Inelastic Analysis of Structures

Milan Jirásek
Swiss Federal Institute of Technology at Lausanne, Switzerland

Zdeněk P. Bažant
Northwestern University, Evanston, Illinois, USA

JOHN WILEY & SONS, LTD

 John Wiley & Sons, Ltd
Baffins Lane, Chichester,
West Sussex, PO19 1UD, England

National 01243 779777
International (+44) 1243 779777

e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on http://www.wiley.co.uk or http://www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158-0012, USA

WILEY-VCH Verlag GmbH
Pappelallee 3, D-69469 Weinheim, Germany

John Wiley & Sons Australia, Ltd
33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Canada) Ltd, 22 Worcester Road
Rexdale, Ontario, M9W 1L1, Canada

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01,
Jin Xing Distripark, Singapore 129809

Library in Congress Cataloging-in-Publication Data

Jirásek, Milan.
Inelastic analysis of structures/Milan Jirásek, Zdenek P. Bazant. – 1st ed.
p. cm.
Includes bibliographical references and index.
1. Plastic analysis (Engineering) 2. Structural dynamics. I. Bazant, Z. P. II. Title.

TA652.J57 2001
624.1'71 – dc21 2001017867

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-471-98716-6

Produced from LaTeX files supplied by the author and processed by Laser Words, (India) Ltd.
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire.
This book is printed on acid-free paper responsibly manufactured from sustainable forestry, in which at least two trees are planted for each one used for paper production.

Cover image: Part of a buckled steel beam. Reproduced by permission of Professor Carlo Castiglioni, Politecnico di Milano, Italy.

Spine inset image: A precast reinforced concrete column near Izmit in Turkey following the August 1999 Kocaeli earthquake. Reproduced by permission of Professor Marc Badoux, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

To Karel Rektorys, Professor Emeritus at the Czech Technical University in Prague, who taught us mathematics and inspired our deepest respect and admiration.

Contents

Preface

Our main objective in writing this book has been to provide a textbook for graduate courses on plasticity, with some ramifications to time-dependent inelastic behavior. In our selection of the topics and the sequence of their exposition, we put a slight emphasis on structural engineering applications. There is nevertheless plenty of material for using the book in postgraduate courses in geotechnical, mechanical, aerospace, naval, petroleum and nuclear engineering. We assume the background at the level of a B.S. degree in civil or mechanical engineering.

Plasticity has already been the subject of many books. So why another one? We hope to provide a book that is unique in many respects. It has been our intention to fill many needs that are not quite met by other books. Being considerably larger than a textbook for a single course, our book provides both a systematic exposition of the fundamentals of plasticity, and an up-to-date introduction to most of the advanced subjects. The courses with the coverage specified below could not be taught fully from some other existing book. We proceed from simple to complex, and introduce examples before generalizing. We try to be systematic and mathematically rigorous while striving, above all, for clarity. We avoid an artificially formalistic presentation that hardly achieves more than impressing by mathematical sophistication. Our book features complete and rigorous mathematical derivations of all the results. Some derivations are more simple and others more rigorous than those found in the previous textbooks. Despite being mainly a textbook, in the advanced chapters our book also covers most of the 'hot' topics of current research, and contains some new research results. A set of problems for the student is included at the end of most chapters. Both simple and hard problems are suggested, the hard ones being marked by an asterisk. It is planned to make the solutions available on the web site `http://www.wiley.co.uk/inelastic`, which will also contain some additional useful information, such as a set of links to sites providing software for the solution of linear programming problems.

We also include a set of six appendices, four of which review, for the reader's convenience, the fundamentals of linear elastic analysis and the mathematical background of linear programming, and two give information on specialized topics the code-type prediction model for creep of concrete and the size effect engendered by softening in plastic hinges.

A special feature, which is not encountered in the basic texts on plasticity and is found only in specialized treatises, is a thorough exposition of the plasticity aspects of concrete and reinforced concrete, including the basic principles of limit state design. Concrete, of course, is not a plastic material *per se*, but plasticity concepts such as the yield surfaces and plastic potentials form a necessary part of models that extend

plasticity to damage. Besides, the theory of plasticity is well suited to reinforced concrete structures that fail by the yielding of steel reinforcement. Similar comments can be made about our inclusion of plasticity models for soils.

To keep with the nature of most civil engineering applications, as well as to make the student's entry into the subject easier, the first two among five parts of our book are restricted to beam structures whose stress state may be simplified as uniaxial. Considerable attention is devoted to shakedown, another classical subject particularly important for structural engineering, but rarely treated consistently in textbooks. The classical topics at the margins of plasticity theory, such as the optimum design and linear programming, are included in our coverage. After digesting the basic concepts in the context of uniaxial stress, the students will find it easier to follow, in the third part, the extension of limit analysis to structures under multiaxial stress.

For the benefit of advanced doctoral students and postdoctoral researchers, we include in the last two parts of the book a number of advanced subjects normally not seen in basic textbooks – numerical algorithms, thermodynamic aspects, plasticity in finite strain, multisurface plasticity, anisotropic plasticity, nonlocal and gradient models for plasticity with strain softening and size effects, viscoplasticity and rate effects, microplane constitutive models, and vertex effects. We also provide a brief survey of polycrystal plasticity. With this scope, we hope to have covered a major part of what today constitutes the modern theory of plasticity. Expositions of the dislocation theory as the micromechanical basis of plasticity, dynamic plasticity, plastic buckling and bifurcations, plasticity of shells and constitutive properties of plastic composites could not be accommodated within the scope of this book.

Another special feature of our book is the inclusion of two chapters (among 29) on the creep of concrete and its effects in structures. Although this kind of inelastic behavior is very important for the durability of civil engineering infrastructure and sometimes affects the safety as well, most structural engineering curricula unfortunately do not have room for a full course devoted to this subject and, deplorably, graduate students leave the university without acquiring any knowledge of concrete creep. Our coverage of this subject provides a feasible compromise – an exposition brief enough not to lose the emphasis on plasticity yet sufficient to acquaint the student with the basic results needed for structural design. Due to space limitations, the treatment of creep is nevertheless much less systematic than that of plasticity, and most intricacies of this vast subject are inevitably skipped.

Our book can serve as a textbook for courses of several types:

- *A Quarter-Length First-Year Graduate Course* with a slight civil engineering emphasis may consist of the following chapters and sections: **1**, **2**, **4–6**, **7.5**, **8.2**, **9**, essentials of **10** (without proofs), **11**, **12.1**, **12.2**, **13.1–13.3**, **14.1.1**, **15.1**, **15.2.1–15.2.3**, **16.1**, **17.1–17.3**, **18**, **19.1**, **19.2.1–19.2.2**, only essential ideas and graphs from **19.5**, **28.1–28.3**, **28.4** without proof, and selections from **29.2.2**.
- *A Quarter-Length First-Year Graduate Course* with a slight mechanical engineering emphasis may consist of the following chapters and sections: **1–7**, **13**, **15**, **16.1**, **16.3**, **17.1**–**17.3**, **18.1–18.3**, **19.1**, **19.2.1**, **19.5**, **20.1**, **25.4**, **27.1**.
- *A Semester-Length First-Year Graduate Course* in structural engineering may fully cover chapters **1**–**19** and **28** and sections **29.1** and **29.2**. In mechanical engineering, one may omit **8**, **11**, **12.3** and **14** and add **20.1–20.3**, **22.1**, **22.2.1**, **22.3.1**, **22.3.2**, **25.4** and **27.1**.

- *A Second Course on Plasticity for Doctoral Students* in structural engineering may cover chapters and sections **20**, **21**, **22.1**, **22.2.1**, **22.2.2**, **22.3.1**, **22.3.2**, **22.4**, **25.2** and **26.1–26.3**. In mechanical engineering one may omit **21**, **25.2.4** and **25.2.5** and add **25.3**, **25.4** and **27.1**. In a computationally oriented course, the entire chapter **22** can be covered.
- *A Short Course for Post-Doctoral Researchers and Advanced Doctoral Students* may start with the three-dimensional formulation of plasticity in chapters **15** and **20** and include the advanced topics in chapters **22–24**. A course with emphasis on structural engineering may also cover chapters and sections **21**, **25.2**, **26.1–26.3**, **28.1**, **28.2** and **29.3**, while a course with emphasis on mechanical engineering may instead cover chapters **26**, **27** and **25** without sections **25.2.4–25.2.5** and **25.4.3**.

The first course outline listed above has been used by the second author in his course on Inelastic Structural Analysis, which he has been regularly teaching at Northwestern University since 1970. The lecture notes that he had prepared for this course during the 1970s served as the point of departure for writing this book – an arduous effort that began in earnest in 1993, right after the first author completed his doctoral study at Northwestern University. The presentations of various advanced subjects in this book have been tried in a number of short courses or advanced graduate courses taught at various institutions[1]. The book was completed during the first author's Visiting Scholar appointment at Northwestern University in the summer of 2000, and the second author's Visiting Professor appointment at the Swiss Federal Institute of Technology at Lausanne (EPFL) in March 2001.

We would like to express our thanks for valuable comments and discussions on the drafts of various chapters to Giulio Maier, professor at Politecnico di Milano; Mircea Z. Cohn, Professor Emeritus at the University of Waterloo; Zuzana Dimitrovová, researcher at the Technical University of Lisbon; Andrzej Truty, associate professor at the Cracow University of Technology; Cino Viggiani, professor at Université Joseph Fourier, Grenoble; and Bořek Patzák and Simon Rolshoven, colleagues of the first author at EPFL.

The first author would like to express his deep gratitude to his wife Vlasta and children Anna and Jakub for their patience and support during the countless hours devoted to the work on this book. He would also like to acknowledge the intellectual stimulation provided to him during his undergraduate and early graduate studies at the Czech Technical University in Prague (ČVUT) by professors Karel Rektorys, Jiří Šejnoha and Zdeněk Bittnar. The outstanding work environment at EPFL has been deeply appreciated.

The second author would like to thank a number of graduate research assistants, postdoctoral researchers, visiting scholars and external collaborators for helping him with various researches that are reflected in this book[2]. The sponsorship of these

[1] First author: Swiss Federal Institute of Technology, Lausanne; Czech Technical University, Prague; Universidad Politecnica de Catalunya, Barcelona; Universität Stuttgart; Rheinisch-Westfälische Technische Hochschule, Aachen; and Universität der Bundeswehr, Munich. Second author: Politecnico di Milano; Swiss Federal Institute of Technology, Lausanne; Royal Institute of Technology, Stockholm; Ecole des ponts et chaussées, Paris; E.N.S. de Cachan; Universidad Politecnica de Catalunya, Barcelona; Universität Stuttgart; University of Mexico; Technische Universität Wien; University of Luleå, Sweden; and University of Palermo, Italy.

[2] M.D. Adley, S. Baweja, M. Brocca, F.C. Caner, I. Carol, L. Cedolin, T.P. Chang, G. Cusatis, J.K.

researches under grants from various U.S. funding agencies, especially from the National Science Foundation, Office of Naval Research, Air Force Office of Scientific Research, U.S. Army Corps of Engineers and Department of Energy, is gratefully acknowledged. The second author wishes to thank his colleagues for many stimulating discussions, and Northwestern University for providing an environment conducive to scholarly pursuits. Last but not least, he wishes to thank his wife Iva for her moral support, and express gratitude to his father, Zdeněk J. Bažant, Professor Emeritus of Foundation Engineering at ČVUT, and to his grandfather Zdeněk Bažant, late Professor of Structural Mechanics at ČVUT, for having stimulated his passion for structural mechanics and engineering.

M.J. and Z.P.B
Lausanne and Evanston
April 2001

Kim, S.S. Kim, F.B. Lin, G. de Luzio, A.M. Marchertas, B.H. Oh, J. Ožbolt, G. Pijaudier-Cabot, S. Prasannan, P.C. Prat, J. Sládek, M.K. Tabbara, T. Tsubaki, Y. Xi, and Y. Xiang.

Introduction

The roots of some elementary ideas of the theory of plasticity can be traced back over three and half centuries. In Galileo's (1638) calculation of the collapse load of a cantilever, one may discern the assumption of a uniform distribution of tensile stresses over the cross section, even though the assumption of a concentrated compression force at the compression face was far from realistic. About a century later, Giovanni Poleni discussed the safety of Michelangelo's dome of Saint Peter's cathedral in a manner in which one could detect the ideas of the static approach to limit analysis (Benvenuto, 1991). In the debates of the stability of masonry arches, vaults and domes among La Hire, Boscovich, Lamé, Clapeyron, Fourier and Pauker, continuing into the early nineteenth century, one could also perceive various elementary ideas of plastic analysis (Benvenuto, 1991). The first realistic and almost complete static analysis of failure, along with the concept of plastic slip and yield condition, is found in Coulomb's (1776) study of earth-retaining walls of military fortifications.

Various elementary ideas of plastic deformation and failure, and the reduction of buckling loads gradually emerged throughout the nineteenth century in the studies of pioneers such as Lüders (1854), Tresca (1868), de St. Vénant (1870), Lévy (1870), Rankine (1876), Bauschinger (1881), Considère (1891), Engesser (1895), Hartmann (1896) and Mohr (1900). The static theorem of limit analysis was anticipated (Maier, Carvelli and Cocchetti, 2000) in the work of Rankine in 1859 and Kötter in 1899, and its intuitive enunciations can be found in the work of Kazinczy (1914) and the inaugural lecture of Kist (1917). During the first quarter of the twentieth century, the basic concepts, such as the yield surfaces, flow rules, slip lines, and plastic friction appeared, principally in the works of von Kármán (1909), von Mises (1913), Hencky (1924) and Reuß (1930). An important milestone was the resolution of the torsional problem (Nádai, 1923) and indentation problem (Hencky, 1923; Prandtl, 1923). The materials science foundation of metal plasticity in the dislocation theory was laid by Taylor (1934) and others.

The static and kinematic theorems of limit analysis were in general first proven in a Russian conference proceedings article by Gvozdev (1938), long unknown in the West. At about the same time, the static shakedown theorem was first proven by Melan (1936), being anticipated a few years earlier by himself and Bleich (1932). The fact that Melan's theorem implies the static theorem of limit analysis was recognized much later.

The general concepts of plasticity, which are expounded in Parts I–III of this book and comprise the general multiaxial stress-strain relations, normality and convexity, maximization of plastic energy dissipation, limit state theorems, shakedown, optimum design, plastic hinges, yield line theory of plates and slip line theory, were established

shortly after World War II by Shanley (1947), Hill (1950), Drucker (1950), Greenberg and Prager (1951), Prager and Hodge (1951), Symonds and Neal (1951), Koiter (1953b), etc.; see Nádai (1950a) and Prager (1959) for additional references.

The second half of the last century was a period of rapid refinement and extensive ramification, which continue at an unrelenting pace until today and are for the most part described in Parts IV and V of this book.

Plasticity concepts began to impact structural analysis and design at the beginning of the last century, although design codes based on limit states were not instituted until the middle of that century. When subjected to the service loads, structures must generally respond in an elastic manner. A century ago, the standard design approach was to calculate the maximum stress according to the theory of elasticity, and make sure that it would not exceed a certain allowable stress, which was set sufficiently smaller than the material strength or yield limit. Later it was recognized that in most design problems (fatigue of metals excepted), this approach often leads to designs that are wasteful to varying degrees. The reason is that only some structures fail at a load at which the material strength or yield limit is exhausted at one point of the structure. Many structures redistribute stresses in such a way that the structure fails at a higher load, sometimes only a little higher but often a much higher load, which is attained only after a large part of the structure has plasticized. Simply setting the allowable stress value higher is not a solution, since the safety of some designs would become inadequate. If the theory of elasticity with allowable stress were still used as the basis of design, many efficient modern structures distinguished by slenderness could not even be built.

A realistic approach to design is to calculate the collapse load of the structure from the minimum expected value of material strength or yield limit, and then make sure that this collapse load would not be exceeded by the actual loads multiplied by a suitable safety factor (which is determined from experience and probabilistic considerations). Depending on the type of material, two different kinds of theories, the first older and more mature than the second, are needed for calculating the collapse load:

- If the material is plastic, as typical of most metals (provided the metal has not been fatigued), then the right approach is the *theory of plasticity*.
- If the material is brittle, then the right approach is either *fracture mechanics*, if the failure is caused by propagation of one or several large cracks, or *damage mechanics*, if the failure is caused by the spread of a zone of cracking or other distributed damage confined to the microscale.

This book deals only with the former.

To help understanding, the first two parts of this book are restricted to structures such as beams, trusses and frames whose stress state may be simplified as uniaxial. The advantage is that the basic concepts and results, such as the limit design theorems, normality and convexity, maximum plastic dissipation and shakedown, are understood more easily. This facilitates understanding of the behavior under multiaxial stress, which is the subject of Part III.

Plastic design of structures requires resolving problems of two basic types:

- Formulation of a realistic material model.
- Calculation of the collapse load if the material model is available.

Both are very rich problems. Most of the first three parts of the book deal with the latter problem, most of the fourth part with the former, and the fifth equally with both.

Although 'brute-force' computational approaches such as the finite element method are nowadays capable of providing numerical answers to most problems of the latter type, much of the present exposition will dwell on analytical or semi-analytical solutions. It is, of course, these solutions that lead to an understanding of the problem, convey insight into the structure behavior, are simple enough to be used in design optimization or probabilistic safety studies, and provide indispensable checks on the correctness of computational approaches.

The constitutive models for plastic materials have proven to be a formidable problem, which has been tackled continuously from the emergence of plasticity until the present. Despite major advances in the past, this is still a very active field of research. Some important phenomena, e.g. the vertex effect, are still ignored by the constitutive models typically used in the current computational practice.

While the first two parts of the book will rest on a simple material model applicable only to uniaxial stress, the third part of the book will expound the basic constitutive models of multiaxial stress, and the fourth part will deal with advanced aspects such as the plastic hardening, anisotropic plasticity, restrictions on constitutive models stemming from thermodynamics, plasticity in large strain, microplane models, models based on the idea of polycrystal, and problems caused by material softening. The fifth part will extend the treatment to time-dependence of inelastic behavior.

As a special feature of this treatise that should be welcome by civil engineers, the fourth part of the book will discuss the plasticity aspects of quasibrittle materials such as concrete. The fifth part will briefly describe the time-dependent inelastic behavior of concrete and its consequences in structures, which are completely different from those of plasticity.

In the aftermath of the excessive enthusiasm of the 1970s which led to the development of plasticity models for all kinds of inelastic response of concrete, this field in recent years has been in a sobering but fruitful period in which the limitations of plasticity have been properly recognized and modeled. Concrete, of course, is not a plastic material, but the theory of plasticity is useful for describing some aspects of its behavior, providing one pays proper attention to the inelastic strain localization engendered by post-peak strain softening, with the inherent size effects and mesh sensitivity of finite element solutions.

Part I

Plastic Analysis of Structures Under Uniaxial Stress

Fundamentals

1

Uniaxial Stress-Strain Relations

The modeling of inelastic behavior of solids is a complex and wide ranging subject. In the first two parts of this book (Chapters 1–14), we will nevertheless need only a few basic facts about the uniaxial stress-strain relations for materials commonly used in civil engineering structures.

The stress-strain diagram of mild steel, the most widely used structural metal, is shown in Figure 1.1(a). As we increase the strain, ε, the stress σ first increases linearly until a certain critical value, the *elasticity limit*, is reached. At that point we see, usually after a small and abrupt 'hump', a rather sharp transition to a horizontal plateau, the level of which is characterized by a stress value σ_0 called the *yield stress*, *plastic stress* or *yield limit*. This yield plateau is much longer than the elastic strain range, but it is eventually terminated at a certain strain after which the stress further increases until the ultimate state is reached.

For many other metals, for example various alloy steels or aluminum alloys, we see no definite yield plateau but a relatively smooth decline of the slope of the uniaxial stress-strain diagram, as shown in Figure 1.1(b). However, even for these metals, the decline of the slope compared to the elastic slope is substantial and, as an approximation, the stress-strain diagram may be treated as one exhibiting a yield plateau.

Thus, for the purposes of structural analysis, we idealize the stress-strain diagrams of all these metals as shown in Figure 1.1(c), which represents the *elastic-perfectly plastic* (or *elastic-ideal plastic*) behavior. The words 'perfect' or 'ideal' refer to the idealized assumption of a perfectly straight and horizontal yield plateau. A similar type of idealized description is used for compression. The stress σ cannot exceed the bounds given by the yield stress in tension, σ_0^+, and the yield stress in compression, $-\sigma_0^-$. Any value of stress satisfying this condition is called *plastically admissible*. For metals, both yield stresses usually have the same magnitude, $\sigma_0^+ = \sigma_0^- = \sigma_0$. The *condition of plastic admissibility* then reads

$$-\sigma_0 \leq \sigma \leq \sigma_0 \tag{1.1}$$

Plastic yielding can take place only if

$$|\sigma| - \sigma_0 = 0 \tag{1.2}$$

which is the uniaxial *yield condition*.

Often the plastic deformations are many times larger than the elastic limit, and in that case the material behavior may be further simplified as *rigid-perfectly plastic*, as shown in Figure 1.1(d). Although for some metals these perfectly plastic idealizations

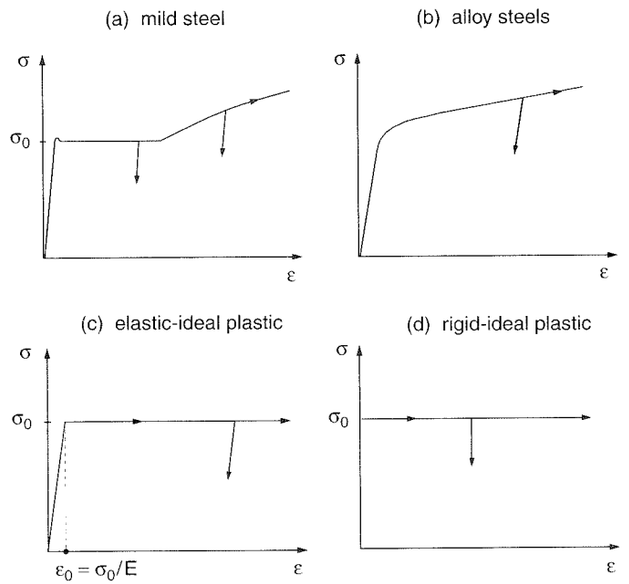


Figure 1.1 Uniaxial stress-strain diagrams: a) mild steel, b) alloy steel, c) elastic-perfectly plastic material model, d) rigid-perfectly plastic material model

may seem to deviate too much from the actual stress-strain diagram, the results of the calculations of the collapse loads of structures are still surprisingly good. The reason is that during collapse, the critical cross sections are at strains much larger than the elastic ones, and so the yield stress gives a close approximation to the actual stress.

On the other hand, more sophisticated constitutive models are needed to take into account the progressive increase of the yield stress after the onset of yielding. This phenomenon is called *hardening*. The material hardens in the sense that the stress needed to induce the plastic flow increases. Note that the current (tangential) material stiffness, defined as the slope of the tangent to the stress-strain diagram, is smaller than the elastic one, and thus the material is softer in terms of stiffness. However, we will reserve the term *softening* for a situation in which the stress decreases with increasing strain, i.e. the tangential stiffness is negative. This behavior is observed in materials experiencing brittle damage due to the growth of microscopic flaws such as voids or cracks and their coalescence into macroscopic fractures. Typical examples of softening materials are rocks, concrete, masonary, ceramics, various composites, and ice.

A suitable approximation of the actual stress-strain diagram is sometimes provided by a model with linear hardening (Figure 1.2(a)) or with linear softening (Figure 1.2(c)). A more realistic description of hardening in metals can be achieved with nonlinear hardening laws (Figure 1.2(b)), e.g. with the Ramberg–Osgood power law discussed in Problem 1.1. Softening curves determined experimentally for quasibrittle materials such as concrete typically exhibit a rather long tail, which can be modeled by a bilinear (Figure 1.2(d)) or exponential softening law (Figure 1.2(e)).

The salient property for which the behavior of a material is called plastic is the *irreversibility of deformation* upon unloading. For ductile materials such as metals or soils, in which the inelastic deformation is mainly due to slip in the microstructure,

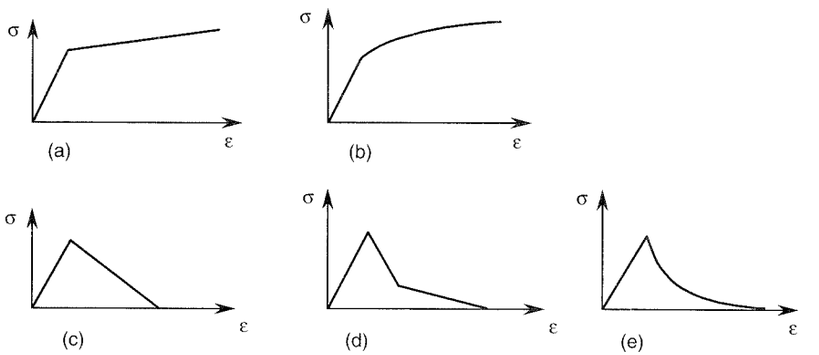


Figure 1.2 Uniaxial stress-strain diagrams: a) linear hardening, b) nonlinear hardening, c) linear softening, d) bilinear softening, e) exponential softening

the slopes of the stress-strain diagram for unloading are essentially the same as the initial (elastic) slope (Figure 1.1). In the context of the small-strain theory, this leads to the additive decomposition

$$\varepsilon = \varepsilon_e + \varepsilon_p \tag{1.3}$$

expressing the total strain, ε, as the sum of the elastic (reversible) strain, ε_e, and the plastic (irreversible) strain, ε_p. The elastic strain is related to the stress by Hooke's law

$$\sigma = E\varepsilon_e \tag{1.4}$$

where E is Young's modulus of elasticity. Equations (1.3) and (1.4) can be combined to give

$$\sigma = E(\varepsilon - \varepsilon_p) = E\varepsilon - E\varepsilon_p \tag{1.5}$$

Note that brittle materials, in which inelastic deformation is mainly due to the growth of microcracks or microvoids, are characterized by a progressive stiffness degradation, and so the unloading slopes are smaller than the initial slope (at sufficiently large strains, stiffness degradation due to the initiation and growth of microvoids is observed even for metals and other ductile materials). An idealized model of this behavior is provided by the continuum damage theory, in which the strain is totally reversible but the elastic modulus decreases (Figure 1.3(a)). In the simplest case, the stress-strain law has the form

$$\sigma = (1-\omega)E\varepsilon = E\varepsilon - E\omega\varepsilon \tag{1.6}$$

where ω is a damage variable growing from 0 (virgin material) to 1 (completely damaged material). Comparing this with (1.5), we note that a plasticity model and a damage model can be tuned up so that they give exactly the same response during loading. It is sufficient to postulate the evolution laws for the plastic strain and for the damage so that the relation $\varepsilon_p = \omega\varepsilon$ is always satisfied. Despite this formal similarity, there are important differences between the two classes of models. For example, during unloading, ε_p and ω remain constant but the total strain ε varies (decreases), and so the constraint $\varepsilon_p = \omega\varepsilon$ cannot be satisfied. For an elastoplastic model, the increments of stress and strain are linked by the initial Young's modulus of elasticity. For a damage model, the material responds elastically; however, with a reduced Young's modulus $E_u = (1-\omega)E$, and no permanent strain is left after complete unloading to

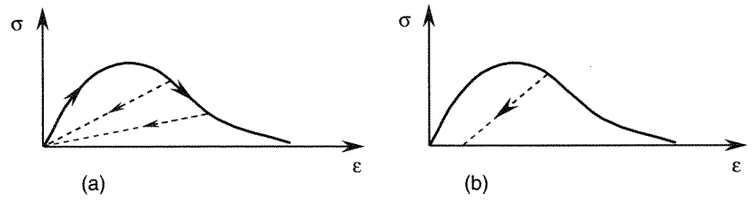


Figure 1.3 Unloading a) to the origin (strain fully reversible), b) with partially irreversible strain

zero stress. The real behavior of materials such as concrete is usually a combination of damage and plasticity. The unloading slopes are smaller than the initial slope, but at the same time a part of the strain is irreversible (Figure 1.3(b)).

In the first three parts of this book (Chapters 1–19), we restrict our attention to perfectly elastoplastic materials. Unless it is stated explicitly that we consider the general case of a material with different behaviors in tension and in compression, we will assume that the tensile and compressive yield stresses have the same magnitude.

During elastic loading (or unloading), the mechanical work done on the material is converted to (or released from) the stored elastic energy.[1] During plastic yielding, a part of this work is dissipated by irreversible plastic processes in the material, usually as heat. For perfectly elastoplastic materials, the *rate of dissipation per unit volume* (or *the dissipation power density*) is given by the product of the stress and the plastic strain rate. Under uniaxial stress, the stress during plastic yielding is either $\sigma = \sigma_0$, and then the plastic strain rate $\dot{\varepsilon}_p$ must be positive, or $\sigma = -\sigma_0$, and then the plastic strain rate $\dot{\varepsilon}_p$ must be negative. In either case, the elastic strain remains constant (see Hooke's law (1.4)). Consequently, $\dot{\varepsilon}_p = \dot{\varepsilon}$ and the rate of dissipation per unit volume is

$$\mathcal{D} = \sigma\dot{\varepsilon}_p = \sigma_0|\dot{\varepsilon}_p| \tag{1.7}$$

The right-hand side of (1.7) depends only on the plastic strain rate and a constant material parameter, and so the dissipation power density is a unique function of the plastic strain rate, $\mathcal{D} = \mathcal{D}(\dot{\varepsilon}_p)$.

Now we make a simple but important observation. Suppose that, at a given point, the material yields under uniaxial stress. Let σ_s be any plastically admissible stress, i.e. a stress satisfying the condition $|\sigma_s| \leq \sigma_0$ (but totally independent of the actual plastic strain rate $\dot{\varepsilon}_p$), then

$$\sigma_s\dot{\varepsilon}_p \leq |\sigma_s|\,|\dot{\varepsilon}_p| \leq \sigma_0|\dot{\varepsilon}_p| = \mathcal{D}(\dot{\varepsilon}_p) \tag{1.8}$$

Consequently, the actual stress σ needed to induce a given plastic strain rate $\dot{\varepsilon}_p$ maximizes the product $\sigma_s\dot{\varepsilon}_p$ among all the plastically admissible stress values σ_s:

$$\mathcal{D}(\dot{\varepsilon}_p) = \sigma\dot{\varepsilon}_p = \max_{\sigma_s\in\bar{\mathcal{E}}}(\sigma_s\dot{\varepsilon}_p) \tag{1.9}$$

where $\bar{\mathcal{E}} = [-\sigma_0, \sigma_0]$ is the set of all the plastically admissible stress states. Its interior, i.e. the open interval $\mathcal{E} = (-\sigma_0, \sigma_0)$, is the *elastic domain* in the stress space.

Statement (1.9) is a special case of the *postulate of maximum plastic dissipation*. We have stated it here for one material point under uniaxial stress, but later we will generalize it to entire bodies under multiaxial stress. In contrast to structures under uniaxial stress, for which this postulate always holds, the generalization to multiaxial stress will require certain assumptions that are violated by some material models.

[1] Strictly speaking, though, the energy balance equation must also take into account a possible exchange of heat; see a more rigorous thermodynamic analysis in Section 23.2.

PROBLEMS

For the following problems, as well as for future use, it is convenient to introduce the *positive part operator*, $\langle . \rangle$, defined by the relations $\langle x \rangle = x$ if $x \geq 0$, $\langle x \rangle = 0$ if $x \leq 0$. Equivalent definitions are $\langle x \rangle = (|x| + x)/2$, or $\langle x \rangle = \max(x, 0)$. Let us also introduce the decomposition of a real variable $x = x^+ - x^-$ into the positive part, $x^+ = \langle x \rangle$, and the negative part, $x^- = \langle -x \rangle$. It is useful to note that $x^+ \geq 0$, $x^- \geq 0$, $x^+x^- = 0$, and $x^+ + x^- = |x|$.

Problem 1.1: A popular model of hardening in metals is the *Ramberg–Osgood power law*. In the original paper (Ramberg and Osgood, 1943), the plastic strain was related by a power law to the stress. However, this would mean that there is no clearly defined yield limit, and an arbitrarily low applied stress always produces some plastic strain. A modified version of the Ramberg–Osgood hardening model preserves a nonvanishing elastic domain by relating the plastic strain to the difference between the current yield stress (after hardening), σ_Y, and the initial yield stress, σ_0; see Lemaitre and Chaboche (1990, p. 167). The Ramberg–Osgood equation reads

$$\sigma_Y = \sigma_0 + K(\varepsilon_p)^m \tag{1.10}$$

where K and m are material parameters. Construct the stress-strain diagram for an alloy with $E = 350$ GPa, $\sigma_0 = 700$ MPa, $K = 900$ MPa and $m = 0.21$. Show also the behavior of the model during unloading. Determine the stress at $\varepsilon = 0.01$ and the corresponding permanent strain after unloading to zero stress.

Problem 1.2: For certain damage models based on equation (1.6), the evolution of damage is driven by the so-called damage energy release rate, $Y = E\varepsilon^2/2$. As damage is irreversible, the current value of ω depends on the maximum value of Y reached in the previous history of the material, denoted as Y_{max}. One of the laws used for concrete (in tension) reads[2]

$$\omega = 1 - \frac{1}{1 + b_1\langle Y_{max} - Y_0\rangle + b_2\langle Y_{max} - Y_0\rangle^2} \tag{1.11}$$

where b_1, b_2 and Y_0 are material parameters. Plot the stress-strain curve for $E = 30$ GPa, $b_1 = 600$ MPa^{-1}, $b_2 = 50 \times 10^3$ MPa^{-2} and $Y_0 = 60 \times 10^{-6}$ MPa, and determine the strength of the material (peak value of stress). Show also the behavior of the model during unloading.

Problem 1.3: Consider a perfectly elastoplastic material with different magnitudes of the yield stress in tension and in compression. The condition of plastic admissibility

[2] It is convenient to use the positive part operator because the initial stage of loading, for which $Y_{max} \leq Y_0$ and $\omega = 0$, can be included in the formula without the need for an 'if' statement.

reads $-\sigma_0^- \leq \sigma \leq \sigma_0^+$ where $\sigma_0^- \neq \sigma_0^+$. Generalize equation (1.7) for the rate of plastic dissipation per unit volume to this class of materials.
Hint: It is helpful to use the decomposition $\dot{\varepsilon}_p = \dot{\varepsilon}_p^+ - \dot{\varepsilon}_p^-$.

Problem 1.4: Show that the postulate of maximum plastic dissipation remains valid for materials with different magnitudes of the yield stress in tension and in compression.

2

Plastic Bars and Yield Hinges

Unless stated otherwise, all the subsequent analyses will be based on the elastic-perfectly plastic stress-strain law (usually with equal magnitudes of the yield stress in tension and in compression). More sophisticated laws will be discussed in Chapters 20 and 21 in the context of general three-dimensional constitutive modeling. With the exception of Chapter 24, it is assumed that the displacements are sufficiently small, so that (1) the state of deformation can be described by the standard small-strain tensor, defined as the symmetric part of the displacement gradient, and (2) the difference between the current and initial geometry is negligible, and has no effect on the equilibrium equations. Inertia effects are not considered and, in the first four parts (Chapters 1–26), the material behavior is assumed to be independent of the rate of loading.

2.1 PLASTIC BAR

The stress-strain diagrams discussed in the first chapter describe the behavior on the level of a material point. Structural analysis of skeletal structures, such as trusses, beams, frames or grillages, requires knowledge of the relations between the relevant internal forces and the corresponding generalized strains. For example, the behavior of a truss member is governed by the axial force, S, which depends on the total extension (change of length) of the bar, e. A more common notation for the bar extension would be Δl, but we reserve the symbol Δ for increments during individual loading stages of incremental structural analysis, to be discussed in the next chapter. The symbol e should remind us of strain ε, to which the extension is closely related.

In the elastic range, Hooke's law $\sigma = E\varepsilon$ combined with the obvious relations $\sigma = S/A$ and $\varepsilon = e/\ell$ results in

$$S = \frac{EA}{\ell} e \tag{2.1}$$

where A is the area of the cross section and ℓ is the length of the bar. When the stress reaches the yield limit, σ_0, the bar starts yielding and the stress cannot grow any more. The axial force therefore remains constant and equal to the *plastic axial force* $S_0 = A\sigma_0$. Similar considerations can be taken into account for compression. The resulting axial force-extension diagram differs from the stress-strain diagram only by the scaling of the axes (Figure 2.1).

According to Chapter 1, the dissipation power density is (for uniaxial stress) given by the product of the stress, σ, and the plastic strain rate, $\dot{\varepsilon}_p$. Expressing the strain rate in terms of the extension rate and integrating over the volume of the bar, we get

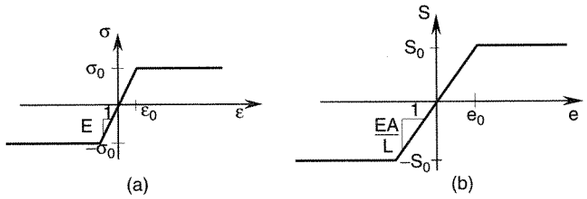


Figure 2.1 a) Stress-strain diagram for an elastic-perfectly plastic material, b) the corresponding force-extension diagram

the rate of dissipation in one bar

$$D_{\text{int}} = \int_V \mathcal{D}\,\mathrm{d}V = \int_V \sigma\dot{\varepsilon}_p\,\mathrm{d}V = V\sigma\dot{\varepsilon}_p = A\ell\sigma\frac{\dot{e}_p}{\ell} = A\sigma\dot{e}_p = S\dot{e}_p \tag{2.2}$$

where $\dot{e}_p$ is the plastic extension rate. Similarly, we could integrate $\mathcal{D} = \sigma_0|\dot{\varepsilon}_p|$, arriving at $D_{\text{int}} = S_0|\dot{e}_p|$. In summary, the dissipation rate in a yielding bar can be expressed in full analogy to (1.7) as

$$D_{\text{int}} = S\dot{e}_p = S_0|\dot{e}_p| \tag{2.3}$$

Clearly, the postulate of maximum plastic dissipation can be generalized to yielding bars. This will be discussed in Chapter 5.

2.2 MOMENT-CURVATURE DIAGRAM

A straight beam subjected to lateral loads exhibits bending, and the relation between the bending moment and the curvature of the deformed middle axis is then of interest. We adhere to the usual assumptions of the bending theory for slender beams, namely that the plane cross sections remain plane, and that they remain normal to the deflected middle axis of the beam. It has been verified by numerous experiments that these assumptions are valid for elastic as well as inelastic bending, provided the beam is sufficiently long compared to its cross section dimensions (for good accuracy, at least ten times longer).

Based on the foregoing assumptions, we have

$$\varepsilon = z\kappa \tag{2.4}$$

in which z is the depth coordinate measured from the middle axis, x (see Figure 2.2), $\kappa = 1/\rho$ = beam curvature, ρ = radius of curvature. The bending moment in the

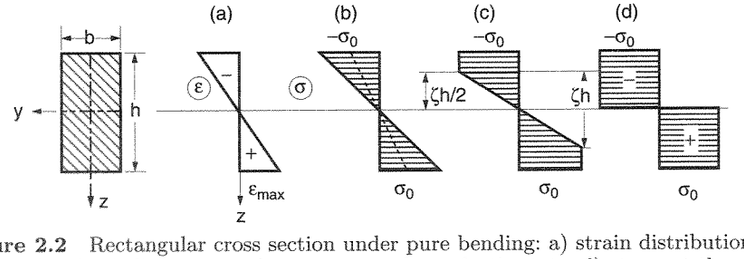


Figure 2.2 Rectangular cross section under pure bending: a) strain distribution, b) stress at the elastic limit state, c) stress at an elastoplastic state, d) stress at the plastic limit state

beam is defined as

$$M = \int_A z\sigma(z)\,dA \tag{2.5}$$

in which $\sigma(z)$ is the normal stress varying as a function of z.

To determine the dependence of bending moment M on curvature κ, we consider various values of small κ, calculate strains from (2.4), then determine stresses σ corresponding to ε from the stress-strain diagram, and finally, evaluate M from (2.5). This procedure will be illustrated for a rectangular cross section.

Example 2.1: Derive the moment-curvature diagram for a rectangular cross section of depth h and width b.

Solution: As long as the strain at the extreme points remains at or below the elastic limit strain $\varepsilon_0 = \sigma_0/E$, the linear stress-strain law $\sigma = E\varepsilon$ is applicable at all points of the cross section, and the stresses vary linearly along the depth of the beam (Figure 2.2(b)). Therefore,

$$M = \int_A z\sigma(z)\,\mathrm{d}A = \int_A zE\kappa z\,\mathrm{d}A = E\kappa\int_A z^2\,\mathrm{d}A = EI\kappa \tag{2.6}$$

where $I = \int_A z^2\,\mathrm{d}A$ is the moment of inertia of the cross section with respect to its horizontal centroidal axis, y. Equation (2.6) is the well known moment-curvature relation according to the linear elastic beam theory. The product EI is called the bending stiffness of the cross section. For a rectangular cross section, we have

$$I = \int_A z^2\,\mathrm{d}A = \int_{-h/2}^{h/2} z^2 b\,\mathrm{d}z = \frac{1}{12}bh^3 \tag{2.7}$$

The elastic limit is reached when the strain in the most extended fibers, ε_{max}, equals σ_0/E, and the corresponding stress thus reaches the yield stress, σ_0 (Figure 2.2(b)). This happens at curvature

$$\kappa_{el} = \frac{\varepsilon_{max}}{z_{max}} = \frac{\sigma_0/E}{h/2} = \frac{2\sigma_0}{Eh} \tag{2.8}$$

The corresponding bending moment

$$M_{el} = EI\kappa_{el} = E\frac{1}{12}bh^3\frac{2\sigma_0}{Eh} = \sigma_0\frac{1}{6}bh^2 \tag{2.9}$$

is called the *elastic limit moment*. Note that the elastic limit moment is a product of the yield stress, σ_0, and the *elastic section modulus*, $W_{el} = bh^2/6$. For a more general cross section (symmetric with respect to the z-axis but otherwise arbitrary), the elastic section modulus is defined as $W_{el} = I/|z|_{max}$, where $|z|_{max}$ is the maximum distance of extreme points of the cross section from the horizontal centroidal axis y.

When the strain at the extreme points exceeds the value of σ_0/E, the elastic part of the cross section, over which the stress distribution is linear, as shown in Figure 2.2(c), has the depth ζh, with $\zeta < 1$. In the remaining (plastic) part of the section, the normal stress is at the positive or negative yield limit. Evaluating the moment from the stress distribution diagram drawn in Figure 2.2(c), we have

$$M = 2\sigma_0 b\frac{1}{2}\left(\frac{\zeta h}{2}\right)\left(\frac{2}{3}\frac{\zeta h}{2}\right) + 2\sigma_0 b\left(\frac{h}{2} - \frac{\zeta h}{2}\right)\frac{1}{2}\left(\frac{\zeta h}{2} + \frac{h}{2}\right) = M_0\left(1 - \frac{\zeta^2}{3}\right) \tag{2.10}$$

in which

$$M_0 = \sigma_0 \frac{bh^2}{4} \tag{2.11}$$

To obtain the meaning of the parameter ζ, we write the condition that at $z_\zeta = \zeta h/2$ the strain is at the elastic limit, i.e. $\varepsilon_0 = \sigma_0/E = \kappa z_\zeta = \kappa\zeta h/2$, from which

$$\zeta = \frac{2\sigma_0}{Eh\kappa} = \frac{\kappa_{el}}{\kappa} \tag{2.12}$$

where κ_{el} is the curvature at the elastic limit given by (2.8). Substituting (2.12) into (2.10) leads to the moment-curvature relation

$$M(\kappa) = M_0 \left(1 - \frac{\kappa_{el}^2}{3\kappa^2}\right) \tag{2.13}$$

which is valid above the elastic limit, i.e. for $\kappa \geq \kappa_{el}$. Plotting (2.6) for $\kappa \leq \kappa_{el}$ and (2.13) for $\kappa \geq \kappa_{el}$, we obtain the moment-curvature diagram drawn in Figure 2.3(a). □

As the curvature κ tends to infinity (which is, of course, a hypothetical case), the bending moment approaches a limiting value M_0 given in (2.10). This represents the so-called *plastic limit moment*, or *yield moment* of the cross section, the largest moment that the cross section can develop. It corresponds to the piecewise constant distribution of normal stress shown in Figure 2.2(d). Note that, similar to the elastic limit moment, the plastic limit moment is the product of the yield stress, σ_0, and a quantity characterizing the geometry of the cross section. We call this quantity the *plastic section modulus*, W_0. From (2.11) it follows that the plastic section modulus for a rectangular section is $W_0 = bh^2/4$. We observe that for the rectangular cross section the ratio of the plastic limit moment to the elastic limit moment, M_0/M_{el}, equals 1.5. As both M_0 and M_{el} are proportional to the yield stress, their ratio $\alpha = M_0/M_{el} = W_0/W_{el}$ depends only on the shape of the cross section, and so we can call it the *shape factor*. For thin-walled cross sections, the shape factor is relatively close to 1; for typical hot-rolled profiles, it is typically about 1.2. We leave it to the reader to analyze some other cross sections as an exercise (see Problems 2.1 and 2.2).

