initial design is infeasible. The weight added to the starting design is at the appropriate locations and is the minimum needed to satisfy the various constraints. Histories, which the user can look at to monitor the progress of the optimization run, are also generated for all the constraints.

Alternative Forging Optimization Formulations

The preceding sections focus on minimizing the weight of the forging. However, optimization problems can be formulated in several different ways besides the minimization of forging weight. Indeed, the goal of a general-purpose optimization program is to systematically and automatically evaluate various input parameters or design variables that fulfill user-defined objectives. Some examples of practical interest are:

- What is the preform shape required to make the strain uniform and greater than a specified minimum value in the final product? The uniformity of deformation and grain size in the part serves to minimize the variation of mechanical properties through the part.
- What are the forging parameters to keep the temperature during forging between specified maximum/minimum limits?
- What are the forging parameters (die speed/energy/load) to minimize the forge time while keeping the maximum instantaneous strain rate less than a specified limit and the forging temperature between specified maximum/minimum limits? What is the combination of maximum load and dwell time under maximum load that minimizes the total forge time and minimizes the creep in the dies at high loads?
- What are the forging parameters (die speed, forge temperature, etc.) to minimize the load and/or energy required (which in turn can minimize die stresses and wear)?
- What are the forging parameters to obtain the desired microstructure in the final product or the desired shape of the final product?

The following examples illustrate the power and versatility of the optimization tools with alternative optimization formulations instead of minimizing the forging weight.

Optimal Preform Design in Forging

An important parameter that characterizes metalforming processes is the initial shape of the workpiece or the shape of the preform. The material microstructure as well as the geometry of the final product are strongly dependent on the shape of the initial workpiece as well as on the preform shapes at each of the subsequent forming stages. It is therefore possible to envision a design procedure by which one can control the shape of the initial workpiece so as to achieve a desired geometry and/or material state in the final product. A formal methodology for this would be to pose the preform design as an optimization problem where the objective function would be a measure of the difference between the desired material microstructure/geometry for the final product and the computed material microstructure/geometry corresponding to a given initial shape of the workpiece. The preform design problem may also be subjected to certain other geometrical or processing constraints by the user.

Some examples of preform die design for axisymmetric forging problems are described in the paragraphs that follow. Figure 15 shows that without proper preform shape definition, the material may not fill the finisher dies completely or may result in the formation of folds or laps. An optimization system has been built into the commercial metalforming software DEFORM (Ref 22). This was used to design an optimal preform die, with which the material filled the finisher die and the distribution of effective strain was made more uniform (Fig. 16). It was also applied to the preform design problem (Fig. 17), and the distribution of effective strain was more uniform with the optimal shape. The methodology has been demonstrated for a few different shapes including more complicated part geometries (Ref 22). In these examples, the shape of the preform was represented by B-spline curves or piecewise linear curves. The design variables are the control points of the B-spline or piecewise linear curve. For example, if the diameter and height of billet are design variables, billet shape can be represented by a rectangle with four control points, and the x and y coordinates of corner point can be design variables. Alternatively, it is also possible to use Bezier representations of the geometry, and the control points of the Bezier curve are the design parameters.

Shape optimization with constraints requires the evaluation of the gradients of all objective and constraint functions with respect to each shape control parameter. The sensitivities of the nodal velocity with respect to the design variables are provided by direct differentiation of the element stiffness equations. By using the sensitivities of the nodal velocity, the sensitivities of the nodal coordinate, effective strain rate and effective strain are calculated and updated at every step. At the end of forging simulation, the objective and constraint function values and their gradients are calculated. In the optimization program, the gradients of objective and constraint functions are used to determine the new search direction.

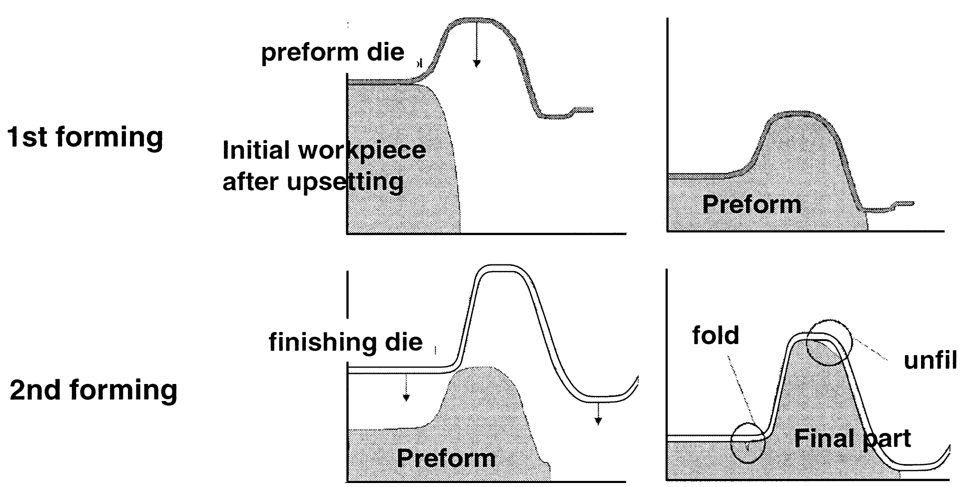


Fig. 15 Fold and underfill defects due to improper preform shape definition

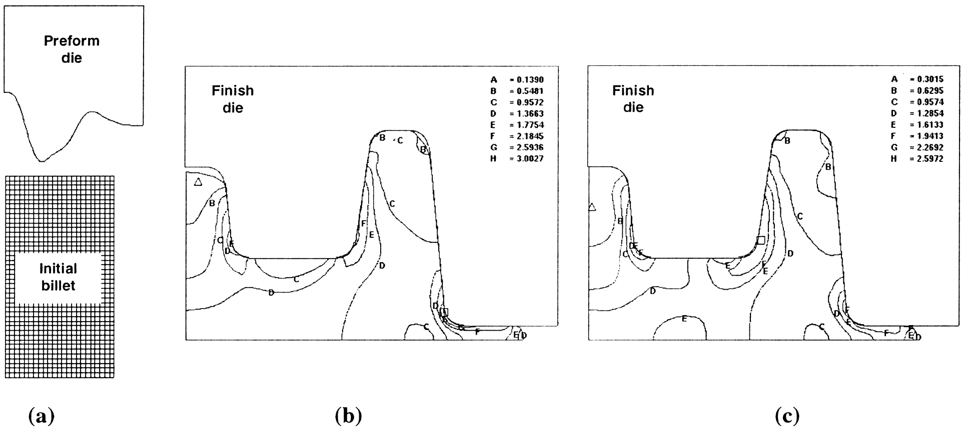


Fig. 16 (a) Optimal preform die shape. (b) Effective plastic-strain distribution without preform die. (c) Effective plastic-strain distribution with optimal preform die

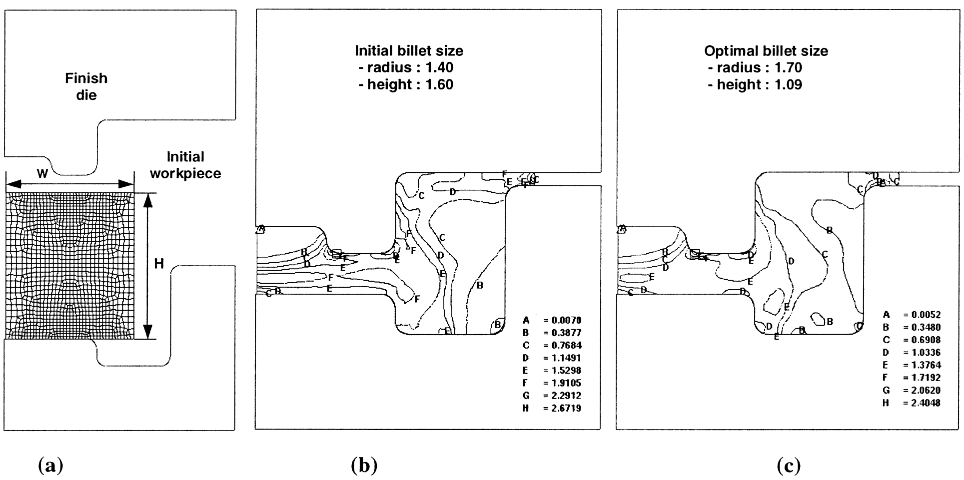


Fig. 17 (a) Initial preform shape. (b) Effective plastic-strain distribution with initial preform shape. (c) Effective plastic-strain distribution with optimal preform shape