As we have seen, the moment-curvature diagram for an elastic-perfectly plastic material is curved, which is an inconvenience for structural analysis. For this reason, we replace it with a bilinear, idealized moment-curvature diagram shown in

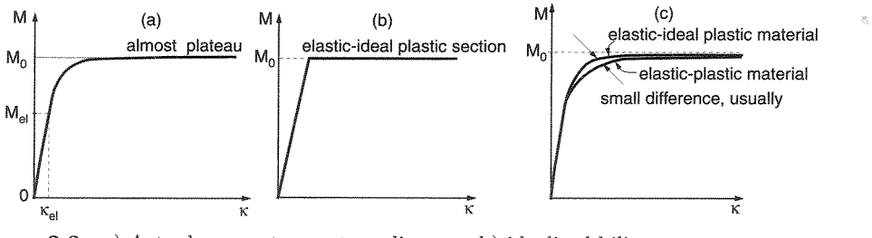


Figure 2.3 a) Actual moment-curvature diagram, b) idealized bilinear moment-curvature diagram, c) comparison of moment-curvature diagrams

Figure 2.3(b), having the same limiting moment, M_0. The difference compared to the actual behavior is only significant for curvatures not much larger than the limiting elastic curvature. For collapse calculations, in which we consider the curvatures in the critical cross sections to be much larger than the elastic curvature, the error incurred due to this idealization is obviously small.

By a similar procedure, it is easy to construct the moment-curvature diagram for a material with any uniaxial inelastic stress-strain relation. The corresponding moment-curvature diagram differs from the curve obtained using the elastic-perfectly plastic idealization (see Figure 2.3(c)). However, for materials with a stress-strain diagram terminated by a plateau, the moment-curvature diagram approaches a horizontal line, and for large curvatures the bilinear idealization from Figure 2.3(b) is acceptable.

2.3 PLASTIC HINGE

Example 2.2: Analyze the failure of a simply supported beam loaded up to collapse by a concentrated load at midspan (Figure 2.4(a)).

Solution: As we increase the load, the cross section at midspan, at which the bending moment is the highest (note the bending moment diagram in Figure 2.4(b)) is plasticized first. However, large deflections are possible nowhere in the beam as long as at least a portion of the cross section depth remains elastic. A sudden change in response is seen when one cross section (in our case at midspan) becomes plastic across its full depth. At that point, an arbitrary further increase of curvature at that cross section is possible (at least within the geometrically linear theory). Since the curvature can increase greatly, mathematically infinitely, we have at midspan an equivalent of a hinge, called a *plastic hinge* or *yield hinge* (Figure 2.4(c)). For the purpose of structural analysis, the yield hinge is idealized as one cross section with

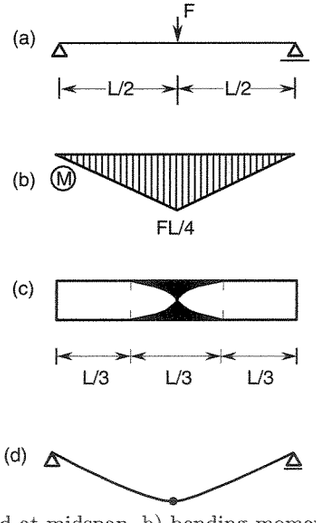


Figure 2.4 a) Beam loaded at midspan, b) bending moment diagram, c) plastic hinge, d) collapse mechanism

infinite curvature surrounded by elastic material. An infinite increase of curvature at this hinge is equivalent to a finite rotation at this cross section, which obviously leads to the collapse of the beam (Figure 2.4(d)). The collapse load is readily determined from the condition $M_{max} = (F_0/2)(L/2) = M_0$, which yields the collapse load as $F_0 = 4M_0/L$. □

We have just made our first calculation of the collapse load. In the case of a simply supported beam this task is trivial. However, the task becomes much more involved in the case of more complicated, statically indeterminate structures and we will devote several subsequent chapters to discussing their behavior. Central to our analysis of beams and frames will be the concept of the yield hinge which we have just explained.

Note that the concept of plastic hinges does not require the plastic rotation to be large. During collapse the load is constant, and so the elastic deformations do not change. Therefore, all the structural parts whose cross section is not fully plasticized behave as a rigid body.

As already stated, the idealized yield hinge lumps all the plastic deformation into a single cross section. The real plasticized zone occupies a certain volume, and its shape can be estimated if the elastoplastic state of the cross sections carrying bending moments between M_{el} and M_0 is taken into account.

Example 2.3: Analyze the shape of the plastic region for the simply supported beam in Figure 2.4(a).

Solution: The distribution of bending moments in the left half of the beam is described by

$$M(x) = \frac{1}{2}Fx, \quad x \in \left[0, \frac{L}{2}\right] \tag{2.14}$$

At collapse, $F = F_0 = 4M_0/L = \sigma_0 bh^2/L$, and the elastic limit moment $M_{el} = \sigma_0 bh^2/6$ is attained at the cross section, located at

$$x = \frac{2M_{el}}{F_0} = \frac{\sigma_0 bh^2/3}{\sigma_0 bh^2/L} = \frac{1}{3}L \tag{2.15}$$

The part of the beam to the left of this cross section is purely elastic since the bending moments are below M_{el}. The plastic zone extends from the section at $x = L/3$ to the section at $x = 2L/3$. All the cross sections except for the fully plastic one have an elastic core around the neutral axis. The boundary between the elastic and plastic zone can be obtained by substituting (2.10) for the bending moment in (2.14), and setting the force F on the right-hand side equal to the collapse load $F_0 = 4M_0/L$. This gives

$$M_0\left(1 - \frac{\zeta^2}{3}\right) = \frac{1}{2}\frac{4M_0}{L}x \tag{2.16}$$

from which

$$\zeta = \pm\sqrt{3\left(1 - \frac{2x}{L}\right)}, \quad x \in \left[\frac{L}{3}, \frac{L}{2}\right] \tag{2.17}$$

Recall that ζ is the ratio of the depth of the elastic core to the total depth of the beam. At $x = L/2$ we get $\zeta = 0$, while at $x = L/3$ we have $\zeta = 1$. The plastic zone

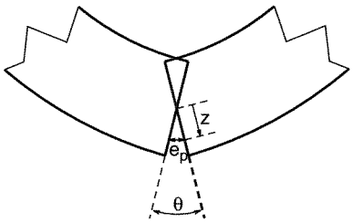


Figure 2.5 Idealized plastic hinge

is plotted in Figure 2.4(c). According to (2.17), the elasto-plastic boundary has a parabolic shape.

The exact length and shape of the plasticized zone has only a small effect on the global response of the structure, and so we can lump the plastic hinge into one single cross section. The total plastic deformation is replaced by a rotation, θ, in an idealized hinge (Figure 2.5). From kinematic considerations, it follows that the plastic extension at an arbitrary point of the cross section can be expressed by a linear function

$$e_p(z) = \theta z \tag{2.18}$$

The power dissipated by a plastic hinge, defined by the volume integral

$$D_{\text{int}} = \int_V \sigma \dot{\varepsilon}_p \, \mathrm{d}V \tag{2.19}$$

can be evaluated as expressed by the integral over the fully plasticized cross section,

$$D_{\text{int}} = \int_A \sigma(z)\dot{e}_p(z) \, \mathrm{d}A \tag{2.20}$$

because the extension $e_p(z)$ corresponds to the plastic strain at z integrated over the length of the plastic hinge. Substituting expression (2.18) for the plastic extension, we can transform (2.20) into

$$D_{\text{int}} = \int_A \sigma(z)\dot{\theta} z \, \mathrm{d}A = \int_A \sigma(z) \, z \, \mathrm{d}A \, \dot{\theta} = M\dot{\theta} \tag{2.21}$$

In analogy to (1.7) and (2.3), the dissipation power in a plastic hinge can be expressed as

$$D_{\text{int}} = M\dot{\theta} = M_0|\dot{\theta}| \tag{2.22}$$

provided that the magnitude of the plastic moment does not depend on the sign of curvature (tension along the bottom face and compression along the top face of cross section, or vice versa). The reader can verify that this condition is satisfied if the yield stress in compression is equal in magnitude to the yield stress in tension.

2.4 MONOSYMMETRIC SECTIONS

So far we have tacitly assumed that the neutral axis remains at its original position, even if a part of the cross section yields. This is certainly true for cross sections symmetric with respect to a horizontal, as well as a vertical, axis such as the

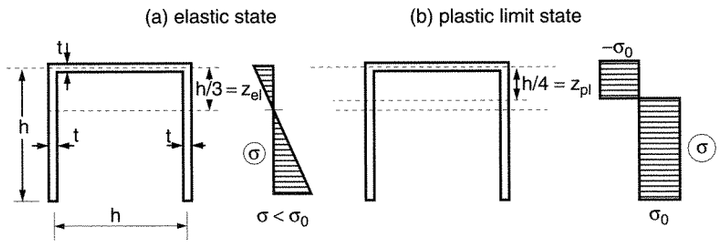


Figure 2.6 Channel cross section a) in an elastic state, b) in the plastic limit state

rectangular cross section or the symmetric I-section. However, if the cross section has only a vertical axis of symmetry, the plastic neutral axis, i.e. the location at which $\varepsilon = 0$, does not have to coincide with the elastic neutral axis, i.e. the centroidal axis. The position of the neutral axis can be determined from the condition

$$\int_A \sigma(z)\,\mathrm{d}A = N \tag{2.23}$$

that links the normal stresses to the normal force, N. For pure bending, the normal force is zero, and so the integral of the normal stresses over the cross section must vanish.

Example 2.4: Determine the position of the neutral axis in the plastic limit state for a channel cross section with one flange and two webs all of the same area and of the same thickness, t, which is negligible compared to the depth of the section, h.

Solution: Let z denote the vertical distance from the flange. The elastic neutral axis (centroid) is readily obtained to be at the depth $z_{el} = h/3$ (Figure 2.6(a)). To determine the plastic neutral axis, we denote its distance from the flange as z_{pl} and write the condition of equality of the compressive and tensile resultants of normal stresses: $2t(h - z_{pl})\sigma_0 = t(h + 2z_{pl})\sigma_0$. The solution $z_{pl} = h/4$ indicates that the plastic neutral axis is closer to the top flange than is the elastic neutral axis (Figure 2.6(b)). □

The result of the previous example can be generalized: the neutral axis divides the cross section into a region under tension and a region under compression. In the plastic limit state, the stress in each region is constant (equal to σ_0 and $-\sigma_0$, respectively). In the absence of a normal force, the condition of equilibrium requires that the areas of both regions be the same, i.e. that the plastic neutral axis divide the area of the cross section into two equal parts. Recall that, for a similar reason, the tensile and compressive regions in the elastic state have equal static moments (with opposite signs) with respect to the neutral axis. This is why, during a gradual plasticization, the neutral axis moves towards the region that is located 'closer' to the center of gravity, i.e. towards the region that has a larger area.

PROBLEMS

Problem 2.1: Determine the shape factor $\alpha = W_0/W_{el}$ for the I-section in Figure 2.7(a), and construct the corresponding 'exact' moment-curvature diagram.

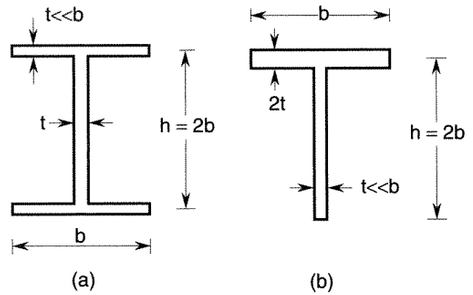


Figure 2.7 a) I-section, b) T-section

Problem 2.2: Consider a T-shaped cross section depicted in Figure 2.7(b). Try to guess (without any calculations) whether the neutral axis moves up or down during the plasticization of the section. Find the elastic limit moment and the yield moment, and construct the 'exact' moment-curvature diagram.

Problem 2.3: Consider a simply supported beam with a T-shaped cross section from Problem 2.2 subjected to a uniform vertical load. Determine the extent of the plastic hinge at the top fibers and at the bottom fibers.

Problem 2.4: Derive and plot the 'exact' dependence of the load-point displacement on the magnitude of the applied load for the simply supported beam from Example 2.2, and compare the result to the idealized model with a plastic hinge. Hint: Similar to elastic analysis, the load-point displacement $w = \int_0^{L/2} x\,\kappa(x)\,\mathrm{d}x$. The curvature $\kappa(x)$ must be expressed in terms of the bending moment $M(x)$ (separately in the elastic region and the elastoplastic region), which in turn depends on the applied force F.

In the following problems consider a material with unequal yield stress in tension and in compression.

Problem 2.5: Derive an expression for plastic dissipation rate in a bar similar to (2.3).
Hint: Look at Problem 1.3.

Problem 2.6: Find the position of the neutral axis in a rectangular cross section in the plastic limit state. Solve a) in particular for $\sigma_0^- = 9\sigma_0^+$, b) in general for $\sigma_0^- = n\sigma_0^+$.

Problem 2.7: Solve Problem 2.6 for the T-section from Figure 2.7(b). Consider only the particular case $\sigma_0^- = 9\sigma_0^+$.

Problem 2.8: Derive an expression for the plastic dissipation rate in a yield hinge similar to (2.22), but covering the case of cross sections for which $M_0^+ \neq M_0^-$ (see Problems 2.6 and 2.7).

3

Incremental Analysis

For a prescribed loading history, the response of a given structure can be traced using an incremental procedure. During the initial stage of loading, the structure behaves elastically, and standard methods of linear elastic analysis can be applied. The corresponding solution is valid until plastic yielding starts at some point of the structure. However, the load carrying capacity is usually not exhausted immediately, and the applied load can further increase while the plastic zone grows. Eventually, a state is reached at which the structure deforms as a kinematic mechanism under a constant applied load, and collapse occurs.

3.1 TRUSSES

For trusses, the stress in each bar is uniform, and so plastic yielding starts simultaneously at all points of a bar.[1] The axial force transmitted by the bar then remains constant, and additional load increments must be balanced by axial forces in the bars that still remain elastic. This means that the structure *incrementally* behaves as if the yielding bar were missing. If the original structure is statically determinate, it turns into a kinematic mechanism immediately after the first bar starts yielding, and thus collapse occurs right at the elasticity limit. However, if the degree of redundancy[2] of the given original structure is $s \geq 1$, then after the onset of yielding in one bar, the degree of redundancy becomes $s - 1$, and so the structure can still transmit an increasing load. When the stress in another bar reaches the yield limit, the degree of redundancy becomes $s - 2$, etc.

Collapse occurs when $s + 1$ bars yield and the structure turns into a mechanism with one degree of freedom (i.e. becomes statically unstable). In fact, this can happen even with less than $s + 1$ yielding bars, if a *partial mechanism* occurs. In that case, a part of the truss remains statically indeterminate, but one or more joints can move without changing the length of any of the bars that are still elastic. For example, the degree of redundancy of the truss shown in Figure 3.1 is $s = 1$, but if the bar

[1] Strictly speaking, this is true only in the ideal case of a bar with a perfectly constant cross sectional area, made of a perfectly homogeneous material, etc. Due to inevitable geometric and material imperfections, yielding starts at the most critical point, but the plastic zone soon expands at least over one entire cross section, and so the assumption of a constant axial force after the onset of yielding is justified.

[2] Recall that the degree of redundancy is the number of redundant, or statically indeterminate internal forces and reactions.

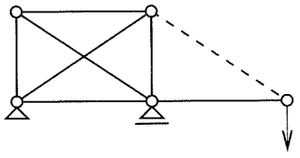


Figure 3.1 Truss with a partial collapse mechanism

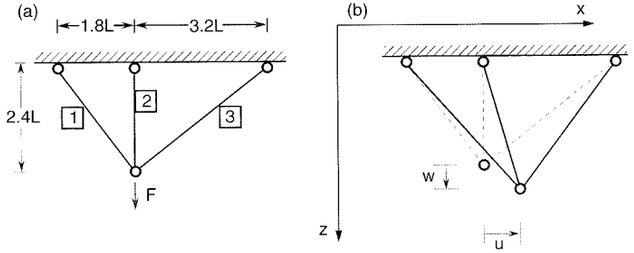


Figure 3.2 Simple three-bar truss: a) geometry and loading, b) coordinate axes and joint displacements

plotted by a dashed line starts yielding, collapse occurs. Additional examples of partial mechanisms are shown in Figure 6.3(b)(c).

During each stage between two subsequent appearances of a new yielding bar, the incremental behavior corresponds to an elastic structure from which the bars that are already yielding have been removed and their action replaced by constant applied forces. This fact can be exploited by the algorithm of incremental elasto-plastic analysis.

Example 3.1: Consider a simple truss consisting of three pin-ended bars connected to a common joint (Figure 3.2(a)). All bars have the same cross sectional area, A, and the same material properties (Young's modulus, E, and yield stress, σ_0). The vertical force F applied at the joint increases in magnitude. Determine the corresponding evolution of displacements and axial forces.

Solution: For a brief review of linear elastic truss analysis, the reader is referred to Appendix A. Let us denote the axial forces by S_1, S_2 and S_3, the bar extensions (changes of length) by e_1, e_2 and e_3, and the displacements of the unsupported joint by u and w (Figure 3.2(b)). The lengths of the bars are $\ell_1 = 3L$, $\ell_2 = 2.4L$ and $\ell_3 = 4L$, where L is a given reference length.[3]

STEP 1: The fundamental equations describing the initial linear elastic stage are as follows:

$$\text{Kinematic equations:} \qquad \begin{aligned} e_1 &= 0.6u + 0.8w \\ e_2 &= w \\ e_3 &= -0.8u + 0.6w \end{aligned} \tag{3.1}$$

[3] In our illustrative examples, we try to avoid substituting numerical values of quantities which have some physical dimensions (i.e. which are not dimensionless). This keeps the calculations easier to follow and independent of a particular system of units.

$$\text{Constitutive equations:} \qquad \begin{aligned} S_1 &= \frac{EA}{3L} e_1 \\ S_2 &= \frac{EA}{2.4L} e_2 \\ S_3 &= \frac{EA}{4L} e_3 \end{aligned} \tag{3.2}$$

$$\text{Equilibrium equations:} \qquad \begin{aligned} 0.6S_1 - 0.8S_3 &= 0 \\ 0.8S_1 + S_2 + 0.6S_3 &= F \end{aligned} \tag{3.3}$$

Following the stiffness method (or displacement method), we substitute (3.1) and (3.2) into (3.3) in order to express the equilibrium equations in terms of unknown displacements:

$$0.28u + 0.04w = 0 \tag{3.4}$$

$$0.04u + 0.72w = FL/EA \tag{3.5}$$

The solution $u = -0.2FL/EA$ and $w = 1.4FL/EA$ can be substituted back into the kinematic and constitutive equations to give

$$S_1 = 0.333F \tag{3.6}$$

$$S_2 = 0.583F \tag{3.7}$$

$$S_3 = 0.250F \tag{3.8}$$

This is the elastic solution, valid as long as the axial forces stay inside the interval $(-S_0, S_0)$, where $S_0 = A\sigma_0$ is the plastic axial force.

Let us denote important states of the loading process by superscripts. From now on, superscript $^{(n)}$ refers to the state in which the number of yielding bars changes from $n-1$ to n. At $F = S_0/0.583 = 1.714\ S_0$, S_2 just reaches S_0, while S_1 and S_3 still remain in the elastic range. Therefore, we have

$$F^{(1)} = 1.714 S_0 \tag{3.9}$$

$$S_1^{(1)} = 0.333 F^{(1)} = 0.571 S_0 \tag{3.10}$$

$$S_2^{(1)} = 0.583 F^{(1)} = S_0 \tag{3.11}$$

$$S_3^{(1)} = 0.25 F^{(1)} = 0.429 S_0 \tag{3.12}$$

$$u^{(1)} = -0.2 \frac{F^{(1)} L}{EA} = -0.343 \frac{S_0 L}{EA} \tag{3.13}$$

$$w^{(1)} = 1.4 \frac{F^{(1)} L}{EA} = 2.4 \frac{S_0 L}{EA} \tag{3.14}$$

STEP 2: During the second load increment from $F^{(1)}$ to $F^{(2)}$ (not yet known), bar 2 is yielding, and so axial force S_2 remains constant. All the fundamental equations are still valid except for the constitutive relation for S_2. In the rate form,[4] the fundamental

[4] To keep the notation short, we will often work with *rates* (denoted by overdots) rather than with *increments* (denoted by Δ preceding the symbol). Of course, in our static calculations time only plays the role of a parameter labeling individual states of the structure. Whenever we write an equation in the rate form, both sides can formally be multiplied by a small time increment Δt, after which all the rates turn into the corresponding increments.

equations for the second increment can be written as follows:

Kinematic equations:
$$\begin{aligned} \dot{e}_1 &= 0.6\dot{u} + 0.8\dot{w} \\ \dot{e}_2 &= \dot{w} \\ \dot{e}_3 &= -0.8\dot{u} + 0.6\dot{w} \end{aligned} \tag{3.15}$$

Constitutive equations:
$$\begin{aligned} \dot{S}_1 &= \frac{EA}{3L}\dot{e}_1 \\ \dot{S}_2 &= 0 \\ \dot{S}_3 &= \frac{EA}{4L}\dot{e}_3 \end{aligned} \tag{3.16}$$

Equilibrium equations:
$$\begin{aligned} 0.6\dot{S}_1 - 0.8\dot{S}_3 &= 0 \\ 0.8\dot{S}_1 + \dot{S}_2 + 0.6\dot{S}_3 &= \dot{F} \end{aligned} \tag{3.17}$$

This is formally the same problem as that of an elastic truss with bar 2 having a zero stiffness. The governing equations of the displacement method now read

$$0.28\dot{u} + 0.04\dot{w} = 0 \tag{3.18}$$
$$0.04\dot{u} + 0.3033\dot{w} = \dot{F}L/EA \tag{3.19}$$

and their solution is $\dot{u} = -0.48\dot{F}L/EA$ and $\dot{w} = 3.36\dot{F}L/EA$. From this we can easily obtain the rates of axial forces and, replacing the rates by increments, we can describe this stage of evolution of the axial forces by

$$S_1 = S_1^{(1)} + \Delta S_1 = 0.571S_0 + 0.8\Delta F \tag{3.20}$$
$$S_2 = S_2^{(1)} = S_0 \tag{3.21}$$
$$S_3 = S_3^{(1)} + \Delta S_3 = 0.429S_0 + 0.6\Delta F \tag{3.22}$$

Axial force S_1 reaches S_0 at $\Delta F = (S_0 - 0.571S_0)/0.8 = 0.536S_0$ while S_3 would reach S_0 only later, at $\Delta F = (S_0 - 0.429S_0)/0.6 = 0.952S_0$. We thus set

$$\Delta F^{(2)} = 0.536S_0 \tag{3.23}$$

and calculate forces and displacements at state $^{(2)}$ when bar 1 starts, yielding

$$F^{(2)} = F^{(1)} + \Delta F^{(2)} = 1.714S_0 + 0.536S_0 = 2.25S_0 \tag{3.24}$$
$$S_1^{(2)} = S_1^{(1)} + \Delta S_1^{(2)} = 0.571S_0 + 0.8\Delta F^{(2)} = S_0 \tag{3.25}$$
$$S_2^{(2)} = S_2^{(1)} = S_0 \tag{3.26}$$
$$S_3^{(2)} = S_3^{(1)} + \Delta S_3^{(2)} = 0.429S_0 + 0.6\Delta F^{(2)} = 0.75S_0 \tag{3.27}$$
$$u^{(2)} = u^{(1)} + \Delta u^{(2)} = -0.343\frac{S_0L}{EA} - 0.48\frac{\Delta F^{(2)}L}{EA} = -0.6\frac{S_0L}{EA} \tag{3.28}$$
$$w^{(2)} = w^{(1)} + \Delta w^{(2)} = 2.4\frac{S_0L}{EA} + 3.36\frac{\Delta F^{(2)}L}{EA} = 4.2\frac{S_0L}{EA} \tag{3.29}$$

STEP 3: A further increase of the applied load is impossible as only one bar remains in the elastic state and the axial force in this bar cannot balance any applied load

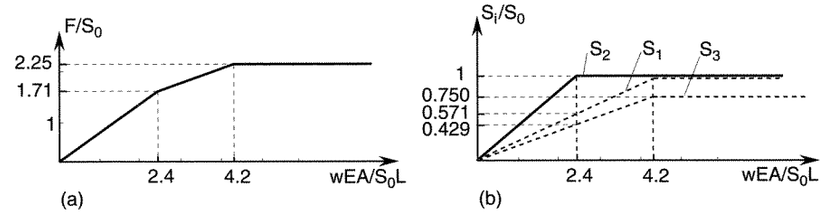


Figure 3.3 Three-bar truss: a) load-displacement diagram, b) evolution of axial forces

increments. With the two yielding bars 'removed', the structure becomes a mechanism with one degree of freedom, and collapses. $F^{(2)}$ thus represents the collapse load, while $F^{(1)}$ is the load at the elasticity limit.

The solution is presented graphically in Figure 3.3. The load-displacement diagram in Figure 3.3(a) illustrates the progressive degradation of the stiffness of the structure. After the plastic limit state, the displacement can grow at a constant applied load. Figure 3.3(b) shows the evolution of the axial forces. □

The incremental solution demonstrated in the preceding example is applicable to any prescribed load evolution. A common assumption is that all the components of the set of forces acting upon the structure change proportionally to a certain *load parameter* μ. This special case is referred to as *proportional loading*. In matrix notation (see Appendix A) we can write

$$\boldsymbol{f} = \mu\bar{\boldsymbol{f}} \tag{3.30}$$

where $\boldsymbol{f}$ is the vector of applied forces (varying in time) and $\bar{\boldsymbol{f}}$ is some fixed reference load vector. The load parameter μ is also called the *load multiplier*. Equation (3.30) describes a straight line passing through the origin in the load space (space of applied forces), and $\bar{\boldsymbol{f}}$ specifies the direction of this line.

The matrix form of the fundamental equations describing a truss structure during a generic loading increment number n is as follows:

$$\text{kinematic equations} \ldots \boldsymbol{B}\dot{\boldsymbol{d}} = \dot{\boldsymbol{e}} \tag{3.31}$$
$$\text{constitutive equations} \ldots \boldsymbol{D}^{(n)}\dot{\boldsymbol{e}} = \dot{\boldsymbol{s}} \tag{3.32}$$
$$\text{equilibrium equations} \ldots \boldsymbol{B}^T\dot{\boldsymbol{s}} = \dot{\boldsymbol{f}} \equiv \dot{\mu}\bar{\boldsymbol{f}} \tag{3.33}$$

Here, $\dot{\boldsymbol{d}}$, $\dot{\boldsymbol{e}}$ and $\dot{\boldsymbol{s}}$ are the vectors (column matrices) of the rates of joint displacements, bar extensions and axial forces, respectively, while $\boldsymbol{B}$, $\boldsymbol{D}^{(n)}$, and $\boldsymbol{B}^T$ are the kinematic matrix, the generalized material stiffness matrix and the static matrix, respectively; see Appendix A. If we neglect geometrically nonlinear effects, $\boldsymbol{B}$ and $\boldsymbol{B}^T$ are constant throughout the entire analysis. On the other hand, $\boldsymbol{D}^{(n)}$ is a diagonal matrix whose diagonal terms are the current stiffnesses of the individual bars; hence, $\boldsymbol{D}^{(n)}$ depends on the current state. In the initial (virgin) state, the stiffness of each bar is its elastic stiffness EA_i/ℓ_i. When a bar starts yielding, its current stiffness drops down to zero.

Combining the fundamental equations, we get the governing equations of the displacement method

$$\boldsymbol{B}^T\boldsymbol{D}^{(n)}\boldsymbol{B}\dot{\boldsymbol{d}} = \dot{\mu}\bar{\boldsymbol{f}} \tag{3.34}$$

or

$$\boldsymbol{K}^{(n)}\dot{\boldsymbol{d}} = \dot{\mu}\bar{\boldsymbol{f}} \tag{3.35}$$

where $\boldsymbol{K}^{(n)} = \boldsymbol{B}^T \boldsymbol{D}^{(n)} \boldsymbol{B}$ is the *current (tangential) stiffness matrix* of the structure.

Using matrix notation, we can present the algorithm of incremental analysis in a clear and concise form:

Algorithm 3.1:

1. Set $\mu^{(0)} = 0$, $\boldsymbol{d}^{(0)} = \boldsymbol{0}$ and $\boldsymbol{s}^{(0)} = \boldsymbol{0}$.
2. Loop over loading steps $n = 1, 2, 3, \ldots$
   For each step:
   (a) Assemble the tangential stiffness matrix $\boldsymbol{K}^{(n)}$.
   (b) Solve $\boldsymbol{K}^{(n)} \bar{\boldsymbol{d}}^{(n)} = \bar{\boldsymbol{f}}$ for the unknown displacement increments $\bar{\boldsymbol{d}}^{(n)}$ that would correspond to a unit increment of the load parameter μ. If $\boldsymbol{K}^{(n)}$ is found to be singular, the structure has become a kinematic mechanism and collapse occurs. In this case, output the results and stop.
   (c) Determine the increments of axial forces $\bar{\boldsymbol{s}}^{(n)} = \boldsymbol{D}^{(n)} \boldsymbol{B} \bar{\boldsymbol{d}}^{(n)}$ that would correspond to a unit increment of the load parameter. Express the axial forces as $\boldsymbol{s} = \boldsymbol{s}^{(n-1)} + \Delta\mu \cdot \bar{\boldsymbol{s}}^{(n)}$, and determine the minimum positive value of $\Delta\mu$ for which the axial force in some bar reaches the plastic axial force (in tension or in compression). Denote this value as $\Delta\mu^{(n)}$.
   (d) Update: $\mu^{(n)} = \mu^{(n-1)} + \Delta\mu^{(n)}$, $\boldsymbol{d}^{(n)} = \boldsymbol{d}^{(n-1)} + \Delta\mu^{(n)} \bar{\boldsymbol{d}}^{(n)}$, $\boldsymbol{s}^{(n)} = \boldsymbol{s}^{(n-1)} + \Delta\mu^{(n)} \bar{\boldsymbol{s}}^{(n)}$.
   (e) End of the loop over loading steps. Increment the step number n by 1 and go to (a).

Example 3.2: Present the solution of Example 3.1 in matrix notation.

Solution: The kinematic matrix

$$\boldsymbol{B} = \begin{bmatrix} 0.6 & 0.8 \\ 0 & 1 \\ -0.8 & 0.6 \end{bmatrix} \tag{3.36}$$

is the same at any loading stage. The load vector $\boldsymbol{f} = \{0, F\}^T$ can be presented as

$$\boldsymbol{f} = F \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \tag{3.37}$$

where F now plays the role of the load parameter μ, and $\{0, 1\}^T$ is the given reference load vector $\bar{\boldsymbol{f}}$. Note that in this case the reference load vector is dimensionless and the load parameter has the dimension of force. Alternatively, we could set $\boldsymbol{f} = \mu \{\bar{F}, 0\}^T$ where $\bar{F}$ is a fixed force, e.g. 10 kN, and μ is a dimensionless parameter.

STEP 1:

$$\boldsymbol{D}^{(1)} = \frac{EA}{L} \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/2.4 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}, \quad \boldsymbol{K}^{(1)} = \boldsymbol{B}^T \boldsymbol{D}^{(1)} \boldsymbol{B} = \frac{EA}{L} \begin{bmatrix} 0.28 & 0.04 \\ 0.04 & 0.72 \end{bmatrix}$$

$$\bar{\boldsymbol{d}}^{(1)} = \frac{L}{EA} \begin{Bmatrix} -0.2 \\ 1.4 \end{Bmatrix}, \quad \bar{\boldsymbol{s}}^{(1)} = \boldsymbol{D}^{(1)} \boldsymbol{B} \bar{\boldsymbol{d}}^{(1)} = \begin{Bmatrix} 0.333 \\ 0.583 \\ 0.25 \end{Bmatrix}$$

For $\Delta F^{(1)} = 1.714\, S_0$, bar 2 yields and we have

$$\boldsymbol{d}^{(1)} = \Delta F^{(1)} \bar{\boldsymbol{d}}^{(1)} = \frac{S_0 L}{EA} \begin{Bmatrix} -0.371 \\ 2.4 \end{Bmatrix}, \quad \boldsymbol{s}^{(1)} = \Delta F^{(1)} \bar{\boldsymbol{s}}^{(1)} = S_0 \begin{Bmatrix} 0.571 \\ 1 \\ 0.429 \end{Bmatrix}$$

STEP 2:
Bar 2 is yielding and thus the second diagonal element of $\boldsymbol{D}$ must be replaced by zero.

$$\boldsymbol{D}^{(2)} = \frac{EA}{L} \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}, \quad \boldsymbol{K}^{(2)} = \boldsymbol{B}^T \boldsymbol{D}^{(2)} \boldsymbol{B} = \frac{EA}{L} \begin{bmatrix} 0.28 & 0.04 \\ 0.04 & 0.3033 \end{bmatrix}$$

$$\bar{\boldsymbol{d}}^{(2)} = \frac{L}{EA} \begin{Bmatrix} -0.48 \\ 3.36 \end{Bmatrix}, \quad \bar{\boldsymbol{s}}^{(2)} = \boldsymbol{D}^{(2)} \boldsymbol{B} \bar{\boldsymbol{d}}^{(2)} = \begin{Bmatrix} 0.8 \\ 0 \\ 0.6 \end{Bmatrix}$$

For $\Delta F^{(2)} = 0.536\, S_0$, bar 1 yields and we have

$$\boldsymbol{d}^{(2)} = \boldsymbol{d}^{(1)} + \Delta F^{(2)} \bar{\boldsymbol{d}}^{(2)} = \frac{S_0 L}{EA} \begin{Bmatrix} -0.6 \\ 4.2 \end{Bmatrix}, \quad \boldsymbol{s}^{(2)} = \boldsymbol{s}^{(1)} + \Delta F^{(2)} \bar{\boldsymbol{s}}^{(2)} = S_0 \begin{Bmatrix} 1 \\ 1 \\ 0.75 \end{Bmatrix}$$

STEP 3:
Bars 1 and 2 are yielding and thus

$$\boldsymbol{D}^{(3)} = \frac{EA}{L} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}, \quad \boldsymbol{K}^{(3)} = \boldsymbol{B}^T \boldsymbol{D}^{(3)} \boldsymbol{B} = \frac{EA}{L} \begin{bmatrix} 0.16 & -0.12 \\ -0.12 & 0.09 \end{bmatrix}$$

The tangential stiffness matrix $\boldsymbol{K}^{(3)}$ is singular ($\det \boldsymbol{K}^{(3)} = 0$), which indicates that the plastic limit state has been reached. □

As is well known from structural analysis (see also Appendix A), the stiffness matrix of the structure can be efficiently obtained by assembling the stiffness matrices of individual bars (referred to as the element stiffness matrices). The assembly procedure makes use of the global degree of freedom numbers, which map rows and columns of the element stiffness matrices onto the rows and columns of the structure stiffness matrix. The element stiffness matrices and the global degree of freedom numbers of bars from the foregoing example are listed in Table 3.1.

The elastic stiffness matrix of the structure is assembled as

$$\boldsymbol{K}^{(1)} = \frac{EA}{L} \begin{bmatrix} 0.12 + 0 + 0.16 & 0.16 + 0 - 0.12 \\ 0.16 + 0 - 0.12 & 0.2133 + 0.4167 + 0.09 \end{bmatrix} = \frac{EA}{L} \begin{bmatrix} 0.28 & 0.04 \\ 0.04 & 0.72 \end{bmatrix}$$

Table 3.1 Element stiffness matrices with the corresponding global degree of freedom numbers

$\boldsymbol{K}_1$	0	0	1	2
0	0.12	0.16	−0.12	−0.16
0	0.16	0.2133	−0.16	−0.2133
1	−0.12	−0.16	0.12	0.16
2	−0.16	−0.2133	0.16	0.2133

$\boldsymbol{K}_2$	0	0	1	2
0	0	0	0	0
0	0	0.4167	0	−0.4167
1	0	0	0	0
2	0	−0.4167	0	0.4167

$\boldsymbol{K}_3$	0	0	1	2
0	0.16	−0.12	−0.16	0.12
0	−0.12	0.09	0.12	−0.09
1	−0.16	0.12	0.16	−0.12
2	0.12	−0.09	−0.12	0.09

During the second loading step, the current stiffness of the yielding bar (bar number 2) is zero, and the stiffness matrix of the structure is assembled only from the contributions of bars 1 and 3:

$$\boldsymbol{K}^{(2)} = \frac{EA}{L}\begin{bmatrix} 0.12+0.16 & 0.16-0.12 \\ 0.16-0.12 & 0.2133+0.09 \end{bmatrix} = \frac{EA}{L}\begin{bmatrix} 0.28 & 0.04 \\ 0.04 & 0.3033 \end{bmatrix}$$

Finally, after bar 1 also starts yielding, the current stiffness matrix of the structure

$$\boldsymbol{K}^{(3)} = \frac{EA}{L}\begin{bmatrix} 0.16 & -0.12 \\ -0.12 & 0.09 \end{bmatrix}$$

is identical with the right lower block of the stiffness matrix of bar 3, which is a singular matrix.

In computerized analysis of large practical problems, the stiffness matrix is always assembled from the contributions of individual bars. The large matrices $\boldsymbol{B}$ and $\boldsymbol{D}$ are never set up explicitly (on the global, structural level). Thus, the formula $\boldsymbol{K} = \boldsymbol{B}^T\boldsymbol{D}\boldsymbol{B}$ is only a formal definition of the stiffness matrix. Similarly, the components of the vector $\bar{\boldsymbol{s}}^{(n)}$ that contains the increments of axial forces due to a unit increment of the load parameter are evaluated for each bar separately, and not by the products $\boldsymbol{D}^{(n)}\boldsymbol{B}\bar{\boldsymbol{d}}^{(n)}$. However, the matrix notation is very useful in theoretical derivations, and it makes it possible to present solution methods for various types of structural models in a unified form.

3.2 BEAMS

Incremental plastic analysis of a statically indeterminate beam or frame can be described by Algorithm 3.1, as can the analysis of trusses explained in the preceding section. The only difference is in the number and nature of components of the vectors and matrices entering the formulation. A brief summary of matrix structural analysis of plane beams and frames is given in Appendix B. The following example demonstrates the basic steps of elastoplastic analysis of a beam.

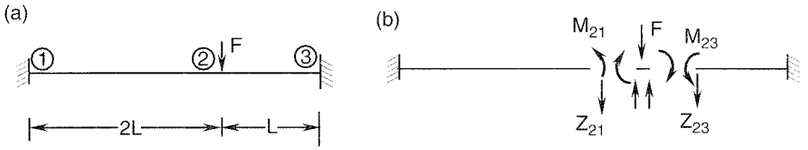


Figure 3.4 Fixed beam loaded by a concentrated force: a) geometry and loading, b) free-body diagram

Example 3.3: Perform incremental elastoplastic analysis of a beam fixed at both ends and subjected to a concentrated force (Figure 3.4(a)). The cross section, characterized by the bending stiffness EI and the plastic limit moment M_0, is constant along the entire beam.

Solution: Let us first identify the critical cross sections, i.e. potential locations of plastic hinges. For a beam having constant properties along the length, plastic hinges can form only at the sections with extreme bending moments. In our example, the bending moment diagram is linear between sections 1 and 2, and between sections 2 and 3, at any stage of the loading process, and so only sections 1, 2, and 3 are critical. In general, it is convenient to set up the structural model such that every critical cross section is a computational joint. We will therefore divide the beam in two members, 1-2 and 2-3. Fundamental unknowns of the displacement method will be the deflection (vertical displacement) w_2 and rotation ϕ_2. Note that the displacements and rotations of joints 1 and 3 are prescribed as zero, and the horizontal displacement of joint 2 does not enter the equations describing bending of the beam (see Appendix B). The fundamental equations of equilibrium easily follow from the free-body diagram in Figure 3.4(b). The equation associated with w_2 reads

$$Z_{21} + Z_{23} = F \tag{3.38}$$

where Z_{21} and Z_{23} are the vertical member end forces shown in Figure 3.4(b), and the equation associated with ϕ_2 reads

$$M_{21} + M_{23} = 0 \tag{3.39}$$

STEP 1: Incremental elastoplastic analysis starts by standard elastic analysis of the model. The unknown forces and moments at the left-hand side of the above equations must be expressed in terms of the fundamental unknowns. Recall the relations between the end forces and displacements for an elastic beam member,

$$M_{ij} = \frac{EI_{ij}}{\ell_{ij}}(2\phi_i + \phi_j - 3\psi_{ij}), \text{ where } \psi_{ij} = \frac{w_i - w_j}{\ell_{ij}} \tag{3.40}$$

$$M_{ji} = \frac{EI_{ij}}{\ell_{ij}}(\phi_i + 2\phi_j - 3\psi_{ij}) \tag{3.41}$$

$$Z_{ij} = -Z_{ji} = -\frac{M_{ij} + M_{ji}}{\ell_{ij}} \tag{3.42}$$

which in matrix notation read

$$\begin{Bmatrix} Z_{ij} \\ M_{ij} \\ Z_{ji} \\ M_{ji} \end{Bmatrix} = \frac{EI_{ij}}{\ell_{ij}^3} \begin{bmatrix} 12 & -6\ell & -12 & -6\ell \\ -6\ell & 4\ell^2 & 6\ell & 2\ell^2 \\ -12 & 6\ell & 12 & 6\ell \\ -6\ell & 2\ell^2 & 6\ell & 4\ell^2 \end{bmatrix}_{ij} \begin{Bmatrix} w_i \\ \phi_i \\ w_j \\ \phi_j \end{Bmatrix} \tag{3.43}$$

In our case, we can write

$$Z_{21} = \frac{EI}{(2L)^3}[12w_2 + 6(2L)\phi_2] = \frac{EI}{L^3}(1.5w_2 + 1.5L\phi_2) \tag{3.44}$$

$$M_{21} = \frac{EI}{(2L)^3}[6(2L)w_2 + 4(2L)^2\phi_2] = \frac{EI}{L^3}(1.5w_2 + 2L\phi_2) \tag{3.45}$$

$$Z_{23} = \frac{EI}{L^3}(12w_2 - 6L\phi_2) \tag{3.46}$$

$$M_{23} = \frac{EI}{L^3}(-6Lw_2 + 4L^2\phi_2) \tag{3.47}$$

Substituting this into (3.38) and (3.39), we obtain

$$\frac{EI}{L^3}(13.5w_2 - 4.5L\phi_2) = F \tag{3.48}$$

$$\frac{EI}{L^3}(-4.5w_2 + 6L^2\phi_2) = 0 \tag{3.49}$$

Alternatively, the elastic stiffness matrix

$$\boldsymbol{K} = \frac{EI}{L^3} \begin{bmatrix} 13.5 & -4.5L \\ -4.5L & 6L^2 \end{bmatrix} \tag{3.50}$$

could be constructed by assembling the elastic stiffness matrices of individual members. This formalized procedure is recorded in Table 3.2, where the common factor EI/L^3 has been left out.

Table 3.2 Assembly of elastic stiffness matrix

$\boldsymbol{K}_{12}$	0	0	1	2
0	.	.	.	.
0	.	.	.	.
1	.	.	1.5	$1.5L$
2	.	.	$1.5L$	$2L^2$

$\boldsymbol{K}_{23}$	1	2	0	0
1	12	$-6L$	.	.
2	$-6L$	$4L^2$	.	.
0	.	.	.	.
0	.	.	.	.

$\boldsymbol{K}$	1	2
1	1.5+12	$1.5L - 6L$
2	$1.5L - 6L$	$2L^2 + 4L^2$

Equations (3.48) (3.49) describe the initial (elastic) stage of the loading process. Substituting the solution

$$\begin{Bmatrix} w_2 \\ \phi_2 \end{Bmatrix} = \begin{Bmatrix} 0.09877L \\ 0.07407 \end{Bmatrix} \frac{FL^2}{EI} \tag{3.51}$$

into the equations for the end moments, we get

$$M_{12} = \frac{EI}{L^3}(1.5Lw_2 + L^2\phi_2) = 0.2222FL \tag{3.52}$$

$$M_{21} = \frac{EI}{L^3}(1.5Lw_2 + 2L^2\phi_2) = 0.2963FL \tag{3.53}$$

$$M_{23} = \frac{EI}{L^3}(-6Lw_2 + 4L^2\phi_2) = -0.2963FL \tag{3.54}$$

$$M_{32} = \frac{EI}{L^3}(-6Lw_2 + 2L^2\phi_2) = -0.4444FL \tag{3.55}$$

The equilibrium equation $M_{21} + M_{23} = 0$ must be satisfied at any stage of the loading process, and so the moment M_{23} will always be equal to $-M_{21}$. In the following calculations, we deal only with M_{21}.

The elastic solution remains valid as long as the bending moments along the entire beam stay below the plastic moment, M_0. Note that the moment distribution along each member (1-2 and 2-3) is linear, and so it is sufficient to check the moments at critical sections 1, 2 and 3. The plastic moment is first reached at section 3. In the elastic limit state we have

$$F^{(1)} = \frac{-M_0}{-0.4444L} = 2.25\frac{M_0}{L} \tag{3.56}$$

$$M_{12}^{(1)} = 0.2222F^{(1)}L = 0.5M_0 \tag{3.57}$$

$$M_{21}^{(1)} = 0.2963F^{(1)}L = 0.6667M_0 \tag{3.58}$$

$$M_{32}^{(1)} = -0.4444F^{(1)}L = -M_0 \tag{3.59}$$

The corresponding moment diagram is shown in Figure 3.5(f).

STEP 2: Let us now proceed to the second stage of incremental plastic analysis, in which the cross section at the right support behaves as a plastic hinge. The moment at this section remains constant, and so the structure *incrementally* behaves as if it had a real hinge at section 3 (Figure 3.5(a)). Assembling the tangential stiffness matrix for this stage of the loading process, we take into account that the stiffness matrix of member 2-3 is now determined according to the formula

$$\begin{Bmatrix} Z_{ij} \\ M_{ij} \\ Z_{ji} \\ M_{ji} \end{Bmatrix} = \frac{EI_{ij}}{\ell_{ij}^3} \begin{bmatrix} 3 & -3\ell & -3 & 0 \\ -3\ell & 3\ell^2 & 3\ell & 0 \\ -3 & 3\ell & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}_{ij} \begin{Bmatrix} w_i \\ \phi_i \\ w_j \\ \phi_j \end{Bmatrix} \tag{3.60}$$

valid for a beam with a hinge connection at the right end. In our case, we have

$$\boldsymbol{K}_{23}^{(2)} = \frac{EI}{L^3} \begin{bmatrix} 3 & -3L & -3 & 0 \\ -3L & 3L^2 & 3L & 0 \\ -3 & 3L & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.61}$$

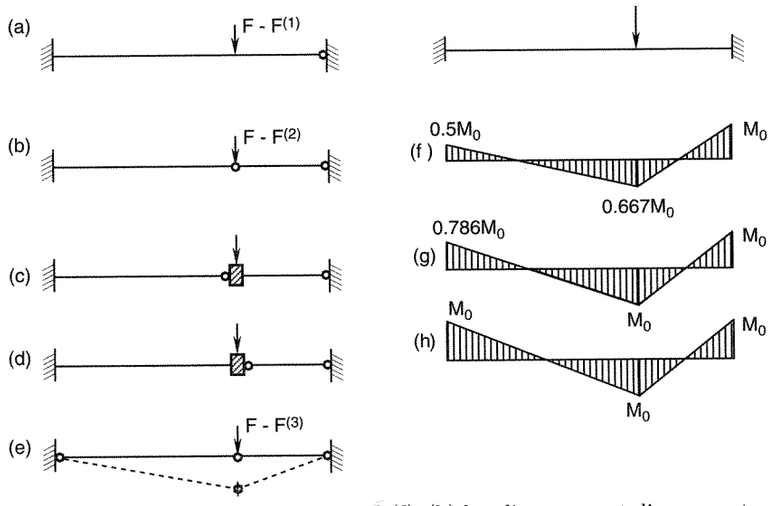


Figure 3.5 (a)–(e) Structural models and (f)–(h) bending moment diagrams at various stages of analysis

and the left upper block of this matrix contributes to the structural stiffness matrix. The contribution of member 1-2 remains unchanged ($\boldsymbol{K}_{12}^{(2)} = \boldsymbol{K}_{12}^{(1)}$), and so the resulting tangential stiffness matrix reads

$$\boldsymbol{K}^{(2)} = \frac{EI}{L^3}\begin{bmatrix} 1.5+3 & 1.5L-3L \\ 1.5L-3L & 2L^2+3L^2 \end{bmatrix} = \begin{bmatrix} 4.5 & -1.5L \\ -1.5L & 5L^2 \end{bmatrix} \tag{3.62}$$

The incremental equations of equilibrium

$$\boldsymbol{K}^{(2)}\dot{\boldsymbol{d}} = \dot{\boldsymbol{f}} \tag{3.63}$$

yield the generalized displacement rates

$$\begin{Bmatrix} \dot{w}_2 \\ \dot{\phi}_2 \end{Bmatrix} = \begin{Bmatrix} 0.24691L \\ 0.07407 \end{Bmatrix} \frac{\dot{F}L^2}{EI} \tag{3.64}$$

and the end moment rates

$$\dot{M}_{12} = \frac{EI}{L^3}(1.5L\dot{w}_2 + L^2\dot{\phi}_2) = 0.4444\dot{F}L \tag{3.65}$$

$$\dot{M}_{21} = \frac{EI}{L^3}(1.5L\dot{w}_2 + 2L^2\dot{\phi}_2) = 0.5185\dot{F}L \tag{3.66}$$

$$\dot{M}_{32} = 0 \tag{3.67}$$

Moment M_{12} would reach M_0 at

$$\Delta F = \frac{M_0 - M_{12}^{(1)}}{0.4444L} = \frac{1-0.5}{0.4444}\frac{M_0}{L} = 1.125\frac{M_0}{L} \tag{3.68}$$

but moment M_{21} reaches M_0 earlier, at

$$\Delta F = \frac{M_0 - M_{21}^{(1)}}{0.5185L} = \frac{1-0.6667}{0.5185}\frac{M_0}{L} = 0.6429\frac{M_0}{L} \tag{3.69}$$

Therefore, we set

$$\Delta F^{(2)} = 0.6429\frac{M_0}{L} \tag{3.70}$$

and evaluate

$$F^{(2)} = F^{(1)} + \Delta F^{(2)} = 2.8929\frac{M_0}{L} \tag{3.71}$$

$$M_{12}^{(2)} = M_{12}^{(1)} + 0.4444\Delta F^{(2)}L = 0.7857M_0 \tag{3.72}$$

$$M_{21}^{(2)} = M_{21}^{(1)} + 0.5185\Delta F^{(2)}L = M_0 \tag{3.73}$$

$$M_{32}^{(2)} = M_{32}^{(1)} = M_0 \tag{3.74}$$

The distribution of bending moments after step 2 is shown in Figure 3.5(g).

STEP 3: During the third stage of the loading process, the structure has two plastic hinges at sections 2 and 3, and incrementally behaves as an elastic structure with ideal hinges (Figure 3.5(b)). This is a statically determinate structure and the moment rate

$$\dot{M}_{12} = 2\dot{F}L \tag{3.75}$$

immediately follows from elementary equilibrium considerations. However, we can also arrive at this results more systematically using the same formalism as before.

The structure in Figure 3.5(b) could be described by a single equation with the vertical displacement w_2 as the only unknown. However, we successively analyze structural models derived from the initial structure by inserting hinges at certain locations and, in that case, it is more convenient from the algorithmic point of view to keep the same set of unknowns throughout the entire analysis. Of course, the rotation of the right end of member 1-2 now differs from the rotation of the left end of member 2-3, and we have to select one of these rotations to represent the second fundamental unknown, ϕ_2. For the final results, it is irrelevant which of the two possible choices is made. We could imagine that ϕ_2 is the rotation of an infinitely small rigid corner that connects the right end of member 1-2 to the left end of member 2-3. Theoretically, moment M_{21} reaches the plastic limit value M_0 at the same time as moment M_{23} reaches $-M_0$. However, due to random imperfections, one of the end cross sections is inevitably slightly weaker than the other one, and the plastic hinge forms at the weaker cross section, while the stronger cross section remains just below the plastic limit.

Assuming that the plastic hinge forms to the left of the joint (Figure 3.5(c)), we would treat both members as rigidly connected at their left end and hinged at their right end. On the other hand, assuming that the plastic hinge forms to the right of the joint (Figure 3.5(d)), we would treat member 1-2 as rigidly connected at both ends and member 2-3 as hinged at both ends. We will demonstrate the solution corresponding to the former choice, leaving it up to the reader to verify that the latter choice would lead to equivalent results (see Problem 3.3).

After a plastic hinge forms at the right end of member 1-2, the tangential stiffness matrix of this member changes to

$$\boldsymbol{K}_{12}^{(3)} = \frac{EI}{L^3}\begin{bmatrix} 0.375 & -0.75L & -0.375 & 0 \\ -0.75L & 1.5L^2 & 0.75L & 0 \\ -0.375 & 0.75L & 0.375 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.76}$$

while the stiffness matrix of member 2-3 remains the same as in the previous step ($\boldsymbol{K}_{23}^{(3)} = \boldsymbol{K}_{23}^{(2)}$, see equation (3.61)). The usual assembly process leads to the tangential stiffness matrix

$$\boldsymbol{K}^{(3)} = \frac{EI}{L^3}\begin{bmatrix} 3.375 & -3L \\ -3L & 3L^2 \end{bmatrix} \tag{3.77}$$

Following the standard procedure, we compute the generalized displacement rates

$$\begin{Bmatrix} \dot{w}_2 \\ \dot{\phi}_2 \end{Bmatrix} = \begin{Bmatrix} 2.667L \\ 2.667 \end{Bmatrix} \frac{\dot{F}L^2}{EI} \tag{3.78}$$

and the end moment rates

$$\dot{M}_{12} = \frac{EI}{L^3} 0.75L\dot{w}_2 = 2\dot{F}L \tag{3.79}$$

$$\dot{M}_{21} = 0 \tag{3.80}$$

$$\dot{M}_{32} = 0 \tag{3.81}$$

Moment M_{12} reaches the plastic moment M_0 at

$$\Delta F^{(3)} = \frac{M_0 - M_{12}^{(2)}}{2L} = \frac{1 - 0.7857}{2}\frac{M_0}{L} = 0.1071\frac{M_0}{L} \tag{3.82}$$

when

$$F^{(3)} = F^{(2)} + \Delta F^{(3)} = 3\frac{M_0}{L} \tag{3.83}$$

$$M_{12}^{(3)} = M_{12}^{(2)} + 2\Delta F^{(3)}L = M_0 \tag{3.84}$$

STEP 4: After the third plastic hinge forms, the structure turns into a mechanism with one degree of freedom (Figure 3.5(e)) and cannot resist any further increments of the applied force. In the numerical solution, this is manifested by the singularity of the tangential stiffness matrix. According to our assumption that the second hinge has formed to the left of joint 2, member 1-2 is now connected by hinges at both ends, and so it has no bending stiffness ($\boldsymbol{K}_{12}^{(4)} = \boldsymbol{O}$). From the contribution of member 2-3 (considered as rigidly connected at the left end and hinged at the right end) we would get

$$\boldsymbol{K}^{(4)} = \frac{EI}{L^3}\begin{bmatrix} 3 & -3L \\ -3L & 3L^2 \end{bmatrix} \tag{3.85}$$

This matrix is indeed singular (the second row is $-L$ times the first row), and so for the present type of loading, the incremental equations of equilibrium have no solution unless $\dot{F} = 0$. This indicates that $F^{(3)}$ is the plastic collapse load.

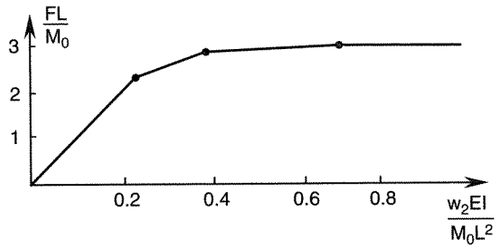


Figure 3.6 Normalized load-displacement diagram

Based on the results of individual steps of our incremental analysis, we can now construct the load-displacement diagram showing the variation of the applied load as a function of the load-point displacement (Figure 3.6). Figure 3.5(h) shows the bending moment diagram at failure (i.e. after step 3).

3.3 FRAMES

3.3.1 Standard Update of Structural Stiffness Matrix

The incremental elastoplastic analysis of plane frames again proceeds according to the general Algorithm 3.1. Every joint now has up to three unknown generalized displacements – horizontal displacement u_i, vertical displacement w_i and rotation ϕ_i. The simplified version of the displacement method based on the assumption of axial inextensibility reduces the number of unknowns by imposing the condition that the length of each member remains constant. The contribution of a beam or frame member to the global stiffness matrix does not completely vanish after the formation of a plastic hinge at one of the end cross sections. If the member has initially been connected in a rigid manner at both ends, then, after the appearance of a plastic hinge, the tangential stiffness matrix becomes identical to the elastic stiffness matrix of a member rigidly connected at the end cross section that remains elastic and hinged at the end cross section that is plastic. After the formation of plastic hinges at both ends, the bending stiffness matrix of the member vanishes, and the member contributes to the global stiffness matrix only by its axial stiffness. In the simplified approach based on axial inextensibility (Appendix B.5), the member shortening is neglected, and the part of the stiffness matrix due to axial stiffness is not considered (it is taken into account by the choice of the independent degrees of freedom). We will first demonstrate that a simple problem can be solved by hand, and then comment on the algorithmic aspects of the method.

Example 3.4: Perform incremental elastoplastic analysis of the frame in Figure 3.7(a). All the members have the same cross section.

Solution:
STEP 1: We will use the simplified method assuming inextensibility of the members, which leads to only three unknown generalized displacements described by the global degree of freedom numbers indicated in Figure 3.7(b). The first global degree of freedom is the horizontal displacement of the floor, $u_2 \equiv u_3$, and the second and third degrees of freedom are the joint rotations ϕ_2 and ϕ_3.

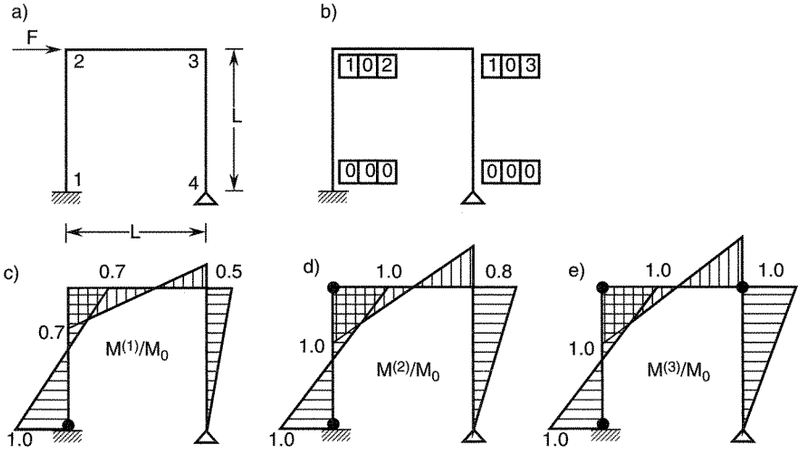


Figure 3.7 Portal frame: a) geometry and loading, b) global degree of freedom numbers at joints, c–e) distribution of bending moments at various stages of analysis

The elastic analysis of this frame is demonstrated in Appendix B (Example B.6), and the elastic bending moment diagram can be found in Figure B.9(d). As all the cross sections have the same plastic moment, M_0, the elastic limit state is reached at

$$F^{(1)} = \frac{M_0}{0.4545L} = 2.2\frac{M_0}{L} \tag{3.86}$$

when a plastic hinge forms at the bottom end of the left column. The corresponding distribution of bending moments is shown in Figure 3.7(c).

STEP 2: After the formation of the first plastic hinge, the contribution of member 1-2 (the left column) to the structure stiffness matrix must be evaluated from formula (B.80). Table 3.3 illustrates the assembly procedure (only the stiffness coefficients with nonzero global degree of freedom numbers are displayed, and the common factor

Table 3.3 Assembly procedure

$\boldsymbol{K}^{hr}_{b,12}$	1	2
1	3	$3L$
2	$3L$	$3L^2$

$\boldsymbol{K}_{b,23}$	1	2	1	3
1	0	0	0	0
2	0	$4L^2$	0	$2L^2$
1	0	0	0	0
3	0	$2L^2$	0	$4L^2$

$\boldsymbol{K}^{rh}_{b,34}$	1	3
1	3	$3L$
3	$3L$	$3L^2$

$\boldsymbol{K}^{(2)}$	1	2	3
1	3+3	$3L$	$3L$
2	$3L$	$3L^2+4L^2$	$2L^2$
3	$3L$	$2L^2$	$4L^2+3L^2$

EI/L^3 is left out). The incremental equations of equilibrium

$$\frac{EI}{L^3}\begin{bmatrix} 6 & 3L & 3L \\ 3L & 7L^2 & 2L^2 \\ 3L & 2L^2 & 7L^2 \end{bmatrix}\begin{Bmatrix} \dot{u}_2 \\ \dot{\phi}_2 \\ \dot{\phi}_3 \end{Bmatrix} = \dot{F}\begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix} \tag{3.87}$$

lead to

$$\begin{Bmatrix} \dot{u}_2 \\ \dot{\phi}_2 \\ \dot{\phi}_3 \end{Bmatrix} = \frac{\dot{F}L^2}{EI}\begin{Bmatrix} 0.25L \\ -0.0833 \\ -0.0833 \end{Bmatrix} \tag{3.88}$$

and

$$\dot{M}_{23} = \frac{2EI}{L}\left(2\dot{\phi}_2 + \dot{\phi}_3\right) = -0.5\dot{F}L \tag{3.89}$$

$$\dot{M}_{32} = \frac{2EI}{L}\left(\dot{\phi}_2 + 2\dot{\phi}_3\right) = -0.5\dot{F}L \tag{3.90}$$

The second plastic hinge forms at joint 2 when the applied force attains the value

$$F^{(2)} = F^{(1)} + \frac{M_0 - 0.7M_0}{0.5L} = 2.2\frac{M_0}{L} + 0.6\frac{M_0}{L} = 2.8\frac{M_0}{L} \tag{3.91}$$

The corresponding distribution of bending moments is shown in Figure 3.7(d). As the moments at cross sections 21 and 23 have the same absolute value, we can select either cross section as the location of the hinge. In reality, the formation of the plastic hinge is also affected by the normal force, which is larger at cross section 21. However, the present analysis neglects this effect, and so the results do not depend on deciding which of the two cross sections plasticizes.

STEP 3: At this stage of analysis, we have several options:

1. First, we can solve the last step just from equilibrium equations because the structure is now (incrementally) statically determinate.
2. We could also stick to the displacement method, but treat only u_2 and ϕ_3 as unknown generalized displacements. In this case, we would use stiffness matrices $\boldsymbol{K}^{hh}_{b,12}$, $\boldsymbol{K}^{hr}_{b,23}$ and $\boldsymbol{K}^{rh}_{b,34}$.
3. In a fully automatic solution, it is simpler to keep the same unknowns throughout the entire analysis:
   (a) Selecting for example 21 as the plasticized cross section, we assemble the tangential stiffness matrix from the contributions of $\boldsymbol{K}^{hh}_{b,12}$, $\boldsymbol{K}_{b,23}$ and $\boldsymbol{K}^{rh}_{b,34}$. The unknown ϕ_2 is then identical with the rotation of the left end of beam 2-3.
   (b) Alternatively, we could select 23 as the plastic cross section, and assemble the tangential stiffness matrix from the contributions of $\boldsymbol{K}^{hr}_{b,12}$, $\boldsymbol{K}^{hr}_{b,23}$ and $\boldsymbol{K}^{rh}_{b,34}$. The unknown ϕ_2 is then identical with the rotation of the top end of column 1-2.

The reader can verify that all the suggested approaches lead to

$$\dot{M}_{34} = \dot{F}L \tag{3.92}$$

and the third plastic hinge is formed at

$$\dot{F}^{(3)} = F^{(2)} + \frac{M_0 - 0.8M_0}{L} = 2.8\frac{M_0}{L} + 0.2\frac{M_0}{L} = 3\frac{M_0}{L} \tag{3.93}$$

After that, the structure fails. Figure 3.7(e) shows the distribution of bending moments at collapse. □

The preceding example illustrates the fundamental aspects of incremental frame analysis. A corresponding computer procedure can easily be developed. In the data describing the structure, we include an array of flags that indicate for each end of every member whether it is connected to the joint in a rigid manner or by a hinge. The initial values of these flags are set to 'hinge' at locations of real structural hinges. After each incremental step, the flag corresponding to the newly formed plastic hinge is modified from 'rigid' to 'hinge'. The tangential stiffness matrix is always assembled based on the current type of end connections.

Note that nothing in the previous procedure prevents us from using the standard displacement method with axially compressible members. We simply choose the full set of generalized displacements, and take into account the contribution of axial member stiffnesses. This approach is algorithmically more convenient because we do not have to construct and satisfy the conditions of axial inextensibility. For rectangular frames, the conditions of axial inextensibility can be imposed by modifying the global degree of freedom numbers. However, for general nonrectangular frames, the structure of these conditions is more complicated and, given the speed and memory capacity of modern computers, the resulting savings are seldom worth the additional programming effort.

Finally, we should comment on the assembly and decomposition of the tangential stiffness matrix. The reader has probably noticed that the stiffness matrix at a given stage of analysis, $\boldsymbol{K}^{(n)}$, differs from the stiffness matrix at the immediately preceding stage, $\boldsymbol{K}^{(n-1)}$, only by the contribution of the member in which a new plastic hinge has formed. We could therefore construct $\boldsymbol{K}^{(n)}$ by correcting $\boldsymbol{K}^{(n-1)}$ by the difference $\boldsymbol{K}_{ij}^{(n)} - \boldsymbol{K}_{ij}^{(n-1)}$, where subscripts ij refer to member i-j with the new plastic hinge. This would save us the evaluation and assembly of all the other member stiffness matrices. The structural stiffness matrix is normally destroyed by the procedure solving the governing equations, and so the solution procedure has to be applied to a copy of the matrix, while the original is stored for future use.

3.3.2 Indirect Update of Structural Stiffness Matrix

For large problems, the time required to solve the equations grows more quickly than the time needed to assemble them. It is therefore tempting to look for a method that could reuse the knowledge of a decomposition of $\boldsymbol{K}^{(n-1)}$ when solving a set of equations with $\boldsymbol{K}^{(n)}$ as the system matrix. Indeed, such a method can be designed by exploiting an interesting concept known as the *theorem of structural variation* (Majid and Celik, 1985), which is closely related to the method of virtual distortions (Holnicki-Szulc and Gierlinski, 1989) and the method of inelastic forces (Jirásek and Bažant, 1995). In general, such theorems make it possible to study the influence of small design changes (e.g. a change of the position of one joint, or of the moment of inertia of one member) without starting the analysis from scratch. The present application of the theorem of structural variation in the context of

incremental elastoplastic analysis is equivalent to the combined elastoplastic and limit analysis via *restricted basis linear programming* (Maier, Giacomini and Paterlini, 1979). Structural variation provides a mechanical interpretation of the mathematical operations performed by such sophisticated algorithms. The essence of this approach is best explained by a simple example.

Example 3.5: Solve the second step of incremental analysis of the frame from Example 3.4 by reusing the elastic stiffness matrix.

Solution: In the first step of the incremental analysis, we set up the elastic stiffness matrix for a frame with the bottom end of the left column rigidly connected to the joint (Figure 3.7(a)). The joint happened to be fixed by an external support, but this is irrelevant for the applicability of the present technique. The fundamental generalized displacements (global degrees of freedom) were u_2, ϕ_2 and ϕ_3. After the formation of the first plastic hinge, we should solve the structure in Figure 3.8(a), which differs from the original one only by the type of connection of the bottom end of column 1-2. Instead of setting up the new stiffness matrix, we can imagine for a while that

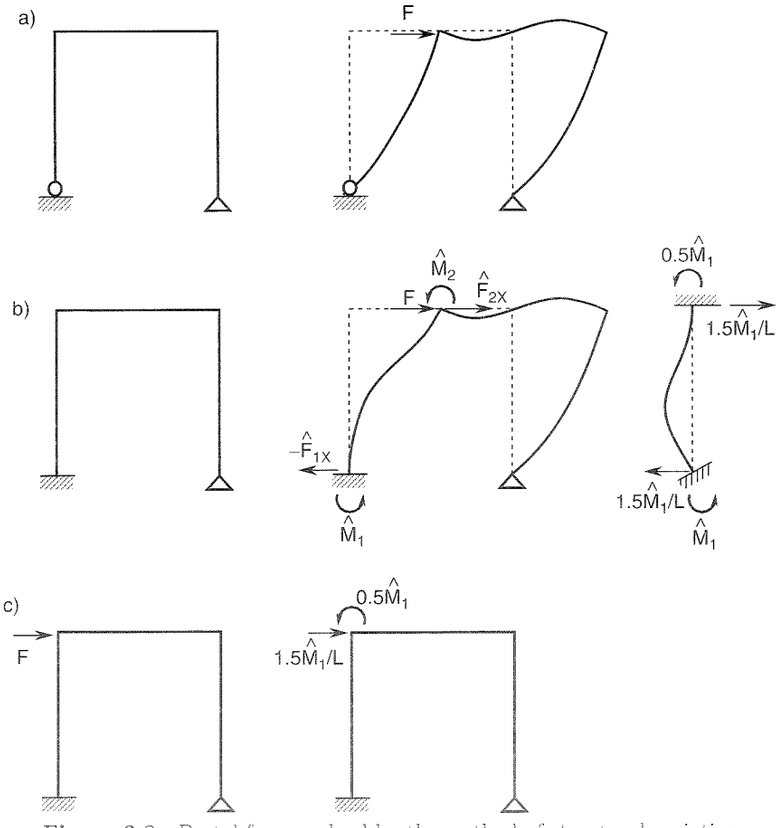


Figure 3.8 Portal frame solved by the method of structural variation

we already know the resulting floor displacement[5] u_2 and joint rotations ϕ_2 and ϕ_3, and apply them to the original structure from Figure 3.7(a). This auxiliary problem is shown in Figure 3.8(b). When we compare it to the actual problem in Figure 3.8(a), we can see that only member 1-2 (the one in which the plastic hinge has formed) is in a different state. For all the other members, the end displacements and end forces are exactly the same for the auxiliary problem as for the actual problem. To ensure equilibrium between the applied and internal forces for the auxiliary problem, we have to add a set of fictitious forces and moments acting as loads on joints 1 and 2. These forces balance the difference between (1) the end forces that would be generated by the actual joint displacements on member 1-2 considered as rigidly connected to joint 1 (in the auxiliary state), and (2) the end forces that would be generated on the same member 1-2 considered as hinge-connected to joint 1 (in the actual state). Consequently, these forces can be expressed in terms of the actual displacements as

$$\hat{\boldsymbol{f}} = \boldsymbol{K}_{12}\boldsymbol{d}_{12} - \boldsymbol{K}_{12}^{hr}\boldsymbol{d}_{12} = (\boldsymbol{K}_{12} - \boldsymbol{K}_{12}^{hr})\boldsymbol{d}_{12} \tag{3.94}$$

$$\begin{Bmatrix} \hat{F}_{1x} \\ \hat{F}_{1z} \\ \hat{M}_1 \\ \hat{F}_{2x} \\ \hat{F}_{2z} \\ \hat{M}_2 \end{Bmatrix} = \frac{EI}{L^3} \begin{bmatrix} 9 & 0 & -6L & -9 & 0 & -3L \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -6L & 0 & 4L^2 & 6L & 0 & 2L^2 \\ -9 & 0 & 6L & 9 & 0 & 3L \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -3L & 0 & 2L^2 & 3L & 0 & L^2 \end{bmatrix} \begin{Bmatrix} u_1 \\ w_1 \\ \phi_1 \\ u_2 \\ w_2 \\ \phi_2 \end{Bmatrix} \tag{3.95}$$

Examining the rows of the matrix $\boldsymbol{K}_{12} - \boldsymbol{K}_{12}^{hr}$, we realize that they are all multiples of one of them (the rank of the matrix is 1). Therefore, the balancing forces can be expressed in terms of one selected component, e.g. $\hat{M}_1$:

$$\hat{F}_{1z} = \hat{F}_{2z} = 0 \tag{3.96}$$

$$\hat{F}_{1x} = -\hat{F}_{2x} = -\frac{1.5}{L}\hat{M}_1 \tag{3.97}$$

$$\hat{M}_2 = \frac{1}{2}\hat{M}_1 \tag{3.98}$$

This is not surprising because the balancing forces are identical with the reactions on member 1-2 fixed at both ends and subjected to a prescribed rotation at end 1 (Figure 3.8(b) right). Such a loading case represents the difference between the auxiliary state of column 1-2 and its actual state.

Now, if we knew the actual displacements, we could easily evaluate the balancing moment from the third equation in (3.95), rewritten as

$$\hat{M}_1 = \frac{2EI}{L}\left(2\phi_1 + \phi_2 + 3\frac{u_2 - u_1}{L}\right) \tag{3.99}$$

On the other hand, if we knew the value of $\hat{M}_1$ (and thus, by virtue of (3.96)–(3.98), also the other components of the balancing load), we could calculate the actual

[5] To keep the notation simple, we do not mark explicitly that what we call the actual displacements and forces are in fact the *increments* of these quantities in the second step of incremental analysis.

displacements by solving the auxiliary problem, for which the stiffness matrix has already been constructed. The basic idea of the method of structural variation is that we solve the auxiliary problem with $\hat{M}_1$ treated as a variable load parameter (this step provides expressions for the generalized displacements in terms of the actual load F and the balancing load $\hat{M}_1$), and then we determine $\hat{M}_1$ from equation (3.99).

Since the stiffness matrix for the auxiliary structure is already available (from the elastic analysis in the first incremental step), the auxiliary problem is easy to solve. We decompose the total loading into two loading cases – one of them corresponds to the actual applied loads, and the other to the fictitious balancing loads. The first loading case (Figure 3.8(c) left) is in fact identical to that already solved in the preceding incremental step, and so we can reuse the solution. The second case (Figure 3.8(c) right) leads to

$$\frac{EI}{L^3}\begin{bmatrix} 15 & 6L & 3L \\ 6L & 8L^2 & 2L^2 \\ 3L & 2L^2 & 7L^2 \end{bmatrix} \begin{Bmatrix} \hat{u}_2 \\ \hat{\phi}_2 \\ \hat{\phi}_3 \end{Bmatrix} = \hat{M}_1 \begin{Bmatrix} 1.5/L \\ 0.5 \\ 0 \end{Bmatrix} \tag{3.100}$$

from which

$$\begin{Bmatrix} \hat{u}_2 \\ \hat{\phi}_2 \\ \hat{\phi}_3 \end{Bmatrix} = \frac{\hat{M}_1 L}{EI} \begin{Bmatrix} 0.1136L \\ -0.0113 \\ -0.0455 \end{Bmatrix} \tag{3.101}$$

Combining this with the first loading case (solved in Example B.6), we obtain

$$u_2 = \frac{L^2}{EI}\left(0.09848FL + 0.1136\hat{M}_1\right) \tag{3.102}$$

$$\phi_2 = \frac{L}{EI}\left(-0.06818FL - 0.0113\hat{M}_1\right) \tag{3.103}$$

$$\phi_3 = \frac{L}{EI}\left(-0.02273FL - 0.0455\hat{M}_1\right) \tag{3.104}$$

Now we substitute expressions (3.102)–(3.103) into (3.99). Of course, the displacement u_1 and rotation ϕ_1 of the fixed joint are prescribed as zero. The resulting equation

$$\hat{M}_1 = 2 \times \left(-0.06818FL - 0.0113\hat{M}_1\right) + 6 \times \left(0.09848FL + 0.1136\hat{M}_1\right) \tag{3.105}$$

yields

$$\hat{M}_1 = 1.333FL \tag{3.106}$$

Substituting this back into (3.102)–(3.104), we finally get the actual generalized displacements

$$u_2 = 0.25\frac{FL^3}{EI} \tag{3.107}$$

$$\phi_2 = -0.0833\frac{FL^2}{EI} \tag{3.108}$$

$$\phi_3 = -0.0833\frac{FL^2}{EI} \tag{3.109}$$

The result exactly agrees with that obtained in the second incremental step in Example 3.4. Note that we have completely avoided assembling and decomposing the tangential stiffness matrix $\boldsymbol{K}^{(2)}$. □

At first glance, the method of structural variation looks complicated, but it can be fully formalized and implemented into a computer code. The relation between the balancing force parameter and the generalized displacements is always simple and independent of the size of the model (number of joints and members), and the solution of the loading case with the actual loads can be taken from the previous step. The only computational effort is involved in the solution of the loading case with balancing loads. However, as the decomposition of the stiffness matrix is already available from the previous step, it is sufficient to perform manipulations on the right-hand side. For large sets of linear equations, the number of right-hand side operations is much smaller than the number of operations needed to decompose the system matrix, and so the solution is substantially speeded up.

Moreover, the procedure can be generalized such that the same stiffness matrix could be reused in several subsequent steps. Each newly formed plastic hinge increases the number of additional unknowns (balancing force parameters) by one, and instead of a single equation (3.99) we have to set up a set of equations with a fully populated (i.e. not sparse) matrix. The governing equations of the displacement method still need to be solved only once per step, always with the same stiffness matrix, and so this approach remains faster than the traditional one (updating and decomposing the stiffness matrix), unless the number of additional unknowns becomes large. If such a situation arises, it is possible to assemble and decompose the current tangential stiffness matrix, and then use it when constructing the auxiliary problem in the subsequent steps. The number of additional unknowns again starts increasing from one, and the efficiency of the method is regained.

Finally, note that, in each step of the incremental analysis, it should be checked whether all the previously formed plastic hinges keep yielding. This can be deduced from the signs of the corresponding plastic rotation increments. If unloading is detected in one or more hinges, the solution is not valid and the step should be recomputed using a modified structural model (unloading plastic hinges replaced by rigid connections).

PROBLEMS

Problem 3.1: Find the failure load for the truss in Figure 3.1.

Problem 3.2: Perform incremental analysis of the truss from Figure 3.2 loaded by an increasing *horizontal* force F. Find the plastic limit value of F, and determine which bars will yield at collapse.

Problem 3.3: Perform the third step of elastoplastic analysis of the beam from Example 3.3 assuming that the second plastic hinge forms to the right of joint 2. Compare the results to those from the example.

Problem 3.4*: Formulate the method of structural variation for trusses, and apply it to the structure from Examples 3.1–3.2. Solve the second step using only the elastic stiffness matrix.

Problem 3.5*: Generalize the method of structural variation to situations in which the actual structure differs from the auxiliary structure by more than a plastic bar

or plastic hinge. Solve the third step of Example 3.4 using only the elastic stiffness matrix.
Hint: You will have to introduce two balancing force parameters (one for each plastic hinge), and construct a set of two equations relating these parameters to the basic unknowns of the displacement method.

4

Elementary Limit Analysis

The incremental analysis traces the entire evolution of structural response to an increasing external loading. When evaluating the ultimate load-carrying capacity of the structure, we are primarily interested in the final stage of the plastic response – plastic collapse. It turns out that the plastic limit load can be assessed without analyzing the entire history of the response. This is the subject of *limit analysis*. Before we prove general theorems which form the theoretical basis of limit analysis, let us demonstrate the basic concepts by some simple examples.

4.1 TRUSSES

In general, any structure must become a kinematic mechanism for plastic collapse to occur. For statically indeterminate trusses, a mechanism is obtained when the number of yielding bars exceeds the degree of static redundancy. As the number of yielding bars increases, we first obtain a mechanism with one degree of freedom. A mechanism with two or more degrees of freedom can never be reached, since the structure collapses before another bar can start yielding. Of course, a mechanism with two degrees of freedom could in theory occur if, at collapse, two bars started yielding simultaneously, e.g. due to symmetry. However, even then there exists a mechanism with only one degree of freedom that gives the same limit load as the 'actual' one, and so it is sufficient to consider only single-degree-of-freedom mechanisms. Besides, there are always inevitable imperfections, which means that one of the two hinges would start yielding before the other.

Example 4.1: Find the load F for which the truss in Figure 4.1(a) collapses, given that the plastic axial force S_0 is the same for all the bars.

Solution: The truss is statically indeterminate to the first degree, and so collapse occurs if two bars yield. As the total number of bars is three, we have only three potential failure mechanisms. During collapse, the length of the elastic bar remains constant. This condition makes it possible to describe the extension rates and the displacement rates in terms of a single variable. For example, assuming that bar 1 remains elastic and bars 2 and 3 yield (Figure 4.1(b)), we must have

$$\dot{e}_1 = 0.6\dot{u} + 0.8\dot{w} = 0 \tag{4.1}$$

from which $\dot{u} = -1.333\ \dot{w}$. Substituting this into the remaining kinematic relations, we express all the extension rates in terms of one selected rate variable, $\dot{w}$:

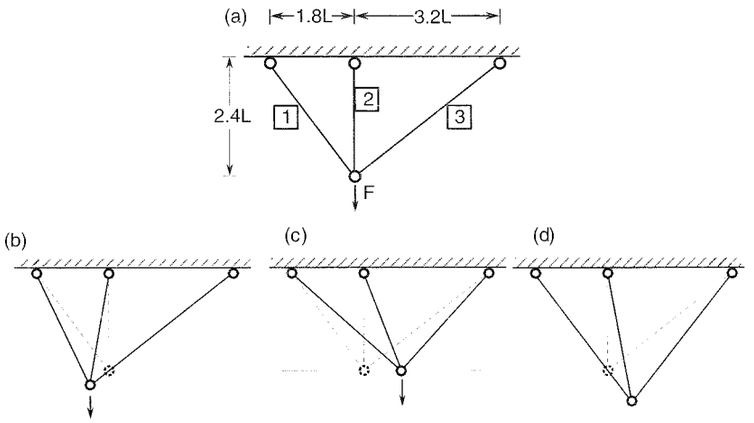


Figure 4.1 Potential failure mechanisms for the three-bar truss

$$\dot{e}_1 = 0 \tag{4.2}$$

$$\dot{e}_2 = \dot{w} \tag{4.3}$$

$$\dot{e}_3 = -0.8\dot{u} + 0.6\dot{w} = 1.667\dot{w} \tag{4.4}$$

We consider that inertial forces during collapse are negligible (which means, for example, that we exclude impact from our considerations). Then, the structure must be in a static equilibrium state during the collapse, and we may calculate the collapse load from equilibrium conditions. The external force F and the internal forces S_1, S_2, S_3 must satisfy two equations of equilibrium

$$0.6S_1 - 0.8S_3 = 0 \tag{4.5}$$

$$0.8S_1 + S_2 + 0.6S_3 = F \tag{4.6}$$

At collapse, the axial forces in the yielding bars are equal to plus or minus the plastic axial force S_0. Their signs must agree with the signs of the respective extension rates. In our case, it might be expected that $\dot{w} > 0$ during collapse, and so $\dot{e}_2 > 0$ and $\dot{e}_3 > 0$. Consequently, $S_2 = S_0$ and $S_3 = S_0$. The equations of equilibrium now contain only two unknowns, S_1 and F, which can be easily calculated:

$$S_1 = \frac{0.8}{0.6} S_3 = 1.333 S_0 \tag{4.7}$$

$$F = 0.8S_1 + S_2 + 0.6S_3 = 0.8 \times 1.333S_0 + S_0 + 0.6S_0 = 2.667S_0 \tag{4.8}$$

Thus, if this were the correct failure mechanism, it would occur at the load $F = 2.667S_0$. However, the axial force $S_1 = 1.333S_0$, calculated from equilibrium, would have to exceed the plastic axial force, which is impossible. This indicates that the considered mechanism is not the actual one.

Before we proceed to other potential failure mechanisms, let us work out a more efficient way of calculating the load from the equations of equilibrium. For the present, very simple structure, the effort involved in solving these equations is negligible. Problems with more unknowns can conveniently be solved by setting up a special equation in which the external load F (in general the load multiplier μ) is the only

unknown. Such an equation can be constructed by applying the *principle of virtual work* (see Appendix D.6). The loads and internal forces at collapse are in equilibrium, and so, according to the principle of virtual work, the work δW_{ext} done by the loads on any virtual displacements must be equal to the work δW_{int} done by the internal forces on the virtual generalized strains compatible with the chosen virtual displacements. If we select as virtual displacements those that correspond to the assumed failure mechanism, nonzero virtual extensions (representing generalized strains) will appear only in the yielding bars, for which the axial forces are known. The only unknown in the *virtual work equality* will be the external load. In our case, we set $\delta u = -1.333\delta w$, $\delta e_1 = 0$, $\delta e_2 = \delta w$, and $\delta e_3 = 1.667\delta w$, and the virtual work equality $\delta W_{\text{ext}} = \delta W_{\text{int}}$ yields, for this particular choice,

$$\begin{aligned} F\,\delta w &= S_1\,\delta e_1 + S_2\,\delta e_2 + S_3\,\delta e_3 \\ &= S_1 \times 0 + S_0 \times \delta w + S_0 \times 1.667\delta w = 2.667\,S_0\,\delta w \end{aligned} \tag{4.9}$$

The result $F = 2.667S_0$ is of course the same as before, but now we have obtained it without solving the complete set of equilibrium equations, and even without setting them up. The principle of virtual work directly leads to a special equilibrium equation that contains only the load parameter as an unknown, and that would otherwise have to be constructed by eliminating the unknown axial forces from the standard equations of equilibrium.

There is also an alternative way of looking at the procedure that leads to the desired equation with only one unknown. During collapse, the external load F is constant, and it does some work on the increasing displacement w. The product of the force, F, and the displacement rate in the sense of F, $\dot{w}$, is the *external power* $\dot{W}_{\text{ext}} = F\dot{w}$. This power is supplied to the structure and, assuming a steady-state collapse with no inertial effects, it must be dissipated by plastic processes in the yielding bars. This fact can be mathematically written as the *power equality*

$$\dot{W}_{\text{ext}} = D_{\text{int}} \tag{4.10}$$

where D_{int} is the rate of dissipation.

According to (2.3), the dissipation rate in a yielding bar is given by $D_{\text{int}} = S_0|\dot{e}_p|$. During plastic collapse, all the elastic deformations remain constant, and so $\dot{e}_p$ can be replaced by $\dot{e}$. Summing the dissipation rates for all the yielding bars, we obtain the total dissipation power of the structure. In our example, extension rates $\dot{e}_2$ and $\dot{e}_3$ are positive (assuming that $\dot{w}$ is positive), and so

$$D_{\text{int}} = S_0\dot{e}_2 + S_0\dot{e}_3 = S_0(\dot{w} + 1.667\dot{w}) = 2.667S_0\dot{w} \tag{4.11}$$

The power equality can now be written as

$$F\dot{w} = 2.667S_0\dot{w} \tag{4.12}$$

from which we again get $F = 2.667S_0$.

The difference between the derivation of (4.9) from the principle of virtual work and the derivation of (4.12) from the power equality is only conceptual – in the former case we deal with the virtual displacements and virtual extensions, in the latter with the displacement rates and extension rates. In the following, we will use the approach

based on the power equality as it is more natural and the corresponding notation is shorter.