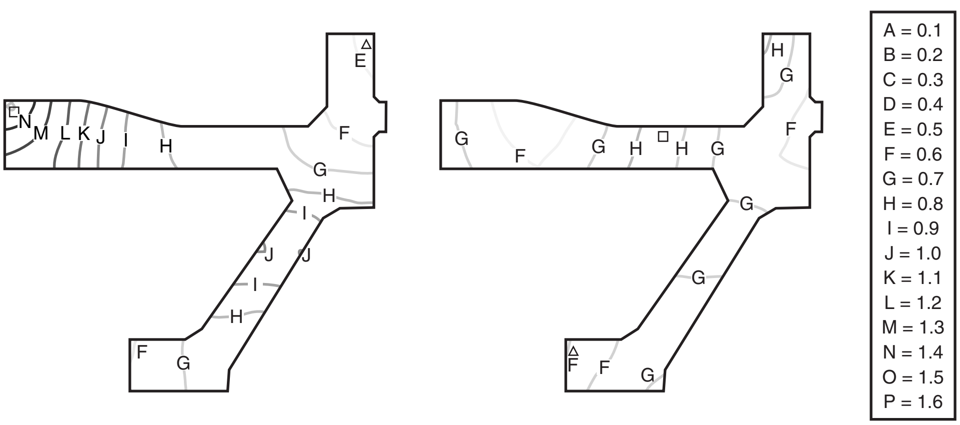


Fig. 18 (a) Effective plastic-strain distribution with initial preform shape. (b) Effective plastic-strain distribution with optimal preform shape

Shape optimization should be carried out at the geometry level in that the constraints are specified in terms of boundary points and geometric entities rather than individual finite elements or nodes. When the shape is perturbed to compute the sensitivity gradients, the boundary nodes can simply be moved to the perturbed boundary without regenerating the entire mesh. In fact, generating a new mesh for the gradient calculation is not recommended because it often leads to inaccurate sensitivity results and other numerical problems due to the fact that the new mesh topology could be very different from the old one, invalidating the design-sensitivity definition. For large shape changes during optimization iterations, it may be necessary to generate an entirely new mesh. Shape changes should avoid kinks and other geometric irregularities. One way to ensure this is to impose additional geometric constraints, for example, the continuity of slopes and curvature signs at boundary points, or to smooth out the irregularities in the geometry at each stage.

Another example of making the strain uniform in a more complicated part geometry is shown in Fig. 18 (Ref 23). Here the preform shapes were defined as simple third-order polynomials and the problem converged in about 25 steps in less than 2 h of HP 3600 computer time using a relatively coarse mesh. For more complicated conditions, preform shapes can be defined by B-splines or Bezier curves rather than lower-order polynomials. In this example, the initial simulations were done with a relatively coarse mesh. The coarse finite-element mesh solution quickly gets to the optimum, and this can serve as the starting point for a fine-mesh solution. This two-step procedure minimizes the number of fine-mesh iterations and enables the fine-mesh optimum solution to be obtained rather quickly. Also, because the user is not concerned with material that is machined off, modeling of complex conditions also may be simplified by not interrogating field variables in the whole volume of forging, but just in the finished part shape. The finished shape is defined as a separate object, and the field variables are remeshed and remapped to this shape after the simulation is completed. The advantage of this procedure is that the field variables from one design to another are evaluated on the same mesh, and this considerably reduces the "noise" in the gradient calculations. Another advantage is that localized peak values at die/forging corners are removed, and it is easier to make the strain uniform. In the limit, the user can make the forging and finished shapes to be the same, but remeshing to the same final mesh does have an advantage.

Multistep Forging Optimization

Optimization of multistep processes may be viewed as the design of the forming sequence that converts the initial workpiece to the final product while meeting the desired manufacturing objectives and satisfying various process, material, equipment, and geometrical constraints. A forming sequence can be viewed at two levels for the purpose of optimization (a) the broad identification of the number, type, and order of forming/heat-treatment operations that make up the sequence (e.g., forward extrusion, open-die forging, etc.) and (b) the specific identification/selection of design variables in each of the forming operations (e.g., die speed/load/energy, die shapes in a preforming stage, etc.).

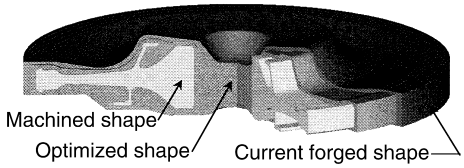


Fig. 19 Typical weight savings between manual (current forging shape) and optimized designs

Fig. 20 Various generic optimized disk shapes

Multistep forging optimization involves the following considerations:

- Minimization of number of forging steps
- Creation of optimization constraints for each forging step
- Development of criteria to decide when to switch to a new forging step
- Use of standardized preform shapes to keep production costs down
- Program control to simulate all forging steps and coordinate data transfer between them

Other examples of both single and multistep metalforming optimization can be found in Ref 24 to 33. Figure 19 shows the typical weight savings between manual and optimized designs, with the final machined shape also shown for reference. Finally, Fig. 20 shows various generic optimized disk shapes.

Heat Treatment Optimization

Of the manufacturing processes to produce gas turbine engine components, heat treatment and machining are two critical operations that have an interacting effect on the finished engine component. For example, nickel-base superalloys are heat treated following forging to precipitate the gamma-prime (γ') strengthening phase; a process that involves quenching from near the solvus temperature. Thermal gradients during quenching cause thermal stresses, which drive localized plastic deformation and residual-stress buildup. Upon cooling to room temperature, residual stresses can exceed half of the alloy tensile strength, which leads to distortion following quench and later during machining.

Residual stresses and associated distortion have a significant effect on manufacturing cost in three distinct ways. First, the forging and intermediate heat treat shapes contain additional material to account for expected distortion. This material, added to ensure a positive material envelope over the finished part shape, represents a raw material cost and increases the machining cost. This also imposes a limit on the benefit of near-net-shape forging, which is being pursued vigorously by the industry. Second, part movement during machining requires that the machining process engineer plan machining operations and fixtures so that distortion does not compromise the finished part shape. Movement during machining or excessive distortion from an inappropriate metal-removal plan can cause machining scrap. To combat this possibility, components such as disks are machined alternately on either side in an attempt to stepwise balance the distortion. The time spent "flipping" parts erodes productivity for thick, stiff components; for thin components the strategy may be inadequate. Third, residual stresses and associated distortion add complexity to machining process development and shop operations. Distortion affects the details of the machining plan and the way the component interfaces with machining fixtures. These effects generally vary between material suppliers and from lot to lot. Distortion thereby not only influences the effort incurred during initial development of machining plans, but may require adjustments after the initial plan has been set.

The buildup of residual stresses during heat treatment (oil quench or fan cool) and subsequent relaxation following metal removal are impossible to assess using intuition, engineering judgment, or empirical methods. The physical interplay of quench heat transfer, elevated-temperature mechanical behavior, and plastic-deformation localization is complex. Subtle changes in processing conditions and component geometry can significantly affect the magnitude and pattern of residual stresses. Determining residual stresses and subsequent distortion requires modeling using finite-element method. This method has been used to evaluate the effect of processing conditions on residual-stress development and the effect of residual stresses upon distortion during machining.

The purpose of the heat treatment process is to develop the necessary mechanical properties in the forged part. This is achieved by heating the part to solution temperature and then cooling it rapidly. During the cooling phase residual stresses are introduced. In the case of nickel-base superalloys, a certain minimum cooling rate has to be maintained to generate the needed creep and tensile properties. On the other hand, the faster the cooling process is, the higher are the resulting residual stresses.

A description of the various heat treat procedures is given in Ref 34. Traditionally, an oil-quenching process has been employed, which ensures fast cooling and thus a high cooling rate, but the oil-quenching process introduces high residual stresses, and, from a process optimization point of view, offers very little room for improvement as there are very few parameters that can be controlled. Therefore, fan cooling is gaining larger acceptance where it is possible to control the airflow on individual sections of the part and thus influence the local surface heat transfer coefficients. Obviously, the heat transfer coefficients that can be achieved with fan cooling are lower than those for oil quenching, so that for thick parts it may not be possible to satisfy cooling-rate requirements, but for moderately thick parts fan cooling offers clear advantages. For very thin parts such as engine seals where machining distortions due to residual stresses are especially critical, fan cooling may be the only process that produces acceptable parts.