To complete the analysis, let us look at the other two potential failure mechanisms:
(1) If bar 2 remains elastic and bars 1 and 3 yield (Figure 4.1(c)), we have the condition

$$\dot{e}_2 = \dot{w} = 0 \tag{4.13}$$

and the external power $\dot{W}_{\text{ext}} = F\dot{w}$ is zero. As the dissipation rate is always positive, the power equality cannot be satisfied for any value of the applied load F, and the present mechanism cannot occur.
(2) If bar 3 remains elastic and bars 1 and 2 yield (Figure 4.1(d)), the condition

$$\dot{e}_3 = -0.8\dot{u} + 0.6\dot{w} = 0 \tag{4.14}$$

implies that

$$\dot{u} = 0.75\dot{w} \tag{4.15}$$
$$\dot{e}_1 = 0.6\dot{u} + 0.8\dot{w} = 1.25\dot{w} \tag{4.16}$$
$$\dot{e}_2 = \dot{w} \tag{4.17}$$

Substituting into the power equality

$$F\dot{w} = S_0\dot{e}_1 + S_0\dot{e}_2 \tag{4.18}$$

we get

$$F\dot{w} = 2.25S_0\dot{w} \tag{4.19}$$

from which

$$F = 2.25S_0. \tag{4.20}$$

Recall that the load calculated from the power equality for the mechanism with bars 2 and 3 yielding is $F = 2.667S_0$, while for the mechanism with bars 1 and 2 yielding it is $F = 2.25S_0$. If the applied load $F < 2.25S_0$, the structure is safe (cannot collapse), because each of the potential failure mechanisms would dissipate more energy than is supplied by the external load. Intuitively, it can be expected that collapse occurs as soon as there is a mechanism for which the dissipation rate equals the external power. Indeed, the incremental analysis of the same problem in the preceding chapter revealed that the structure collapses at $F = 2.25S_0$ by yielding of bars 1 and 2. We can verify this fact by checking the axial force in bar 3, which should be in an elastic state. This force can be calculated from one of the equilibrium equations, e.g. from (4.5):

$$0.6S_1 - 0.8S_3 = 0 \quad \ldots \quad S_3 = \frac{0.6}{0.8}S_1 = 0.75S_0 \tag{4.21}$$

The value of S_3 is indeed below the yield limit. The collapse load is in this case the lowest one among the loads calculated for all the potential failure mechanisms from the power equality. This statement is quite general, and it will be proven in the next chapter as the so-called Upper Bound Theorem. □

The method of limit analysis discussed so far has been based on the analysis of potential failure mechanisms, and it represents the *kinematic approach*. Let us

now outline the *static approach*, based on the analysis of admissible states of stress (represented by axial forces). We know that the axial forces at collapse, as well as at any other intermediate state of the loading process, must satisfy the conditions of equilibrium,

$$0.6S_1 - 0.8S_3 = 0 \tag{4.22}$$
$$0.8S_1 + S_2 + 0.6S_3 = F \tag{4.23}$$

and the conditions of plastic admissibility

$$-S_0 \le S_i \le S_0, \qquad i = 1, 2, 3 \tag{4.24}$$

One of the axial forces, say S_1, can be selected as the statically redundant one and used to express the other axial forces,

$$S_3 = 0.75S_1 \tag{4.25}$$
$$S_2 = F - 0.8S_1 - 0.6S_3 = F - 1.25S_1 \tag{4.26}$$

Substituting this into the conditions of plastic admissibility, we get

$$-S_0 \le \quad S_1 \quad \le S_0 \tag{4.27}$$
$$-S_0 \le F - 1.25S_1 \le S_0 \tag{4.28}$$
$$-S_0 \le \quad 0.75S_1 \quad \le S_0 \tag{4.29}$$

These six inequalities are satisfied by many different pairs of values of F and S_1, each of them representing (along with the corresponding values of S_2 and S_3) a *statically admissible state*. The maximum statically admissible value of F can be found by combining the right inequality in (4.28) with the right inequality in (4.27):

$$F \le S_0 + 1.25S_1 \le S_0 + 1.25S_0 = 2.25S_0 \tag{4.30}$$

It can be easily verified that, for $F = 2.25S_0$ and $S_1 = S_0$, all the inequalities in (4.27) through (4.29) are satisfied, and $F = 2.25S_0$ is thus the maximum statically admissible value of F. At the same time, it happens to be equal to the collapse load, and the corresponding axial forces $S_1 = S_0$, $S_2 = F - 1.25S_1 = S_0$ and $S_3 = 0.75S_1 = 0.75S_0$ represent the state at collapse. Again, the above statement can be formulated as a general theorem, the Lower Bound Theorem, which will be derived in the next chapter.

4.2 BEAMS AND FRAMES

Beams and frames typically fail after a sufficient number of plastic hinges form in the most exposed cross sections and the structure turns into a mechanism. We assume that the effect of normal and shear forces on the formation of the plastic hinges is negligible, i.e. that a plastic hinge forms when the bending moment reaches the plastic limit moment of the cross section. This assumption greatly simplifies the analysis, and the error that results is usually small. More accurate analysis taking into account the effect of normal forces will be presented in Chapter 13.

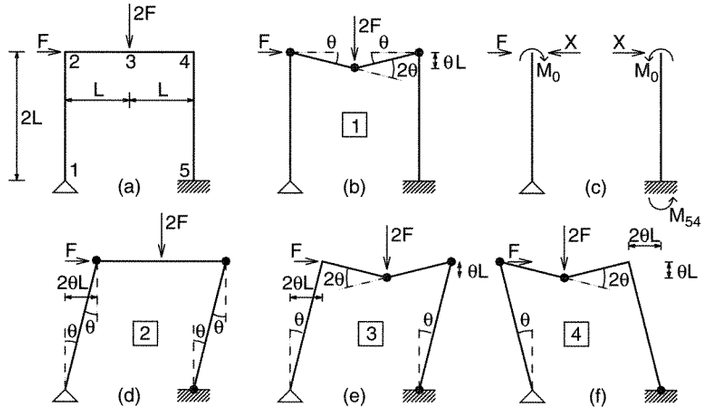


Figure 4.2 Portal frame with potential failure mechanisms

Example 4.2: Consider the portal frame in Figure 4.2(a), loaded by a horizontal load F and a vertical midspan load $2F$. The plastic bending moment for all the cross sections of this frame is the same, M_0. Find the value of F at which the structure collapses.

Solution: First we observe that between each two of the five cross sections[1] marked as 1, 2, 3, 4 and 5 in Figure 4.2(a) the variation of the bending moment is linear. Therefore, local extremes of the bending moment can only occur at these cross sections, called *critical*. The plastic hinges may, therefore, form only at four locations, 2, 3, 4, 5. Since the redundancy of the structure is two, we need to insert three hinges to create a mechanism. This allows for four possible combinations of the plastic hinges shown in Figure 4.2(b) and Figure 4.2(d)–(f), i.e. we have four potential failure mechanisms. To determine the collapse load, we should in principle analyze all these mechanisms and determine the collapse load for each of them. The actual collapse mechanism is that for which the plastic moment is not exceeded in any cross section of the structure.

The rotation rates in all the hinges of a single degree-of-freedom mechanism can be expressed in terms of one chosen rotation rate $\dot{\theta}$. When analyzing the kinematics of potential failure mechanisms, we use the linearized expressions of small-displacement theory. For example, for mechanism 1 (Figure 4.2(b)) we write $\dot{w}_3 = L\dot{\theta}$, or for mechanism 2 (Figure 4.2(d)) we write $\dot{u}_2 = \dot{u}_4 = 2L\dot{\theta}$.

According to (2.22), the dissipation rate in a plastic hinge is given by the product of the plastic moment and the magnitude of the rotation rate. Writing the power equality $\dot{W}_{\text{ext}} = D_{\text{int}}$ for mechanism 1, we obtain the relation

$$2F \times \dot{w}_3 = M_0 \times \dot{\theta} + M_0 \times 2\dot{\theta} + M_0 \times \dot{\theta} \tag{4.31}$$

[1] Strictly speaking, we should distinguish joints 1 to 5 from end sections 12, 21, 23, etc. However, in the present simple frame each joint connects at most two members, and the connected end sections have the same plastic moment. In such a case, we can select only one cross section per joint as the critical one, and refer to it simply by the joint number. However, when referring to the end moments, we still use two subscripts because such a notation permits a unique and objective definition of the sign of each moment.

or, after substituting the kinematic relation $\dot{w}_3 = L\dot{\theta}$,

$$2FL\dot{\theta} = 4M_0\dot{\theta} \tag{4.32}$$

The rotation rate $\dot{\theta}$ cancels out from this relation, and the (potential) collapse load is obtained as

$$F = 2\frac{M_0}{L} \tag{4.33}$$

We must now check whether the equilibrium state corresponding to this mechanism is admissible, for which it suffices to check that at cross section 5 the magnitude of the bending moment, $|M_{54}|$, does not exceed the yield moment, M_0. One, although not the fastest, way to determine M_{54} is to consider the equilibrium conditions for individual members of the frame, as indicated in Figure 4.2(c). The moment reaction in the right support can be expressed as $M_{54} = 2L \times X - M_0$ where X is the horizontal force transmitted by members 2-3 and 3-4. This force can be computed from the condition of moment equilibrium of the left column, $2L \times X - 2L \times F - M_0 = 0$. Solving for $X = F + M_0/2L = 2M_0/L + 0.5M_0/L = 2.5M_0/L$ and substituting back, we finally obtain $M_{54} = 2L \times 2.5M_0/L - M_0 = 4M_0$; this exceeds M_0, and so the equilibrium state corresponding to the mechanism is inadmissible.

In general, it is much easier to check the bending moment in cross sections other than plastic hinge locations using the principle of virtual work. Note that the virtual work equality must hold for the work of an equilibrated system of forces on *any* compatible displacements and generalized strains (in our case, hinge rotations), not only on the mechanism to which these forces correspond. For example, we may write the virtual work of the external and internal forces from mechanism 1 on the displacements and rotations corresponding to mechanism 2 from Figure 4.2(d). This leads to the equation

$$F \times 2L\dot{\theta} = -M_0\dot{\theta} + M_0\dot{\theta} + M_{54}\dot{\theta} \tag{4.34}$$

which yields

$$M_{54} = 2FL = 4M_0 \tag{4.35}$$

confirming our previous result that the internal forces corresponding to mechanism 1 are inadmissible.

When expressing the virtual work of forces of one mechanism on displacements on another mechanism, one must pay attention to the sign of the work of internal forces. If the signs of moment and rotation agree (the second and third term on the right-hand side of (4.34)) the work is positive, and if these signs do not agree (the first term on the right-hand side of (4.34) corresponding to the left corner of the frame), then the work is negative. On the other hand, when we consider the dissipation rate for a certain mechanism, the sign of the moment is always the same as the sign of the rotation rate; in that case, we do not need to bother figuring out the signs of the individual contributions to the dissipation rate and we simply take the product of the plastic moment and the magnitude of the rotation rate, cf. (2.22).

We now proceed to consider mechanism 2 (Figure 4.2(d)). The power equality is $F \times 2L\dot{\theta} = 3M_0\dot{\theta}$, from which we have

$$F = \frac{3}{2}\frac{M_0}{L} \tag{4.36}$$

We must again check that the bending moment at cross section 3 does not exceed M_0. To determine M_{32}, we write the virtual work of the internal forces from this mechanism on the displacements from mechanism 1,

$$2F\dot{\theta}L = -M_0\dot{\theta} + M_{32} \times 2\dot{\theta} + M_0\dot{\theta} \tag{4.37}$$

from which we have

$$M_{32} = FL = \frac{3}{2}M_0 > M_0 \tag{4.38}$$

We see that the equilibrium state corresponding to this mechanism is again inadmissible. The reader can check that, by considering the equilibrium of individual members in the frame, one arrives at the same value of M_{32}, although in a more tedious way.

For mechanism 3 (Figure 4.2(e)), the power equality is $F \times 2\dot{\theta}L + 2F \times \dot{\theta}L = M_0 \times 2\dot{\theta} + M_0 \times 2\dot{\theta} + M_0 \times \dot{\theta}$, which yields

$$F = \frac{5}{4}\frac{M_0}{L} \tag{4.39}$$

To check the bending moment at cross section 2, we write the virtual work of internal forces from this mechanism on the displacements from mechanism 1,

$$2F \times \dot{\theta}L = -M_{21}\dot{\theta} + M_0 \times 2\dot{\theta} + M_0\dot{\theta} \tag{4.40}$$

from which we have

$$M_{21} = -2FL + 3M_0 = -\frac{5}{2}M_0 + 3M_0 = \frac{M_0}{2} < M_0 \tag{4.41}$$

Now we see that this equilibrium state is admissible, and the present potential failure mechanism leads to a statically admissible state of internal forces. Later, we will see that the load corresponding to such a situation is guaranteed to be the actual collapse load, F_0.

Nevertheless, we will still check mechanism 4 (Figure 4.2(f)). The power equality yields $2F \times \dot{\theta}L - F \times 2\dot{\theta}L = M_0 \times 2\dot{\theta} + M_0 \times 2\dot{\theta} + M_0\dot{\theta}$, from which we have

$$5M_0\dot{\theta} = 0 \tag{4.42}$$

This mechanism cannot occur because no external power would be supplied and therefore no energy could be dissipated. Obviously, plastic yielding is impossible without energy dissipation. Mathematically speaking, equation (4.42) implies $\dot{\theta} = 0$, which means that rotation in the assumed hinges is impossible.

As we have exhausted all potential failure mechanisms, we have completely solved the problem. The collapse load $F_0 = 1.25M_0/L$ is given by (4.39). □

In the previous example we found only one potential failure mechanism leading to a statically admissible distribution of internal forces; therefore, the collapse load was uniquely defined. This has not been by chance; we will see later that there exists indeed only one admissible solution to problems of this type. We should also observe that the correct collapse load was the smallest one among all those obtained by equating the external power to the internal dissipation rate. This, too, has not been by chance. We will

prove in Chapter 5 that the true collapse load is always the smallest one among those determined by the power equality for the individual potential failure mechanisms.

It is also of interest to determine loads for which the structure does *not* collapse. For mechanism 1, the largest moment is $M_{54} = 4M_0$, and so an admissible distribution of bending moments is obtained by reducing the collapse load $F = 2M_0/L$ from (4.33) four times, i.e. the load $F_s = M_0/2L$ is safe. For mechanism 2, the largest moment is $M_{32} = 3M_0/2$ (see (4.38)) and the corresponding collapse load $F = 3M_0/2L$ reduced 3/2 times gives another safe estimate $F_s = M_0/L$ of the true collapse load given by (4.39). For mechanism 3, no scaling down is necessary since the corresponding moments do not exceed M_0 anywhere, and therefore the load $F_s = 5M_0/4L$ is safe. We observe that among all these safe loads based on the condition $|M| \leq M_0$, the true collapse load is the largest one.

It is proper to comment now on the applicability of limit analysis when there are significant deviations of the shape of the moment-curvature diagram from the elastic-perfectly plastic bilinear idealization. The power equality $\dot{W}_{\text{ext}} = D_{\text{int}}$ holds whenever there is a state in which each plastic hinge rotates at a constant moment. What the variation of the bending moment prior to or after this state is does not matter, as long as one such state common to all hinges exists. If we plot the bending moment as a function not of curvature, κ, but of the load-point displacement, w, then the histories of the moment from the individual hinges may look as exemplified in Figure 4.3. The conditions of plastic limit analysis are realized in case (a), whereas in case (b) the plastic limit analysis is inapplicable. The latter situation may often happen in columns or otherwise in reinforced concrete structures if they are either over-reinforced or prestressed. In that case, the bending moment as a function of curvature physically decreases.

The fact that the members of the structure are deformed elastically and inelastically prior to the collapse state is in itself irrelevant. During the collapse the internal forces within the structure remain constant and, therefore, the deformation of each member

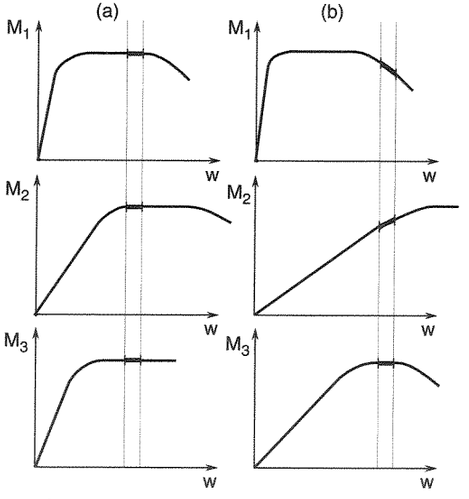


Figure 4.3 Evolution of moments in critical cross sections: a) limit analysis applicable, b) limit analysis inapplicable

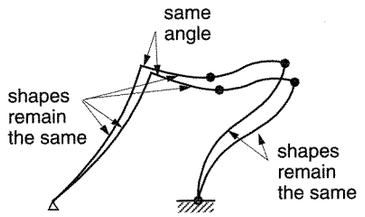


Figure 4.4 Deformed shape prior to collapse and displacement increment during collapse

also remains constant. So the displacement increments for each member during collapse are the same as for a mechanism consisting of rigid parts (see Figure 4.4). Nevertheless, we require the deformations to be sufficiently small for the geometrically linear analysis of the kinematics to be applicable. If the deformations are large, then, for example, we would have to consider the axial shortening of the members due to lateral deflections.

To sum up, the basic hypotheses for plastic limit analysis are:

- a horizontal yield plateau must exist for all the critical hinges;
- all the hinges must be at the plateau simultaneously (at least at some stage of the loading process);
- the deformations must be sufficiently small prior to collapse.

PROBLEMS

Problem 4.1: Perform limit analysis of the truss in Figure 4.1 loaded by an increasing horizontal force F. Use a) the kinematic approach, b) the static approach.

Problem 4.2: Find the collapse load for the truss in Figure 4.5(a) using the kinematic approach. All three bars have the same plastic axial force S_0.

Problem 4.3: Solve Problem 4.2 using the static approach.

Problem 4.4: Analyze potential failure mechanisms for the continuous beam in Figure 4.5(b), and find the collapse load and moment distribution at collapse.

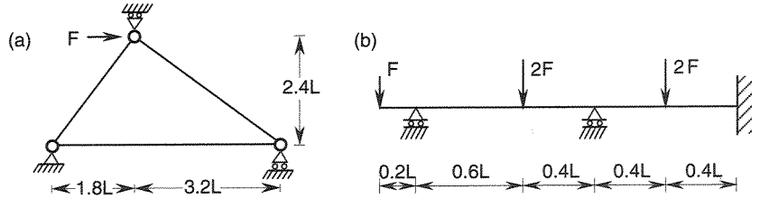


Figure 4.5 a) Truss to be analyzed in Problems 4.2 and 4.3, b) continuous beam to be analyzed in Problem 4.4

5

Theorems of Limit Analysis

We will now prove two general theorems which represent the most important results of plasticity theory, serving as a foundation of all limit analysis. First we need to give precise definitions of the basic concepts. Consider proportional loading, described by (3.30). The value of the load multiplier μ for which the structure collapses is called the *plastic limit load multiplier* or the *safety factor* and is denoted by μ_0. The plastic limit multiplier has the meaning of the safety factor in the true sense of the word only if the reference loading represents the actual service load (or design load). Sometimes we define the reference loading by unit (dimensionless) forces, and then it is more appropriate to call μ_0 the *collapse load* rather than the safety factor.

A state of stress at a material point is called *plastically admissible* if it does not exceed the yield limits. For uniaxial stress states, the yield limits are given by the tensile yield stress, σ_0^+, and compressive yield stress, σ_0^-. For simplicity, we assume in this chapter that they have the same magnitude, σ_0. In this case, the condition of plastic admissibility of stress σ reads

$$-\sigma_0 \le \sigma \le \sigma_0 \tag{5.1}$$

In Chapter 15 we will generalize this condition to multiaxial stress states. The yield limits will then be described by the so-called yield function.

The stress state of a structure is characterized by the stress field, i.e. the spatial distribution of stress. The stress field will be called plastically admissible if the stress at every point of the structure is plastically admissible. For an idealized truss model, the stress field in each bar is uniaxial and homogeneous (constant in space). The plastic admissibility conditions (5.1) are then equivalent to conditions

$$-S_{0,i} \le S_i \le S_{0,i}, \qquad i = 1, 2, \ldots n \tag{5.2}$$

where S_i is the axial force in bar number i, $S_{0,i} = \sigma_0 A_i$ is the axial force at yielding (plastic axial force) in bar number i, and n is the number of bars. In the matrix notation, these conditions can be written as

$$-\boldsymbol{s}_0 \le \boldsymbol{s} \le \boldsymbol{s}_0 \tag{5.3}$$

where $\boldsymbol{s}$ is the vector of axial forces, and $\boldsymbol{s}_0$ is the vector of plastic axial forces. The inequality signs between vectors in (5.3) are rather unusual, but they greatly simplify notation; they imply a set of inequalities for the corresponding components.

When the effects of the shear and normal forces on the formation of a plastic hinge are neglected, the plastic admissibility conditions for a beam or frame read

$$-M_{0,i} \leq M_i \leq M_{0,i}, \qquad i = 1, 2, \ldots n \tag{5.4}$$

where M_i is the moment at cross section number i, $M_{0,i} = \sigma_0 W_{0,i}$ is the plastic moment of cross section number i, and n is the number of critical cross sections. In the matrix notation, conditions (5.4) have the same general form (5.3) as for trusses, but vector $\boldsymbol{s}$ now contains the bending moments at all the critical cross sections of a beam or frame, and vector $\boldsymbol{s}_0$ contains the plastic limit moments at the critical cross sections. In general, $\boldsymbol{s}$ is called the vector of internal forces and $\boldsymbol{s}_0$ the vector of *plastic capacities*. If inequalities (5.3) are satisfied in the strict sense,

$$-\boldsymbol{s}_0 < \boldsymbol{s} < \boldsymbol{s}_0 \tag{5.5}$$

the state of internal forces is said to be *inside the elastic domain.*

A *statically admissible state* (of a structure) is any plastically admissible stress field that is in equilibrium with a certain multiple of the reference loading and that satisfies the yield conditions. The corresponding load multiplier μ_s is called a *statically admissible load multiplier.* Formally, a statically admissible state is described by a vector of internal forces $\boldsymbol{s}$ and by a load multiplier μ_s satisfying the plastic admissibility conditions (5.3) and the equilibrium conditions

$$\boldsymbol{B}^T \boldsymbol{s} = \mu_s \bar{\boldsymbol{f}} \tag{5.6}$$

Matrix $\boldsymbol{B}^T$ is the static matrix defined in Appendices A and B, and $\bar{\boldsymbol{f}}$ is the reference load vector.

Furthermore, a *kinematically admissible state* is defined as any potential failure mechanism for which the external power is positive. The corresponding load multiplier μ_k, determined from the power equality, is called the *kinematically admissible multiplier.* A potential failure mechanism can be specified by the (plastic) extension rates of the bars of a truss, or by the rotation rates at the critical cross sections of a beam or frame. In a unified description, the vector of *generalized strains*, $\boldsymbol{e}$, is introduced as the vector containing the bar extensions or the critical cross section rotations. With this notation, a kinematically admissible state is formally described by a vector of displacement rates $\dot{\boldsymbol{d}}$ and a vector of generalized strain rates $\dot{\boldsymbol{e}}$ such that

$$\dot{e} = \boldsymbol{B}\dot{\boldsymbol{d}} \tag{5.7}$$

and

$$\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} > 0 \tag{5.8}$$

The left-hand side of (5.8) has the meaning of external power supplied by the reference loading $\bar{\boldsymbol{f}}$. The actual external power can be expressed as

$$\dot{W}_{\text{ext}} = \boldsymbol{f}^T \dot{\boldsymbol{d}} = \mu_k \bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} \tag{5.9}$$

Each component of $\boldsymbol{e}$ is work-conjugate with the corresponding component of the vector of internal forces, $\boldsymbol{s}$, i.e. with the corresponding axial force or bending moment.

Making use of (2.3) and (2.22), the dissipation rate is obtained by summing the contributions of all bars of a truss,

$$D_{\text{int}} = \sum_i S_i \dot{e}_i = \sum_i S_{0,i} |\dot{e}_i| \tag{5.10}$$

or by summing the contributions of all the critical cross sections (potential yield hinges) in a beam or frame,

$$D_{\text{int}} = \sum_i M_i \dot{\theta}_i = \sum_i M_{0,i} |\dot{\theta}_i| \tag{5.11}$$

In a unified, compact notation we can write

$$D_{\text{int}} = \boldsymbol{s}^T \dot{\boldsymbol{e}} = \boldsymbol{s}_0^T |\dot{\boldsymbol{e}}| \tag{5.12}$$

where $|.|$ is not a norm but the absolute value operator (applied on each scalar entry of its vectorial argument separately). The kinematically admissible load multiplier μ_k corresponding to a kinematically admissible mechanism can be determined from the power equality $\dot{W}_{\text{ext}} = D_{\text{int}}$. Substituting from (5.9) and (5.12), we get

$$\mu_k = \frac{\boldsymbol{s}_0^T |\dot{\boldsymbol{e}}|}{\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}}} \tag{5.13}$$

Let us look more carefully at expression (5.10) for the dissipation rate in a truss. In complete analogy to the considerations in Chapter 1, we can show that if S_s is any plastically admissible axial force (totally independent of the actual extension rate $\dot{e}$), i.e. a force satisfying the conditions $-S_0 \leq S_s \leq S_0$, then $S_s \dot{e} \leq D_{\text{int}}(\dot{e})$ where $D_{\text{int}}(\dot{e}) = S_0 |\dot{e}|$ is the rate of dissipation in the bar subjected to the plastic extension rate $\dot{e}$. Summing up the contributions of individual bars, we get the following theorem (which we state here in a general form, because exactly the same considerations are valid for frames).

Theorem

If $\boldsymbol{s}_s$ is any plastically admissible vector of internal forces, i.e. a vector satisfying the conditions

$$-\boldsymbol{s}_0 \leq \boldsymbol{s}_s \leq \boldsymbol{s}_0 \tag{5.14}$$

and $\dot{\boldsymbol{e}}$ is any vector of generalized plastic strain rates, then

$$\boldsymbol{s}_s^T \dot{\boldsymbol{e}} \leq D_{int}(\dot{\boldsymbol{e}}) \tag{5.15}$$

where

$$D_{int}(\dot{\boldsymbol{e}}) = \boldsymbol{s}_0^T |\dot{\boldsymbol{e}}| \tag{5.16}$$

is the rate of dissipation produced by $\dot{\boldsymbol{e}}$.

To explain the essence of the above theorem, let us assume that we know the vector $\dot{\boldsymbol{e}}$ for the actual collapse mechanism. The actual internal forces $\boldsymbol{s}$ associated with this mechanism maximize the scalar product $\boldsymbol{s}^T\dot{\boldsymbol{e}}$ among all the plastically admissible vectors. To see this, it suffices to realize that $\boldsymbol{s}^T\dot{\boldsymbol{e}} = \boldsymbol{s}_0^T|\dot{\boldsymbol{e}}| = D_{\text{int}}(\dot{\boldsymbol{e}})$. Therefore, the preceding theorem is often referred to as the postulate of maximum plastic dissipation, because it can be mathematically written as

$$D_{\text{int}}(\dot{\boldsymbol{e}}) = \boldsymbol{s}^T\dot{\boldsymbol{e}} = \max_{\boldsymbol{s}_s \in \bar{\mathcal{E}}} \boldsymbol{s}_s^T\dot{\boldsymbol{e}} \tag{5.17}$$

and verbally stated as follows.

Postulate of Maximum Plastic Dissipation
For given generalized plastic strain rates, the actual internal forces maximize the plastic dissipation rate among all the plastically admissible internal forces.

Now we are ready to derive the basic theorems of limit analysis. We will work with the following two states:

- An arbitrary statically admissible state given by μ_s and $\boldsymbol{s}_s$ such that $\boldsymbol{B}^T\boldsymbol{s}_s = \mu_s\bar{\boldsymbol{f}}$ and $-\boldsymbol{s}_0 \le \boldsymbol{s}_s \le \boldsymbol{s}_0$.
- An arbitrary kinematically admissible state given by $\dot{\boldsymbol{d}}_k$ and $\dot{\boldsymbol{e}}_k$ such that $\dot{\boldsymbol{e}}_k = \boldsymbol{B}\dot{\boldsymbol{d}}_k$ and $\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}_k > 0$, for which we can evaluate the corresponding kinematically admissible multiplier $\mu_k = D_{\text{int}}(\dot{\boldsymbol{e}}_k)/(\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}_k)$.

Since internal forces $\boldsymbol{s}_s$ are plastically admissible, the postulate of maximum plastic dissipation applied to $\boldsymbol{s}_s$ and $\dot{\boldsymbol{e}}_k$ gives

$$\boldsymbol{s}_s^T\dot{\boldsymbol{e}}_k \le D_{\text{int}}(\dot{\boldsymbol{e}}_k) \tag{5.18}$$

According to the definition of the kinematically admissible multiplier μ_k corresponding to the given rates $\dot{\boldsymbol{d}}_k$ and $\dot{\boldsymbol{e}}_k$, the right-hand side of (5.18) can be written as

$$D_{\text{int}}(\dot{\boldsymbol{e}}_k) = \mu_k\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}_k \tag{5.19}$$

Using the kinematic equations $\dot{\boldsymbol{e}}_k = \boldsymbol{B}\dot{\boldsymbol{d}}_k$ and the equations of equilibrium $\boldsymbol{B}^T\boldsymbol{s}_s = \mu_s\bar{\boldsymbol{f}}$, the left-hand side of (5.18) can be transformed into

$$\boldsymbol{s}_s^T\dot{\boldsymbol{e}}_k = \boldsymbol{s}_s^T\boldsymbol{B}\dot{\boldsymbol{d}}_k = (\boldsymbol{B}^T\boldsymbol{s}_s)^T\dot{\boldsymbol{d}}_k = \mu_s\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}_k \tag{5.20}$$

Substituting (5.19) and (5.20) into (5.18), we obtain

$$\mu_s\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}_k \le \mu_k\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}_k \tag{5.21}$$

The scalar product $\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}_k$ is positive (see the conditions of kinematic admissibility), and so (5.21) implies

$$\mu_s \le \mu_k \tag{5.22}$$

This is a very important result, expressed by the following theorem.

Fundamental Theorem of Limit Analysis
No statically admissible multiplier is larger than any kinematically admissible multiplier.

The plastic limit state (actual state at incipient collapse) is defined as a state of static equilibrium at which nonvanishing displacement rates can appear even if the load remains constant. For structures under uniaxial stress, the existence of such a state is almost self-evident. A reader wishing to have a formal proof can obtain it as a special case of the general uniqueness theorem for triaxial stress states given in Section 16.3. By definition, the plastic limit state is both statically and kinematically admissible, and so we can write

$$\mu_s \le \mu_0 \le \mu_k \tag{5.23}$$

where μ_0 is the safety factor. This immediately leads to two important corollaries.

Lower Bound Theorem
The safety factor is the largest statically admissible multiplier.

Upper Bound Theorem
The safety factor is the smallest kinematically admissible multiplier.

The names of these theorems refer to the fact that any statically (kinematically) admissible multiplier provides a lower (an upper) bound on the safety factor. Sometimes they are called the *static theorem* and the *kinematic theorem* of limit analysis. Alternatively, they can be stated in the following form:

Upper Bound Theorem
If there is a potential failure mechanism for which the dissipation rate is smaller than the external power supplied by the given loads, the structure **will collapse** *under such loads.*

Lower Bound Theorem
If there is a state of stress that is inside the elastic domain and in equilibrium with the given applied loads, the structure **will not collapse** *under such loads.*

In this formulation of the lower bound (static) theorem, we require the stress state to be in the interior of the elastic domain and not on its boundary, i.e. the plastic admissibility conditions must be satisfied as strict inequalities (5.5). Only in this case can we guarantee that the structure would not collapse. At incipient collapse, the stress state is plastically admissible, but some of the conditions (5.3) are satisfied with an equality sign. Of course, from the practical point of view, it makes little difference whether the boundary of the safe domain is considered as 'safe' or not.

Even though some related ideas date back to the 18th century, plastic limit analysis was initiated by Kazinczy (1914), who calculated the failure load of a clamped beam and verified the result experimentally. Similar concepts were proposed by Kist (1917, 1920) and Grüning (1926). However, early work relied mainly on engineering

intuition. The fundamental theorem of limit analysis was discussed by Gvozdev (1938) in the context of generalized stresses and strains, which covered structural models (beams, frames, plates) as well as three-dimensional continua. Gvozdev described the derivation of the bound theorems only verbally and his work, presented at a local Russian conference in 1936 and published in the proceedings two years later, remained unnoticed in the Western world until it was translated into English by Haythornthwaite and published in 1960 in the *International Journal of Mechanical Sciences* (Gvozdev, 1960). In the meantime, a formal proof of the bound theorems for beams and frames was given by Horne (1949), and by Greenberg and Prager (1951). Generalizations for bodies under arbitrary triaxial stress states (Drucker, Prager and Greenberg, 1952) will be presented in Chapter 16.

PROBLEMS

Problem 5.1*: Generalize the dissipation rate expression (5.12) to structures for which the plastic capacities (plastic axial forces, or plastic limit moments) depend on the signs of the corresponding deformation rates, i.e. to trusses for which $S_0^+ \neq S_0^-$, and to frames for which $M_0^+ \neq M_0^-$.
Hint: Use the results of Problems 2.5 and 2.8.

Problem 5.2*: Formulate and prove the postulate of maximum plastic dissipation for the type of structures mentioned in Problem 5.1. Is limit analysis applicable to these structures?

6

Methods of Limit Analysis

As already explained, there are two basic approaches to limit analysis corresponding to the two fundamental theorems. The kinematic approach is based on the upper bound theorem, according to which the safety factor can be obtained by looking for the minimum kinematically admissible load multiplier. The static approach is based on the lower bound theorem, according to which the safety factor can be obtained by looking for the maximum statically admissible load multiplier.

To simplify the notation, we will leave out the subscripts s and k referring to statically and kinematically admissible states, whenever no confusion can arise. Thus, we will write $\dot{\boldsymbol{d}}$ and $\dot{\boldsymbol{e}}$ instead of $\dot{\boldsymbol{d}}_k$ and $\dot{\boldsymbol{e}}_k$ (or $\dot{\boldsymbol{e}}_{pk}$), and $\boldsymbol{s}$ instead of $\boldsymbol{s}_s$. We will nevertheless keep the subscripts attached to the load multiplier, to indicate whether the multiplier is statically or kinematically admissible.

6.1 KINEMATIC APPROACH

Example 6.1: The truss in Figure 6.1(a) is loaded by an increasing horizontal force F. Given that the yield axial force S_0 is the same for all the bars, determine the plastic limit value of F using the kinematic approach to limit analysis.

Solution: As the structure is statically indeterminate to the first degree, two bars out of five have to yield at collapse and the number of potential failure mechanisms is $\binom{5}{2} = 10$. However, not all of them are kinematically admissible. Recall that a kinematically admissible mechanism must provide a positive external power. For our problem, this means that the horizontal displacement velocity $\dot{u}$ at the loaded node must not be zero, and so one of bars 1 and 3 must yield. On the other hand, these bars cannot yield both at the same time (for such a mechanism, the loaded node would only move vertically). These conditions are satisfied by only six admissible mechanisms with one degree of freedom. We do not need to analyze all of them. Let us examine, for example, the mechanism, with bars 2 and 3 yielding (Figure 6.1(b)). The kinematic relations

$$\dot{e}_2 = -\dot{u} \tag{6.1}$$

$$\dot{e}_3 = -0.8\dot{u} \tag{6.2}$$

can be substituted into the power equality

$$F_k \dot{u} = S_0|\dot{e}_2| + S_0|\dot{e}_3| \tag{6.3}$$

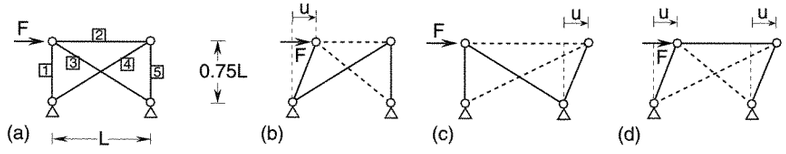


Figure 6.1 Truss analyzed by the kinematic approach

to give

$$F_k \dot{u} = S_0 \dot{u} + 0.8 S_0 \dot{u} \tag{6.4}$$

Knowing the kinematically admissible value corresponding to this mechanism, $F_k = 1.8S_0$, we can calculate the axial forces in the non-yielding bars from the equations of equilibrium, substituting $S_3 = -S_0$ and $S_2 = -S_0$:

$$S_1 + 0.6S_3 = 0 \quad \ldots \quad S_1 = -0.6S_3 = 0.6S_0 \tag{6.5}$$
$$-S_2 - 0.8S_4 = 0 \quad \ldots \quad S_4 = -1.25S_2 = 1.25S_0 \tag{6.6}$$
$$0.6S_4 + S_5 = 0 \quad \ldots \quad S_5 = -0.6S_4 = -0.75S_0 \tag{6.7}$$

Axial force S_4 is not plastically admissible and thus the present mechanism cannot be the actual one. The calculated kinematically admissible external force $F_k = 1.8S_0$ is only an upper bound on the collapse load. Making use of the lower bound theorem, we can also obtain a lower bound at no extra cost. Note that we have constructed a vector of axial forces $\boldsymbol{s} = \{0.6S_0, -S_0, -S_0, 1.25S_0, -0.75S_0\}^T$ that is in equilibrium with the external load $F_k = 1.8S_0$. If we proportionally scale down all the forces to plastically admissible values, the equations of equilibrium still hold. We thus get a statically admissible state and the corresponding load provides a lower bound on the collapse load. To this end, we multiply the forces by a reduction factor $1/1.25 = 0.8$, which leads to $\boldsymbol{s} = \{0.48S_0, -0.8S_0, -0.8S_0, S_0, -0.6S_0\}^T$ and $F_s = 1.44S_0$. After this step, we know that the collapse load F_0 is between the bounds

$$1.44S_0 \leq F_0 \leq 1.8S_0 \tag{6.8}$$

To improve this estimate, let us modify the assumed failure mechanism. For the mechanism considered so far, plastic admissibility has been violated by axial force S_4, and so it is natural to expect that an improvement of the collapse load estimate can be obtained if we allow bar 4 to yield and treat either bar 2 or bar 3 as a non-yielding one. The mechanism with bars 2 and 4 yielding (Figure 6.1(c)) would give a zero displacement rate at the loaded node, and is thus kinematically inadmissible. The mechanism with bars 3 and 4 yielding (Figure 6.1(d)) can be analyzed as follows:

$$\dot{e}_3 = -0.8\dot{u} \tag{6.9}$$
$$\dot{e}_4 = 0.8\dot{u} \tag{6.10}$$
$$F_k \dot{u} = S_0|-0.8\dot{u}| + S_0|0.8\dot{u}| \tag{6.11}$$
$$F_k = 1.6S_0 \tag{6.12}$$

Indeed, we have improved the upper bound on F_0. The axial forces in the yielding bars must be $S_3 = -S_0$, $S_4 = S_0$. Using the equations of equilibrium, the remaining

axial forces can be obtained:

$$S_2 + 0.8S_3 + F_k = 0 \quad \ldots \quad S_2 = -0.8S_3 - F_k = -0.8S_0 \tag{6.13}$$
$$S_1 + 0.6S_3 = 0 \quad \ldots \quad S_1 = -0.6S_3 = 0.6S_0 \tag{6.14}$$
$$0.6S_4 + S_5 = 0 \quad \ldots \quad S_5 = -0.6S_4 = -0.6S_0 \tag{6.15}$$

As all these forces are plastically admissible, the state is not only kinematically but also statically admissible, and thus it is the actual collapse state and $F_0 = 1.6S_0$ is the collapse load. □

Let us now consider a somewhat different structure with bars transmitting only axial forces – a rigid plate supported by five pin-ended bars as shown in Figure 6.2(a). The structure is statically indeterminate to the second degree, and so it could be expected that three bars must yield at collapse. However, if bars 1 and 3 yield, the plate can rotate around joint 2 and a mechanism with only two yielding bars is obtained; see Figure 6.2(b). This is a special case called a *partial mechanism*. The adjective 'partial' refers to the fact that only a part of the structure becomes a mechanism while some other part remains statically indeterminate. This is better illustrated by the large truss in Figure 6.3(a), which is five times statically indeterminate but yielding of as few as two bars can lead to collapse (Figures 6.3(b,c)).

Example 6.2: Find the collapse load for the structure in Figure 6.2(a).

Solution: Let us determine the kinematically admissible value F_k associated with the partial mechanism in Figure 6.2(b). The extension rates can be expressed in terms of the vertical displacement rate $\dot{w}_1$ at joint 1 as

$$\dot{e}_1 = \dot{w}_1 \tag{6.16}$$
$$\dot{e}_3 = 0.6\dot{w}_1 \tag{6.17}$$

The power equality

$$2.5F_k\dot{w}_1 = S_0\dot{e}_1 + S_0\dot{e}_3 \tag{6.18}$$

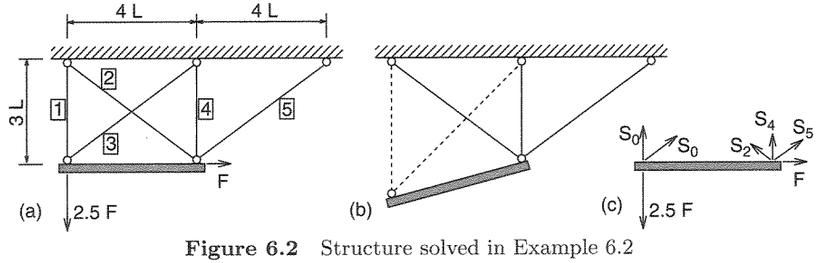


Figure 6.2 Structure solved in Example 6.2

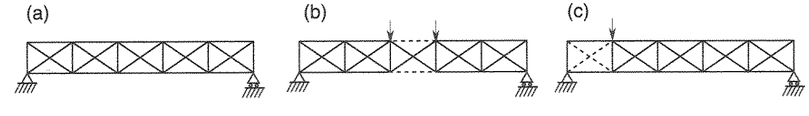


Figure 6.3 Truss with partial collapse mechanisms

yields a kinematically admissible load

$$F_k = \frac{1+0.6}{2.5} S_0 = 0.64 S_0 \tag{6.19}$$

Let us check whether this is the actual collapse load. Unlike the standard situation of a mechanism with $s+1$ yielding bars, here the axial forces cannot be uniquely determined from the equations of equilibrium. For the structure under consideration, we have a total of three independent equations of equilibrium, corresponding to three degrees of freedom of the rigid plate. These equations can be constructed by applying the principle of virtual work to three virtual mechanisms. If we use the assumed failure mechanism in Figure 6.2(b) as one of them, the virtual work equality takes a form analogous to the power equality (6.18) but with displacement rate $\dot{w}_1$ and extension rates $\dot{e}_1, \dot{e}_3$ replaced by virtual displacement δw_1 and virtual extensions δe_1, δe_3:

$$2.5 F \delta w_1 = S_1 \delta e_1 + S_3 \delta e_3 \tag{6.20}$$

Since the relations between δw_1 and δe_1, δe_3 are directly analogous to relations (6.16) and (6.17) between $\dot{w}_1$ and $\dot{e}_1, \dot{e}_3$, the resulting equation of equilibrium has the form

$$2.5F = S_1 + 0.6 S_3 \tag{6.21}$$

This corresponds to the condition of moment equilibrium of the rigid plate around joint 2 divided by the distance between joints 1 and 2. If bars 1 and 3 are yielding, then $S_1 = S_0$, $S_3 = S_0$ and (6.21) implies $F = 0.64 S_0$. Therefore, the fundamental equation of the kinematic approach, serving for determination of a kinematically admissible load multiplier, can be interpreted either as the power equality derived from the principle of energy conservation, or as an equation of equilibrium derived from the principle of virtual work. The latter interpretation simply means that, assuming a certain failure mechanism, we can construct an equation of equilibrium which does not contain axial forces in the bars that are not yielding. Because the axial forces in the yielding bars are known, such an equation is sufficient for calculating the load multiplier. Once exploited, this equation cannot be used when calculating the unknown axial forces in the non-yielding bars. The number of independent equations of equilibrium thus decreases by one.

For our problem, we are left with only two equations for three unknown axial forces S_2, S_4 and S_5. The free-body diagram in Figure 6.2(c) implies

$$0.8 S_2 - 0.8 S_5 = F + 0.8 S_0 \tag{6.22}$$
$$0.6 S_2 + S_4 + 0.6 S_5 = 2.5F - S_0 - 0.6 S_0 \tag{6.23}$$

Substituting $F = F_k = 0.64 S_0$, we get

$$0.8 S_2 - 0.8 S_5 = 1.44 S_0 \tag{6.24}$$
$$0.6 S_2 + S_4 + 0.6 S_5 = 0 \tag{6.25}$$

This set of equations cannot be solved uniquely but the only thing that matters for our purpose is whether there exists at least one plastically admissible solution. Let us

express S_2 and S_4 in terms of S_5. The conditions of plastic admissibility can then be written as

$$-S_0 \le S_2 = \quad 1.8 S_0 + S_5 \quad \le S_0 \tag{6.26}$$
$$-S_0 \le S_4 = -1.08 S_0 - 1.2 S_5 \le S_0 \tag{6.27}$$
$$-S_0 \le S_5 = \quad S_5 \quad \le S_0 \tag{6.28}$$

from which

$$-2.8 S_0 \le S_5 \le -0.8 S_0 \tag{6.29}$$
$$-1.733 S_0 \le S_5 \le -0.067 S_0 \tag{6.30}$$
$$-S_0 \le S_5 \le S_0 \tag{6.31}$$

Clearly, any value of S_5 from the interval $[-S_0, -0.8 S_0]$ represents (along with the corresponding axial forces in the other bars) a statically admissible state that is in equilibrium with $F_s = 0.64 S_0$. The actual state at collapse cannot be uniquely determined unless we combine the equations of equilibrium with the kinematic and constitutive equations. However, the sole fact that at least one statically admissible state with $F_s = 0.64 S_0 = F_k$ exists is sufficient to prove that $F_0 = 0.64 S_0$ is the collapse load. □

6.2 DISTRIBUTED LOADS

In all the preceding analyses we were dealing only with external loading represented by concentrated forces. When a structure is subjected to distributed loads, the basic theorems of limit analysis remain valid because the distributed loads can be regarded as a limiting case of concentrated loads. However, the number of potential hinge locations is now infinite, which causes a difficulty when considering various potential failure mechanisms.

Example 6.3: Consider a two-span continuous beam loaded by uniform loads p in the left span (Figure 6.4(a)). Determine the collapse load and the position of plastic hinges.

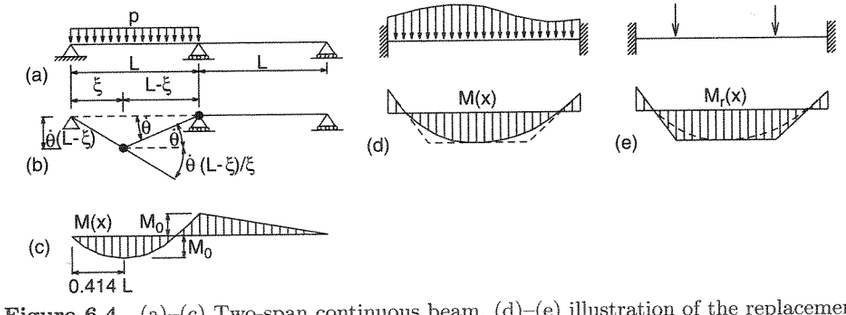


Figure 6.4 (a)–(c) Two-span continuous beam, (d)–(e) illustration of the replacement theorem

Solution: From the distribution of bending moments, it is clear that one plastic hinge will form at the middle support and the other at some section in the left span, described by the (so far unknown) distance ξ from the left support; see Figure 6.4(b). Let us select the rotation rate at the hinge above the support, $\dot{\theta}$, as the independent degree of freedom. Using the kinematic relations for small rotations, the rotation rate at the other hinge can be expressed as $\dot{\theta}(L-\xi)/\xi + \dot{\theta} = \dot{\theta}L/\xi$, and the deflection rate in the left span is described by the function

$$\dot{w}(x) = \begin{cases} \dot{\theta}(L-\xi)x/\xi & \text{if } 0 \le x \le \xi \\ \dot{\theta}(L-x) & \text{if } \xi \le x \le L \end{cases} \tag{6.32}$$

The external power is evaluated as

$$\dot{W}_{\text{ext}} = \int_0^L p\dot{w}(x)\,\mathrm{d}x = p\int_0^L \dot{w}(x)\,\mathrm{d}x = p\frac{1}{2}L\dot{\theta}(L-\xi) \tag{6.33}$$

Now it is easy to write the power equality

$$\frac{1}{2}p_k L\dot{\theta}(L-\xi) = M_0\dot{\theta} + M_0\left(\dot{\theta} + \dot{\theta}\frac{L-\xi}{\xi}\right) \tag{6.34}$$

from which

$$p_k = \frac{2}{L-\xi}\left(1+\frac{L}{\xi}\right)\frac{M_0}{L} = \frac{2}{\xi}\frac{L+\xi}{L-\xi}\frac{M_0}{L} \tag{6.35}$$

Each value of $\xi \in (0, L)$ corresponds to a potential failure mechanism, and $p_k(\xi)$ from (6.35) is the corresponding kinematically admissible load. According to the upper bound theorem, the actual collapse mechanism is that for which $p_k(\xi)$ is minimized. Therefore, $\mathrm{d}p_k/\mathrm{d}\xi = 0$ must hold for the actual collapse load. Carrying out the differentiation of the right-hand side in (6.35), we obtain the equation

$$\frac{\mathrm{d}p_k}{\mathrm{d}\xi} = \frac{2M_0}{L}\frac{\xi^2 + 2L\xi - L^2}{\xi^2(L-\xi)^2} = 0 \tag{6.36}$$

There are two solutions, $\xi_{1,2} = (-1 \pm \sqrt{2})L$, but only one of them lies inside the interval $(0, L)$ and is therefore of physical interest. The correct position of the plastic hinge is thus given by $\xi_0 = (\sqrt{2}-1)L$. Substituting this into (6.35), we get the collapse load

$$p_0 = \min_{\xi\in(0,L)} p_k(\xi) = p_k(\xi_0) = (6+4\sqrt{2})\frac{M_0}{L^2} = 11.66\frac{M_0}{L^2} \tag{6.37}$$

The moment distribution at collapse is shown in Figure 6.4c. Note that if we do not bother to get the exact position of the plastic hinge and assume that $\xi/L = 0.5$, we get $p_k(0.5) = 12M_0/L^2$. In the present case, this is only 2.9% higher than the exact collapse load. □

In large structures it would be a nuisance to consider an arbitrary unknown location of the yield hinge under each distributed load. In this regard, the following theorem, which allows the replacement of distributed loads by concentrated ones, is useful.

Replacement Theorem
If any distributed load having the same sign across the entire span is replaced by a statically equivalent (also said equipollent) concentrated load, the structure is not made stronger, i.e. the computed collapse load will not be larger than the actual one.

The theorem guarantees the replacement of a distributed load by an equipollent concentrated load to yield a safe estimate of the collapse load.

Proof: If the distributed load is of one sign, the curvature of the bending moment diagram is also of one sign, for example convex as in Figure 6.4(d). For the statically equivalent load, the corresponding polygonal bending moment diagram must always be a polygon circumscribed to the actual curved bending moment diagram, i.e., the lines of the polygonal diagram are tangents to the curved bending moment diagram; see Figure 6.4(e). Thus, it is clear that if the condition $-M_0 \le M_r(x) \le M_0$ (in which $M_r(x)$ is the bending moment distribution for the replacement loads) is verified, then also the condition $-M_0 \le M(x) \le M_0$ is guaranteed to hold for the actual bending moment distribution $M(x)$. Thus, if distribution $M_r(x)$ is statically admissible, then also distribution $M(x)$ is statically admissible, which proves the previous theorem.

6.3 COMBINATION OF MECHANISMS

For large frames it is helpful to systematize the choice of various potential failure mechanisms. Since we consider only small deflections, all the potential mechanisms form a linear space, in which linear combinations can be constructed. Figure 6.5 gives two examples of the superposition of mechanisms. Among all the potential mechanisms, one can select a set of the so-called *elementary mechanisms* having the following properties:

1. **Independence:** No elementary mechanism can be obtained as a linear combination of the others.
2. **Completeness:** Any potential failure mechanism can be obtained as a linear combination of the elementary ones.

In mathematical terms, the elementary mechanisms form a basis in the space of all potential failure mechanisms.

For a truss, any potential failure mechanism is uniquely described by a vector of generalized displacement rates, $\dot{\boldsymbol{d}}$. The number of independent mechanisms is thus the

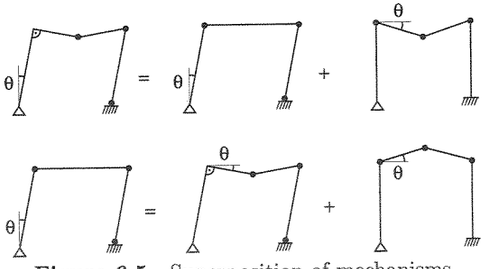


Figure 6.5 Superposition of mechanisms

same as the number of unknowns in the displacement method, n. For a beam or frame, the displacement rates describing a potential failure mechanism must satisfy the conditions of axial inextensibility, otherwise some members would change their length during plastic collapse, which is not consistent with the behavior of standard plastic hinges.[1] Consequently, the number of independent potential failure mechanisms is the same as the number of unknowns that is used in the simplified displacement method neglecting the effect of axial extension.

Alternatively, we could describe any potential failure mechanism by a vector of generalized strain rates $\dot{\boldsymbol{e}}$ satisfying the conditions of compatibility. The number of independent compatibility conditions is equal to the degree of static redundancy of the structure, s. This means that the number of elementary mechanisms is $n = m - s$, where m is the number of generalized strains that can be nonzero during plastic flow. For a truss, m is the number of members (bars), while for a beam or frame it is the number of critical sections.

For rectangular frames, elementary mechanisms are typically chosen such that each of them is associated with one fundamental unknown of the simplified displacement method. Floor (horizontal) displacements are related to *panel mechanisms*, joint rotations to *joint mechanisms*, and deflections (vertical displacements) under concentrated forces acting on a beam between the columns are related to *beam mechanisms*. The joint mechanisms as such do not usually represent actual collapse mechanisms (unless there is a concentrated moment applied directly at the joint), however, they are useful in superposing them onto other mechanisms.

Example 6.4: Construct the set of elementary mechanisms for the frame in Figure 6.6(a).

Solution: The redundancy of the frame is $s = 6$. Regarding the critical cross sections, we have at least two options:

- In a fully automated analysis, both end cross sections of each frame member are considered as critical (unless the end is connected to the joint by an internal hinge). For convenience, we introduce 'computational' joints at the point of application of each external force. In the present case, the structural model has eight joints (see Figure 6.6(b)) and eight members, which gives $m = 2\times 8 = 16$ critical cross sections. The number of elementary mechanisms is then $n = m - s = 16 - 6 = 10$.
- When solving the problem 'by hand', we try to minimize the number of elementary mechanisms. Provided that a joint connects only two members and is not subjected to an external moment (only to external forces), it follows from the moment equation of equilibrium that the end moments are equal in magnitude (they have opposite signs). In such a case, it suffices to consider only one of the end cross sections connected by the joint as a critical one. If the yield moments of the cross sections are different, we have to select the cross section with the lower resistance, while if the yield moments are equal, the choice is arbitrary. Let us emphasize that this simplification is not admissible for joints connecting three or more members. In the present case, joints 1, 2, 3 and 5 satisfy the above conditions.

[1] In the theory of generalized plastic hinges that takes into account the effect of the normal forces, the members can experience plastic axial deformations, and the number of independent potential failure mechanisms increases; see Chapter 13.

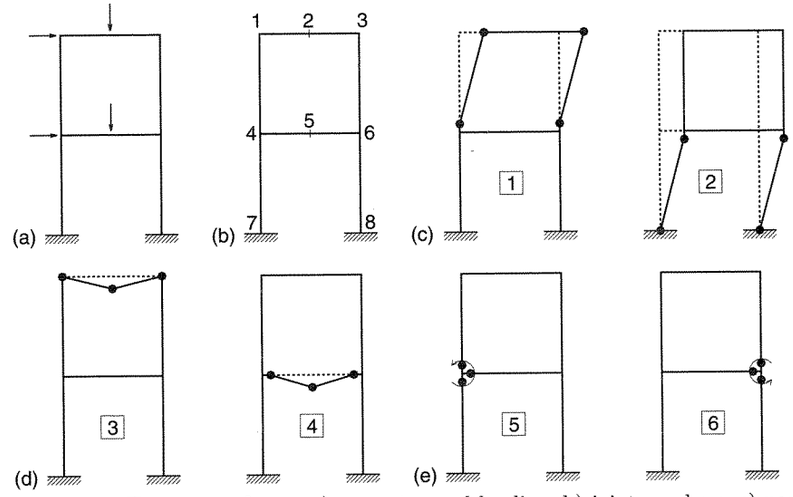


Figure 6.6 Two-storey frame: a) geometry and loading, b) joint numbers, c) panel mechanisms, d) beam mechanisms, e) joint mechanisms

Consequently, we can merge the critical cross sections 12 and 14, 21 and 23, 32 and 36 and 54 and 56. The number of critical sections is reduced to $m = 12$, and the number of elementary mechanisms is only $n = m - s = 12 - 6 = 6$. Two of them are panel mechanisms (Figure 6.6(c)), associated with the floor displacements $u_1 = u_2 = u_3$ and $u_4 = u_5 = u_6$. There are also two beam mechanisms (Figure 6.6(d)), associated with the vertical displacements w_2 and w_5. Finally, two joint mechanisms (Figure 6.6(e)) are associated with the rotations of the joints that link more than two members, i.e. with ϕ_4 and ϕ_6.

Example 6.5: Obtain the collapse load for the two-storey frame in Figure 6.7a by combining the elementary mechanisms. Note that the frame consists of parts with three different cross sections, characterized by plastic limit moments M_0, $2M_0$ and $3M_0$.

Solution: For convenience let us introduce a dimensionless load factor μ as the actual load F multiplied by L/M_0, i.e. $F = \mu M_0/L$. Let us first evaluate the kinematically admissible values of μ for the elementary mechanisms, plotted in Figures 6.6(c)–(e) and labeled by boxed numbers 1 to 6. For the panel mechanism 1 in Figure 6.6(c), the power equality $(2\mu_k M_0/L) \times (2L\dot{\theta}) = 4M_0\dot{\theta}$ leads to $\mu_k = 1$. Note that the plastic moment at the top corners (hinges 1 and 3), which link members with different cross sectional characteristics, has been considered to have the smaller value, M_0. Writing the power equalities for panel mechanism 2 and for beam mechanisms 3 and 4, we obtain kinematically admissible multipliers 1.2, 6.0, and 1.5; all of them are larger than the first one. The joint mechanisms 5 and 6 are not admissible, because the corresponding external power is zero.

Next, we consider combinations of panel mechanisms 1 and 2. If we simply take their sum, we obtain mechanism 7 in Figure 6.7(b), with two plastic hinges around each of the multiple joints, leading to $\mu_k = 8/7 = 1.143$. By adding the joint mechanisms with

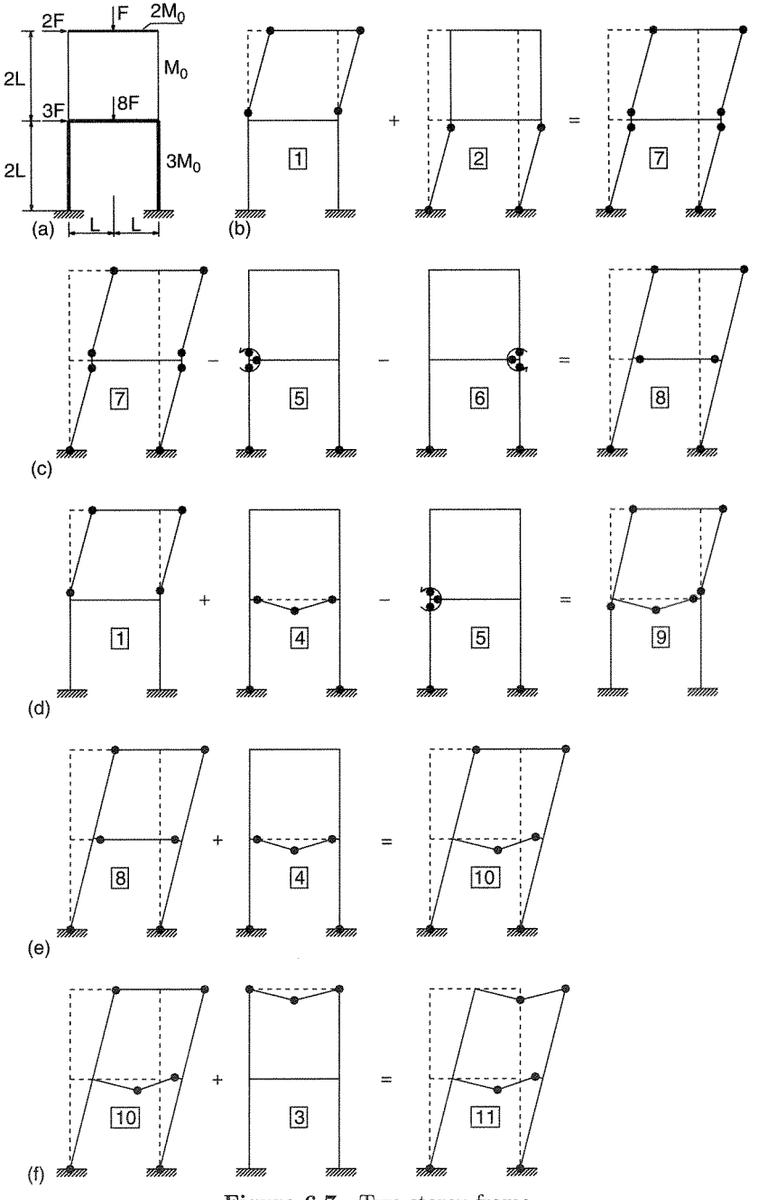


Figure 6.7 Two-storey frame

a negative sign, i.e. by rotating the joints clockwise, we reduce the number of plastic hinges, which results into a reduction of the dissipation power while the external power remains the same. Consequently, the kinematically admissible multiplier for mechanism 8 in Figure 6.7(c) is smaller, $\mu_k = 1$. This is the same value as previously obtained for mechanism 1. We now try to guess other combinations such that the dissipation power is reduced or the external power increases, so as to obtain a lower collapse load. We combine panel mechanism 1 with beam mechanism 4 and joint mechanism 5, obtaining mechanism 9 in Figure 6.7(d) with $\mu_k = 1.25$, which is worse than before. Now we try combining the best mechanism so far, 8, with the lower beam mechanism 4 obtaining mechanism 10 with $\mu_k = 10/11 = 0.909$; see Figure 6.7(e).

We have thus improved the previous best upper bound. Further, we also try combining the upper beam mechanism 3 with the last mechanism. This yields mechanism 11 in Figure 6.7(f), for which we however get $\mu_k = 24/23 = 1.043$, larger than the previous result. Intuitively, we can see now no way of combining further mechanisms in order to either reduce the dissipation power or increase the external power. Thus, although we have considered only some of the potential failure mechanisms, we are confident that $F_0 = 0.909 M_0/L$ is the true collapse load. □

For computational economy, it may be observed that in combining mechanisms we may directly sum the previously established expressions for the internal and external power for the mechanisms that are being superimposed. A systematic approach of this kind is provided by the methods of linear programming, which will be discussed in the next chapter. To get some idea of how good our previously calculated collapse load is, we might find some statically admissible moment distribution and the corresponding load giving the lower bound. This would lend certainty to our results without exhausting all potential failure mechanisms.

6.4 NONRECTANGULAR FRAMES

In the case of nonrectangular frames, the analysis of the kinematics is more tedious. To determine the relations between the rotations and displacements in the individual members of a potential failure mechanism, one must first figure out the instantaneous centers of rotation, O, for each element of the mechanism.

Example 6.6: Determine the collapse load for the gable frame in Figure 6.8(a).

Solution: The cross sectional properties are constant and there are no multiple joints (linking more than two members), so there is no need to distinguish between the end sections 21 and 23, 32 and 34, or 43 and 45. It is sufficient to consider 5 critical sections labeled 1 to 5, same as the joints. The frame is statically indeterminate to the third degree, and so the collapse mechanism must have four plastic hinges (unless it is a partial mechanism). This means that the actual collapse mechanism is one of five potential single degree-of-freedom mechanisms.

Let us first try the floor mechanism with plastic hinges at joints 1, 2, 4 and 5, because its kinematic analysis is quite simple. As shown in Figure 6.8(b), the rotation rate at all hinges is the same, $\dot{\theta}$, the vertical displacement rate at joint 3 is $\dot{w}_3 = 0$, and the horizontal displacement rate at joint 4 is $\dot{u}_4 = \dot{\theta} L$. From the power equality

$$F_k \dot{\theta} L = M_0 \dot{\theta} + M_0 \dot{\theta} + M_0 \dot{\theta} + M_0 \dot{\theta} \tag{6.38}$$

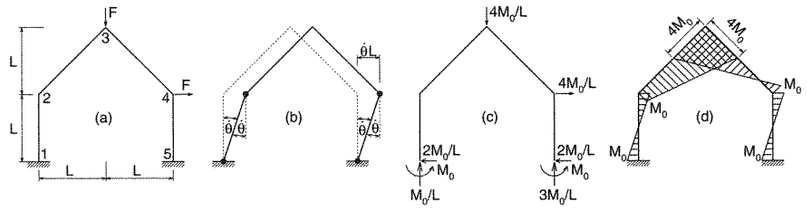


Figure 6.8 Gable frame – first mechanism

we obtain $F_k = 4M_0/L$. Now we have to check the moment M_3, to see whether the current solution is also statically admissible. The end moments in both columns (members 12 and 45) are known (equal to the plastic limit moment), and so it is easy to calculate the horizontal reactions, which equal $2M_0/L$ at both supports. Since the moment reactions M_0 are also known, it is easy to evaluate the vertical reactions from global equations of equilibrium; see Figure 6.8(c). Knowing the reactions, we can easily plot the bending moment diagram (Figure 6.8(d)), and it turns out that the moment at joint 3 exceeds the plastic limit moment four times. So from the analysis of this mechanism we obtain a rather poor estimate $M_0/L \leq F_0 \leq 4M_0/L$. The lower bound has been obtained by scaling down the bending moment distribution from Figure 6.8(d) and the corresponding external forces from Figure 6.8(c) by a factor 1/4.

Further we need to analyze the mechanism with a yield hinge at joint 3. Consider for example the mechanism shown in Figure 6.9(a). It consists of three rigid parts, denoted by letters A, B and C. The motion of each part is characterized by the center of rotation, O_I, and the rotation rate, $\dot{\theta}_I$, $I = A, B, C$. In the linearized approximation, valid for infinitesimally small rotations, each point of a rigid part is moving in the direction perpendicular to the line connecting that point with the center of rotation. A plastic hinge connecting two neighboring rigid parts is attached to both of them and must follow their motion. This is possible only if the hinge is located on the line connecting the two corresponding centers of rotation. Realizing these basic concepts from kinematics of small rotations, we can analyze our particular mechanism easily.

The center of rotation O_A of part A is at hinge 1 (the base of the left column), and the center of rotation O_C of part C is at hinge 5 (the base of the right column). Parts A and B are connected by hinge 3, and parts B and C are connected by hinge 4. This means that the center of rotation O_B must lie at the intersection of the line connecting hinges 1 and 3 with the line connecting hinges 5 and 4.

Now we know the centers of rotation and we can look at the rates (Figure 6.9(b)). Let us select the rotation rate $\dot{\theta} \equiv \dot{\theta}_A$ as the independent one. The distance of joint 3 from center O_A is $\sqrt{5}L$, and so the displacement rate at that joint (in the direction perpendicular to line 1-3) is $\dot{\theta} \times \sqrt{5}L$. This rate must be equal to the rotation rate of part B multiplied by the distance of O_B from joint 3, which also happens to be $\sqrt{5}L$. From that we obtain the rotation rate $\dot{\theta}_B = \dot{\theta}$, and we can express the rate at which joint 4 travels in the horizontal direction as $\dot{\theta} \times 3L$. Dividing this by the distance of center O_C from joint 4, we obtain the rotation rate $\dot{\theta}_C = \dot{\theta} \times 3L/L = 3\dot{\theta}$. Once we know the rotation rate of each rigid part, it is easy to get the rotation rate at each plastic hinge as the sum or difference of the rotation rates of the rigid parts connected by that hinge. Here we must take the orientation of each rotation rate into account. In our case, we obtain $\dot{\theta}_1 = \dot{\theta}_A = \dot{\theta}$, $\dot{\theta}_3 = \dot{\theta}_A + \dot{\theta}_B = 2\dot{\theta}$, $\dot{\theta}_4 = \dot{\theta}_B + \dot{\theta}_C = 4\dot{\theta}$, and

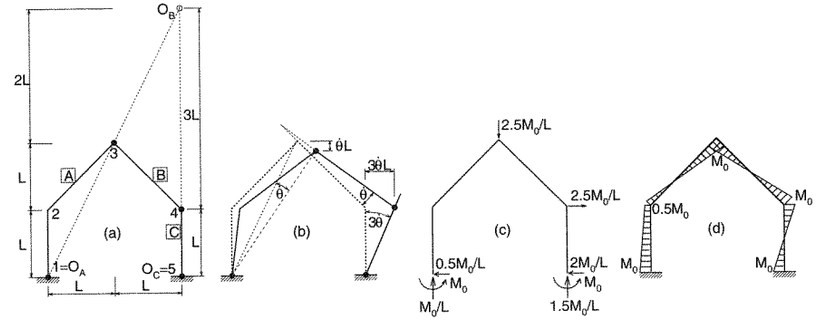


Figure 6.9 Gable frame – second mechanism

$\dot{\theta}_5 = \dot{\theta}_C = 3\dot{\theta}$. We need also the displacement rates at the points of application of the external forces, $\dot{w}_3 = \dot{\theta}L$ and $\dot{u}_4 = 3\dot{\theta}L$. After substitution of the above values, the power equality

$$F\dot{w}_3 + F\dot{u}_4 = M_0\dot{\theta}_1 + M_0\dot{\theta}_3 + M_0\dot{\theta}_4 + M_0\dot{\theta}_5 \tag{6.39}$$

$$F\dot{\theta}L + 3F\dot{\theta}L = M_0\dot{\theta} + 2M_0\dot{\theta} + 4M_0\dot{\theta} + 3M_0\dot{\theta} \tag{6.40}$$

provides the value of the kinematically admissible load $F_k = 2.5M_0/L$, which is smaller than that for the floor mechanism. From equilibrium of the right column, we obtain the horizontal reaction at the right support, $2M_0/L$, and the horizontal reaction at the left support then follows from global equilibrium; see Figure 6.9(c). This is all we need to compute the bending moment at critical section 2, $M_{23} = 0.5M_0$, and as this moment is plastically admissible, the bending moment distribution is statically admissible. We have found the actual collapse load, $F_0 = 2.5M_0/L$. The bending moment diagram at collapse is drawn in Figure 6.9(d). □

The preceding example demonstrated a practical approach to the solution of a relatively simple problem by hand calculations. To simplify the notation, we were not fussy about the signs, and we did not distinguish between end sections framing into the same joint. This was possible because every joint was connecting at most two members, and the moments at the end sections were thus equal in magnitude. More complicated problems are usually solved by computer programs, which have to be based on a systematic and formalized procedure. This will be explained in Chapter 7.

6.5 STATIC APPROACH – METHOD OF INEQUALITIES

Let us now show an application of the static approach to limit analysis of trusses and frames. The theoretical basis is provided by the lower bound theorem, i.e. the safety factor is evaluated as the largest statically admissible multiplier.

Example 6.7: Solve the problem from Example 6.1 using the static approach.

Solution: The truss is reproduced in Figure 6.10(a). Any statically admissible state is represented by the applied load F_s and five axial forces $S_1, S_2, \ldots S_5$ satisfying four independent equilibrium conditions. We can select one of the axial forces, say S_2, as the statically redundant one and express the remaining axial forces in terms of the applied

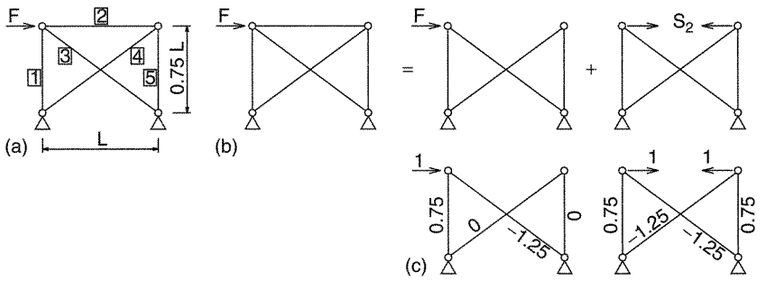


Figure 6.10 Truss analyzed by the static approach

force F_s and the statically redundant quantity S_2. This could be done by solving the equations of equilibrium with terms containing F_s and S_5 transferred to the right-hand side. As an alternative, we can think of the actual state as of a superposition of two equilibrium states of the corresponding statically determinate structure as indicated in Figure 6.10(b). Axial forces for each of these states can be easily calculated. Figure 6.10(c) shows the axial forces caused by unit values of F_s and S_2, respectively. These dimensionless forces represent the coefficients in the expressions for the actual axial forces. In addition to the equations of equilibrium, a statically admissible state must satisfy the conditions of plastic admissibility, $-S_0 \le S_i \le S_0$, $i = 1, 2, \ldots 5$. We thus have

$$\begin{aligned} -S_0 \le S_1 &= 0.75F_s + 0.75S_2 \le S_0 \\ -S_0 \le S_2 &= S_2 \le S_0 \\ -S_0 \le S_3 &= -1.25F_s - 1.25S_2 \le S_0 \\ -S_0 \le S_4 &= -1.25S_2 \le S_0 \\ -S_0 \le S_5 &= 0.75S_2 \le S_0 \end{aligned} \tag{6.41}$$

According to the lower bound theorem, the collapse load is the maximum possible value of F_s for which there exists some S_2 such that the above inequalities are satisfied. The second and the fifth inequality are obviously weaker than the fourth one, which is equivalent to

$$-0.8S_0 \le S_2 \le 0.8S_0 \tag{6.42}$$

Similarly, the third inequality is stronger than the first one and it yields

$$F_s \le \frac{1}{1.25}(S_0 - 1.25S_2) = 0.8S_0 - S_2 \tag{6.43}$$

Combining this with the previous inequality for S_2 we obtain

$$F_s \le 0.8S_0 - S_2 \le 0.8S_0 - (-0.8S_0) = 1.6S_0 \tag{6.44}$$

The maximum possible value

$$\max F_s = 1.6S_0 \tag{6.45}$$

is reached for $S_2 = -0.8S_0$. This is in agreement with the result obtained by the kinematic approach. Substituting into the relations for axial forces, we get

$$S_1 = 0.75 \times 1.6S_0 + 0.75 \times (-0.8S_0) = 0.6S_0 \tag{6.46}$$

$$S_3 = -1.25 \times 1.6S_0 - 1.25 \times (-0.8S_0) = -S_0 \tag{6.47}$$

$$S_4 = -1.25 \times (-0.8S_0) = S_0 \tag{6.48}$$

$$S_5 = 0.75 \times (-0.8S_0) = -0.6S_0 \tag{6.49}$$

As might have been expected, two of the axial forces are just at the yield limit. This confirms that the present statically admissible state is the actual state at collapse. □

Example 6.8: Consider the frame in Figure 6.11, with critical sections 2, 3, 4. Find the collapse load using the static approach.

Solution: Just for the purpose of setting up the equilibrium conditions, we consider the virtual work on the mechanisms shown in Figure 6.11, which yields

$$\begin{aligned} 2M_{32} - 2M_{21} &= 3F_sL - 2F_sL \\ -M_{43} + M_{21} &= 2F_sL \end{aligned} \tag{6.50}$$

Furthermore, for the moment field to be statically admissible, we must have

$$\begin{aligned} -M_0 \le M_{21} \le M_0 \\ -M_0 \le M_{32} \le M_0 \\ -M_0 \le M_{43} \le M_0 \end{aligned} \tag{6.51}$$

The objective is now to find the maximum load F_s that satisfies the foregoing system of equations and inequalities. An elementary approach, which is feasible only for very small structures, is as follows: From (6.50) we have $M_{32} = M_{21} + F_sL/2, M_{43} = M_{21} - 2F_sL$, and a substitution in (6.51) yields

$$\begin{aligned} -M_0 \le M_{21} &\le M_0 \\ -M_0 \le M_{21} + F_sL/2 &\le M_0 \\ -M_0 \le M_{21} - 2F_sL &\le M_0 \end{aligned} \tag{6.52}$$

To be able to eliminate the unknown M_{21} we must isolate it, rewriting these inequalities as follows

$$\begin{aligned} -M_0 \le M_{21} \le M_0 \\ -M_0 - F_sL/2 \le M_{21} \le M_0 - F_sL/2 \\ -M_0 + 2F_sL \le M_{21} \le M_0 + 2F_sL \end{aligned} \tag{6.53}$$

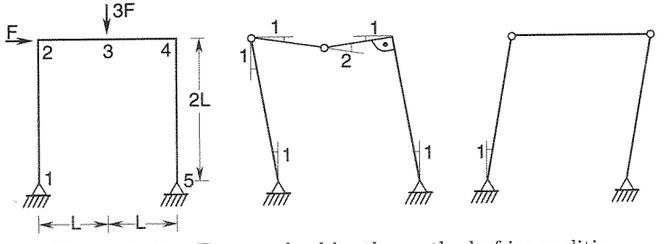


Figure 6.11 Frame solved by the method of inequalities

Now we look for the largest possible load F_s for which there is a value of moment M_{21} satisfying all conditions (6.53). Obviously, the second line represents the most severe restriction on M_{21} from the right, while the third line gives the most severe restriction from the left. Conditions $-M_0 + 2F_sL \le M_{21} \le M_0 - F_sL/2$ can be satisfied only if $-M_0+2F_sL \le M_0-F_sL/2$, i.e., if $2.5F_sL \le 2M_0$. So, the largest statically admissible load is

$$F_0 = \max F_s = \frac{0.8M_0}{L} \tag{6.54}$$

This is the exact solution of the critical load. The reader may check by the kinematic approach that F_0 is indeed the correct collapse load. □

PROBLEMS

Problem 6.1: Analyze all kinematically admissible mechanisms with one degree of freedom for the structure in Figure 6.2(a) and show that all the corresponding kinematically admissible values of F are larger than or equal to the collapse load determined in Example 6.2.

Problem 6.2: Consider the truss in Figure 6.12. Given that $S_{0,b} = 2S_{0,d}$ where $S_{0,d}$ = plastic axial force in the diagonals and $S_{0,b}$ = plastic axial force in the other bars, determine the kinematically admissible multiplier corresponding to the partial mechanism with the diagonals in the first field yielding, and prove that this is the actual collapse mechanism.

Problem 6.3: A beam grillage is a planar framework loaded perpendicular to its plane. The collapse load on beam grillages may be approximately analyzed, neglecting the plastic torsional resistance of the beams, in which case the plastic bending represents the only dissipative mechanism. Figure 6.13 shows examples of plastic hinge locations for various potential failure mechanisms of two simple rectangular grillages. Write the power equality for these mechanisms. The joints on the boundary are supported against the out-of-plane displacement (but they can rotate freely), and all the other joints are loaded by the same out-of-plane force, F.

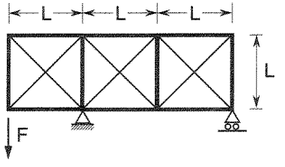


Figure 6.12 Truss to be analyzed in Problem 6.2

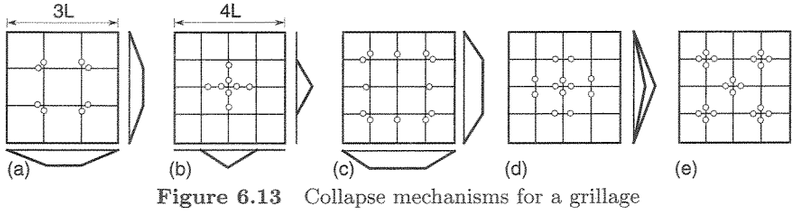


Figure 6.13 Collapse mechanisms for a grillage

Problem 6.4: For the gable frame in Figure 6.8(a), analyze three potential failure mechanisms that were not considered in Example 6.6, determine the corresponding kinematically admissible multipliers and verify that they are larger than the collapse load.

Problem 6.5: Solve Example 6.1 for $S_0^- = 2S_0^+$, i.e. for the case when the plastic axial force in compression is twice the plastic axial force in tension (for all the bars).

7

Linear Programming in Limit Analysis

The solutions presented in the preceding chapter were based on the solid foundation of the theorems of limit analysis. However, the methods we have used to search for the maximum statically admissible multiplier or for the minimum kinematically admissible multiplier have so far been unsystematic, and partially relying on intuition. Such an approach would not be feasible in the design of large real-world structures. Fortunately, it turns out that plastic limit analysis can be formulated as a problem of linear programming, and powerful methods developed in the mathematical theory of optimization can then be applied.

The basic definitions used in the theory of linear programming and a brief description of the simplex algorithm are given in Appendix C. The present chapter is meant as a simple introduction to the application of linear programming methods in plasticity. A systematic treatment of this wide subject, including extensions to nonlinear programming, can be found in Cohn, Maier and Grierson (1979) and Smith (1990).

7.1 KINEMATIC APPROACH

Consider first the kinematic approach. The upper bound theorem states that the safety factor can be obtained by minimizing the ratio

$$\mu_k = \frac{\boldsymbol{s}_0^T |\dot{\boldsymbol{e}}|}{\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}}} \tag{7.1}$$

where $\boldsymbol{s}_0$ and $\bar{\boldsymbol{f}}$ are given and $\dot{\boldsymbol{e}}$ and $\dot{\boldsymbol{d}}$ must satisfy the conditions

$$\dot{\boldsymbol{e}} = \boldsymbol{B}\dot{\boldsymbol{d}} \tag{7.2}$$

$$\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} > 0 \tag{7.3}$$

Any kinematically admissible mechanism is characterized by certain displacement rates $\dot{\boldsymbol{d}}$ and deformation rates $\dot{\boldsymbol{e}}$. If we multiply both $\dot{\boldsymbol{d}}$ and $\dot{\boldsymbol{e}}$ by the same positive number, we get another kinematically admissible mechanism with the same corresponding load multiplier μ_k. We can thus restrict our attention to solutions normalized by the condition

$$\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} = 1 \tag{7.4}$$

For such mechanisms, expression (7.1) for the kinematically admissible load multiplier simplifies to

$$\mu_k = \boldsymbol{s}_0^T |\dot{\boldsymbol{e}}| \tag{7.5}$$

The constraints (7.2) and (7.4) are linear in terms of the unknown quantities $\dot{\boldsymbol{d}}$ and $\dot{\boldsymbol{e}}$, but the objective function (7.5) is not, due to the presence of the absolute value operator. This can be easily circumvented by representing the extension rates as

$$\dot{\boldsymbol{e}} = \dot{\boldsymbol{e}}^+ - \dot{\boldsymbol{e}}^- \tag{7.6}$$

where $\dot{\boldsymbol{e}}^+ = (|\dot{\boldsymbol{e}}| + \dot{\boldsymbol{e}})/2 \geq \boldsymbol{0}$ is the positive part and $\dot{\boldsymbol{e}}^- = (|\dot{\boldsymbol{e}}| - \dot{\boldsymbol{e}})/2 \geq \boldsymbol{0}$ is the negative part of $\dot{\boldsymbol{e}}$. The absolute value can obviously be expressed as $|\dot{\boldsymbol{e}}| = \dot{\boldsymbol{e}}^+ + \dot{\boldsymbol{e}}^-$, and so the objective function is linear in terms of the new variables $\dot{\boldsymbol{e}}^+$ and $\dot{\boldsymbol{e}}^-$. The problem

$$\text{Minimize } \mu_k(\dot{\boldsymbol{e}}^+, \dot{\boldsymbol{e}}^-, \dot{\boldsymbol{d}}) \equiv \boldsymbol{s}_0^T \dot{\boldsymbol{e}}^+ + \boldsymbol{s}_0^T \dot{\boldsymbol{e}}^- \tag{7.7}$$

subject to

$$\dot{\boldsymbol{e}}^+ - \dot{\boldsymbol{e}}^- - \boldsymbol{B}\dot{\boldsymbol{d}} = \boldsymbol{0} \tag{7.8}$$

$$\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} = 1 \tag{7.9}$$

$$\dot{\boldsymbol{e}}^+ \geq \boldsymbol{0} \tag{7.10}$$

$$\dot{\boldsymbol{e}}^- \geq \boldsymbol{0} \tag{7.11}$$

is a problem of linear programming.

7.2 TRANSFORMATION TO STANDARD FORM

The simplex method (Appendix C) is applicable to problems written in the so-called *standard form* with nonnegative variables and equality constraints:

$$\text{Minimize } \boldsymbol{c}^T \boldsymbol{x} \tag{7.12}$$

subject to

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{7.13}$$

$$\boldsymbol{x} \geq \boldsymbol{0} \tag{7.14}$$

The only reason why the present formulation (7.7)–(7.11) is not in the standard form is that the variables $\dot{\boldsymbol{d}}$ are not necessarily nonnegative. One way of transforming the problem to the standard form is to represent $\dot{\boldsymbol{d}}$ as a difference of two nonnegative vectors:

$$\dot{\boldsymbol{d}} = \dot{\boldsymbol{d}}^+ - \dot{\boldsymbol{d}}^-, \ \dot{\boldsymbol{d}}^+ \geq \boldsymbol{0}, \ \dot{\boldsymbol{d}}^- \geq \boldsymbol{0} \tag{7.15}$$

Constraints (7.8) then change into

$$\dot{\boldsymbol{e}}^+ - \dot{\boldsymbol{e}}^- - \boldsymbol{B}\dot{\boldsymbol{d}}^+ + \boldsymbol{B}\dot{\boldsymbol{d}}^- = \boldsymbol{0} \tag{7.16}$$

and the problem (7.7) with nonnegative variables $\dot{\boldsymbol{e}}^+$, $\dot{\boldsymbol{e}}^-$, $\dot{\boldsymbol{d}}^+$ and $\dot{\boldsymbol{d}}^-$, and with equality constraints (7.9) and (7.16), is in the standard form.

Another way of transforming problem (7.7)–(7.11) to the standard form is to eliminate the displacement rates from the constraints. For a statically indeterminate structure, the number of generalized strains, m, is larger than the number of generalized displacements, n. The difference $s = m - n$ is the degree of statical indeterminacy (or redundancy). We can use n kinematic equations to express the generalized displacements in terms of the generalized strains, and then eliminate the generalized displacements from the remaining $m - n$ kinematic equations. The resulting relations are the *equations of compatibility*, which can alternatively be obtained by the principle of virtual work. Both techniques will be demonstrated in the following example.

Example 7.1: Set up the equations of compatibility for the truss in Figure 7.1(a).