Previous work related to heat treat optimization can be found in Ref 35 to 37, Batista and Kosel (Ref 35) developed the analytical sensitivities of the residual stresses after quenching. The sensitivities of the relative error of residual stresses from the estimated errors of material data during the quenching process were reported. Karthikeyan et al. (Ref 36) developed mathematical models to optimize the heat treatment conditions for maximum yield strength and ductility of aluminum/silicon-carbide particulate composites. Saigal and Leisk (Ref 37) optimized the tensile properties of alumina/aluminum metal-matrix composites using Taguchi analysis.

Formulation of Heat Treatment Optimization. The heat treatment optimization tool (Ref 38) optimizes the heat treatment of engine disks by changing the heat transfer coefficients (the design variables) around the disk surface in order to:

- Achieve a residual-stress distribution for minimum distortion during machining
- Minimize cooling-rate nonuniformity and hence residual stresses
- Achieve cooling rates to meet mechanical property requirements
- Improve part-dimensional stability during machining and during service

As in the forging optimization software described earlier, the heat treat optimizer is based on a general procedure to generically and non-invasively integrate commercially available software. The integration and optimization environment is similar to that described for forging optimization except that the UG CAD system is not used since the geometry is not part of the heat treat optimization process. iSIGHT has been adopted as the framework of the system and DEFORM as its simulation code that supplies the objective and constraint information. The heat treat module (Ref 39) of DEFORM is used to simulate the heat treat process.

The challenge here is to formulate an optimization problem without actually having to execute a combined heat-transfer/stress-analysis each time the optimizer needs a new design point. An accurate heat transfer analysis requires small time steps in the simulation, and a stress analysis, in turn, requires a fine finite-element mesh; therefore the combination of both is the most computationally expensive analysis possible. In general, though, the stress analysis is much more time consuming than the heat transfer analysis alone. Since it is known that spatially uniform cooling reduces residual stresses, the idea is to formulate an objective function that penalizes nonuniform cooling and at the same time ensures cooling at or above the target cooling rate. These are obviously two conflicting objectives since fast cooling always means uneven cooling as the heat can only be extracted at the surface of the part. Therefore, the objective function for the heat treatment optimization problem is formulated as a quadratic that penalizes the deviation from the cooling-rate target, which is a material-dependent value:

$$\text{obj} = \sum_{\text{nodes}} \begin{cases} w \cdot \left(\bar{t}_{\text{target}} - \bar{t}\right)^2 & \text{if } \bar{t} < \bar{t}_{\text{target}} \\ (1-w) \cdot \left(\bar{t}_{\text{target}} - \bar{t}\right)^2 & \text{if } \bar{t} \geq \bar{t}_{\text{target}} \end{cases} \qquad \text{(Eq 18)}$$

where w is a user-defined weighting factor between 0 and 1 that penalizes under- and overachievement of the target cooling rate differently, $\bar{t}$ is the cooling rate, and $\bar{t}_{\text{target}}$ is the target cooling rate to achieve the desired properties. A value w of close to 1 seems to give the best results. The objective function is really an integral over the volume of the part or the cross-sectional area for the 2D analysis for the axisymmetric parts considered here, but since temperature values are only available at the finite-element nodes, it is formulated here as a sum over all nodes.

This quadratic formulation of the objective ensures a smooth convergence behavior. It captures both the desire for uniform cooling and the fulfillment of the cooling-rate target by penalizing the objective with the square of the deviation from the target. Additionally, with the weight w that is to be chosen between 0 and 1, the user has the option to differentiate between over- and underachievement of the cooling-rate target. Since the objective of the heat treatment optimization is to at least reach the cooling-rate target, the user usually wants to penalize underachievement far more severely than overachievement. A weighting factor w of 0.98 (98% penalty on cooling-rate underachievement) has been successfully applied.

The design variables are the surface heat transfer coefficients, h_i, which can be related back to a certain airflow produced by the fan-cooling apparatus. Usually, a breakpoint between two design variables should be chosen wherever the geometry of the part changes, for example, as the thickness changes from bore to web or web to rim, and so forth, as shown in Fig. 21.

The procedure gives the user a choice in terms of optimization constraints. A hard constraint c_{cr} on the cooling rate could be imposed such that:

$$c_{\text{cr}} = \frac{1}{nodes} \sum_{nodes} \begin{cases} \bar{t}_{\text{target}} - \bar{t} & \text{if } \bar{t} < \bar{t}_{\text{target}} \\ 0 & \text{if } \bar{t} \geq \bar{t}_{\text{target}} \end{cases} \qquad \text{(Eq 19)}$$

This constraint has a discontinuity at 0, exactly where it is active, and will never assume a value less than 0, that is satisfied and not active. This discontinuity leads to problems with gradient-based optimizers, which will always see a zero constraint gradient for a satisfied or active constraint; therefore, in the case of constraint satisfaction the constraint value of 0 is replaced with the difference of the target cooling rate $\bar{t}_{\text{target}}$ and the minimum $\bar{t}_{\text{min}}$ of all nodal cooling rates:

$$c'_{\text{cr}} = \bar{t}_{\text{target}} - \bar{t}_{\text{min}} \qquad \text{(Eq 20)}$$

Thus, at least the sign of the constraint gradient that the optimizer sees above and below a constraint value of 0 will be equal. An additional constraint can be placed on the nodal fraction that fulfills the cooling-rate target, which has to be equal to 1.0 if the target is met everywhere. The two constraints may seem somewhat redundant, but depending on the optimization strategy used, one or the other or a combination of both leads to the best convergence.

Heat Treatment Optimization of a Turbine Disk. The heat treatment optimization procedure described previously was applied to a generic turbine disk. Figure 21 shows the heat treatment geometry and the distribution of the nine design variables employed. Since each finite-element analysis of the heat transfer problem took about 11 to 15 min on an HP-715 workstation, execution time per iteration was in the range of 3 to 4 h, including computational overhead, using finite differencing to obtain gradients for the optimizer. In order to cut down on these times, the optimization was started with all heat transfer coefficients linked to only one design variable. This problem was executed for six iterations, using the sequential linear programming technique from ADS (Ref 18), which is one of the packages in iSIGHT, until both constraints were active. The full convergence history of the objective function is depicted in Fig. 22.

For this problem, the modified method of feasible directions, also from ADS, was chosen as the optimization technique. The deviation function was initially reduced from a value of 1.4 to about 0.6 and then further down to under 0.2. These numbers as such have no physical meaning, but the significance can be seen in the reduction of the standard deviation of the nodal cooling rates, indicating a much more uniform cooling than at the starting point. This is also evidenced by a comparison of the initial cooling-rate distribution (Fig. 23) and the optimized cooling-rate distribution (Fig. 24), all normalized with respect to the target value.

Figure 25 shows the history of the design variables for the second segment in the optimization process where all design variables are active, normalized with respect to the starting value. The result is large heat transfer coefficients around the thick portions of the disk (h_1, h_2, h_3, h_7, h_8) and significantly lower coefficients for the thinner sections, leading to a more uniform cooling while still maintaining the target cooling rate.

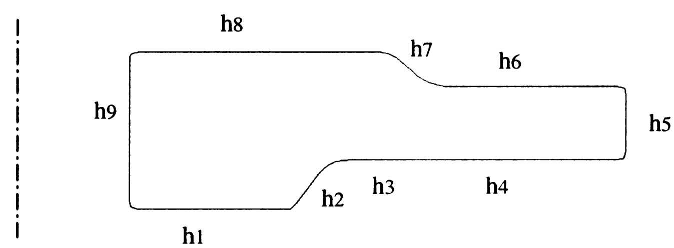


Fig. 21 Generic turbine disk geometry and heat transfer coefficient distribution

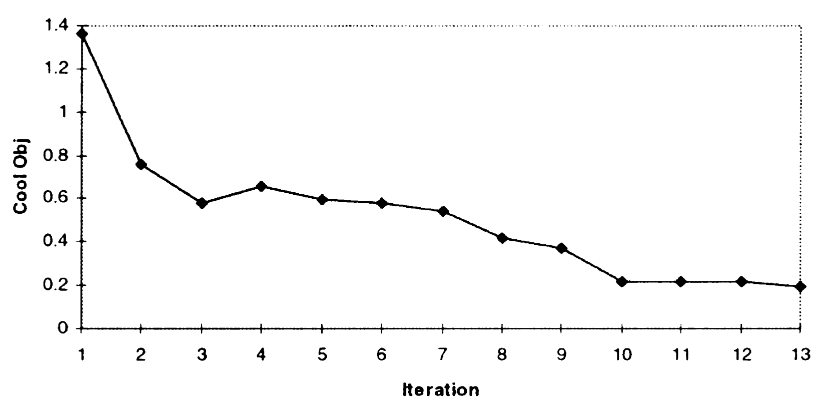


Fig. 22 Objective function history

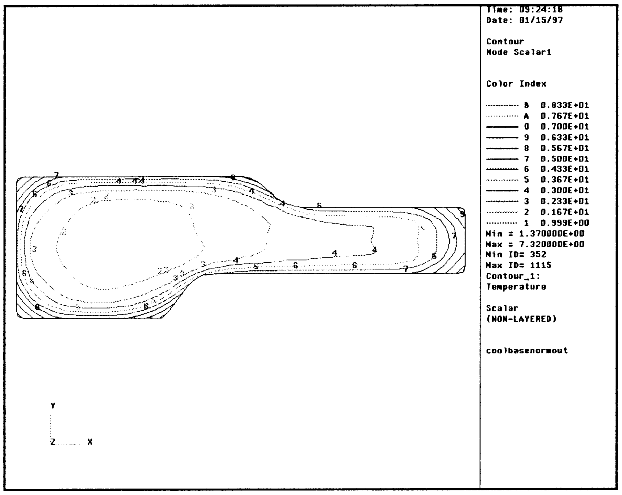


Fig. 23 Initial cooling-rate distribution: note the large surface/center non-uniformity

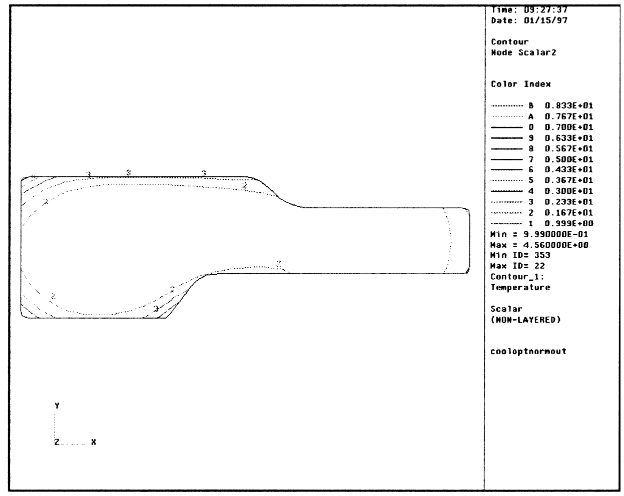


Fig. 24 Optimized cooling-rate distribution: note uniformity in cooling rate compared to Fig. 23

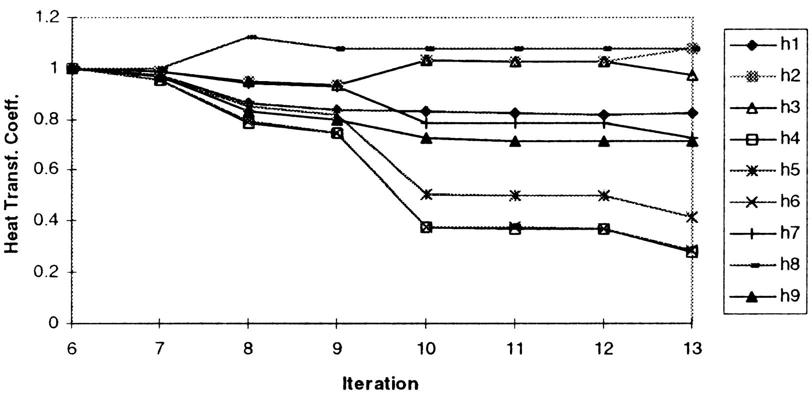


Fig. 25 Normalized design variable history

The question still to be answered is what effect this optimization procedure, which is based on heat transfer analysis only, has on the residual stresses of the part which is of ultimate interest. Therefore, a combined heat-transfer/stress-analysis was performed on both the starting configuration and on a disk with the optimized heat transfer coefficient distribution. For comparison purposes, an analysis of a typical oil-quenching process was also performed.

Figures 26 and 27 show the resulting hoop and radial residual stresses, respectively, normalized with respect to the maximum tensile stress of the oil-quenched part. Residual stresses are highest for the oil-quenched disk, which was closely followed by the nonoptimized fan-cooled disk with uniform high fan blowing all around. The residual stresses for the optimized disk, in turn, are considerably lower, almost by one order of magnitude compared to the oil-quenched part in terms of tensile stresses (Fig. 26, 27). The reductions in compressive stresses are not quite that large, but still by a factor of between six and seven. Table 1 shows the stress reductions of the maximum tensile and compressive hoop and radial stresses compared to the oil-quenching process. These results clearly show the advantage of a numerically optimized fan-cooling process compared to the traditional oil quenching. One can claim, however, that the starting point for the optimization was not very realistic with almost maximum fan cooling all around the disk, and an experienced person would have been able to set up a better fan-cooling process manually. However, this argument misses the point since the starting point for the optimization was deliberately chosen to be far away from the optimum in order to show that the optimizer will find the optimal process—within convergence margins, of course, no matter where it starts. This reduces the need for prior experience in arriving at the optimized solution, which was confirmed during multiple runs with different starting points on actual geometries, which are of proprietary nature and cannot be shown here. The formulation of the objective function as a quadratic clearly aids in this behavior.

Therefore, to summarize, the optimized solution results in:

- A cooling-rate distribution that is much more uniform, especially in the thinner disk sections
- The cooling-rate target is met throughout the part
- Maximum stresses reduced by 89% (tension) and 84 % (compression) as compared to a traditional oil-quenched part

An integrated engine disk design and manufacturing simulation and optimization scenario has been formulated. As an initial step, a heat treatment optimization procedure was developed, taking advantage of the lower computational cost of a pure heat transfer analysis compared to a combined heat-transfer/stress-analysis. The formulation of the objective as a quadratic cooling-rate deviation function has shown to produce considerably lower residual stresses in the generic turbine disk analyzed in this chapter. In other words, with this formulation it is possible to lower residual stresses in forged parts without the cost of a coupled thermal-stress analysis during the actual optimization.

Alternative Heat Treatment Optimization Formulations. The previous examples have focused on minimizing the cooling-rate variations and the residual stresses in the forging. The general framework of iSIGHT permits the optimization problem to be formulated in several different ways. The goal of a general-purpose optimization program is to systematically and automatically determine the input parameters or design variables that fulfill the user-defined objectives. Examples of alternative heat treatment optimization formulations of practical interest that illustrate the power and versatility of these tools include:

- The inverse problem of determining heat transfer coefficients from measured temperature-time data
- The problem of not minimizing residual stresses but of creating a residual-stress distribution that minimizes the subsequent machining distortions and ensures optimal dimensional stability during service of the part