Solution: The horizontal and vertical displacement of the unsupported node are denoted u and w, respectively. The kinematic equations for this truss,

$$0.6u + 0.8w = e_1 \tag{7.17}$$

$$w = e_2 \tag{7.18}$$

$$-0.8u + 0.6w = e_3 \tag{7.19}$$

were constructed in Example 3.1. Using the first two kinematic equations, the displacements can be expressed in terms of the extensions as

$$w = e_2 \tag{7.20}$$

$$u = \frac{1}{0.6}(e_1 - 0.8w) = 1.667e_1 - 1.333e_2 \tag{7.21}$$

Substituting into the last kinematic equation, we get the equation of compatibility

$$-1.333e_1 + 1.667e_2 = e_3 \tag{7.22}$$

Alternatively, we can consider a set of virtual axial forces δS_1, δS_2 and δS_3 that are in equilibrium with zero external loads, i.e. that satisfy the conditions of equilibrium

$$0.6\,\delta S_1 - 0.8\,\delta S_3 = 0 \tag{7.23}$$

$$0.8\,\delta S_1 + \delta S_2 + 0.6\,\delta S_3 = 0 \tag{7.24}$$

Setting formally $\delta S_3 = 1$, we get $\delta S_1 = 1.333$ and $\delta S_2 = -1.667$. The physical interpretation is that we prescribe a unit axial force in bar 3 and calculate the axial

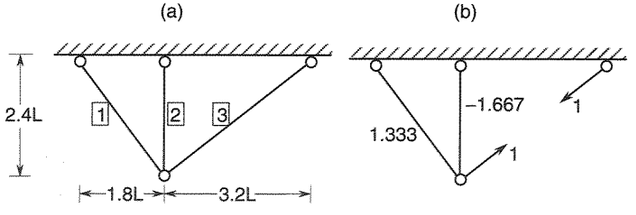


Figure 7.1 Truss analyzed in Examples 7.1 and 7.2

forces in the other bars by solving the statically determinate problem in Figure 7.1(b). As the virtual applied forces are zero, the work done by the virtual axial forces on the actual bar extensions must be zero:

$$\delta W^*_{\text{int}} = \boldsymbol{\delta s}^T \boldsymbol{e} = \sum_{i=1}^{3} \delta S_i \, e_i = 0 \tag{7.25}$$

$$1.333e_1 - 1.667e_2 + 1 \times e_3 = 0 \tag{7.26}$$

This condition of compatibility is the same as before; cf. (7.22). After rescaling, we can write it as

$$e_1 - 1.25e_2 + 0.75e_3 = 0 \tag{7.27}$$

Of course, the same condition must hold for the extension rates. This could be formally proven by differentiating (7.27) with respect to time. □

In general, the conditions of compatibility can be derived by applying Gauss transformations simultaneously to the kinematic matrix $\boldsymbol{B}$ and to a unit matrix $\boldsymbol{I}$ having the same number of rows as $\boldsymbol{B}$. Initially, these matrices represent the kinematic equations

$$\boldsymbol{B}\boldsymbol{d} = \boldsymbol{I}\boldsymbol{e} \tag{7.28}$$

or, in the rate form,

$$\boldsymbol{B}\dot{\boldsymbol{d}} = \boldsymbol{I}\dot{\boldsymbol{e}} \tag{7.29}$$

As already mentioned, in the kinematic matrix $\boldsymbol{B}$, the number of rows, m, is larger than the number of columns, n. So we can transform it by the usual row operations to a matrix consisting of a unit block $\boldsymbol{I}$ of size (n, n) and a zero block $\boldsymbol{O}$ of size $(m-n, n)$ below it. The unit matrix of size (m, m) on the right-hand side of (7.29) is changed by the same transformations to a general matrix represented by blocks $\boldsymbol{P}^T$ and $\boldsymbol{S}^T$ (Figure 7.2). The corresponding equations

$$\boldsymbol{I}\,\dot{\boldsymbol{d}} = \boldsymbol{P}^T\dot{\boldsymbol{e}} \tag{7.30}$$

$$\boldsymbol{O}\,\dot{\boldsymbol{d}} = \boldsymbol{S}^T\dot{\boldsymbol{e}} \tag{7.31}$$

are respectively the representation of $\dot{\boldsymbol{d}}$ in terms of $\dot{\boldsymbol{e}}$,

$$\dot{\boldsymbol{d}} = \boldsymbol{P}^T\dot{\boldsymbol{e}} \tag{7.32}$$

and the equations of compatibility,

$$\boldsymbol{S}^T\dot{\boldsymbol{e}} = \boldsymbol{0} \tag{7.33}$$

Using (7.32), it is easy to eliminate the displacements from the normalizing condition $\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}} = 1$, which can now be replaced by $\boldsymbol{p}^T\dot{\boldsymbol{e}} = 1$ where $\boldsymbol{p} = \boldsymbol{P}\bar{\boldsymbol{f}}$. The final form of the linear programming problem (7.7)–(7.11) is then

$$\text{Minimize } \mu_k(\dot{\boldsymbol{e}}^+, \dot{\boldsymbol{e}}^-) \equiv \boldsymbol{s}_0^T\dot{\boldsymbol{e}}^+ + \boldsymbol{s}_0^T\dot{\boldsymbol{e}}^- \tag{7.34}$$

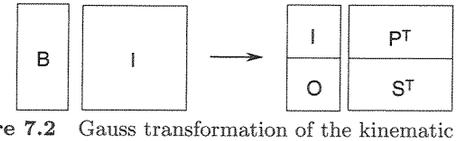


Figure 7.2 Gauss transformation of the kinematic matrix

subject to

$$\boldsymbol{S}^T\dot{\boldsymbol{e}}^+ - \boldsymbol{S}^T\dot{\boldsymbol{e}}^- = \boldsymbol{0} \tag{7.35}$$

$$\boldsymbol{p}^T\dot{\boldsymbol{e}}^+ - \boldsymbol{p}^T\dot{\boldsymbol{e}}^- = 1 \tag{7.36}$$

$$\dot{\boldsymbol{e}}^+ \geq \boldsymbol{0},\ \dot{\boldsymbol{e}}^- \geq \boldsymbol{0} \tag{7.37}$$

This problem can be readily solved by the simplex method.

7.3 APPLICATION TO TRUSSES

Example 7.2: Formulate the kinematic approach to limit analysis of the truss in Figure 7.1 as a problem of linear programming in the standard form, and solve it using the simplex method. The truss is subjected to an increasing vertical load, and all the three bars are characterized by the same plastic axial force, S_0.

Solution: The equation of compatibility

$$\dot{e}_1 - 1.25\dot{e}_2 + 0.75\dot{e}_3 = 0 \tag{7.38}$$

has been derived in Example 7.1. Let us represent the external loading as

$$\boldsymbol{f} = \begin{Bmatrix} 0 \\ F \end{Bmatrix} = \mu \begin{Bmatrix} 0 \\ S_0 \end{Bmatrix} \tag{7.39}$$

where μ is the load multiplier, and $\bar{\boldsymbol{f}} = \{0, S_0\}^T$ is the reference load vector. As $\dot{\boldsymbol{d}} = \{\dot{u}, \dot{w}\}^T$, the normalizing condition (7.9) reads $S_0\dot{w} = 1$ and, by substitution $\dot{w} = \dot{e}_2$ (which follows from (7.20)), this can be transformed into $S_0\dot{e}_2 = 1$. The resulting problem of linear programming (7.34)–(7.37) can be stated as follows:

$$\text{Minimize } \mu_k(\dot{\boldsymbol{e}}^+, \dot{\boldsymbol{e}}^-) \equiv S_0(\dot{e}_1^+ + \dot{e}_1^- + \dot{e}_2^+ + \dot{e}_2^- + \dot{e}_3^+ + \dot{e}_3^-) \tag{7.40}$$

subject to

$$\dot{e}_1^+ - \dot{e}_1^- - 1.25\dot{e}_2^+ + 1.25\dot{e}_2^- + 0.75\dot{e}_3^+ - 0.75\dot{e}_3^- = 0 \tag{7.41}$$

$$S_0\dot{e}_2^+ - S_0\dot{e}_2^- = 1 \tag{7.42}$$

$$\dot{e}_i^+ \geq 0,\ \dot{e}_i^- \geq 0, \qquad i = 1, 2, 3 \tag{7.43}$$

To simplify the normalizing condition and the objective function, we can divide both the left-hand side of (7.42) and the definition of the objective function (7.40) by the same factor, S_0. The corresponding simplex tableau is presented in Table 7.1. The first two lines represent the constraints, and the last one defines the objective function.

The first step of the simplex algorithm consists in constructing any feasible solution. Let us select e.g. $\dot{e}_2^+$ and $\dot{e}_3^+$ as the basic variables. The tableau is then modified by suitable Gauss transformations. The first row is divided by -1.25 and then subtracted from the second row and from the third row in order to eliminate the coefficients corresponding to the basic variable $\dot{e}_2^+$. Then, the second row is divided by 0.6, its 0.6-multiple is added to the first row and its 1.6-multiple is subtracted from the third

Table 7.1 Initial simplex tableau (not feasible)

$-\mu_k$	$\dot{e}_1^+$	$\dot{e}_2^+$	$\dot{e}_3^+$	$\dot{e}_1^-$	$\dot{e}_2^-$	$\dot{e}_3^-$	RHS
0	1	−1.25	0.75	−1	1.25	−0.75	0
0	0	1	0	0	−1	0	1
1	1	1	1	1	1	1	0

Table 7.2 Simplex tableau with an initial feasible solution

$-\mu_k$	$\dot{e}_1^+$	$\dot{e}_2^+$	$\dot{e}_3^+$	$\dot{e}_1^-$	$\dot{e}_2^-$	$\dot{e}_3^-$	RHS
0	0.000	1.000	0.000	0.000	−1.000	0.000	1.000
0	1.333	0.000	1.000	−1.333	0.000	−1.000	1.667
1	−0.333	0.000	0.000	2.333	2.000	2.000	−2.667

row in order to eliminate the coefficients corresponding to the basic variable $\dot{e}_3^+$. The modified tableau in Table 7.2 indicates that the initial feasible solution is

$$\dot{e}_1^+ = \dot{e}_1^- = \dot{e}_2^- = \dot{e}_3^- = 0 \tag{7.44}$$
$$\dot{e}_2^+ = 1 \tag{7.45}$$
$$\dot{e}_3^+ = 1.667 \tag{7.46}$$
$$-\mu_k = -2.667 \ \ldots \ F_k = 2.667 S_0 \tag{7.47}$$

This solution represents the mechanism with bars 2 and 3 yielding in tension. The last line of the tableau describes the variation of μ_k with respect to the nonbasic variables:

$$\mu_k = 2.667 - 0.333\dot{e}_1^+ + 2.333\dot{e}_1^- + 2\dot{e}_2^- + 2\dot{e}_3^- \tag{7.48}$$

The presence of a negative coefficient at $\dot{e}_1^+$ indicates that the kinematically admissible multiplier can be decreased by increasing $\dot{e}_1^+$ from zero to a positive value, i.e., by considering a mechanism with bar 1 yielding in tension. This means that $\dot{e}_1^+$ will enter the basis, i.e. it will be added to the set of basic variables. The column below $\dot{e}_1^+$ is therefore the key column. The key row is in general determined by looking at the ratios of the elements in the key column to the elements on the right hand side (see Appendix C). The key row is that which maximizes this ratio among the rows having a positive element in the key column. For our problem, the only row satisfying this condition is the second one. This means that $\dot{e}_3^+$ will be removed from the set of basic variables, as the corresponding column has a nonzero entry in the key row. Physically, we shift from a mechanism with bars 2 and 3 yielding to a mechanism with bars 1 and 2 yielding. After dividing the second row by 1.333 and adding its 0.333-multiple to the last row, we get the tableau in Table 7.3.

As all the columns corresponding to the nonbasic variables have positive entries in the last row representing the objective function, the solution

$$\dot{e}_3^+ = \dot{e}_1^- = \dot{e}_2^- = \dot{e}_3^- = 0 \tag{7.49}$$

Table 7.3 Simplex tableau with the optimal solution

$-\mu_k$	$\dot{e}_1^+$	$\dot{e}_2^+$	$\dot{e}_3^+$	$\dot{e}_1^-$	$\dot{e}_2^-$	$\dot{e}_3^-$	RHS
0	0.000	1.000	0.000	0.000	−1.000	0.000	1.000
0	1.000	0.000	0.750	−1.000	0.000	−0.750	1.250
1	0.000	0.000	0.250	2.000	2.000	1.750	−2.250

$$\dot{e}_2^+ = 1 \tag{7.50}$$
$$\dot{e}_1^+ = 1.25 \tag{7.51}$$
$$-\mu_k = -2.25 \ \ldots \ F_k = 2.25 S_0 \tag{7.52}$$

is the optimal one. This means that the collapse load is $F_0 = \min F_k = 2.25 S_0$ and collapse occurs by yielding of bars 1 and 2, both in tension.

It can be proven (see Section 7.6) that the coefficients in the last row of the simplex tableau have a very interesting and useful physical meaning – they represent the elastic reserves, i.e. the differences between the plastic axial force and the axial force calculated from equilibrium under the assumption that the axial forces in the presumably yielding bars are equal to the plastic forces. This is consistent with the fact that the values in the columns corresponding to the basic variables are zero, and it also explains why negative values indicate that the solution is not optimal. Figure 7.3 illustrates the definition of the positive elastic reserve, $S_p = S_0 - S$, and the negative elastic reserve, $S_n = S - (-S_0) = S_0 + S$. A negative value of either S_p or S_n means that the axial force S would exceed the yield limit, and so the solution is not statically admissible; therefore, it cannot be optimal. Moreover, it is very helpful that all the actual axial forces at collapse easily follow from the values found in the last row of the optimal simplex tableau. In our case, the value 0.25 in the column corresponding to $\dot{e}_3^+$ represents the positive elastic reserve $S_{p,3} = S_0 - S_3 = 0.25 S_0$, and so $S_3 = 0.75 S_0$. The same result can be obtained using the value in the column under $\dot{e}_3^-$, which represents the negative elastic reserve $S_{n,3} = S_3 + S_0 = 1.75 S_0$. □

The present, simple example could have been solved by the elementary analysis presented in the previous chapter. However, for larger problems the simplex method is superior as it provides a systematic way of decreasing the value of the kinematically admissible multiplier and deciding whether its minimum possible value has been obtained. Such an algorithm can easily be programmed.

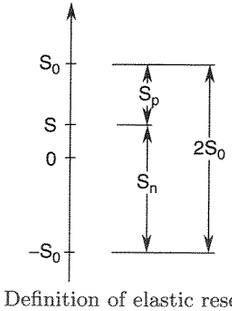


Figure 7.3 Definition of elastic reserves S_p and S_n

In each simplex tableau, the columns below $\dot{e}_i^+$ and $\dot{e}_i^-$ (where i is any integer between 1 and m) differ only by their sign, except for the coefficients in the last row. This is quite natural since the initial tableau has this particular property (see the matrices multiplying $\dot{\boldsymbol{e}}^+$ and $\dot{\boldsymbol{e}}^-$ in (7.35)–(7.36)), and all the subsequent manipulations performed with entire rows preserve it. For similar reasons, the sum of the coefficients in the last row corresponding to $\dot{e}_i^+$ and $\dot{e}_i^-$ always equals $2S_{0,i}$ (in our example we have divided the coefficients of the objective function by S_0, and so the sum of $\dot{e}_i^+$ and $\dot{e}_i^-$ equals 2). Also, the very first column below $-\mu_k$ remains unchanged throughout the calculation.

It thus suffices to store only one half of the tableau in the computer memory – the columns corresponding to $\dot{e}_i^+, i = 1, 2, \ldots m$, and the right-hand side. The total storage requirement is $(s+2)(m+1)$ real numbers. Note that the initial feasible solution can be constructed by taking almost any $s+1$ bar extensions as the basic variables (of course, we have to avoid combinations which lead to a singular submatrix of $\boldsymbol{B}$). If the resulting right-hand side corresponding to some $\dot{e}_i^+$ is negative, the entire row is multiplied by -1, and $\dot{e}_i^-$ is taken as one of the basic variables.

7.4 APPLICATION TO FRAMES

Similar to trusses, the standard form of the linear programming problem resulting from the kinematic approach to limit analysis of frames is again given by (7.7)–(7.11). Let us first demonstrate the derivation of the compatibility conditions, which will later be used while converting the problem to the standard form (7.34)–(7.37).

Example 7.3: Construct the compatibility conditions for the frame in Figure 7.4(a).

Solution: We start from the kinematic equations describing the relation between the generalized strains and the generalized displacements. Deformation of an elastic plane frame member is described by three variables – the extension e_{ij} and the bending deformations τ_{ij} and τ_{ji}. However, during plastic collapse, the elastic strains

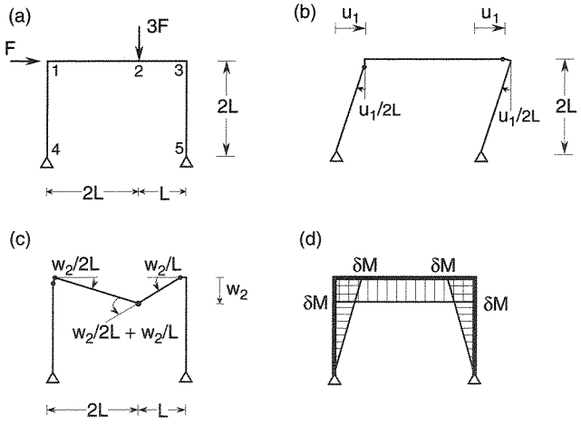


Figure 7.4 Frame solved in Examples 7.3 and 7.5: a) geometry and loading, b) panel mechanism, c) beam mechanism, d) self-equilibrated virtual moments

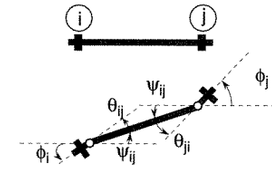


Figure 7.5 Kinematic relations

remain constant and all the deformation increments are entirely plastic. In the present simplified theory, neglecting the effect of normal and shear forces on the formation of a plastic hinge, all the plastic deformations are modeled by rotations of the yield hinges. This means that the plastic deformation of a member is described by, at most, two variables – plastic rotations θ_{ij} and θ_{ji} of the yield hinges that have formed at the ends, which represent the plastic components of the generalized bending strains τ_{ij} and τ_{ji}; see Figure 7.5. No plastic extension of a member is taken into account, and so the members can be treated as inextensible during collapse. This condition reduces the number of independent generalized displacements.

Our frame, for example, is subdivided into four members, and its elastic deformation is described by 10 variables – extensions e_{14}, e_{12}, e_{23} and e_{35}, and bending deformations τ_{14}, τ_{12}, τ_{21}, τ_{23}, τ_{32} and τ_{35}. Recall that the angles τ_{41} and τ_{53} are not considered as independent variables because the corresponding ends are connected to the respective joints by hinges. The elastic generalized displacements are the displacements and rotations of joints 1, 2 and 3. During plastic collapse, the additional (plastic) extensions are assumed to be zero, and the plastic deformation is described by rotations θ_{14}, θ_{12}, θ_{21}, θ_{23}, θ_{32} and θ_{35}. Their rates (increments) are related to the displacement and rotation rates by the kinematic equations that follow from the rate form of equations (B.5)–(B.6) in Appendix B, with $\dot{\tau}$ replaced by $\dot{\theta}$:

$$\dot{\theta}_{ij} = \dot{\phi}_i - \dot{\psi}_{ij} \tag{7.53}$$

$$\dot{\theta}_{ji} = \dot{\phi}_j - \dot{\psi}_{ij} \tag{7.54}$$

where

$$\dot{\psi}_{ij} = \frac{\sin\alpha_{ij}}{\ell_{ij}}(\dot{u}_j - \dot{u}_i) - \frac{\cos\alpha_{ij}}{\ell_{ij}}(\dot{w}_j - \dot{w}_i) \tag{7.55}$$

is the rate of the chord rotation; cf. (B.83). The inextensibility condition

$$\dot{e}_{ij} = \cos\alpha_{ij}(\dot{u}_j - \dot{u}_i) + \sin\alpha_{ij}(\dot{w}_j - \dot{w}_i) = 0 \tag{7.56}$$

can be used to reduce the number of independent joint displacement rates. In our case, we write the inextensibility conditions as

$$\dot{e}_{14} = \dot{\bar{w}}_4 - \dot{w}_1 = 0 \tag{7.57}$$

$$\dot{e}_{12} = \dot{u}_2 - \dot{u}_1 = 0 \tag{7.58}$$

$$\dot{e}_{23} = \dot{u}_3 - \dot{u}_2 = 0 \tag{7.59}$$

$$\dot{e}_{35} = \dot{\bar{w}}_5 - \dot{w}_3 = 0 \tag{7.60}$$

and, taking into account that the displacement rates at the supports, $\dot{\bar{w}}_4$ and $\dot{\bar{w}}_5$, are zero, we get

$$\dot{w}_1 = 0, \quad \dot{u}_2 = \dot{u}_1, \quad \dot{u}_3 = \dot{u}_2, \quad \dot{w}_3 = 0 \tag{7.61}$$

Consequently, we can express all displacement rates in terms of two independent rates, $\dot{u}_1$ and $\dot{w}_2$. By virtue of inextensibility, the number of generalized displacements has thus been reduced from nine to five, and the number of generalized strains from ten to six.

Based on (7.53)–(7.55), we can construct the kinematic equations:

$$\dot{\theta}_{14} = \dot{\phi}_1 - \dot{\psi}_{14} = \dot{\phi}_1 - (\dot{\bar{u}}_4 - \dot{u}_1)/2L = \dot{\phi}_1 + \dot{u}_1/2L \tag{7.62}$$
$$\dot{\theta}_{12} = \dot{\phi}_1 - \dot{\psi}_{12} = \dot{\phi}_1 - (\dot{w}_1 - \dot{w}_2)/2L = \dot{\phi}_1 + \dot{w}_2/2L \tag{7.63}$$
$$\dot{\theta}_{21} = \dot{\phi}_2 - \dot{\psi}_{12} = \dot{\phi}_2 - (\dot{w}_1 - \dot{w}_2)/2L = \dot{\phi}_2 + \dot{w}_2/2L \tag{7.64}$$
$$\dot{\theta}_{23} = \dot{\phi}_2 - \dot{\psi}_{23} = \dot{\phi}_2 - (\dot{w}_2 - \dot{w}_3)/L = \dot{\phi}_2 - \dot{w}_2/L \tag{7.65}$$
$$\dot{\theta}_{32} = \dot{\phi}_3 - \dot{\psi}_{23} = \dot{\phi}_3 - (\dot{w}_2 - \dot{w}_3)/L = \dot{\phi}_3 - \dot{w}_2/L \tag{7.66}$$
$$\dot{\theta}_{35} = \dot{\phi}_3 - \dot{\psi}_{35} = \dot{\phi}_3 - (\dot{\bar{u}}_5 - \dot{u}_3)/2L = \dot{\phi}_3 + \dot{u}_1/2L \tag{7.67}$$

These equations represent a kinematic description of the frame in the case of a general loading applied on the unsupported joints 1, 2 and 3. A further reduction of the number of variables is possible if we consider a special case in which only external forces but no external moments are applied at the joints. At joints that connect only two members, it suffices to consider only one of the end cross sections as the critical one. In our problem, the number of generalized strains can be reduced from six to three. At the same time, the number of generalized displacements must also be reduced by three. Setting $\dot{\theta}_{12} = 0$, $\dot{\theta}_{23} = 0$ and $\dot{\theta}_{35} = 0$ and using the corresponding kinematic equations, we can express the rotation rates from (7.63), (7.65) and (7.67) as follows:

$$\dot{\phi}_1 = -\dot{w}_2/2L \tag{7.68}$$
$$\dot{\phi}_2 = \dot{w}_2/L \tag{7.69}$$
$$\dot{\phi}_3 = -\dot{u}_1/2L \tag{7.70}$$

Then we can eliminate them from the remaining kinematic equations and get

$$\dot{\theta}_{14} = \dot{u}_1/2L - \dot{w}_2/2L \tag{7.71}$$
$$\dot{\theta}_{21} = \dot{w}_2/2L + \dot{w}_2/L = 1.5\dot{w}_2/L \tag{7.72}$$
$$\dot{\theta}_{32} = -\dot{u}_1/2L - \dot{w}_2/L \tag{7.73}$$

or, in matrix notation,

$$\begin{Bmatrix} \dot{\theta}_{14} \\ \dot{\theta}_{21} \\ \dot{\theta}_{32} \end{Bmatrix} = \frac{1}{L} \begin{bmatrix} 0.5 & -0.5 \\ 0 & 1.5 \\ -0.5 & -1 \end{bmatrix} \begin{Bmatrix} \dot{u}_1 \\ \dot{w}_2 \end{Bmatrix} \tag{7.74}$$

Equations (7.74) have been derived by a formal mathematical procedure but they can be constructed directly from kinematic considerations. Note that every independent generalized displacement is associated with one of the elementary mechanism. In our example, $\dot{u}_1$ corresponds to the panel mechanism in Figure 7.4(b), and $\dot{w}_2$ corresponds to the beam mechanism in Figure 7.4(c). From the figures it is clear that the hinge rotation rates for the panel mechanism are $\dot{\theta}_{14} = \dot{u}_1/2L$, $\dot{\theta}_{21} = 0$ and $\dot{\theta}_{32} = -\dot{u}_1/2L$, while for the beam mechanism they are $\dot{\theta}_{14} = -\dot{w}_2/2L$,

$\dot{\theta}_{21} = 3\dot{w}_2/2L$ and $\dot{\theta}_{32} = -\dot{w}_2/L$. This provides us with the coefficients of the kinematic matrix,

$$\boldsymbol{B} = \frac{1}{L} \begin{bmatrix} 0.5 & -0.5 \\ 0 & 1.5 \\ -0.5 & -1 \end{bmatrix} \tag{7.75}$$

Each column of the matrix corresponds to an elementary mechanism, and each row to a critical cross section.

Having derived the kinematic equations, we can transform them into equations of compatibility using the general elimination procedure explained in Example 7.1. As the structure is statically indeterminate to the first degree, there is only one independent compatibility condition. This is also indicated by the difference between the number of generalized strains and the number of generalized displacements, $m - n = 3 - 2 = 1$. The Gauss elimination procedure

$$\left[\begin{array}{cc|ccc} 0.5 & -0.5 & 1 & 0 & 0 \\ 0 & 1.5 & 0 & 1 & 0 \\ -0.5 & -1 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|ccc} 1 & 0 & 2 & 0.667 & 0 \\ 0 & 1 & 0 & 0.667 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] \tag{7.76}$$

leads to the inverted kinematic relations

$$\dot{u}_1 = 2L\dot{\theta}_{14} + 0.667L\dot{\theta}_{21} \tag{7.77}$$
$$\dot{w}_2 = 0.667L\dot{\theta}_{21} \tag{7.78}$$

and to the compatibility condition

$$\dot{\theta}_{14} + \dot{\theta}_{21} + \dot{\theta}_{32} = 0 \tag{7.79}$$

As for trusses, the compatibility condition can alternatively be derived using the principle of virtual work. Considering the work of self-equilibrated virtual moments from Figure 7.4(d) on the actual rotation rates, we get

$$\delta M\ \dot{\theta}_{14} + \delta M\ \dot{\theta}_{21} + \delta M\ \dot{\theta}_{32} = 0 \tag{7.80}$$

which yields the compatibility condition (7.79). □

The procedure illustrated by the previous example can be formalized and conveniently written in the matrix notation. The complete kinematic equations for a frame can again be written as $\dot{\boldsymbol{e}} = \boldsymbol{B}\dot{\boldsymbol{d}}$. However, unlike the case of trusses, the generalized strains $\boldsymbol{e}$ can now be divided into two groups – axial extensions, e_{ij}, and bending angles, τ_{ij} and τ_{ji}. Let us formally rewrite the kinematic equations as

$$\dot{\boldsymbol{e}}_a = \boldsymbol{B}_a\dot{\boldsymbol{d}} \tag{7.81}$$
$$\dot{\boldsymbol{e}}_b = \boldsymbol{B}_b\dot{\boldsymbol{d}} \tag{7.82}$$

where vector $\boldsymbol{e}_a$ collects the axial extensions, and vector $\boldsymbol{e}_b$ consists of the bending deformations.

In the standard theory of yield hinges that neglects the effect of normal forces, the axial extension rates during collapse are zero. The rates of bending angles during collapse correspond to the rotation rates in the plastic hinges. This means that a

kinematically admissible mechanism is described by displacement rates $\dot{\boldsymbol{d}}$ and hinge rotation rates $\dot{\boldsymbol{e}}_b$ satisfying conditions

$$\boldsymbol{B}_a\dot{\boldsymbol{d}} = \boldsymbol{O}\dot{\boldsymbol{e}}_b \tag{7.83}$$

$$\boldsymbol{B}_b\dot{\boldsymbol{d}} = \boldsymbol{I}\dot{\boldsymbol{e}}_b \tag{7.84}$$

The conditions of inextensibility (7.83) are used when reducing the number of generalized displacements from the full set containing all the unknown joint displacements and rotations to the set used by the simplified displacement method. In general, the reduction can be based on Gauss elimination. For rectangular frames it is usually possible to make use of the particularly simple form of the inextensibility conditions, and express appropriately selected dependent generalized displacements directly, i.e. without any Gauss transformations. This has been demonstrated in the preceding example. For general frames, we apply the transformation procedure shown in Figure 7.6, which yields expressions for the displacement rates in terms of the hinge rotation rates,

$$\dot{\boldsymbol{d}} = \boldsymbol{P}^T\dot{\boldsymbol{e}}_b \tag{7.85}$$

and equations of compatibility,

$$\boldsymbol{S}^T\dot{\boldsymbol{e}}_b = \boldsymbol{0} \tag{7.86}$$

The only difference compared to trusses is that, on the right-hand side, we do not start with a unit matrix but with a matrix consisting of a zero block and a unit block; compare Figure 7.6 with Figure 7.2.

Example 7.4: Construct the compatibility conditions for the frame from Figure 7.7(a) by the systematic approach, starting from the kinematic equations.

Solution: Before starting the general procedure, we inspect the structure and realize that member 2–5 is a statically determinate cantilever that can be solved by elementary structural analysis. It follows from equilibrium that $M_{25} = 0.75FL$.

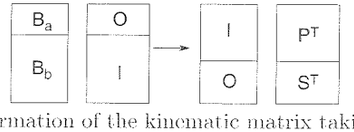


Figure 7.6 Gauss transformation of the kinematic matrix taking into account constraints due to axial inextensibility

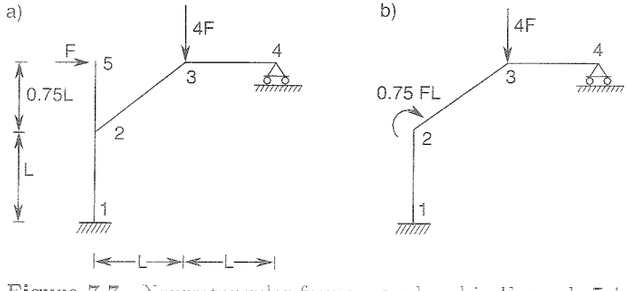


Figure 7.7 Nonrectangular frame – analyzed in Example 7.4

and so we can replace the original frame by the simpler one shown in Figure 7.7(b), in which member 2–5 has been left out and its action on the remaining structure has been replaced by an external moment $M_2 = -0.75FL$.

In the standard displacement method, we would introduce seven unknown generalized displacements, u_2, w_2, ϕ_2, u_3, w_3, ϕ_3 and u_4. The structure has five critical sections, 12, 21, 23, 32 and 34.[1] Based on the general formulae (7.53)–(7.56), the kinematic equations relating the generalized displacements and the generalized strains are set up easily. We present them in the format already suggested by (7.83)–(7.84):

$$\left[\begin{array}{ccccccc} 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -0.8 & 0.6 & 0 & 0.8 & -0.6 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ -0.48 & -0.64 & 1 & 0.48 & 0.64 & 0 & 0 \\ -0.48 & -0.64 & 0 & 0.48 & 0.64 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \end{array}\right] \left\{\begin{array}{c} \dot{u}_2/L \\ \dot{w}_2/L \\ \dot{\phi}_2 \\ \dot{u}_3/L \\ \dot{w}_3/L \\ \dot{\phi}_3 \\ \dot{u}_4/L \end{array}\right\} = \left[\begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \left\{\begin{array}{c} \dot{\theta}_{12} \\ \dot{\theta}_{21} \\ \dot{\theta}_{23} \\ \dot{\theta}_{32} \\ \dot{\theta}_{34} \end{array}\right\} \tag{7.87}$$

The horizontal lines separate the inextensibility conditions (7.83) from the kinematic equations for bending angles (7.84). Note that we have written the equations in terms of displacement rates divided by a fixed length L (such as $\dot{u}_2/L$) in order to obtain a matrix with dimensionless coefficients.

Applying Gauss elimination with pivoting, we can transform (7.87) into

$$\left[\begin{array}{ccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \left\{\begin{array}{c} \dot{u}_2/L \\ \dot{w}_2/L \\ \dot{\phi}_2 \\ \dot{u}_3/L \\ \dot{w}_3/L \\ \dot{\phi}_3 \\ \dot{u}_4/L \end{array}\right\} = \left[\begin{array}{ccccc} 1 & 0 & 0 & 0.000 & 0.000 \\ 0 & 0 & 0 & 0.000 & 0.000 \\ -1 & 1 & 0 & 0.000 & 0.000 \\ 1 & 0 & 0 & 0.375 & -0.375 \\ 0 & 0 & 0 & 0.500 & -0.500 \\ 0 & 0 & 0 & 0.500 & 0.500 \\ 1 & 0 & 0 & 0.375 & -0.375 \\ \hline 1 & -1 & 1 & -0.500 & 0.500 \end{array}\right] \left\{\begin{array}{c} \dot{\theta}_{12} \\ \dot{\theta}_{21} \\ \dot{\theta}_{23} \\ \dot{\theta}_{32} \\ \dot{\theta}_{34} \end{array}\right\} \tag{7.88}$$

The expressions for generalized displacements in terms of the generalized strains (7.85) are now separated by horizontal lines from the compatibility condition (7.86). Any set of rotation rates satisfying the compatibility condition

$$\dot{\theta}_{12} - \dot{\theta}_{21} + \dot{\theta}_{23} - 0.5\dot{\theta}_{32} + 0.5\dot{\theta}_{34} = 0 \tag{7.89}$$

describes a potential failure mechanism. To be kinematically admissible, the mechanism must not only satisfy compatibility but also lead to a positive external power. □

Now we are ready to demonstrate limit analysis of frames by linear programming.

[1] Of course, it would be sufficient to consider only one of sections 32 and 34 as critical. On the other hand, sections 21 and 23 have to be considered separately, as there is an external moment applied on joint 2. In a systematic analysis it is convenient to treat all end sections that are rigidly connected to the corresponding joints as critical, and this is what we will do in the present example.

Example 7.5: Formulate the kinematic approach to limit analysis of the frame from Figure 7.4(a) as a problem of linear programming in the standard form, and solve this problem using the simplex method. All members have the same plastic limit moment, M_0.

Solution: Any kinematically admissible collapse mechanism is described by the rotation rates $\dot{\theta}_{14}$, $\dot{\theta}_{21}$ and $\dot{\theta}_{32}$ satisfying the compatibility condition (7.79). Let us introduce a dimensionless load multiplier $\mu = FL/M_0$, so that the reference load (corresponding to $\mu = 1$) consists of a horizontal force M_0/L and a vertical force $3M_0/L$. The normalizing condition

$$\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} = \frac{M_0}{L}\dot{u}_1 + \frac{3M_0}{L}\dot{w}_2 = 1 \tag{7.90}$$

can be (after a substitution from (7.77) and (7.78)) expressed in terms of the rotation rates as

$$\frac{M_0}{L}(2L\dot{\theta}_{14} + 0.667L\dot{\theta}_{21}) + \frac{3M_0}{L}0.667L\dot{\theta}_{21} = 1 \tag{7.91}$$

or

$$M_0(2\dot{\theta}_{14} + 2.667\dot{\theta}_{21}) = 1 \tag{7.92}$$

Under this condition, the kinematically admissible load multiplier associated with the mechanism is given by

$$\mu_k = M_0|\dot{\theta}_{14}| + M_0|\dot{\theta}_{21}| + M_0|\dot{\theta}_{32}| \tag{7.93}$$

It is convenient to replace the normalizing condition (7.92) by

$$2\dot{\theta}_{14} + 2.667\dot{\theta}_{21} = 1 \tag{7.94}$$

after which the expression for the kinematically admissible load multiplier (7.93) simplifies to

$$\mu_k = |\dot{\theta}_{14}| + |\dot{\theta}_{21}| + |\dot{\theta}_{32}| \tag{7.95}$$

As usual, we have to split the generalized strains into the positive and negative parts in order to replace the absolute values by linear expressions. The resulting problem of linear programming is presented in Table 7.4. The first row of the simplex tableau represents the compatibility condition (7.79), the second row is the normalizing condition (7.94), and the last row defines the objective function (7.95).

The initial choice of basic variables can be guided by their physical meaning. Any feasible solution of the linear programming problem corresponds to a certain collapse

Table 7.4 Initial simplex tableau (not feasible)

$-\mu_k$	$\dot{\theta}_{14}^+$	$\dot{\theta}_{21}^+$	$\dot{\theta}_{32}^+$	$\dot{\theta}_{14}^-$	$\dot{\theta}_{21}^-$	$\dot{\theta}_{32}^-$	RHS
0	1.000	1.000	1.000	−1.000	−1.000	−1.000	0.000
0	2.000	2.667	0.000	−2.000	−2.667	0.000	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	0.000

Table 7.5 Initial feasible solution

$-\mu_k$	$\dot{\theta}_{14}^+$	$\dot{\theta}_{21}^+$	$\dot{\theta}_{32}^+$	$\dot{\theta}_{14}^-$	$\dot{\theta}_{21}^-$	$\dot{\theta}_{32}^-$	RHS
0	1.000	1.333	0.000	−1.000	−1.333	0.000	0.500
0	0.000	0.333	−1.000	0.000	−0.333	1.000	0.500
1	0.000	−0.667	2.000	2.000	2.667	0.000	−1.000

Table 7.6 Optimal solution

$-\mu_k$	$\dot{\theta}_{14}^+$	$\dot{\theta}_{21}^+$	$\dot{\theta}_{32}^+$	$\dot{\theta}_{14}^-$	$\dot{\theta}_{21}^-$	$\dot{\theta}_{32}^-$	RHS
0	0.750	1.000	0.000	−0.750	−1.000	0.000	0.375
0	−0.250	0.000	−1.000	0.250	0.000	1.000	0.375
1	0.500	0.000	2.000	1.500	2.000	0.000	−0.750

mechanism. Let us start, for example, with the panel mechanism from Figure 7.4(b). For this mechanism, the rotation rate $\dot{\theta}_{14}$ is positive and the rotation rate $\dot{\theta}_{32}$ is negative. We therefore select $\dot{\theta}_{14}^+$ and $\dot{\theta}_{32}^-$ as the initial basic variables. An appropriate transformation of the tableau then provides a feasible solution shown in Table 7.5 with a kinematically admissible value $\mu_k = 1$.

The presence of a negative coefficient −0.667 in the last row indicates that this solution is not optimal. After another simplex transformation, during which $\dot{\theta}_{14}^+$ is replaced in the basis by $\dot{\theta}_{21}^+$, we arrive at the optimal solution in Table 7.6. Physically this means that we move the hinge from joint 1 to joint 2. The minimal value of the kinematically admissible multiplier is $\mu_0 = \min \mu_k = 0.75$, which corresponds to the collapse load $F_0 = 0.75M_0/L$. Similar to trusses, the values in the last row of the simplex tableau have the meaning of elastic reserves (this will be proven in Section 7.6). So we can find the moments at all the critical sections directly from the final tableau. The value 0.5 in the column under $\dot{\theta}_{14}^+$ indicates that $M_0 - M_{14} = 0.5M_0$, i.e. $M_{14} = 0.5M_0$. The same result could be derived from the value 1.5 in the column under $\dot{\theta}_{14}^-$: $M_{14} - (-M_0) = 1.5M_0$, i.e. $M_{14} = 0.5M_0$. The collapse mechanism and the corresponding loading and bending moments are shown in Figure 7.8.

7.5 STATIC APPROACH

Let us now explore the linear programming formulation of limit analysis based on the static approach. The lower bound theorem states that the safety factor μ_0 can be obtained as the largest statically admissible multiplier. A statically admissible state

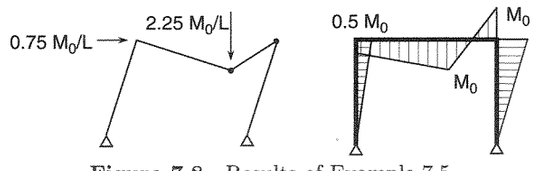


Figure 7.8 Results of Example 7.5

is characterized by a load multiplier μ_s and a vector of internal forces $\boldsymbol{s}$ satisfying the conditions of equilibrium and plastic admissibility. The corresponding linear programming problem can be stated as

$$\text{Maximize } f(\boldsymbol{s}, \mu_s) \equiv \mu_s \tag{7.96}$$

subject to

$$\boldsymbol{B}^T \boldsymbol{s} = \mu_s \bar{\boldsymbol{f}} \tag{7.97}$$

$$-\boldsymbol{s}_0 \leq \boldsymbol{s} \leq \boldsymbol{s}_0 \tag{7.98}$$

To convert this problem to the standard form, we have to eliminate the inequality constraints and make sure that all the unknown variables are nonnegative. The most straightforward way to do this is to introduce new variables

$$\boldsymbol{s}_p = \boldsymbol{s}_0 - \boldsymbol{s} \tag{7.99}$$

$$\boldsymbol{s}_n = \boldsymbol{s}_0 + \boldsymbol{s} \tag{7.100}$$

which have the physical meaning of positive and negative elastic reserves; see Figure 7.3. Clearly, the conditions $\boldsymbol{s} \leq \boldsymbol{s}_0$ and $\boldsymbol{s} \geq -\boldsymbol{s}_0$ are, respectively, equivalent to $\boldsymbol{s}_p \geq \boldsymbol{0}$ and $\boldsymbol{s}_n \geq \boldsymbol{0}$.

We have eliminated the inequality constraints and at the same time obtained variables that are nonnegative. Of course, $\boldsymbol{s}_p$ and $\boldsymbol{s}_n$ cannot be totally independent as $\boldsymbol{s}$ can be expressed alternatively in terms of $\boldsymbol{s}_p$ or $\boldsymbol{s}_n$, and both expressions must lead to the same result. Summing (7.99) and (7.100) we get a condition linking $\boldsymbol{s}_p$ and $\boldsymbol{s}_n$,

$$\boldsymbol{s}_p + \boldsymbol{s}_n = 2\boldsymbol{s}_0 \tag{7.101}$$

This condition also naturally follows from the physical meaning of $\boldsymbol{s}_p$ and $\boldsymbol{s}_n$ as elastic reserves; see Figure 7.3. Subtracting (7.99) from (7.100) we obtain a 'symmetric' expression for $\boldsymbol{s}$ in terms of the new variables,

$$\boldsymbol{s} = \frac{1}{2}(\boldsymbol{s}_n - \boldsymbol{s}_p) \tag{7.102}$$

This relation can be substituted into (7.97) in order to completely eliminate the 'old' variables $\boldsymbol{s}$ from the formulation. The resulting linear programming problem

$$\text{Minimize } \hat{f}(\boldsymbol{s}_p, \boldsymbol{s}_n, \mu_s) \equiv -\mu_s \tag{7.103}$$

subject to

$$-\boldsymbol{B}^T \boldsymbol{s}_p + \boldsymbol{B}^T \boldsymbol{s}_n - 2\bar{\boldsymbol{f}}\mu_s = \boldsymbol{0} \tag{7.104}$$

$$\boldsymbol{s}_p + \boldsymbol{s}_n = 2\boldsymbol{s}_0 \tag{7.105}$$

$$\boldsymbol{s}_p \geq \boldsymbol{0}, \quad \boldsymbol{s}_n \geq \boldsymbol{0}, \quad \mu_s \geq 0 \tag{7.106}$$

is in the standard form. Note that we have added the condition that the load multiplier μ_s be nonnegative, which was not present in the original formulation. However, we know that the maximum statically admissible load multiplier is positive due to its physical meaning, and so this additional constraint (required by the definition of the standard form) does not alter the solution of the problem.

The transformed problem (7.103)–(7.106) is in the standard form, but its initial basic feasible solution is not easy to construct (see Problem 7.3). We could use the starting procedure mentioned in Appendix C that can generate a basic feasible solution for any linear programming problem in the standard form. This procedure introduces additional variables and uses a modified cost function to drive them out of the basis.