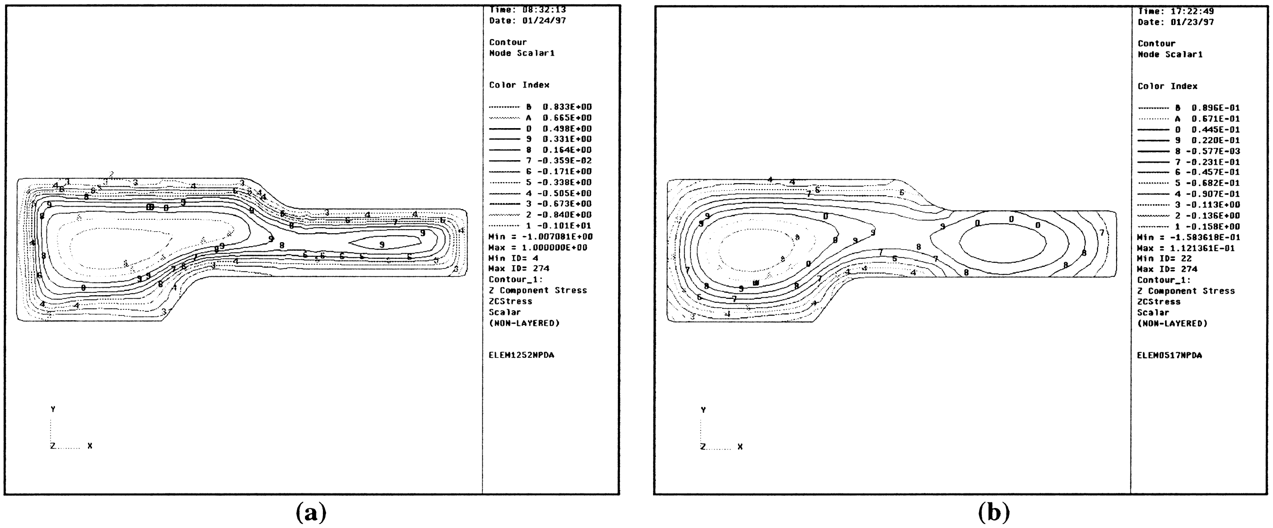


Fig. 26 Hoop stress. (a) Oil-quench process. (b) Optimized process

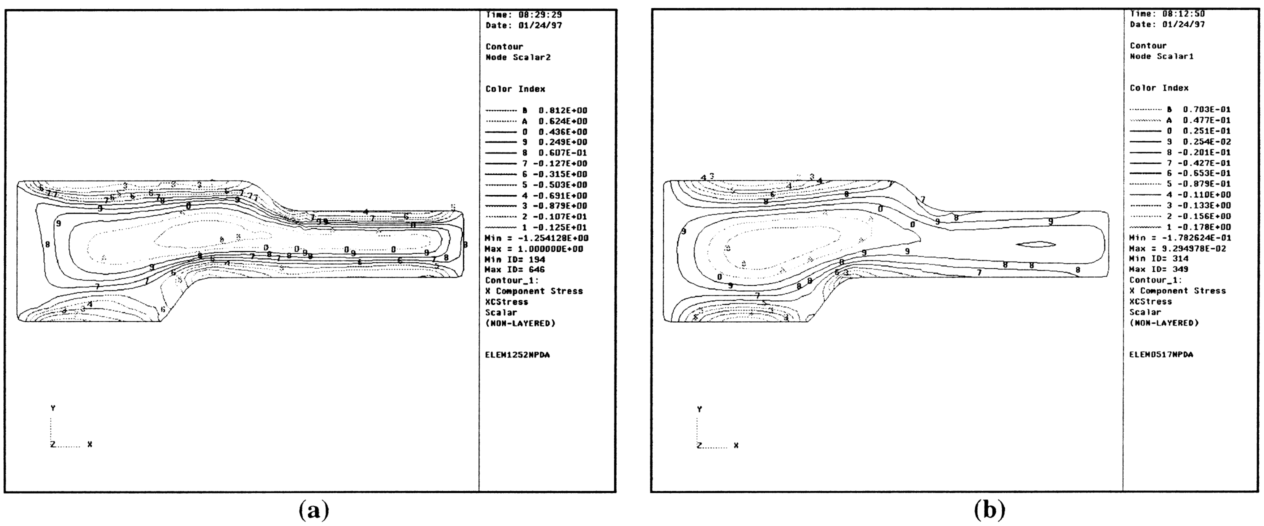


Fig. 27 Radial stress. (a) Oil-quench process. (b) Optimized process

Table 1 Stress reductions in tension and compression after heat treatment optimization

	Traditional oil quench	Initial optimizer guess	Final optimized solution
Hoop stress			
Tension	1.000	0.861	0.112
Compression	1.007	0.854	0.159
Radial stress			
Tension	1.000	0.850	0.093
Compression	1.254	0.908	0.178

- The problem of obtaining a cooling-rate distribution that in turn results in an optimal distribution of mechanical properties (e.g., high tensile strength in the disk bore and high creep strength in the disk rim) and a minimum weight of the finished disk shape

Of these, the inverse problem is described in some detail. The inverse problem of obtaining surface heat transfer coefficients from measured temperature-time data during the heat treatment of a part has received considerable attention. Several inverse methods (for example, Ref 40) have been developed and used to solve this problem and have met with varying degrees of success. The inverse problem can be cast as an optimization problem. The objective function is the sum of the squares of the deviation of predicted and measured temperatures with the surface heat transfer coefficients as the design variables. The optimization problem then is to determine the set of design variables that minimize the objective function. Two examples of the solution of this problem are described.

In the first example, the heat transfer coefficient during a gas-quenching process is optimized to obtain minimum distortion satisfying the two constraints of maximum stress (to avoid quench cracking) and the average surface hardness requirements of the final product (Ref 41). The design variables are the heat transfer coefficients that are functions of the surface and ambient temperatures or time. This functionality is represented by spline functions making the coefficients defining the splines as the design variables First, the response surface method is used to obtain the analytical models of the objective function and constraints in terms of the design variables. Then, some of the less important terms are eliminated to improve the fit of the response models. Next, with the closed-form response surface equations, the sequential quadratic programming method in the design optimization tool DOT (Ref 18) is used to obtain the optimal design point. The finite-element package, DEFORM is used to predict the material responses during the quenching process.

Phase transformations, deformation, and heat transfer are integrated during the simulation process. Due to the fitting error of the response surfaces, the optimal design point obtained by the response surfaces may not satisfy the constraint with the finite-element method even though the constraints are satisfied by the response surface method. Therefore, the optimal design point obtained by the response surface method is added to the design of experimental points to improve the fitting accuracy of the response surfaces until the relative errors of the objective and constraints obtained by the response surface method and finite-element method are in good agreement. The response surface method provides an effective way to optimize this process.

The second example described here deals with the more conventional inverse problem of obtaining heat surface transfer coefficients from temperature-time data (Ref 42). In this case, the disk geometry shown in Fig. 28 was used. The temperature-time data during quenching is prescribed at two points, and the boundary of the disk is divided into segments, each with its own heat transfer coefficient, which is specified as a function of surface temperature. The control points in the functional dependence of heat transfer coefficient on temperature are the design variables. The optimization procedure is incorporated into the heat treat module of DEFORM, and a converged solution is obtained in 20 iterations matching the predicted and measured temperature profiles very closely. For a more realistic problem, there would be several temperature-measurement points, and the surface of the disk would be divided into several segments, each with its own heat transfer coefficient.

Another optimization-based inverse method, which determines the thermal boundary conditions (heat transfer coefficients) during the heat treatment of aircraft engine disks has been developed (Ref 55). A nonlinear optimization methodology has been developed that provides the mean and confidence bounds for the predicted heat transfer coefficients. It is shown that the heat-transfer coefficients are strongly dependent on the number and location of the thermocouples. In particular it is shown that the placement of the thermocouples strongly influences the confidence bounds for the estimated parameters. The procedure has been successfully applied to selected examples.

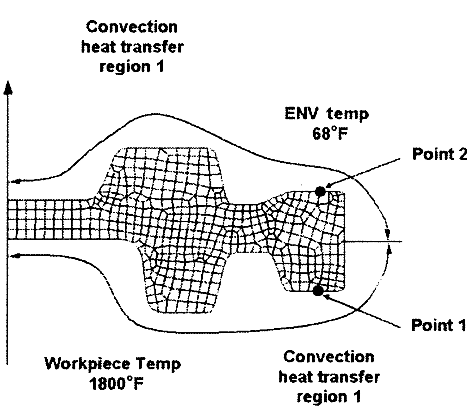


Fig. 28 The determination of heat transfer coefficients from temperature-time data

These examples have served to illustrate that with general-purpose optimization tools, the user has a wide variety of options of selecting the objective function and constraints for any particular situation to obtain the optimal solution quickly and effectively.

Summary of Optimization Techniques

Optimization Methods

No one optimization approach works well for all problems, particularly for MDO problems. To avoid convergence or other problems, special techniques have to be developed on a case-by-case basis. All the techniques described in this section are numerical optimization techniques. They generally assume the design space is unimodal, convex, and continuous in nature. For the forging optimization problem, all these schemes produced results (optimized forging weight) within 1 to 2% of each other, but the number of iterations required varied significantly, depending on the optimization scheme used. The various methods are briefly described in this section; more details may be found in the iSIGHT manuals (Ref 9).

Exterior Penalty (EP). This method is widely used for constrained optimization. It is usually reliable and has a relatively good chance of finding the true optimal solution if a relative minimum exists. The EP method approaches the optimum from an infeasible region, becoming feasible in the limit as the penalty parameter approaches infinity.

DONLP—Sequential Quadratic Programming. This technique uses a slightly modified version of Pantoja-Mayne update for the Hessian of the Lagrangian, variable scaling and an improved step-size algorithm. With this technique, bounds on the variables are treated in a projected gradientlike fashion.

Hook-Jeeves Direct-Search Method. This algorithm begins with a starting guess and searches for a local minimum. This optimization technique does not require the objective function to be continuous. Because the algorithm does not use derivatives of the objective function, the function does not need to be differentiable.

Sequential Linear Programming (SLP). This is a strategy used to solve constrained optimization problems. This technique is easily coded and applicable to many practical engineering problems.

Modified Method of Feasible Directions (MMFD). This is a direct numerical optimization technique used to solve constrained optimization problems. This method iteratively finds a search direction and performs a one-dimensional search along this direction. The emphasis is to reduce the objective while maintaining a feasible design. This technique attempts to rapidly obtain an optimal design, handles inequality and equality constraints, and satisfies constraints with high precision at the optimal point.

Successive Approximation Method (SAM). This method lets one specify a nonlinear problem as a linearized problem. The SAM method is a general one, which uses a simplex algorithm in addition to sparse matrix methods for linearized problems. If one of the variables is declared an integer, the simplex algorithm is iterated with a branch and bound algorithm until the desired optimal solution is found.

Response surface methodology (RSM) is one of the approaches for function approximation in design optimization and is especially suitable when the sensitivity information is not available or difficult to obtain. The response surface model is a simplified Taylor series with only lower-order terms. The advantage of RSM is its wide applicability since any real function can be expanded as a Taylor series. Since the accuracy of Taylor series is usually restrained by the truncation error in highly nonlinear cases or in approximating a large domain, the RSM faces an accuracy and complexity challenge. If too many terms are included, then the process of model identification is complicated and computationally expensive. To find a balance between the accuracy and complexity, in practical applications, a second-order polynomial is a widely used formulation for RSM. Higher-order terms can be selectively dropped, depending on their importance as determined from initial analyses. Unlike gradient-based approximations, the RSM usually does not match the function values at trial points, and model parameters change gradually with new points added into the data set. Therefore, the use of RSM in design optimization (especially with discrete variables) could create traps or oscillations in convergence. Further details of the RSM method and its application are provided in Ref 43 and 44.

Optimization Exploration Techniques

The use of MDO optimization systems such as iSIGHT can combine multiple exploration capabilities. The use of these design exploration methods in conjunction with approximation models can be used to greatly reduce the computations required to obtain an optimized solution.

Design of experiments (DOE) allows for a larger number of design parameters to be used to examine the effects on the objective. DOE is then able to estimate a near-optimal design from these data. Engineers can gain a lot of insight from a minimum number of simulations. This insight can be used in subsequent optimizations.

Exploitive optimization techniques are very good at improving a design in a local area. These techniques are commonly known as hill-climbing techniques since they require gradient calculations to "learn" which way to improve the design.

Explorative Optimization Techniques. These optimization techniques are very good at exploring the complete design space. They usually require 500 designs to begin to find optimal designs.

Approximation Models. With a few well-chosen design points, an approximation model of the real simulation can be created. Any of iSIGHT's optimization techniques can then use this approximation model for calculating an optimal design. These approximation models are automatically updated as more simulations are performed.

Sensitivity Analysis

As a by-product of optimization, a sensitivity analysis showing how sensitive the outputs are to the process variables is useful to help identify which variables to control and how tightly. The results of this analysis can help set tolerances and controls on the significant process variables and lead to an effective process control strategy and ensure a robust process and product by reducing the effects of variability that is inherent in any process. Several papers have been published on sensitivity analysis, and examples may be found in Ref 43 and Ref 45 to 52. These describe shape and process optimization procedures to design preforms based on sensitivity analysis and genetic (adaptive) algorithm techniques.

A gradient-based approach can be used to solve the constrained optimization problem efficiently. This necessitates the evaluation of the gradient of the objective function and constraints with respect to the design variables. These gradients or sensitivities are quantitative measures of changes in the objective function and constraints as a result of perturbations in the design variables. Gradient computations can be performed using finite-difference approximations and the results of the direct analysis for two nearby design variables. In addition to significant computer resources required for solving the direct problem multiple times, difficulties arise in such calculations from the fact that many direct-analysis tools are insensitive to infinitesimal changes in the design variables (e.g., the die surface) and cannot provide accurate sensitivity fields. In addition, the calculated sensitivities are dependent on the step size used to compute the perturbed direct-deformation problem. This is particularly true when the computed sensitivity fields are of the same order of magnitude as the numerical error in the solution of the direct analysis. However, an efficient design methodology necessitates the accurate computation of the design derivatives of various deformation-related parameters. Sensitivity analysis is a method that is widely used to evaluate these gradients.

A continuum sensitivity analysis was developed (Ref 45, 46) for large inelastic deformations to include the effects of contact and friction for both steady-state and non-steady-state metalforming processes. This publication also contains references to other work in the area of sensitivity analysis for various forming processes with respect to shape and material parameters for isothermal as well as nonisothermal deformations, steady and non-steady processes. One needs to calculate the sensitivity of the material state and geometry at various stages of deformation with respect to infinitesimal changes in each of the design variables (process parameters and die shape). The formulation is based on the differentiation of the governing field equations of the direct problem and development of weak forms for the corresponding field-sensitivity equations that are consistent with the kinematic analysis, material behavior, and the contact/friction subproblem, and the transfer of design sensitivities between meshes during remeshing. Different schemes of design differentiation can be envisioned depending on the level at which the design differentiation is performed. The field-sensitivity equations are linear and can be efficiently solved simultaneously with the solution of the direct-deformation problem. To avoid issues of nondifferentiability of the contact conditions, appropriate regularizing assumptions are introduced. Finite-dimensional gradients of objective functions are then computed using the results of the shape-sensitivity analysis. The accuracy and effectiveness of the method are demonstrated with representative extrusion and forging problems. The results of the continuum sensitivity analysis are validated by a comparison with those obtained by finite-difference approximations (i.e., using the solution of a perturbed deformation problem).

The procedure has been extended to multistage forming processes that involve the computation of both shape as well as parameter sensitivities (Ref 46). However, unlike single-stage, shape-sensitivity analysis, where the initial workpiece shape depends explicitly on shape design variables, the intermediate preform shape in a generic forming stage of a multistage process depends implicitly on the design variables (nonshape parameters) that define the processing history of the intermediate preform. Using analytical gradients instead of numerical derivatives, ANSYS (Ref 53) has currently introduced the calculation of the high-order derivatives of displacement, stresses, and reaction forces in a product called FEM Explorer. It provides the first-order sensitivities, but also the full Taylor expansion. The main issue is to reduce computational time by using analytical derivatives as opposed to finite differencing.

Conclusions

The material modeling methods and finite-element modeling techniques developed in the 1980s and 1990s for metalworking processes are accepted, routine practices in the automotive, aerospace, and other industries. These process models have reduced shop-floor trial and error. The next step is to reduce modeling trial and error and use optimization tools to obtain a process and product that is not just acceptable but optimal. In this chapter, the application of MDO techniques to two manufacturing processes, forging and heat treatment, has been described in detail. The forging optimization system minimizes the total forging cost subject to the constraints of satisfying property requirements and equipment capability. The heat treatment optimizer results in a part optimized for properties and stable dimensions both during machining and during engine service.

The optimization method was formulated as a parametric geometry and detailed forging and heat treatment analysis based problem. An averaging scheme was introduced to compute global constraint functions that are smooth and work with unstructured finite-element meshes. The methodology was demonstrated on the isothermal and nonisothermal forging shape and heat treatment optimization of turbine disks. The results showed both significant process and cost improvements. The results were obtained on workstations in a reasonable period of time and with minimum user interaction. The upfront optimization effort is small compared to the manufacturing savings it results in.

The optimization formulation was incorporated into an integrated system involving three commercial software packages: iSIGHT, Unigraphics, and DEFORM. A general procedure was developed to noninvasively integrate external CAD and CAE systems for geometry and detailed analysis-based optimization. iSIGHT was used as the integration framework. A client-server architecture was established for the iSIGHT and CAD integration so that iSIGHT can drive the CAD system through interprocess communication to perform geometry operations such as updating models, computing geometry properties, exporting geometry for mesh generation, monitoring shape modification, and so forth. The noninvasive coupling of a CAE tool such as DEFORM with a general-purpose optimization software like iSIGHT permits the user to construct a variety of objective functions and define the appropriate constraints and design variables to optimize special-purpose customized processes. The examples given here merely serve as illustrations of the power of optimizing tools.

Material models for microstructure and property evolution and material behavior need to be used in conjunction with optimization algorithms to optimize workability, mechanical properties, achieve efficient material flow, and precise dimensional control. MDO can overcome the multitudinous challenges associated with achieving simultaneous microstructure/property control, part quality (durability), shorter delivery time, and reduced manufacturing costs. It is in the interest of the component supplier industries to adopt MDO to help them increase their productivity.