We prefer to follow an alternative approach, in which all the variables have a clear physical meaning, and no auxiliary objective function has to be introduced. Instead of introducing the elastic reserves $\boldsymbol{s}_p$ and $\boldsymbol{s}_n$, we decompose the internal forces $\boldsymbol{s}$ into their positive and negative parts. In terms of new variables

$$\boldsymbol{s}^+ = (|\boldsymbol{s}| + \boldsymbol{s})/2, \qquad \boldsymbol{s}^- = (|\boldsymbol{s}| - \boldsymbol{s})/2 \tag{7.107}$$

we can express $\boldsymbol{s}$ and its absolute value as

$$\boldsymbol{s} = \boldsymbol{s}^+ - \boldsymbol{s}^-, \qquad |\boldsymbol{s}| = \boldsymbol{s}^+ + \boldsymbol{s}^- \tag{7.108}$$

and transform the static equations (7.97) into

$$\boldsymbol{B}^T \boldsymbol{s}^+ - \boldsymbol{B}^T \boldsymbol{s}^- - \bar{\boldsymbol{f}}\mu_s = \boldsymbol{0} \tag{7.109}$$

Note that the conditions of plastic admissibility (7.98) are fully equivalent with $|\boldsymbol{s}| \leq \boldsymbol{s}_0$, and so we can rewrite them as

$$\boldsymbol{s}^+ + \boldsymbol{s}^- \leq \boldsymbol{s}_0 \tag{7.110}$$

To transform the inequality constraints (7.110) into equality constraints as required by the standard form, we now introduce a vector of *slack variables*

$$\boldsymbol{\Delta s} = \boldsymbol{s}_0 - (\boldsymbol{s}^+ + \boldsymbol{s}^-) \tag{7.111}$$

Physically, the slack variable ΔS_i associated with a certain internal force S_i represents the 'smaller elastic reserve', i.e. it equals the positive reserve $S_0 - S_i$ if S_i is positive, and the negative reserve $S_i + S_0$ if S_i is negative. The basic unknowns of the problem now consist of the variables $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$ that have replaced $\boldsymbol{s}$, and of the slack variables, $\boldsymbol{\Delta s}$. Note that conditions (7.110) are equivalent to $\boldsymbol{\Delta s} \geq \boldsymbol{0}$, and that the variables $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$ are nonnegative due to their definition (7.107). Again, the condition $\mu_s \geq 0$ can be added without altering the solution.

In summary, we have the following linear programming problem in the standard form:

$$\text{Minimize } \tilde{f}(\boldsymbol{s}^+, \boldsymbol{s}^-, \boldsymbol{\Delta s}, \mu_s) \equiv -\mu_s \tag{7.112}$$

subject to

$$\boldsymbol{B}^T \boldsymbol{s}^+ - \boldsymbol{B}^T \boldsymbol{s}^- - \bar{\boldsymbol{f}}\mu_s = \boldsymbol{0} \tag{7.113}$$

$$\boldsymbol{s}^+ + \boldsymbol{s}^- + \boldsymbol{\Delta s} = \boldsymbol{s}_0 \tag{7.114}$$

$$\boldsymbol{s}^+ \geq \mathbf{0}, \quad \boldsymbol{s}^- \geq \mathbf{0}, \quad \boldsymbol{\Delta s} \geq \mathbf{0}, \quad \mu_s \geq 0 \tag{7.115}$$

The main advantage of this formulation is that, due to our insight into the underlying physical problem, we are able to systematically select a basis that provides a feasible solution. We will first demonstrate the basic idea by a simple example, and then describe the procedure in general terms.

Example 7.6: Formulate the static approach to limit analysis of the truss in Figure 7.1 as a problem of linear programming in the standard form, and solve it using the simplex method. The truss is subjected to an increasing horizontal load.

Solution: The truss is reproduced in Figure 7.9(a). The equations of equilibrium have already appeared in several examples; e.g. see Example 3.1. We can easily set up the linear programming problem in the 'natural' form (7.96)–(7.98) with maximization of μ_S replaced by minimization of $-\mu_S$:

$$\text{Minimize } f^*(S_1, S_2, S_3, F_s) \equiv -F_s \tag{7.116}$$

subject to

$$\begin{bmatrix} 0.6 & 0 & -0.8 \\ 0.8 & 1 & 0.6 \end{bmatrix} \begin{Bmatrix} S_1 \\ S_2 \\ S_3 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} F_s \tag{7.117}$$

$$-\begin{Bmatrix} S_0 \\ S_0 \\ S_0 \end{Bmatrix} \leq \begin{Bmatrix} S_1 \\ S_2 \\ S_3 \end{Bmatrix} \leq \begin{Bmatrix} S_0 \\ S_0 \\ S_0 \end{Bmatrix} \tag{7.118}$$

Introducing new variables $S_i^+ = (|S_i| + S_i)/2$, $S_i^- = (|S_i| - S_i)/2$, and $\Delta S_i = S_0 - |S_i|$, $i = 1, 2, 3$, we can replace the constraints by

$$\begin{bmatrix} 0.6 & 0 & -0.8 \\ 0.8 & 1 & 0.6 \end{bmatrix} \begin{Bmatrix} S_1^+ \\ S_2^+ \\ S_3^+ \end{Bmatrix} - \begin{bmatrix} 0.6 & 0 & -0.8 \\ 0.8 & 1 & 0.6 \end{bmatrix} \begin{Bmatrix} S_1^- \\ S_2^- \\ S_3^- \end{Bmatrix} - \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} F_s = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{7.119}$$

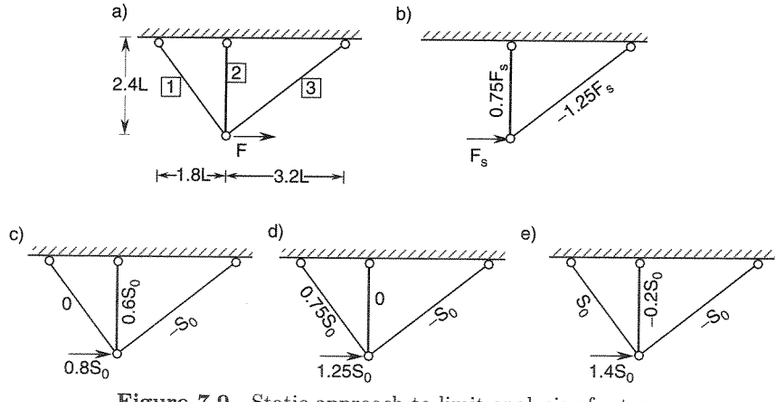


Figure 7.9 Static approach to limit analysis of a truss

Table 7.7 Initial simplex tableau (not feasible)

S_1^+	S_2^+	S_3^+	S_1^-	S_2^-	S_3^-	ΔS_1	ΔS_2	ΔS_3	F_s	RHS/S_0
0.6	0.0	−0.8	−0.6	0.0	0.8	0	0	0	−1	0
0.8	1.0	0.6	−0.8	−1.0	−0.6	0	0	0	0	0
1.0	0.0	0.0	1.0	0.0	0.0	1	0	0	0	1
0.0	1.0	0.0	0.0	1.0	0.0	0	1	0	0	1
0.0	0.0	1.0	0.0	0.0	1.0	0	0	1	0	1
0.0	0.0	0.0	0.0	0.0	0.0	0	0	0	−1	0

$$\begin{Bmatrix} S_1^+ \\ S_2^+ \\ S_3^+ \end{Bmatrix} + \begin{Bmatrix} S_1^- \\ S_2^- \\ S_3^- \end{Bmatrix} + \begin{Bmatrix} \Delta S_1 \\ \Delta S_2 \\ \Delta S_3 \end{Bmatrix} = \begin{Bmatrix} S_0 \\ S_0 \\ S_0 \end{Bmatrix} \tag{7.120}$$

and by the conditions that all variables be nonnegative. The initial simplex tableau corresponding to this linear programming problem in the standard form is shown in Table 7.7. The tableau has been simplified by leaving out the leftmost column, which does not carry any essential information and remains the same during all transformations.

Now comes the crucial step – we have to find a basic feasible solution. Recall that the aim is to select the basic variables such that, when we set the remaining (nonbasic) variables equal to zero, the basic variables can be uniquely computed from the constraints, and their values become nonnegative. In our particular problem, we have 10 variables and five constraints, and therefore we need to set $10-5 = 5$ variables to zero.

Let us start by examining the first two constraints that represent the equations of equilibrium. They are satisfied by the axial forces evaluated on any primary structure (obtained from the actual structure by leaving out one statically redundant bar). Consider, for example, the primary structure obtained by leaving out bar number 1; see Figure 7.9(b). The axial forces in the remaining bars are easily evaluated as $S_2 = 0.75F_s$ and $S_3 = -1.25F_s$. To end up with positive values of the computed basic variables, we select S_2^+ and S_3^- as the first two members of the basis while the variables S_1^+, S_3^+, S_1^- and S_2^- stay out of the basis.

Now we have to specify additional three basic variables. Clearly, the load multiplier will definitely be a member of the final basis, and so we include it right away in the initial basis. Consequently, one of the slack variables must remain out of the initial basis, which means that its value is set to zero. In physical terms, one of the axial forces must just reach the yield limit, while the two remaining ones stay below the yield limit. In our case, where the plastic axial force is the same for all the bars, it is clear that the force reaching the plastic limit must be that with the largest absolute value of the coefficient at F_s, i.e. the force S_3. For $F_s = S_0/1.25 = 0.8S_0$ we have $S_3 = -1.25F_s = -S_0$, and the associated slack variable $\Delta S_3 = S_0 - |S_3|$ is equal to zero. The remaining axial forces $S_1 = 0$ and $S_2 = 0.75F_s = 0.6S_0$ stay below the yield limit, and the associated slack variables $\Delta S_1 = S_0 - |S_1| = S_0$ and $\Delta S_2 = S_0 - |S_2| = 0.4S_0$ are positive.

The previous considerations lead to the initial basis consisting of S_2^+, S_3^-, ΔS_1, ΔS_2, and F_s. Indeed, Table 7.8 indicates that the corresponding simplex tableau provides

Table 7.8 Initial feasible solution

S_1^+	S_2^+	S_3^+	S_1^-	S_2^-	S_3^-	ΔS_1	ΔS_2	ΔS_3	F_s	RHS/S_0
−0.6	0.0	1.6	0.6	0.0	0.0	0.0	0.0	0.8	1.0	0.8
0.8	1.0	1.2	−0.8	−1.0	0.0	0.0	0.0	0.6	0.0	0.6
1.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0
−0.8	0.0	−1.2	0.8	2.0	0.0	0.0	1.0	−0.6	0.0	0.4
0.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	1.0
−0.6	0.0	1.6	0.6	0.0	0.0	0.0	0.0	0.8	0.0	0.8

Table 7.9 Improved feasible solution

S_1^+	S_2^+	S_3^+	S_1^-	S_2^-	S_3^-	ΔS_1	ΔS_2	ΔS_3	F_s	RHS/S_0
0.0	0.75	2.5	0.0	−0.75	0.0	0.0	0.0	1.25	1.0	1.25
1.0	1.25	1.5	−1.0	−1.25	0.0	0.0	0.0	0.75	0.0	0.75
0.0	−1.25	−1.5	2.0	1.25	0.0	1.0	0.0	−0.75	0.0	0.25
0.0	1.00	0.0	0.0	1.00	0.0	0.0	1.0	0.00	0.0	1.00
0.0	0.00	1.0	0.0	0.00	1.0	0.0	0.0	1.00	0.0	1.00
0.0	0.75	2.5	0.0	−0.75	0.0	0.0	0.0	1.25	0.0	1.25

a feasible solution. Note that the basic variables have values that we predicted when constructing the initial basis. The solution corresponds to a statically admissible state $S_1 = 0$, $S_2 = 0.6S_0$, $S_3 = -S_0$, and $F_s = 0.8S_0$; see Figure 7.9(c).

After the first cycle of the simplex method, the tableau is transformed to that in Table 7.9. Note that S_1^+ entered the basis and S_2^+ left it. The solution now corresponds to the statically admissible state $S_1 = 0.75S_0$, $S_2 = 0$, $S_3 = -S_0$ and $F_s = 1.25S_0$; see Figure 7.9(d). The last row of the tableau indicates that the solution is not yet optimal. This is also clear from the fact that only one of the axial forces is at the yield limit.

In the second cycle of the simplex method, S_2^- enters the basis and ΔS_1 leaves it. This necessarily means that we arrive at the collapse state, because two axial forces will be at the yield limit, and this generates a collapse mechanism. Indeed, the last row of the simplex tableau in Table 7.10 contains only nonnegative terms, and the corresponding statically admissible state $S_1 = S_0$, $S_2 = -0.2S_0$, $S_3 = -S_0$, and $F_s = 1.4S_0$ is the actual state at collapse; see Figure 7.9(e). □

The strategy used in the previous example when constructing the initial basic feasible solution is quite general. It can be applied to any linear programming problem resulting from the static approach to limit analysis and converted to the standard form (7.112)–(7.115). In physical terms, the basic idea can be described as follows:

1. Select an appropriate number of statically redundant internal forces and construct the corresponding primary structure.
2. Determine the values of the statically determinate internal forces on the primary structure subjected to the given reference loading.

Table 7.10 Optimal solution

S_1^+	S_2^+	S_3^+	S_1^-	S_2^-	S_3^-	ΔS_1	ΔS_2	ΔS_3	F_s	RHS/S_0
0.0	0.0	1.6	1.2	0.0	0.0	0.6	0.0	0.8	1.0	1.4
1.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0
0.0	−1.0	−1.2	1.6	1.0	0.0	0.8	0.0	−0.6	0.0	0.2
0.0	2.0	1.2	−1.6	0.0	0.0	−0.8	1.0	0.6	0.0	0.8
0.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	1.0
0.0	0.0	1.6	1.2	0.0	0.0	0.6	0.0	0.8	0.0	1.4

3. Find the largest possible value of the load multiplier for which all the internal forces computed on the primary structure are plastically admissible. At least one internal force satisfies the yield condition.
4. The initial basis consists of
   (a) the load multiplier,
   (b) the positive parts of those internal forces on the primary structure that are positive under the reference loading,
   (c) the negative parts of those internal forces on the primary structure that are negative under the reference loading, and
   (d) the slack variables (elastic reserves) associated with all the internal forces except for the one that satisfies the yield condition.

The LP formulation resulting from the static approach was illustrated by an example of a truss, but all the general relations and procedures are directly applicable to beams and frames. If the structural model has n unknown generalized displacements and m generalized strains, the number of internal forces is also m, and the redundancy is $s = m - n$. The total number of variables in the linear programming problem (7.112)–(7.115) is $3m + 1$, and the number of constraints is $n + m$. The initial basis provided by the above algorithm has $1+n+m-1 = n+m$ members, which is exactly the number of constraints.

Compared to problem (7.34)–(7.37), resulting from the kinematic approach after elimination of the generalized displacements, which has only $2m$ variables and $m-n+1$ constraints, the static approach is substantially more expensive. However, later it will turn out that the static approach has certain advantages in optimum design, to be discussed in Chapter 12. After a simple modification, the strategy leading to the initial basic feasible solution will be applicable to optimum design problems as well.

7.6 DUALITY

It is interesting to observe the close relationship between the problems of linear programming (7.7)–(7.11) and (7.96)–(7.98), resulting from the kinematic and static approaches. The link between the two formulations is obvious from Table 7.11.

In the horizontal direction, we can read the kinematic formulation. For this purpose, the leftmost column can be ignored. The top row lists the variables, the rows in the central part of the tableau represent the kinematic equations and the normalizing

Table 7.11 Duality of the static and kinematic approach

	$\dot{e}^+$	$\dot{e}^-$	$\dot{d}$		
s	I	$-I$	$-B$	$=$	0
μ_s	0^T	0^T	$\bar{f}^T$	$=$	1
	$\leq$	$\leq$	$=$		
	s_0^T	s_0^T	0^T		

condition of unit external power, and the bottom row displays the coefficients of the objective function (to be minimized).

In the vertical direction, we can read the static formulation. For this purpose, the top row can be ignored. The left column lists the variables, the columns in the central part of the tableau show the conditions of plastic admissibility and of equilibrium, and the right column displays the coefficients of the objective function (to be maximized).

The above relationship between the static and kinematic approach is called *duality*, and is studied in its general form by the theory of linear programming. One of the problems is called the *primal problem*, and the other is the *dual problem*. Note that the variables of the primal problem correspond to the constraints of the dual problem, and vice versa. A nonnegative variable always corresponds to an inequality constraint while a free (i.e. not necessarily nonnegative) variable always corresponds to an equality constraint. The reader can verify this property in Table 7.11 as an exercise. Also, the coefficients of the objective function of the primal problem represent the right-hand side of the dual problem, and vice versa. In one of the problems, the objective function must be minimized, while in the other it must be maximized.

Based on the aforementioned rules, it is possible to construct the dual problem to any linear programming problem (not necessarily in the standard form). The dual variables are at first introduced only as formal symbols, but it is usually possible to identify their physical meaning from the structure of the dual problem.

Example 7.7: Construct the dual problem to (7.34)–(7.37), and identify its physical meaning.

Solution: Let us write the primal problem into Table 7.12(a). All constraints have the form of an equality, and so the dual variables are free.

Let us denote the dual variables associated with constraints (7.35) and (7.36) by $\boldsymbol{u}$ and v, respectively. The primal variables are nonnegative, and so the dual constraints

Table 7.12 a) Primal problem, b) primal problem with its dual

(a)

	$\dot{e}^+$	$\dot{e}^-$		
	S^T	$-S^T$	$=$	0
	p^T	$-p^T$	$=$	1
	s_0^T	s_0^T		

(b)

	$\dot{e}^+$	$\dot{e}^-$		
u	S^T	$-S^T$	$=$	0
v	p^T	$-p^T$	$=$	1
	$\leq$	$\leq$		
	s_0^T	s_0^T		

must be inequalities. This leads us to Table 7.12(b). The dual constraints read

$$\boldsymbol{S}\boldsymbol{u} + \boldsymbol{p}v \leq \boldsymbol{s}_0 \tag{7.121}$$

$$-\boldsymbol{S}\boldsymbol{u} - \boldsymbol{p}v \leq \boldsymbol{s}_0 \tag{7.122}$$

and the dual objective function

$$g(\boldsymbol{u}, v) \equiv \boldsymbol{0}^T\boldsymbol{u} + 1 \times v = v \tag{7.123}$$

must be maximized. It is easy to guess that v is the statically admissible load multiplier, and that constraints (7.121)–(7.122) represent the conditions of plastic admissibility. To confirm this, recall that in Section 7.2 the matrices $\boldsymbol{S}$ and $\boldsymbol{P}$ were constructed such that

$$\boldsymbol{P}^T\boldsymbol{B} = \boldsymbol{I} \tag{7.124}$$

$$\boldsymbol{S}^T\boldsymbol{B} = \boldsymbol{O} \tag{7.125}$$

Also, recall that the vector $\boldsymbol{p}$ was defined as

$$\boldsymbol{p} = \boldsymbol{P}\bar{\boldsymbol{f}} \tag{7.126}$$

Multiplying the expression $\boldsymbol{S}\boldsymbol{u} + \boldsymbol{p}v$ by $\boldsymbol{B}^T$ from the left we obtain

$$\boldsymbol{B}^T(\boldsymbol{S}\boldsymbol{u} + \boldsymbol{p}v) = \boldsymbol{B}^T\boldsymbol{S}\boldsymbol{u} + \boldsymbol{B}^T\boldsymbol{P}\bar{\boldsymbol{f}}v = \boldsymbol{O}\boldsymbol{u} + \boldsymbol{I}\bar{\boldsymbol{f}}v = \bar{\boldsymbol{f}}v \tag{7.127}$$

Recalling the usual form of the equilibrium equations, $\boldsymbol{B}^T\boldsymbol{s} = \mu\bar{\boldsymbol{f}}$, we realize that $\boldsymbol{S}\boldsymbol{u} + \boldsymbol{p}v$ are the internal forces and v is the load multiplier, which confirms that constraints (7.121)–(7.122) indeed represent the conditions of plastic admissibility.

The physical meaning of $\boldsymbol{u}$ is more difficult to find. Note that $\boldsymbol{B}^T(\boldsymbol{S}\boldsymbol{u}) = (\boldsymbol{B}^T\boldsymbol{S})\boldsymbol{u} = \boldsymbol{O}\boldsymbol{u} = \boldsymbol{0}$, and so $\boldsymbol{S}\boldsymbol{u}$ are self-equilibrated internal forces. Also, $\boldsymbol{B}^T\boldsymbol{p} = \boldsymbol{B}^T\boldsymbol{P}\bar{\boldsymbol{f}} = \boldsymbol{I}\bar{\boldsymbol{f}} = \bar{\boldsymbol{f}}$, and so $\boldsymbol{p}$ are internal forces in equilibrium with the reference loading $\bar{\boldsymbol{f}}$. The size of $\boldsymbol{u}$ is equal to the number of independent compatibility conditions, i.e. to the redundancy of the structure. Therefore, the components of $\boldsymbol{u}$ are generalized redundants. □

Duality is exploited for example by the so-called dual simplex method, which does not require an initial feasible solution as a starting point. For our purpose, it is important that the dual formulation helps one to identify the meaning of the coefficients in the last row of the transformed simplex tableau. Consider a linear programming problem in the standard form

$$\text{Minimize } f(\boldsymbol{x}) \equiv \boldsymbol{c}^T\boldsymbol{x} \tag{7.128}$$

subject to

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{7.129}$$

$$\boldsymbol{x} \geq \boldsymbol{0} \tag{7.130}$$

and its dual problem

$$\text{Maximize } g(\boldsymbol{y}) \equiv \boldsymbol{b}^T\boldsymbol{y} \tag{7.131}$$

Table 7.13 Partitioning of the LPP in standard form and of its dual

	$\boldsymbol{x}_B$	$\boldsymbol{x}_N$		
$\boldsymbol{y}$	$\boldsymbol{A}_B$	$\boldsymbol{A}_N$	$=$	$\boldsymbol{b}$
	$\leq$	$\leq$		
	$\boldsymbol{c}_B^T$	$\boldsymbol{c}_N^T$		

subject to

$$\boldsymbol{A}^T\boldsymbol{y} \leq \boldsymbol{c} \tag{7.132}$$

Consider a given choice of basic variables $\boldsymbol{x}_B$ and nonbasic variables $\boldsymbol{x}_N$; the corresponding partitioning of the problem is schematically shown in Table 7.13. According to (C.10) in Appendix C, the objective function can be rewritten in the form

$$f_N(\boldsymbol{x}_N) = \boldsymbol{c}_B^T\boldsymbol{A}_B^{-1}\boldsymbol{b} + (\boldsymbol{c}_N - \boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B)^T\boldsymbol{x}_N \tag{7.133}$$

Suppose that the given choice of basic variables leads to a basic feasible solution of the primal problem. The solution is optimal if all the coefficients at the nonbasic variables are nonnegative, i.e. if

$$\boldsymbol{c}_N - \boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B \geq \boldsymbol{0} \tag{7.134}$$

From Appendix C we know that the left-hand side of (7.134) appears in the bottom row of the transformed simplex tableau; cf. Table D.1.

For the same basis we can rewrite the dual constraints as

$$\boldsymbol{A}_B^T\boldsymbol{y} \leq \boldsymbol{c}_B \tag{7.135}$$

$$\boldsymbol{A}_N^T\boldsymbol{y} \leq \boldsymbol{c}_N \tag{7.136}$$

Let us construct a vector $\boldsymbol{y}$ such that the 'basic constraints' (7.135) are satisfied as equalities, i.e. let $\boldsymbol{y} = \boldsymbol{A}_B^{-T}\boldsymbol{c}_B$. This vector represents a feasible solution of the dual problem if it satisfies the 'nonbasic' constraints (7.136), i.e. if

$$\boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B \leq \boldsymbol{c}_N \tag{7.137}$$

Comparing (7.137) to (7.134) we realize that a basic feasible solution of the primal problem is optimal if and only if the corresponding dual solution is dual feasible (i.e. satisfies the constraints of the dual problem). Moreover, each coefficient in the bottom row of the transformed simplex tableau is equal to the difference between the right-hand side and the left-hand side in the corresponding dual inequality constraint.

Example 7.8: Explain the physical meaning of the coefficients in the last row of a simplex tableau obtained by simplex transformations of problem (7.34)–(7.37).

Solution: As explained in Example 7.7, the constraints of the problem dual to (7.34)–(7.37) can be interpreted as the conditions of plastic admissibility. Each primal variable (element of $\dot{\boldsymbol{e}}^+$ or $\dot{\boldsymbol{e}}^-$) represents the positive or negative part of a certain deformation rate (plastic bar extension rate, or plastic hinge rotation rate). The

corresponding dual constraint (one of constraints (7.121) or (7.122)) is the 'positive' or 'negative' condition of plastic admissibility, e.g. $M_i \leq M_0$ or $-M_i \leq M_0$ in the case of a plastic hinge. It has been shown that the coefficients in the last row of the simplex tableau are equal to the difference between the right-hand side and the left-hand side of the dual constraint. In our case, this means that these coefficients are the positive and negative elastic reserves, $M_0 - M_i$ and $M_0 + M_i$. By specifying a basis for the primal solution we select a certain assumed collapse mechanism. In terms of the dual formulation, this means that we specify the values of the internal forces in the (assumed) plastic bars or hinges to be at the yield limit, and we compute the internal forces in the remaining bars or hinges from equilibrium. If all of the computed internal forces are plastically admissible, the solution is not only kinematically admissible but also statically admissible, and we thus have the actual collapse mechanism. If one or more conditions of plastic admissibility are violated, the solution is only kinematically admissible but not statically admissible, and we then must modify the basis. According to the standard simplex algorithm, the new member of the basis will be the variable for which the negative coefficient in the last row has the largest absolute value. In physical terms, this means that we create a new plastic hinge at the critical cross section in which the bending moment exceeds the plastic moment by the largest amount. Of course, another plastic hinge must be removed in order to keep a collapse mechanism with one degree of freedom. The hinge to be removed is automatically selected by the procedure looking for the key row in the simplex algorithm. □

The relations explained in the preceding example find their most important application in the evaluation of the actual internal forces at collapse. We have already exploited them (without any proof) in Examples 7.2 and 7.5. Let us now explain the general procedure. Suppose that we find the collapse mechanism by the simplex method applied to the kinematic formulation. For the variables representing the positive parts of deformation rates, the coefficients in the last row of the transformed simplex tableau are the positive elastic reserves $M_0 - M_i$ (or $S_0 - S_i$ in the case of bars). Thus, the bending moment M_i is evaluated simply by subtracting this reserve from the plastic moment. The same result would be obtained by adding the negative reserve $M_0 + M_i$ (located in the column that corresponds to the negative part of the same deformation rate) to the negative plastic moment, $-M_0$. The coefficients corresponding to the basic variables are zero, and so the resulting internal forces are at the yield limit (positive or negative). The coefficients corresponding to the nonbasic variables are positive, and so the resulting internal forces are below the yield limit.

The number of arithmetic operations needed to solve a linear programming problem in the standard form by the simplex method is proportional to m^3n, where m is the number of constraints and n is the number of variables. Therefore, the cost of solving the primal problem and the dual problem may be dramatically different. Advanced software tools for linear programming problems should automatically select the formulation that minimizes the computational cost, and should provide the solution of both the primal and the dual problem.

PROBLEMS

Problem 7.1: Set up the linear programming problem resulting from the kinematic approach to limit analysis of the truss with the geometry from Figure 7.1(a) subjected

to a *horizontal* force F. Solve the problem using the simplex method, taking as the initial feasible solution the one that corresponds to the yielding of bars 1 and 2.

Problem 7.2: Set up the linear programming problem resulting from the kinematic approach to limit analysis of the structure in Figure 6.2, and solve it using the simplex method. Explain how the final simplex tableau reflects the fact that the collapse mechanism is a partial one.

Problem 7.3*: Explain why the initial basic feasible solution of (7.103)–(7.106) is not easy to construct.
Hint: Consider two cases: a) the load multiplier is a member of the initial basis, b) the load multiplier is not a member of the initial basis. For each case, imagine a basic solution and find its physical interpretation. What would we have to know in order to select the members of the basis so that the solution be feasible?

Problem 7.4*: Construct the dual problem to (7.112)–(7.115), and detect the physical meaning of the dual variables.

Problem 7.5*: Try to derive (or guess) the meaning of the coefficients in the last row of a simplex tableau corresponding to problem (7.112)–(7.115).

Problem 7.6*: Generalize the kinematic approach to limit analysis for structures in which the plastic capacities depend on the signs of the corresponding deformation rates. Denote the vectors of plastic capacities in 'tension' and in 'compression' as $\boldsymbol{s}_0^+$ and $\boldsymbol{s}_0^-$, respectively.

Problem 7.7*: Solve Example 7.2 for the case $S_0^- = 2S_0^+$.

Problem 7.8*: Generalize the static approach to limit analysis for structures in which the plastic capacities depend on the signs of the corresponding deformation rates. Try to convert the problem to a standard form similar to (7.112)–(7.115).

8

Displacements at Incipient Collapse

The methods of limit analysis explained in the preceding chapters can serve as an efficient tool for determining the load carrying capacity of structures. Compared to incremental analysis, the efficiency is achieved by looking directly at the final state at collapse, without any need to follow the intermediate loading steps. However, the plastic limit modes can be used in design only if the deflections corresponding to the safe states are not prohibitively large, and if the ductility of the structure is not exhausted. One important advantage of the incremental approach is that it provides complete information on the evolution of the displacements and generalized strains throughout the loading process. In limit analysis, it is possible to determine the displacements when the collapse state is just reached, i.e. at *incipient collapse*. This information is useful for determining the serviceability of the structure, as well as for assessing the applicability of the geometrically linear theory.

The displacements at incipient collapse can be evaluated relatively easily if it is known which bar yields last (for trusses), or which plastic hinge is formed last (for beams and frames). We will discuss the problem in the context of truss analysis, but an analogous reasoning can be applied to beams and frames.

8.1 TRUSSES

At incipient collapse, the last bar just begins to yield, and so its plastic extension is still zero and its total extension equals the elastic one, which can be easily determined from Hooke's law. The elastic law can also be applied to all the non-yielding bars as they are still in an elastic state. The number of known bar extensions is then the same as the number of unknown displacements (if we exclude a partial collapse mechanism). Therefore, the displacements can be uniquely solved from the kinematic equations. This procedure is equivalent to analyzing a statically determinate structure consisting of all the non-yielding bars and the bar that yields last, while the other yielding bars are replaced by forces applied at the nodes as external loads. Instead of setting up and solving the kinematic equations, the principle of virtual work can be exploited.

Example 8.1: Find the displacements at incipient collapse for the truss from Example 7.2.

Solution: The collapse force $F_0 = 2.25S_0$ and the collapse mechanism (bars 1 and 2 yielding) have been determined in Example 7.2. The kinematic equations read

$$0.6u + 0.8w = e_1 \tag{8.1}$$

$$w = e_2 \tag{8.2}$$

$$-0.8u + 0.6w = e_3 \tag{8.3}$$

We also know the axial forces at collapse, $S_1 = S_0$, $S_2 = S_0$ and $S_3 = 0.75S_0$. Bar 3 is in an elastic state, and so

$$e_3 = \frac{S_3}{k_3} = 0.75S_0\frac{4L}{EA} = 3\frac{S_0L}{EA} \tag{8.4}$$

1. Suppose that bar 1 yields last. Then

$$e_{p,1} = 0,\ e_1 = e_{e,1} = \frac{S_1}{k_1} = 3\frac{S_0L}{EA} \tag{8.5}$$

and the displacements u and w can be calculated from

$$0.6u + 0.8w = 3\frac{S_0L}{EA} \tag{8.6}$$

$$-0.8u + 0.6w = 3\frac{S_0L}{EA} \tag{8.7}$$

Solving this system of equations, we get

$$u = -0.6\frac{S_0L}{EA} \tag{8.8}$$

$$w = 4.2\frac{S_0L}{EA} \tag{8.9}$$

Alternatively, we can use the principle of virtual work. To determine the horizontal displacement u, consider the truss subjected to a horizontal virtual force δF. According to the principle of virtual work, we have

$$\delta F \times u = \delta S_1 \times e_1 + \delta S_2 \times e_2 + \delta S_3 \times e_3 \tag{8.10}$$

where δS_i, $i = 1, 2, 3$, is any set of virtual axial forces that are in equilibrium with the virtual external force. As the truss is statically indeterminate, there are infinitely many sets of virtual axial forces satisfying this condition, all of them giving the same result on the right-hand side of (8.10). The extensions e_1 and e_3 are known but e_2 is not, and so it is convenient to set $\delta S_2 = 0$ and calculate δS_1 and δS_3 from equilibrium. This is equivalent to solving the statically determinate truss in Figure 8.1(a). Substituting $\delta S_1 = 0.6\,\delta F$, $\delta S_2 = 0$, and $\delta S_3 = -0.8\,\delta F$ into (8.10), we get

$$\delta F \times u = (0.6e_1 + 0 \times e_2 - 0.8e_3)\delta F \tag{8.11}$$

from which

$$u = 0.6e_1 - 0.8e_3 = 0.6\frac{3S_0L}{EA} - 0.8\frac{3S_0L}{EA} = -0.6\frac{S_0L}{EA} \tag{8.12}$$

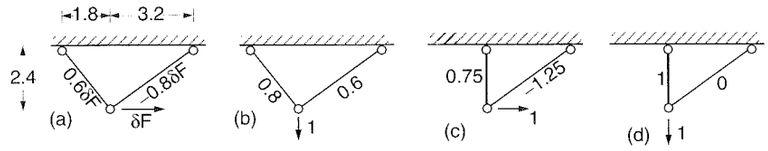


Figure 8.1 Virtual states

Similarly, the vertical displacement w can be obtained by using the virtual state in Figure 8.1(b). Since both sides of the virtual work equality are proportional to δF and this symbol cancels out from the final result, we can immediately replace δF by a unit force. The corresponding axial forces are then $\delta S_1 = 0.8$, $\delta S_2 = 0$ and $\delta S_3 = 0.6$, and the virtual work equality yields

$$1 \times w = 0.8\frac{3S_0L}{EA} + 0.6\frac{3S_0L}{EA} = 4.2\frac{S_0L}{EA} \tag{8.13}$$

2. Now suppose that bar 2 yields last. Then the known extensions at incipient collapse are

$$e_2 = e_{e,2} = \frac{S_2}{k_2} = 2.4\frac{S_0L}{EA} \tag{8.14}$$

$$e_3 = e_{e,3} = \frac{S_3}{k_3} = 3\frac{S_0L}{EA} \tag{8.15}$$

Using the virtual states in Figures 8.1(c) and (d), we obtain

$$u = 0.75e_2 - 1.25e_3 = -1.95\frac{S_0L}{EA} \tag{8.16}$$

$$w = 1 \times e_2 - 0 \times e_3 = 2.4\frac{S_0L}{EA} \tag{8.17}$$

Therefore, the vertical displacement under the applied force at incipient collapse could be either $4.2S_0L/EA$, or $2.4S_0L/EA$, depending on the actual order in which the bars yield. To find which answer is correct, we can check the extensions of the yielding bars for either assumption. They should exceed the elastic limit extensions of the respective bars. If the calculated value is smaller than the extension at the onset of yielding, the assumption as to which bar yields last is not correct. For our problem, the following results are obtained:

$$1.\quad e_2 = w = 4.2\frac{S_0L}{EA} > 2.4\frac{S_0L}{EA} = e_{e,2} \quad \ldots \text{ correct} \tag{8.18}$$

$$2.\quad e_1 = 0.6u + 0.8w = 0.6(-1.95)\frac{S_0L}{EA} + 0.8 \times 2.4\frac{S_0L}{EA} = \\ = 0.75\frac{S_0L}{EA} < 3\frac{S_0L}{EA} = e_{e,1} \quad \ldots \text{ impossible} \tag{8.19}$$

This means that assumption 1 is correct, bar 1 yields last, and the displacements at incipient collapse are $u = -0.6S_0L/EA$ and $w = 4.2S_0L/EA$. The plastic extension of bar 2 is $e_{p,2} = e_2 - e_{e,2} = 4.2S_0L/EA - 2.4S_0L/EA = 1.8S_0L/EA$. □

The method used in the previous example was based on the assumption that the bars that do not yield during collapse have never yielded before and thus their deformation is only elastic. However, there are some situations in which a bar starts yielding but later the extension rate reverses its sign and the bar unloads, even under a proportionally increasing load. Such a situation is rare but possible (see Problem 8.2). Excluding this type of behavior from our considerations, we can prove the following theorem.

Displacement Theorem
Let $\bar{\boldsymbol{f}}$ be the reference load vector and $\boldsymbol{d}$ the displacement vector at incipient collapse. Consider displacements $\boldsymbol{d}^*$ predicted under the assumption that a certain bar yields last. Then

$$\bar{\boldsymbol{f}}^T\boldsymbol{d}^* \le \bar{\boldsymbol{f}}^T\boldsymbol{d} \tag{8.20}$$

provided that no bar has been unloaded after previous yielding. In other words, the actual displacement vector $\boldsymbol{d}$ maximizes the work done by the reference loading vector.

Proof: The actual vector of extensions at incipient collapse, $\boldsymbol{e}$, is the sum of the elastic extensions, $\boldsymbol{e}_e$, and the plastic extensions, $\boldsymbol{e}_p$. Select one of the yielding bars, say bar number k, and assume that it starts yielding last. This bar along with the non-yielding ones forms a statically determinate structure. Construct a vector $\boldsymbol{s}^*$ containing the axial forces generated in this structure by the reference load $\bar{\boldsymbol{f}}$, with added zeros on the positions corresponding to the yielding bars except bar number k. As $\boldsymbol{s}^*$ is in equilibrium with $\bar{\boldsymbol{f}}$ we have

$$\boldsymbol{B}^T\boldsymbol{s}^* = \bar{\boldsymbol{f}} \tag{8.21}$$

and thus

$$\bar{\boldsymbol{f}}^T\boldsymbol{d} = \boldsymbol{s}^{*T}\boldsymbol{B}\boldsymbol{d} = \boldsymbol{s}^{*T}\boldsymbol{e} = \boldsymbol{s}^{*T}\boldsymbol{e}_e + \boldsymbol{s}^{*T}\boldsymbol{e}_p \tag{8.22}$$

The principle of virtual work applied to the statically determinate structure under consideration implies that

$$\boldsymbol{s}^{*T}\boldsymbol{e}_e = \bar{\boldsymbol{f}}^T\boldsymbol{d}^* \tag{8.23}$$

Due to the assumption that no bar unloads after previous yielding, the elements of $\boldsymbol{e}_p$ are zero for the non-yielding bars. Furthermore, the elements of $\boldsymbol{s}^*$ are zero for the yielding bars, except for bar number k (the one assumed to yield last). The only nonzero contribution to the product $\boldsymbol{s}^{*T}\boldsymbol{e}_p$ can thus be $S_k^* e_{p,k}$. If the assumption that bar k yields last is correct, then $e_{p,k} = 0$ and consequently $\boldsymbol{s}^{*T}\boldsymbol{e}_p = 0$. If the assumption is incorrect, then bar number k has already accumulated some plastic deformation and $e_{p,k}$ is nonzero. Since S_k^* can be determined by applying the principle of virtual work to the collapse mechanism, S_k^* and $e_{p,k}$ must have the same sign and their product is positive. In either case, we have

$$\boldsymbol{s}^{*T}\boldsymbol{e}_p \ge 0 \tag{8.24}$$

and combining (8.22), (8.23) and (8.24) we get

$$\bar{\boldsymbol{f}}^T\boldsymbol{d} = \boldsymbol{s}^{*T}\boldsymbol{e}_e + \boldsymbol{s}^{*T}\boldsymbol{e}_p = \bar{\boldsymbol{f}}^T\boldsymbol{d}^* + \boldsymbol{s}^{*T}\boldsymbol{e}_p \ge \bar{\boldsymbol{f}}^T\boldsymbol{d}^* \tag{8.25}$$

which proves the theorem.

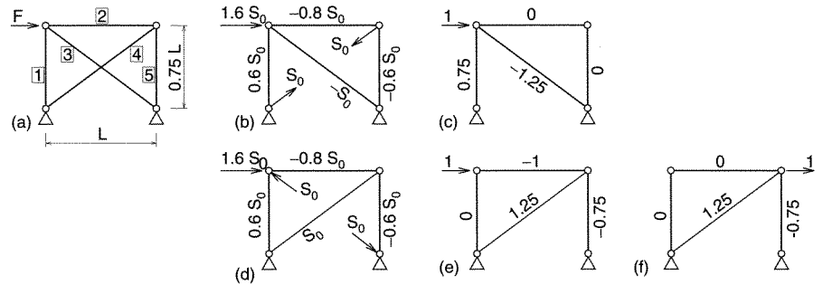


Figure 8.2 Application of displacement theorem

Note that if some bars unload after previous yielding, the corresponding elements of $\boldsymbol{e}_p$ are nonzero and they are multiplied by nonzero forces from $\boldsymbol{s}^*$. In that case, (8.25) may not hold for some vector $\boldsymbol{d}^*$, and we only have an upper bound on the projection of $\boldsymbol{d}$:

$$\bar{\boldsymbol{f}}^T\boldsymbol{d} \le \max_{\boldsymbol{d}^*} \bar{\boldsymbol{f}}^T\boldsymbol{d}^* \tag{8.26}$$

Example 8.2 Determine the displacements at incipient collapse for the truss from Example 6.1, reproduced in Figure 8.2(a).

Solution: The truss collapses at $F_0 = 1.6S_0$ by yielding of bars 3 and 4. Assume that bar 3 yields last. Then the displacements can be determined on the elastic, statically determinate structure in Figure 8.2(b). Note that $\bar{\boldsymbol{f}}^T\boldsymbol{d}$ in this case corresponds to the horizontal displacement of the upper left node, u_1. We can calculate it using the virtual state in Figure 8.2(c):

$$1 \times u_1 = 0.75L\frac{0.6S_0 \times 0.75}{EA} + 1.25L\frac{(-S_0)(-1.25)}{EA} = 1.9\frac{S_0 L}{EA} \tag{8.27}$$

Now assume that bar 4 yields last. The displacement u_1 is then calculated on the structure in Figure 8.2(d) using the virtual state in Figure 8.2(e):

$$1\times u_1 = L\frac{(-0.8S_0)(-1)}{EA} + 1.25L\frac{S_0 \times 1.25}{EA} + 0.75L\frac{(-0.6S_0)(-0.75)}{EA} = 2.7\frac{S_0 L}{EA} \tag{8.28}$$

This value is larger than that obtained before and thus it is the correct one. Bar 3 yields first and bar 4 last. The other components of $\boldsymbol{d}$ can be obtained on the structure in Figure 8.2(d) using appropriate virtual states. For example, the horizontal displacement of the upper right joint is obtained using the virtual state in Figure 8.2(f):

$$1 \times u_2 = 1.25L\frac{S_0 \times 1.25}{EA} + 0.75L\frac{(-0.6S_0)(-0.75)}{EA} = 1.9\frac{S_0 L}{EA} \tag{8.29}$$

8.2 BEAMS AND FRAMES

The proof of the Displacement Theorem can easily be extended to beams and frames. The mathematical steps are exactly the same, only the physical interpretation of

individual symbols changes. For example, the yielding bars are replaced by the plastic hinges, the axial forces by the bending moments, the bar extensions by the bending angles, and the plastic bar extensions by the plastic hinge rotations. Application of the theorem to a beam will be presented in the next example. Application to frames is conceptually the same.

Example 8.3: Consider a continuous beam of two spans L_1 and L_2, loaded by a concentrated force at midspan of the right field (Figure 8.3(a)). Determine the collapse load and displacements at incipient collapse.

Solution: As the beam is singly statically indeterminate and has only two critical sections (sections 2 and 3), the collapse mechanism is obvious (Figure 8.3(b)) and the collapse load $F_0 = 6M_0/L_2$ is easily determined from the power equality

$$F_0 \times \frac{L_2\dot{\theta}}{2} = M_0 \times \dot{\theta} + M_0 \times 2\dot{\theta} \tag{8.30}$$

The bending moment distribution $M(x)$ at collapse is plotted in Figure 8.3(c).

Consider first the case when hinge 2 forms first. Then the structure deforms as sketched in Figure 8.3(d), and at incipient failure it is equivalent to an elastic structure with a hinge at section 2 and the plastic moment M_0 applied as a load in that hinge. Deflection w_3 can be calculated using the principle of virtual work, with the virtual state shown in Figure 8.3(e). We get

$$w_3^{(3)} = \int_0^{L_1+L_2} \frac{M(x)\overline{M}^{(3)}(x)}{EI}\,dx = \frac{M_0 L_2^2}{16EI} \tag{8.31}$$

where the superscript $^{(3)}$ means that the deflection is determined under the assumption that hinge 3 forms last.