Even after a process design has gone through computer analysis and optimization, some shop floor tweaking of the process is often necessary. Thermal shrinkage/expansion, lubrication, die deformations, material property variability, and other uncertainties that cannot be incorporated into a deterministic process model require fine tuning the process on the shop floor. Computer analysis is a cost-effective way to weed out poor designs that will not pass a manufacturing trial or will pass only after a great deal of effort. This is because the costs of computer analysis have dropped several orders of magnitude in the past decades, while the costs of labor, plant space, and other factors have increased. The costs of computer analysis will continue to drop an order of magnitude every five years, while the costs of shop-floor trials will continue to rise. Someday, it will be cost effective to optimize die designs so that they are less sensitive to the process variables over which one has limited control. After this happens, when the process is put into production, it will result in one good part after another, even with the inherent variations in the process parameters over which one has little or no control. The potential cost savings will be large. Deterministic models current at the time of publication assume that the material data, operating conditions, and other parameters are well specified. In reality, many defects and part rejections occur due to uncertainties in some or all of these parameters, and the computer models need to incorporate these statistical data variations in order to provide confidence levels on the results.

Most of the real problems are inverse in nature; that is, they involve the determination of conditions that will result in a given final situation. Analysis tools are direct in nature; that is, they tell the user what the final outcome will be for a given set of conditions. Instead of manually running direct simulations for a range of given conditions, the optimization tools provide an alternative to determine the conditions that will provide not just an acceptable but an optimal product. The quantity to be optimized (the objective function), the independent parameters (the design variables), and the bounds within which the user can operate (optimization constraints) can all be decided by the user. These are the advantages of MDO.

Future Work

Future work should be directed at transitioning a robust, user-friendly, fast-acting tool to a production environment for routine application. Some of the items to be considered are:

- More comprehensive optimization of manufacturing processes:
  a. Inclusion of coupled die-stress analysis with die life/wear as a constraint
  b. Generalized formulation of multiple-step forging operations
  c. Exploration of a larger design space by incorporating more geometric parameters as design variables
  d. Incorporation of additional optimization constraints, for example, material-dependent optimization strategy; develop processing windows and microstructural models for various materials in terms of strain, strain-rate, temperature constraints used by the optimizer
  e. Capture manufacturing equipment constraints more realistically and optimize parameters only to the tolerance to which they can be controlled and integrate with process control
  f. Incorporate soft versus hard constraint differentiation: some constraints can be approximately satisfied (soft) as opposed to others that have to be satisfied (hard)
  g. Heat treatment: include shape optimization, more generalized heat transfer coefficient dependence on temperature/time, quench crack criteria, and coupled thermal/stress analyses optimization
  h. Conduct more extensive applications and validation
- Methods to reduce computational effort:
  a. Subgroup design variables: establish parametric relationships in CAD system to reduce the total number of design variables to be optimized.
  b. Screen shapes in CAD system: identify the shape design variables that have the most significant influence on forging weight and ignore the low-influence variables for optimization.
  c. Use approximate models and response surfaces to speed up computations—as optimization problems grow in size, algorithms such as sequential quadratic programming require large amounts of computer memory. The optimizer itself begins to use considerable computer time, relative to the function and gradient computations. Several methods are in the research stage to overcome such problems, and these need to be more robust. One strategy is to identify potential designs with screening DOE runs in a "large" design space and then zoom into smaller design spaces, rerun DOEs, construct response surfaces, validate results with extra runs, and conduct optimization on response surfaces.
  d. Develop adaptive meshing to reduce number of finite-elements without sacrificing accuracy.
  e. Improve optimization robustness: discrete optimization, other optimization techniques.
  f. Conduct rapid optimization on a coarse finite-element mesh followed by fine tuning on a fine mesh using the coarse mesh results as a good starting point.
  g. Use response surface mapping, sensitivity coefficients in conjunction with design of experiments and design for six sigma methodologies.
- CAD related:
  a. Set up initial geometry to ensure optimization compatibility. The die designer should use feature-based parametric modeling, keeping in mind the optimization requirements when defining the die geometries. This will ensure smooth geometry transition to optimization and simplify the task of setting up the optimizer. The way geometries are normally created, some of the effort has to be redone to convert the geometry from the die designer into optimization usable format.
  b. Check for geometry errors in CAD-FEM translation.
  c. Use robust parametric CAD representation of geometries (especially in 3D).
- Programmatic:
  a. Perform seamless integration of various tools.
  b. Improve user-friendliness (graphical user interface).
  c. Keep updated with latest releases of optimization, CAD and CAE tools.
- Communication:
  a. Incorporate new interprocess communication mechanisms to improve data passing between processes.
  b. Develop systems to permit distributed computing on a network using diverse computer platforms using the latest web-based technology.
- Extend to other manufacturing processes:
  a. Machining: minimize number of machining passes for acceptable distortion.
  b. Inertia welding: minimize energy and material requirements.
  c. Inspection: minimize time for sonic and number of scans; include sonicability as a geometric constraint during forging optimization.
  d. Determination of hard to measure materials properties. Use the materials properties as the design variables and their optimal values are obtained by minimizing the deviations of modeling predictions from the measured values of selected response variables.
- Develop fast-acting low-fidelity optimization models for preliminary design trade-offs for new parts.

As these problems and issues are addressed, one should see multidisciplinary optimization being applied to an increasingly broader range of processes, first in isolation to each process and then in a coupled manner to interacting processes. The ultimate goal is integrating the simultaneous optimization of all the steps involved in product design and manufacturing as a multidisciplinary optimization problem. This will result in the overall cost and performance optimization within an acceptable turnaround time.

ACKNOWLEDGMENT

The development of forging and heat treatment optimization techniques was jointly funded by DARPA through the RaDEO (Rapid Design Exploration and Optimization) project and the General Electric Company. Engineous Software Inc. (ESI) was a team member of this project. The author would like to thank Peter Finnigan, Ramesh Gambheera, Beichang He, Peter Röhl, and Youdong Zhou at the GE Global Research Center for many helpful discussions and their contributions to this project.

REFERENCES

1. Proceedings of AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Optimization, 6–8 Sept 2000
2. S. Kodiyalam, V. Kumar, and P.M. Finnigan. Constructive Solid Geometry to Three-Dimensional Structural Shape Optimization. *AIAA J.,* Vol 30 (No. 5), 1992, p 1408–1415
3. E. Hardee, K.-H. Chang, K.K. Choi, X. Yu, and I. Grindeanu, A CAD-Based Design Sensitivity Analysis and Optimization for Structural Shape Design Applications, *Proc. Sixth AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization* (Bellevue, WA), 4–6 Sept 1996
4. L. Fourment and J.L. Chenot, Optimal Design for Non-Steady-State Metal Forming Processes—I. Shape Optimization Method, *Int. J. Numer. Methods Eng.,* Vol 39 (No. 1), 1996, p 33–50
5. L. Fourment and J.L. Chenot, Optimal Design for Non-Steady-State Metal Forming Processes—II. Application of Shape Optimization in Forging, *Int. J. Numer. Methods Eng.,* Vol 39 (No. 1), 1996, 51–66
6. E. Wright and R.V. Grandhi, A Shape Optimization Technique for Controlling Deformation Parameters in Forging, *Proc. 38th AIAA/ASME/ASCE/AHS/ASC Structural, Dynamics, and Materials Conference* (Kissimmee, FL), 7–10 April 1997, p 1404–1414
7. Z. Gao, E. Wright, and R.V. Grandhi. Thermo-Mechanical Sensitivity Analysis for Preform Design in Metal Forming, *Proc. 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization* (St. Louis, MO), 2–4 Sept 1998, p 1019–1029
8. D.J. Powell, "Inter-GEN, A Hybrid Approach to Engineering Optimization," Technical Report, General Electric Co., Feb 1991
9. *iSIGHT Developer's Guide Version 2.2.,* Engineous Software, Inc., Raleigh, NC, 1997
10. *Engineous User Manual,* General Electric Corporate Research and Development, Schenectady, NY, 1995
11. H. Lee, S. Goel, et al., "Toward Modeling the Concurrent Design of Aircraft Engine Turbines," presented at the International Gas Turbine and Aeroengine Congress and Exposition (Cincinnati, OH), May 1993
12. *DEFORM User's Manual,* Scientific Forming Technologies Corp., Columbus, OH, 1999
13. *ABAQUS/Standard User's Manual,* Hibbit, Karlsson, and Sorensen, Inc., Pawtucket, RI, 1997
14. MARC manuals, MARC Analysis Research Corporation, Palo Alto, CA, 1994
15. V. Kumar, M.D. German, and S. Srivatsa, "Design Optimization of Thermomechanical Processes with Application to Heat Treatment for Turbine Disks," presented at the Manufacturing International Conference (Atlanta, GA), 1990
16. *Unigraphics User Function Programming Manual Version 11,* Electronic Data Systems Corp., Unigraphics Div., St. Louis, MO, 1996
17. D. Libes, *Exploring Expect: A Tcl-Based Toolkit for Automating Interactive Programs,* O'Reilly and Associates, 1995
18. G.N. Vanderplaats, *ADS—A FORTRAN Program for Automated Design Synthesis Version 3.00,* VMA Engineering, March 1988
19. B. He, P.J. Röhl, R. Irani, S. Thamboo, and S. Srivatsa, CAD/CAE Integration with Application to the Optimal Design of Turbine Disks, *Proc. 39th AIAA/ASME/ASCE/AHS/ASC Structure, Structural Dynamics and Materials Conference* (Long Beach, CA), April 1998, p 2742–2749
20. B. He, Y. Zhou, R. Gambheera, and S. Srivatsa. "Turbine Disk Forging Process Optimization," 1999 ASME Design Engineering Technical Conf. (Las Vegas, NV), 12–16 Sept 1999
21. B. He, Y. Zhou, R. Gambheera, and S. Srivatsa. "A Practical Approach to Forging Optimization," paper AIAA-2000-4940, AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Optimization, 6–8 Sept 2000
22. J.B. Yang and W.T. Wu, Scientific Forming Technology Corporation, personal communication of unpublished work
23. S.K. Srivatsa, GE Aircraft Engines, unpublished work
24. A. Srikanth and N. Zabaras, Preform Design and Shape Optimization in Metal Forming Processes, *Comput. Methods Appl. Mech. Eng.,* Vol 190, 2000, p 1859–1901
25. A. Srikanth and N. Zabaras, A Gradient Based Optimization Approach for the Design of Single and Multi-Stage Metal Forming Processes, *Proc. ASME Manufacturing Engineering Division,* R.J. Furness, Ed., MED-Vol 11, 2000, p 495–507; presented in the Symposium on Advances in Metal Forming, International Mechanical Engineering Congress and Exposition (Orlando, FL), 5–10 Nov 2000
26. G. Zhao, E. Wright, and R.V Grandhi, Preform Die Shape Design in Metal Forming Using an Optimization Method, *Int. J. Numer. Methods Eng.,* Vol 40, 1997, p 1213–1230
27. J.S. Chung and S.M. Hwang, Application of a Genetic Algorithm to the Optimal Design of the Die Shape in Extrusion, *J. Mater. Process. Technol.,* Vol 72, 1998, p 69–77
28. G. Zhao, Z. Zhao, T. Wang, and R.V. Grandhi, Preform Design of a Generic Turbine Disk Forging Process, *J. Mater. Process. Technol.,* Vol 84, 1998, p 193–201
29. J.S. Chung and S.M. Hwang, Application of a Genetic Algorithm to Process Optimal Design in Non-Isothermal Metal Forming, *J. Mater. Process. Technol.,* Vol 80–81, 1998, p 136–143
30. M.S. Joun and S.M. Hwang, Die Shape Optimal Design in Three Dimensional Shape Metal Extrusion by the Finite Element Method, *Int. J. Numer. Methods Eng.,* Vol 41, 1998, p 311–335
31. S. H. Chung and S.M. Hwang, Optimal Process Design in Non-Isothermal Non-steady Metal Forming by the Finite Element Method, *Int. J. Numer. Methods Eng.,* Vol 42, 1998, p 1343–1390
32. S.M. Byon and S.M. Hwang, Process Optimal Design in Non-Isothermal Steady State Metal Forming by the Finite Element Method. *Int. J. Numer. Methods Eng.,* Vol 46, 1999, p 1075–1100
33. S. Roy, S. Ghosh, and R. Shivpuri, Optimal Design of Process Variables in Multipass Wire Drawing by Genetic Algorithms, *J. Manuf. Sci. Eng.,* Vol 118, 1996, p 244–251
34. G.E. Totten, C.E. Bates, and N.A. Clinton, *Handbook of Quenchants and Quenching Technology,* ASM International, 1993, p 35–68
35. M. Batista and F. Kosel, "Sensitivity Analysis of Heat Treatment of Steel," AIAA-96-4152, 1996, p 1452–1460
36. R. Karthikeyan, P.R.L. Narayanan, and R.S. Naagarazan, Heat Treatment Optimization for Tensile Properties of Al/SiCp Metal Matrix Composites Using Design of Experiments, Process, *Fabr. Adv. Mater.,* Vol 5, p 703–711
37. A. Saigal and G. Leisk, Taguchi Analysis of Heat Treatment Variables on the Mechanical Behavior of Alumina/Aluminum Metal Matrix Composites, *Compos. Eng.,* Vol 5 (No. 2), 1995, p 129–142
38. P.J. Röhl and S.K. Srivatsa. A Comprehensive Approach to Engine Disk IPPD, *Proc. 38th AIAA/ASME/ASCE/AHS/ASC Structural, Dynamics, and Materials Conference and Exhibit and AIAA/ASME/AHS Adaptive Structures Forum* (Kissimmee, FL), 7–10 April 1997, p 1250–1257
39. T.C. Tszeng, W.T. Wu, and J.P. Tang, Prediction of Distortion During Heat Treating and Machining Processes, *Proc. 16th ASM Heat Treating Society Conf. Exposition* (Cincinnati, OH) 1998, p 9–15
40. J.V. Beck, B. Blackwell, and C.R. St. Clair, Jr., *Inverse Heat Conduction: Ill-Posed Problems,* John Wiley & Sons, 1985
41. Z. Li, R.V. Grandhi, and R Shivpuri, Optimum Design of Heat Transfer Coefficient During Gas Quenching Process

by Using Response Surface Method, *Int. J. Mach. Tools Manuf.,* Vol 42, 2002, p 549–558
42. J.B. Yang and W.T. Wu, Scientific Forming Technologies Corp., personal communication of unpublished work
43. J.J. Tsay, J.E.B. Cardosa, and J.S. Arora, Nonlinear Structural Design Sensitivity Analysis for Path Dependent Problems, Part 2: Analytical Examples, *Comput. Meth. Appl. Mech. Eng.,* Vol 81, 1990, 209–228
44. R.H. Myers and D.C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments,* John Wiley & Sons, 1995, p 641–653
45. S. Badrinarayanan and N. Zabaras, A Sensitivity Analysis for the Optimal Design of Metal Forming Processes, *Comput. Meth. Appl. Mech. Eng.,* Vol 129, 1996, p 319–348
46. N. Zabaras, Y. Bao, A. Srikanth, and W.G. Frazier, A Continuum Lagrangian Sensitivity Analysis for Metal Forming Processes with Application to Die Design Problems, *Int. J. Numer. Meth. Eng.,* Vol 48, 2000, p 679–720
47. Z.Y. Gao and R.V. Grandhi, Sensitivity Analysis and Shape Optimization for Preform Design in Thermo-Mechanical Coupled Analysis, *Int. J. Numer. Meth. Eng.,* Vol 45, 1999, p 1349–1373
48. H.J. Antunez and M. Kleiber, Sensitivity Analysis of Metal Forming Processes Involving Frictional Contact in Steady State, *J. Mater. Process. Technol.,* Vol 60, 1996, p 485–491
49. M.A. Ghouali, G. Duvaut, S. Ortola, and A. Oster, Local Analytical Design Sensitivity Analysis of the Forging Problem Using FEM, *Comput. Meth. Appl. Mech. Eng.,* Vol 163, 1998, p 55–70
50. I. Doltsinis and T. Rodic, Process Design and Sensitivity Analysis in Metal Forming, *Int. J. Numer. Meth. Eng.,* Vol 45, 1999, p 661–692
51. M. Kleiber and H.J. Antunez, T.D. Hien, and P. Kowalczyk, *Parameter Sensitivity in Non-Linear Mechanics—Theory and Finite Element Computations,* John Wiley & Sons, 1997
52. J.J. Tsay and J.S. Arora, Nonlinear Structural Design Sensitivity Analysis for Path Dependent Problems. Part 1: General Theory, *Comput. Meth. Appl. Mech. Eng.,* Vol 81, 1990, p 183–208
53. G. Venter, R.T. Haftka, and J.H. Starnes, Construction of Response Surfaces Approximations for Design Optimization, *AIAA J.,* 36 (No. 12), 1998, p 2242–2249
54. Ansys 7.0 Manuals, Ansys Inc., 2002
55. S. Akkaram, A. Makinde, Y. Zhou, and S.K. Srivatsa, personal communication

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

N

Q

R

T

W

Y

Z