Now consider the case that hinge 2 forms last. Then the structure deforms as sketched in Figure 8.3(f), and at incipient failure it is equivalent to an elastic structure with a hinge at section 3 and the plastic moment M_0 applied as external loading in that hinge. Deflection w_3 can be calculated using the virtual state shown in Figure 8.3(g). We get

$$w_3^{(2)} = \int_0^{L_1+L_2} \frac{M(x)\overline{M}^{(2)}(x)}{EI}\,dx = \frac{M_0 L_2^2}{24EI}\left(1 + 4\frac{L_1}{L_2}\right) \tag{8.32}$$

According to the displacement theorem, hinge 3 forms last if $w_3^{(3)} > w_3^{(2)}$ (note that w_3 is the displacement that is work-conjugate to the applied force F). Comparing (8.31) to (8.32), we see that this condition holds when $L_1/L_2 < 1/8$. When $L_1/L_2 = 1/8$ both hinges form simultaneously, and when $L_1/L_2 > 1/8$ hinge 2 forms last. One may check this conclusion by calculating which hinge forms first. This can be done by comparing the moments at sections 2 and 3 computed for the beam as an elastic structure.

For structures with a high degree of static redundancy, we would have to perform the entire incremental solution, i.e. carry out the elastic solutions for structures with progressively more and more hinges, and in this way trace the sequence of formation of the hinges from the first to the last. The displacement theorem allows us to dispense

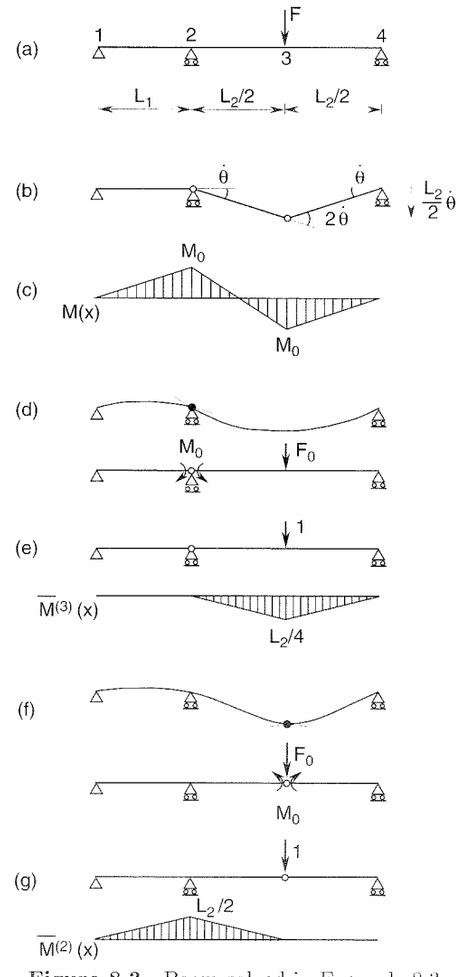


Figure 8.3 Beam solved in Example 8.3

with these tedious calculations, and replace them with the shorter calculations of statically determinate structures corresponding to various possible locations of the hinge to form last.

As indicated by the result of our example, the plastic hinge above the support forms first only if the left span is very small compared to the right span. For the purpose of analysis, the left span could be replaced by a rotational spring of stiffness $k = 3EI/L_1$ acting against the rotation of the left end section of a beam of span L_2; see Figure 8.4(a). Consider L_2 to be kept constant and L_1 to vary. In the limit $L_1 \to 0$ we obtain a beam clamped at its left end: see Figure 8.4(b). In the elastic range, the moment at the support is 20% larger than the moment below the load: that is why hinge 2 forms first.

The other limit, $L_1 \to \infty$, should intuitively correspond to a rotational spring with a zero stiffness, i.e. to a simply supported beam: see Figure 8.4(c). However, the

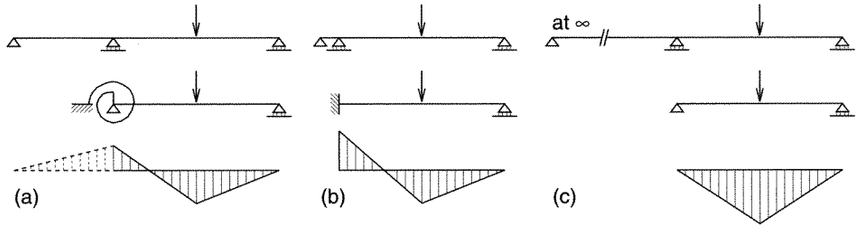


Figure 8.4 Two-span beam replaced by a single-span beam: a) general case, b) $L_1 \to 0$, c) $L_1 \to \infty$

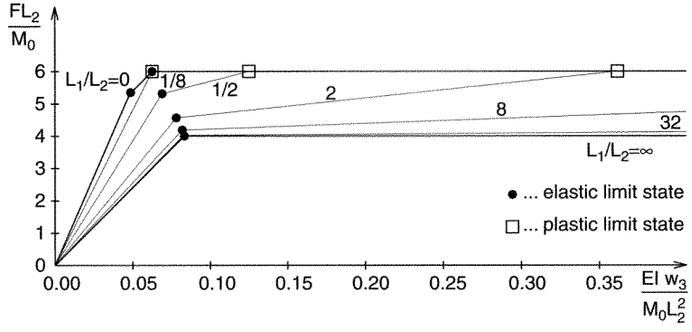


Figure 8.5 Load-deflection diagram for various span ratios

collapse load of the original two-span beam, $F_0 = 6M_0/L_2$, is independent of L_1 and does not converge to the collapse load of a simply supported beam, $F_{0,ss} = 4M_0/L_2$. This is known as the Stüssi–Kollbrunner paradox (Stüssi and Kollbrunner, 1935). Its explanation is quite simple if one looks at the load-deflection diagrams corresponding to various values of the ratio L_1/L_2; see Figure 8.5. The point at which the slope of the diagram changes corresponds to the elastic limit, and it can be observed that the elastic limit load converges to $F_{0,ss} = 4M_0/L_2$ as $L_1/L_2 \to \infty$. The deflection at the elastic limit converges to the value that corresponds to a simply supported beam; see the circular markers in Figure 8.5. On the other hand, the deflection at collapse converges to infinity; see the square markers in Figure 8.5. So, even though the theoretical collapse load is equal to $6M_0/L_2$ for an arbitrarily large (but finite) value of L_1, in practice it is impossible to exploit the plastic reserve because the deflections become prohibitively large. For example, when $L_1/L_2 = 20$, the deflection at incipient plastic collapse becomes more than 40 times larger than that at the elastic limit. In this case, it would obviously be foolish to design against the plastic limit state. If we limit the deflections by an arbitrarily large but fixed value, the load at which this deflection is reached converges to $F_{0,ss} = 4M_0/L_2$ as $L_1/L_2 \to \infty$. □

PROBLEMS

Problem 8.1: Determine the collapse load and displacements at incipient collapse for the truss in Figure 7.1 loaded by an inclined force with components $F_x = \mu S_0$ and $F_z = 2\mu S_0$.

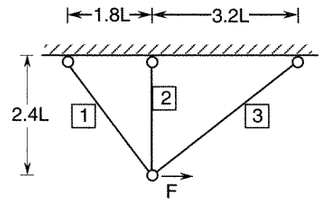


Figure 8.6 Truss to be solved in Problem 8.2

Problem 8.2: Perform incremental analysis of the truss in Figure 8.6 with $A_1 = A_2 = A_3 = A$ and $\sigma_{1,0} = 0.7\sigma_0$, $\sigma_{2,0} = 0.05\sigma_0$ and $\sigma_{3,0} = \sigma_0$. Show that bar 2 is the bar that yields first but later on, after bar 1 starts yielding, the sign of the extension rate in bar 2 changes, i.e. the bar unloads. Try to calculate the displacements at incipient collapse knowing only the collapse mechanism and axial forces, and show that inequality (8.20) from the displacement theorem does not hold.

Part II

Plastic Analysis of Structures Under Uniaxial Stress

Further Topics

9

Nonproportional and Cyclic Loads

9.1 NONPROPORTIONALLY VARYING LOADS

Certain types of loads on structures are far from proportional. For instance, the structure may first be subjected to a vertical dead load and subsequently to alternating lateral wind forces that vary without any relation to the vertical load. In such a case, simple proportional loading is not adequate. In a more general, multi-parametric representation, the load vector is described as a linear combination

$$\boldsymbol{f} = \sum_{i=1}^{p} \mu_i \overline{\boldsymbol{f}}_i \tag{9.1}$$

where $\overline{\boldsymbol{f}}_i,\ i = 1, 2, \ldots p$, are fixed vectors representing independent types of loading (dead load, live load, wind, etc.), and $\mu_i,\ i = 1, 2, \ldots p$, are the corresponding load factors. The conditions of failure must be expressed in terms of the load factors.

The principal goal is to determine the set of loads that can be resisted by the structure, i.e. to construct the *safe domain* in the *load space* (the set of all combinations of the load factors). In general, a necessary condition for a structure to be safe against plastic collapse is that, for all the kinematically admissible mechanisms, the external power produced by the given loads be smaller than the rate of dissipation. Obviously, it suffices to consider only the mechanisms with one degree of freedom. For each mechanism, the power inequality $\dot{W}_{\text{ext}} < D_{\text{int}}$ translates into two linear inequalities defining an infinite layer (intersection of two half spaces with parallel boundaries) in the space of load factors. Let us define the *elastoplastic domain* as the intersection of all such layers corresponding to all kinematically admissible mechanisms. The boundary of the elastoplastic domain is called the *plastic limit surface*. Similarly, we can define the *elastic domain* as the set of load combinations for which all points of the structure remain elastic (assuming proportional loading from an initial stress-free state), and the *elastic limit surface* as the boundary of the elastic domain, or the interface between the elastic and elastoplastic domains. The reason why we distinguish between the safe domain and the elastoplastic domain is that, as we will see in the next section, the structure can fail under nonproportional loads even if the loads remain inside the elastoplastic domain.

Example 9.1: Determine the elastoplastic domain for the three-bar truss in Figure 9.1(a) subjected to independent loads F_x and F_z. The plastic axial force, S_0, is the same for all three bars.

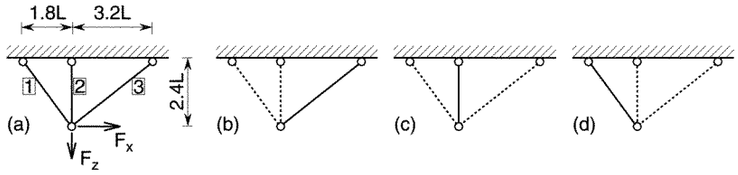


Figure 9.1 Truss under nonproportional loading

Solution: In this simple case, the load components F_x and F_z themselves are considered as the load factors. For the given structure, there are three possible collapse mechanisms with one degree of freedom, shown in Figures 9.1(b)–(d).

1. The mechanism in Figure 9.1(b) is characterized by

$$\begin{aligned} \dot{e}_3 &= -0.8\dot{u} + 0.6\dot{w} = 0 \quad \ldots \quad \dot{u} = 0.75\dot{w} \\ \dot{e}_1 &= 0.6\dot{u} + 0.8\dot{w} = 1.25\dot{w} \\ \dot{e}_2 &= \dot{w} \end{aligned} \tag{9.2}$$

The structure is safe against this mode of collapse if $\dot{W}_{\text{ext}} < D_{\text{int}}$, i.e. if

$$\begin{aligned} F_x\dot{u} + F_z\dot{w} &< S_0|\dot{e}_1| + S_0|\dot{e}_2| \\ (0.75F_x + F_z)\dot{w} &< (1.25S_0 + S_0)|\dot{w}| \end{aligned} \tag{9.3}$$

This is equivalent to two inequalities corresponding to $\dot{w} < 0$ and to $\dot{w} > 0$, respectively;

$$-2.25S_0 < 0.75F_x + F_z < 2.25S_0 \tag{9.4}$$

Graphically, this condition is represented by an infinite strip in the plane of load factors (F_x, F_z); see Figure 9.2.

2. The mechanism in Figure 9.1(c) may be analyzed in a similar way:

$$\begin{aligned} \dot{e}_2 &= \dot{w} = 0 \quad \ldots \quad \dot{w} = 0 \\ \dot{e}_1 &= 0.6\dot{u} + 0.8\dot{w} = 0.6\dot{u} \\ \dot{e}_3 &= -0.8\dot{u} + 0.6\dot{w} = -0.8\dot{u} \end{aligned} \tag{9.5}$$

$$\begin{aligned} F_x\dot{u} + F_z\dot{w} &< S_0|\dot{e}_1| + S_0|\dot{e}_2| \\ F_x\dot{u} &< (0.6S_0 + 0.8S_0)|\dot{u}| \\ -1.4S_0 < F_x &< 1.4S_0 \end{aligned} \tag{9.6}$$

3. Finally, the mechanism in Figure 9.1(d) may be analyzed as follows:

$$\begin{aligned} \dot{e}_1 &= 0.6\dot{u} + 0.8\dot{w} = 0 \quad \ldots \quad \dot{w} = -0.75\dot{u} \\ \dot{e}_2 &= \dot{w} = -0.75\dot{u} \\ \dot{e}_3 &= -0.8\dot{u} + 0.6\dot{w} = -1.25\dot{u} \end{aligned} \tag{9.7}$$

$$\begin{aligned} F_x\dot{u} + F_z\dot{w} &< S_0|\dot{e}_2| + S_0|\dot{e}_3| \\ (F_x - 0.75F_z)\dot{u} &< (0.75S_0 + 1.25S_0)|\dot{u}| \\ -2S_0 < F_x - 0.75F_z &< 2S_0 \end{aligned} \tag{9.8}$$

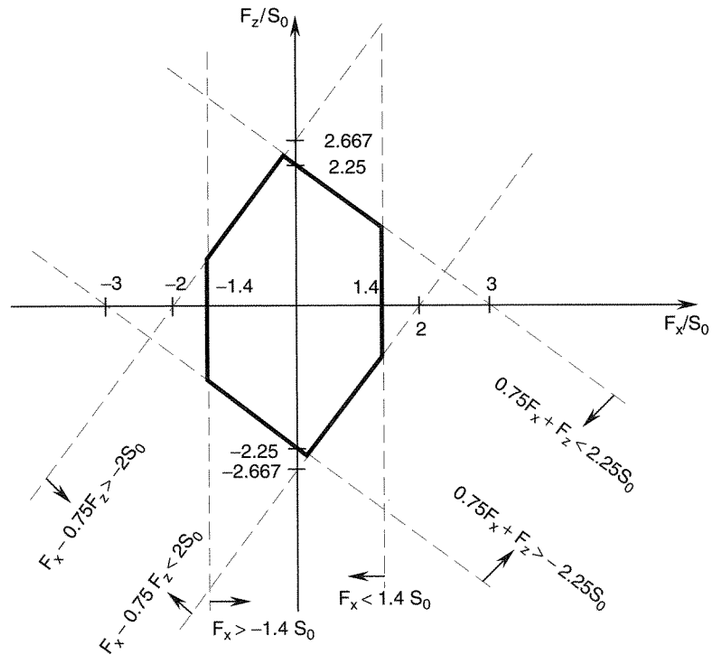


Figure 9.2 Elastoplastic domain for the truss from Example 9.1

The elastoplastic domain is graphically represented by the hexagon in Figure 9.2. The boundary of this hexagon (the plastic limit surface) represents the collapse states for proportional loading, the interior of the hexagon are the safe states, and the exterior represents the loading states that cannot be resisted by the structure. Each side of the hexagon corresponds to a certain collapse mechanism with one degree of freedom and each vertex to a mechanism with two degrees of freedom. The intersection of the F_x-axis with the boundary of the safe domain lies on the side corresponding to the mechanism in Figure 9.1(c). This means that a horizontal load $F_x = 1.4S_0$ (or $F_x = -1.4S_0$) causes collapse by yielding of bars 1 and 3. Similarly, a vertical load $F_z = 2.25S_0$ (or $F_z = -2.25S_0$) causes collapse by yielding of bars 1 and 2 (cf. Example 4.1 and Problem 4.1).

Example 9.2: Determine the elastoplastic domain for the frame in Figure 9.3(a) subjected to a horizontal load F_1 and a vertical load F_2. The plastic moment for any cross section is M_0.

Solution: Consider first the collapse mechanism in Figure 9.3(b), characterized by rotation rate $\dot{\theta}$. The external power and the power dissipated in plastic hinges may be expressed for this mechanism as follows:

$$\dot{W}_{\text{ext}} = -F_1 \times 4\dot{\theta}L + F_2 \times 2\dot{\theta}L, \quad D_{\text{int}} = (3+3)M_0|\dot{\theta}| \tag{9.9}$$

The power inequality $\dot{W}_{\text{ext}} < D_{\text{int}}$ is satisfied if

$$-3M_0/L < 2F_1 - F_2 < 3M_0/L \tag{9.10}$$

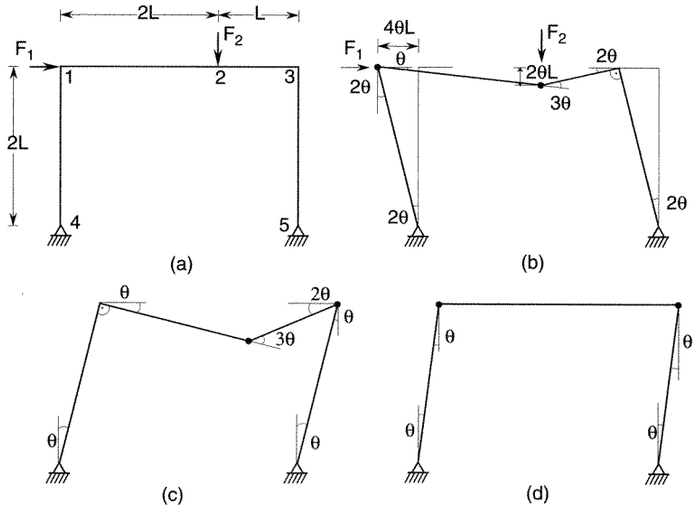


Figure 9.3 Frame under nonproportional loading

Next, consider the mechanism in Figure 9.3(c). Expressing again $\dot{W}_{\text{ext}}$ and D_{int}, we obtain the inequalities

$$-3M_0/L < F_1 + F_2 < 3M_0/L \tag{9.11}$$

Finally, considering the mechanism in Figure 9.3(d), we get

$$-M_0/L < F_1 < M_0/L \tag{9.12}$$

The system of six inequalities (9.10)–(9.12) represents the conditions of a safe state. The elastoplastic domain in Figure 9.4 again has the form of a hexagon. □

If the structure possesses k collapse mechanisms with one degree of freedom, the elastoplastic domain in a two-dimensional load space is a polygon with at most $2k$

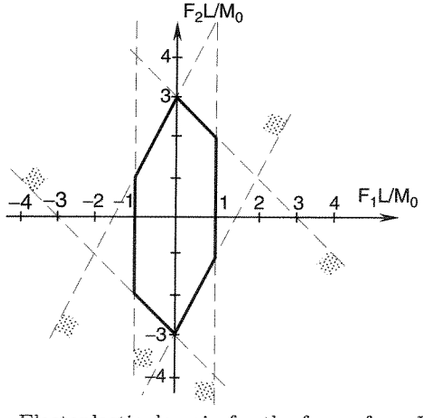


Figure 9.4 Elastoplastic domain for the frame from Example 9.2

sides (in a general multi-dimensional load space, the elastoplastic domain is a so-called *simplex* with at most $2k$ faces, i.e. an intersection of $2k$ hyper-halfspaces). If the magnitudes of the yield moments and plastic axial forces do not depend on their sign, the polygon is centrally symmetric with respect to the origin.

The most important general property of the elastoplastic domain is that it is always convex. Convexity of the limit state envelope is a rather general property which arises naturally in plasticity in many situations although under multiaxial stress certain hypotheses are required.

9.2 CYCLIC LOADING AND SHAKEDOWN

From the previous analysis, it might seem that a structure is safe for all the load variations that fall within the elastoplastic domain. This is not so, however, in the case of cyclic loads repeated many times. The reason is that the interior of the elastoplastic domain includes not only purely elastic states, but also partially plastic states for which there exist yielding bars, or plastic hinges, although not enough of them to form a mechanism. Cyclic loading that produces in the bars or hinges a certain kind of plastic deformation may lead to collapse, as we will now explain.

Example 9.3: Consider again the truss in Figure 9.1(a). We will analyze the following cycle of loading: first, F_x increases to $0.8S_0$, second, F_z increases to $1.6S_0$, third, F_x is removed, and fourth, F_z is removed. During the entire cycle OABCO, the loads remain inside the elastoplastic domain (Figure 9.5). Perform incremental analysis of the response to this cycle.

Solution: Let us first prepare the relations to be used at various stages of the loading process. The displacement increments $\boldsymbol{\Delta d}^{(n)}$ and the increments of axial forces $\boldsymbol{\Delta s}^{(n)}$ at a generic step number n can be calculated according to the following formulae:

$$\boldsymbol{\Delta d}^{(n)} = (\boldsymbol{K}^{(n)})^{-1}\boldsymbol{\Delta f}^{(n)}$$

$$\boldsymbol{\Delta s}^{(n)} = \boldsymbol{D}^{(n)}\,\boldsymbol{B}\,\boldsymbol{\Delta d}^{(n)} = \boldsymbol{D}^{(n)}\,\boldsymbol{B}\,(\boldsymbol{K}^{(n)})^{-1}\boldsymbol{\Delta f}^{(n)}$$

The matrix of current bar stiffnesses $\boldsymbol{D}^{(n)}$ depends on the state of individual bars. It affects the tangential stiffness matrix $\boldsymbol{K}^{(n)} = \boldsymbol{B}^T\boldsymbol{D}^{(n)}\boldsymbol{B}$ and the matrix used to calculate the increments of axial forces. Three different types of response will occur in the present example.

1. ELASTIC RESPONSE:

$$\frac{EA}{L}\begin{bmatrix}0.28 & 0.04\\ 0.04 & 0.72\end{bmatrix}\begin{Bmatrix}\dot{u}\\ \dot{w}\end{Bmatrix} = \begin{Bmatrix}\dot{F}_x\\ \dot{F}_z\end{Bmatrix}$$

$$\begin{Bmatrix}\dot{u}\\ \dot{w}\end{Bmatrix} = \frac{L}{EA}\begin{bmatrix}3.6 & -0.2\\ -0.2 & 1.4\end{bmatrix}\begin{Bmatrix}\dot{F}_x\\ \dot{F}_z\end{Bmatrix}$$

$$\begin{Bmatrix}\dot{S}_1\\ \dot{S}_2\\ \dot{S}_3\end{Bmatrix} = \frac{EA}{24L}\begin{bmatrix}8 & 0 & 0\\ 0 & 10 & 0\\ 0 & 0 & 6\end{bmatrix}\begin{bmatrix}0.6 & 0.8\\ 0 & 1\\ -0.8 & 0.6\end{bmatrix}\begin{Bmatrix}\dot{u}\\ \dot{w}\end{Bmatrix} = \begin{bmatrix}0.667 & 0.333\\ -0.083 & 0.583\\ -0.750 & 0.250\end{bmatrix}\begin{Bmatrix}\dot{F}_x\\ \dot{F}_z\end{Bmatrix}$$

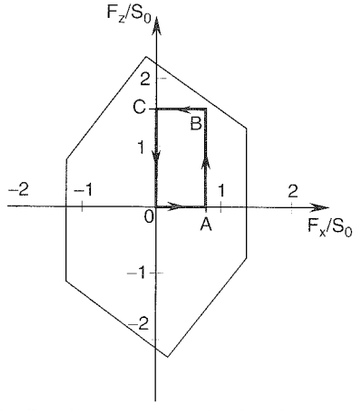


Figure 9.5 Loading cycle inside the elastoplastic domain

2. BAR 1 YIELDING, bars 2 and 3 elastic:

$$\frac{EA}{L}\begin{bmatrix} 0.16 & -0.12 \\ -0.12 & 0.507 \end{bmatrix}\begin{Bmatrix} \dot{u} \\ \dot{w} \end{Bmatrix} = \begin{Bmatrix} \dot{F}_x \\ \dot{F}_z \end{Bmatrix}$$

$$\begin{Bmatrix} \dot{u} \\ \dot{w} \end{Bmatrix} = \frac{L}{EA}\begin{bmatrix} 7.6 & 1.8 \\ 1.8 & 2.4 \end{bmatrix}\begin{Bmatrix} \dot{F}_x \\ \dot{F}_z \end{Bmatrix}$$

$$\begin{Bmatrix} \dot{S}_1 \\ \dot{S}_2 \\ \dot{S}_3 \end{Bmatrix} = \begin{bmatrix} 0 & 0 \\ 0.75 & 1 \\ -1.25 & 0 \end{bmatrix}\begin{Bmatrix} \dot{F}_x \\ \dot{F}_z \end{Bmatrix}$$

3. BAR 2 YIELDING, bars 1 and 3 elastic:

$$\frac{EA}{L}\begin{bmatrix} 0.28 & 0.04 \\ 0.04 & 0.303 \end{bmatrix}\begin{Bmatrix} \dot{u} \\ \dot{w} \end{Bmatrix} = \begin{Bmatrix} \dot{F}_x \\ \dot{F}_z \end{Bmatrix}$$

$$\begin{Bmatrix} \dot{u} \\ \dot{w} \end{Bmatrix} = \frac{L}{EA}\begin{bmatrix} 3.64 & -0.48 \\ -0.48 & 3.36 \end{bmatrix}\begin{Bmatrix} \dot{F}_x \\ \dot{F}_z \end{Bmatrix}$$

$$\begin{Bmatrix} \dot{S}_1 \\ \dot{S}_2 \\ \dot{S}_3 \end{Bmatrix} = \begin{bmatrix} 0.6 & 0.8 \\ 0 & 0 \\ -0.8 & 0.6 \end{bmatrix}\begin{Bmatrix} \dot{F}_x \\ \dot{F}_z \end{Bmatrix}$$

The evolution of axial forces and displacements during the first and a part of the second loading cycle is shown in Table 9.1. Each step can easily be evaluated using the previous relations. The initial stage of the loading process is elastic until bar 1 starts yielding at $F_x = 0.8S_0$ and $F_z = 1.4S_0$ (state $n = 2$ in Table 9.1). Unloading from $F_x = 0.8S_0$ and $F_z = 1.6S_0$ (load B, state 3) is again elastic until bar 2 starts yielding at $F_x = 0.2S_0$ and $F_z = 1.6S_0$ (state 4). After F_x has been completely removed (load C, state 5), F_z starts to unload and bar 2 again responds in an elastic way.

During the second loading cycle, bar 1 starts yielding at $F_x = 0.8S_0$ and $F_z = 1.56S_0$ (state 8), i.e. later than in the first cycle. The distribution of axial forces at

Table 9.1 Incremental solution of a truss under cyclic loading

load	n	$\frac{F_x^{(n)}}{S_0}$	$\frac{F_z^{(n)}}{S_0}$	$\frac{S_1^{(n)}}{S_0}$	$\frac{S_2^{(n)}}{S_0}$	$\frac{S_3^{(n)}}{S_0}$	$\frac{u^{(n)}EA}{S_0L}$	$\frac{w^{(n)}EA}{S_0L}$
O	0	0	0	0	0	0	0	0
A	1	0.8	0	0.533	−0.067	−0.6	2.88	−0.16
	2	0.8	1.4	1	0.75	−0.25	2.6	1.8
B	3	0.8	1.6	1	0.95	−0.25	2.96	2.28
	4	0.2	1.6	0.6	1	0.2	0.8	2.4
C	5	0	1.6	0.48	1	0.36	0.072	2.496
O	6	0	0	−0.053	0.067	−0.04	0.392	0.256
A	7	0.8	0	0.48	0	−0.64	3.272	0.096
	8	0.8	1.56	1	0.91	−0.25	2.96	2.28
B	9	0.8	1.6	1	0.95	−0.25	3.032	2.376

this state is the same as in the first loading cycle and from now on, the evolution of axial forces is periodically repeated. However, the horizontal displacement u increases in each cycle by $0.072S_0L/EA$ and the vertical displacement w by $0.096S_0L/EA$. This growth is due to the plastic straining of bars 1 and 2. In each cycle, bar 1 yields while F_z is changing from $1.56S_0$ to $1.6S_0$, and bar 2 yields while F_x is changing from $0.2S_0$ to 0. The two bars never yield simultaneously, and so no collapse mechanism actually forms. However, the plastic extension of bar 1 increases in each cycle by $0.12S_0L/EA$ and the plastic extension of bar 2 by $0.096S_0L/EA$. After a sufficient number of cycles, the ductility of the material is exhausted and the structure fails.

Example 9.4: Consider again the frame from Example 9.2, reproduced in Figure 9.6(a). Suppose that the load F_2 first increases to $2.6M_0/L$, and subsequently a lateral load $F_1 = 0.3M_0/L$ is repeatedly applied and removed; see Figure 9.6(b). Analyze the response of the structure.

Solution: The problem can be analyzed by an incremental technique similar to the previous example. We need to know the elastic response of three structures – the actual one in Figure 9.6(a), and the auxiliary ones in Figures 9.6(c),(d). The auxiliary structures represent the incremental behavior during the load steps in which one plastic hinge is active (at either section 2 or 3). In the displacement method, we would have to use five unknowns (or at least three if we handle 1-3 as a single member with load F_2 applied between the end joints). The flexibility method is more practical for hand calculation because it requires only one redundant for the actual structure. The auxiliary structures are statically determinate and can be solved just from equilibrium. Having evaluated the bending moment distributions, we can express the load-point displacements u_1 and w_2 using the principle of virtual work. Without going into details (some of which can be found in Hodge (1959)), we list the main results.

1. ACTUAL STRUCTURE (Figure 9.6(a)):

$$\dot{M}_{11} = \dot{F}_1 L - \frac{3}{13}\dot{F}_2 L = \dot{F}_1 L - 0.2308\dot{F}_2 L \tag{9.13}$$

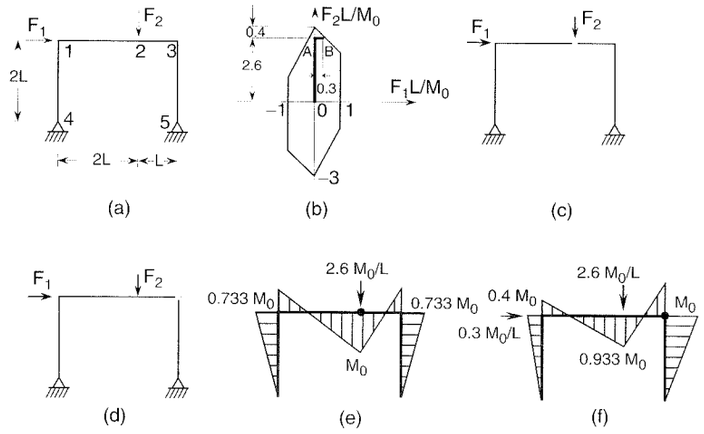


Figure 9.6 Frame under cyclic loading: a) geometry and loading, b) load path inside the elastoplastic domain, c-d) statically determinate structures obtained by adding a hinge, e) moment distribution at A, f) moment distribution at B

$$\dot{M}_{21} = -\frac{1}{3}\dot{F}_1 L + \frac{17}{39}\dot{F}_2 L = -0.3333\dot{F}_1 L + 0.4359\dot{F}_2 L \tag{9.14}$$

$$\dot{M}_{32} = -\dot{F}_1 L - \frac{3}{13}\dot{F}_2 L = -\dot{F}_1 L - 0.2308\dot{F}_2 L \tag{9.15}$$

$$\dot{u}_1 = \frac{L^3}{EI}\left(\frac{7}{3}\dot{F}_1 - \frac{1}{9}\dot{F}_2\right) = \frac{L^3}{EI}\left(2.3333\dot{F}_1 - 0.1111\dot{F}_2\right) \tag{9.16}$$

$$\dot{w}_2 = \frac{L^3}{EI}\left(-\frac{1}{9}\dot{F}_1 + \frac{25}{117}\dot{F}_2\right) = \frac{L^3}{EI}\left(-0.1111\dot{F}_1 + 0.2137\dot{F}_2\right) \tag{9.17}$$

2. Auxiliary structure with a hinge at section 2 (Figure 9.6(c)):

$$\dot{M}_{11} = \frac{1}{3}\dot{F}_1 L - \frac{2}{3}\dot{F}_2 L = 1.3333\dot{F}_1 L - 0.6667\dot{F}_2 L \tag{9.18}$$

$$\dot{M}_{21} = 0 \tag{9.19}$$

$$\dot{M}_{32} = -\frac{2}{3}\dot{F}_1 L - \frac{2}{3}\dot{F}_2 L = -0.6667\dot{F}_1 L - 0.6667\dot{F}_2 L \tag{9.20}$$

$$\dot{u}_1 = \frac{L^3}{EI}\left(\frac{76}{27}\dot{F}_1 - \frac{20}{27}\dot{F}_2\right) = \frac{L^3}{EI}\left(2.8148\dot{F}_1 - 0.7407\dot{F}_2\right) \tag{9.21}$$

$$\dot{w}_2 = \frac{L^3}{EI}\left(-\frac{20}{27}\dot{F}_1 + \frac{28}{27}\dot{F}_2\right) = \frac{L^3}{EI}\left(-0.7407\dot{F}_1 + 1.0370\dot{F}_2\right) \tag{9.22}$$

3. Auxiliary structure with a hinge at section 3 (Figure 9.6(d)):

$$\dot{M}_{11} = 2\dot{F}_1 L \tag{9.23}$$

$$\dot{M}_{21} = \frac{2}{3}\dot{F}_1 L + \frac{2}{3}\dot{F}_2 L = 0.6667\dot{F}_1 L + 0.6667\dot{F}_2 L \tag{9.24}$$

$$\dot{M}_{32} = 0 \tag{9.25}$$

Table 9.2 Incremental solution of a frame under cyclic loading

load	n	$\frac{F_1^{(n)}L}{M_0}$	$\frac{F_2^{(n)}L}{M_0}$	$\frac{M_{21}^{(n)}}{M_0}$	$\frac{M_{32}^{(n)}}{M_0}$	$\frac{u_1^{(n)}EI}{M_0L^2}$	$\frac{w_2^{(n)}EI}{M_0L^2}$
O	0	0	0	0	0	0	0
	1	0	2.294	1	−0.529	−0.255	0.490
A	2	0	2.6	1	−0.733	−0.481	0.807
	3	0.267	2.6	0.911	−1	0.141	0.778
B	4	0.300	2.6	0.933	−1	0.363	0.807
	5	0.100	2.6	1	−0.800	−0.104	0.830
A	6	0	2.6	1	−0.733	−0.385	0.903

$$\dot{u}_1 = \frac{L^3}{EI}\left(\frac{20}{3}\dot{F}_1 + \frac{8}{9}\dot{F}_2\right) = \frac{L^3}{EI}\left(6.6667\dot{F}_1 + 0.8889\dot{F}_2\right) \tag{9.26}$$

$$\dot{w}_2 = \frac{L^3}{EI}\left(\frac{8}{9}\dot{F}_1 + \frac{4}{9}\dot{F}_2\right) = \frac{L^3}{EI}\left(0.8889\dot{F}_1 + 0.4444\dot{F}_2\right) \tag{9.27}$$

The sequence of states of moments and deflections produced by the loading path OABA (Figure 9.6(b)) is summarized in Table 9.2. We start by increasing load F_2 alone and reach the elastic limit, $M_{21} = M_0$, at load $F_2 = 2.294M_0/L$ (state $n = 1$ in the table). If F_2 is further increased to $2.6M_0/L$ (load A, state 2), M_{32} increases magnitude but does not quite reach the yield value $-M_0$: see the moment distribution in Figure 9.6(c).

From now on we keep load F_2 constant, and we start increasing load F_1. The frame may either be elastic or have one hinge at M_{21}. Assuming first the elastic behavior, which means unloading in hinge 2, we obtain a solution for which indeed M_{21} decreases; see the negative coefficient at $\dot{F}_1$ in (9.14). At the same time, however, the magnitude of M_{32} grows until it reaches the plastic limit, which occurs for $F_1 = 0.267M_0/L$ (state 3). As F_1 is further increased to $0.3M_0/L$, we have a plastic hinge rotation at joint 3, and M_{21} grows until it reaches $0.933M_0$ at $F_1 = 0.3M_0/L$ (load B, state 4); see Figure 9.6(f).

For the subsequent decrease of F_1 we assume unloading at joint 3, i.e. an elastic increment, and from (9.15) we see that this is indeed the case. At the same time, M_{21} increases until, at load $F_1 = 0.1M_0/L$ (state 5), it reaches the yield moment. As F_1 is further decreased, the plastic hinge at section 2 rotates and M_{32} decreases until it reaches $-0.733M_0$ at $F_1 = 0$, which represents the end of the first load cycle (load A, state 6).

During the cycle ABA, i.e. between states $n = 2$ and $n = 6$ in Table 9.1, the load-point displacements have undergone the increments of $+0.096M_0L^2/EI$. These increments are due entirely to the plastic rotations at hinges 2 and 3. After each subsequent cycle of applying and removing horizontal load F_1, the moments return to the distribution in Figure 9.6(e), while the displacements increase by $+0.096M_0L^2/EI$ per cycle; see the deflection history in Figure 9.7(a). The history of bending moments M_{21} and M_{32} for this loading is plotted in Figure 9.7(b). We can see that the moment evolution is periodic, i.e. during each cycle the moment state describes a closed loop. On the other hand, the evolution of displacements u_1 and w_2 is a sum

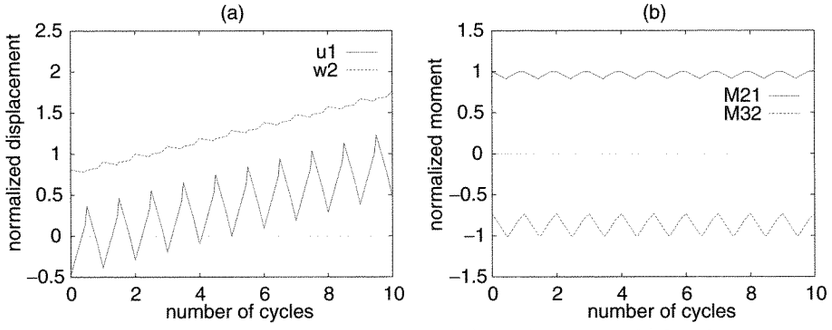


Figure 9.7 Incremental collapse: evolution of a) displacements, b) bending moments

of a periodic function and a linear function. Obviously, if we have many cycles, e.g. several thousand, very large deformations can accumulate, which surely results in a collapse of the structure. □

The type of failure produced by cyclic loading in Examples 9.3 and 9.4 is called the *incremental collapse*. Another type of cyclic loading that remains within the elastoplastic domain, yet produces collapse, is obtained when F_2 is increased to $2.6M_0/L$, then decreased to $-2.4M_0/L$, then increased back to $2.6M_0/L$, etc., while $F_1 = 0$ all the time. In this case, we do not find any accumulation of deflections during the cycles; see the deflection history in Figure 9.8(a). However, we find that bending moments M_{14} and M_{21} alternate between positive and negative yield moments, and the hinges undergo alternate positive and negative plastic rotations; see the moment history in Figure 9.8(b).

Experiments show that alternating plastic deformations always leads to failure of the material after a relatively small number of cycles (never more than about 100 cycles). This kind of collapse, due to *alternating plasticity*, is called the *plastic fatigue* (or cyclic collapse). Similar to the previous case (Figure 9.7), the bending moment state periodically hits the yield boundary but, for plastic fatigue, moment M_{21} is alternately at the positive and negative yield limit (Figure 9.8(b)), while previously each moment was repeatedly hitting only the plastic limit of one sign.

Another clear illustration of the difference between incremental collapse and plastic fatigue is provided by the moment-rotation diagrams for moment M_{21}, plotted in

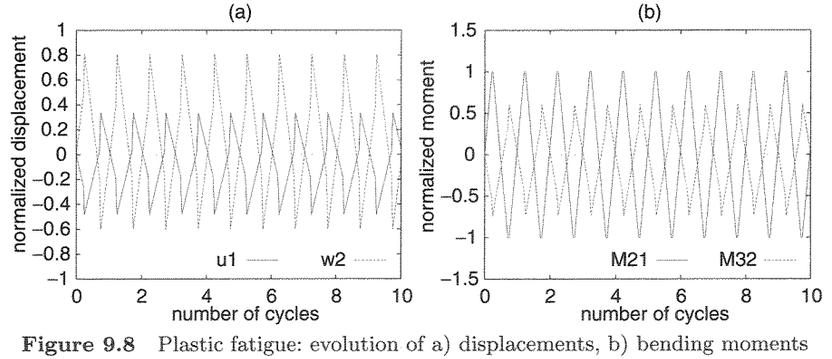


Figure 9.8 Plastic fatigue: evolution of a) displacements, b) bending moments

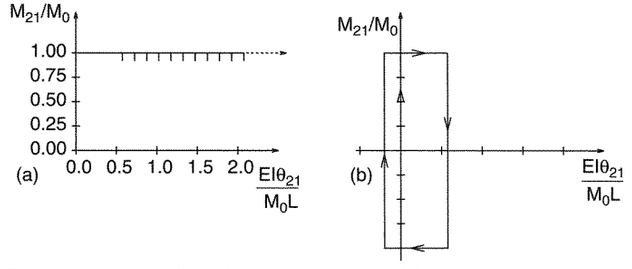


Figure 9.9 Moment-rotation diagram for plastic hinge at section 2: a) incremental collapse, b) plastic fatigue

Figure 9.9. For plastic fatigue, the state of plastic hinges traces a closed loop, while for incremental collapse, the rotation of the plastic hinge keeps increasing without bounds. In both cases, finite plastic work is being done within each cycle, and after an infinite number of cycles the total dissipated plastic work would become infinite.

The previous observation leads us to a safety concept for cyclic loads. We may expect the structure to be safe under cyclic loads with many repetitions only if there is a limit on the plastic work done by any allowable load history, up to an infinite number of cycles. For this to happen, the structure must obviously find a way to resist the cyclic load, beginning with a certain cycle, in a purely elastic fashion. Such behavior, which must obviously be attained for a structure to be safe, was called the *shakedown* (the structure 'shakes down' to a purely elastic behavior) by Prager (1948).

PROBLEMS

Problem 9.1: Construct the elastic domain and the elastoplastic domain for the structure in Figure 9.10 subjected to independent forces F_x and F_z. From the graphical representation of the elastic domain, determine which bar will start yielding first if the loading is proportional such that $F_x = F_z$, and from the graphical representation of the elastoplastic domain, determine the collapse mechanism and the collapse load for such loading.

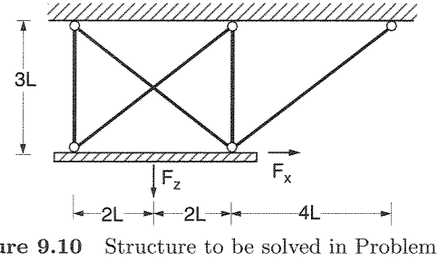


Figure 9.10 Structure to be solved in Problem 9.1

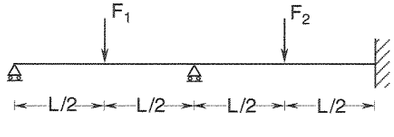


Figure 9.11 Continuous beam to be solved in Problem 9.2

Problem 9.2: Construct the elastic domain and the elastoplastic domain for the continuous beam in Figure 9.11 subjected to independent forces F_1 and F_2.

Problem 9.3: Which of the properties of the elastoplastic domain ceases to hold in the case of structures with different plastic capacities in tension and in compression? Symmetry or convexity?

10

Theorems of Shakedown Analysis

10.1 BASIC CONCEPTS

In the previous chapter we have seen that a structure subjected to general variable loads can collapse even if the loads remain inside the elastoplastic domain of the load space. Thus the elastoplastic domain represents the safe domain only if the loading is monotonic, i.e. if the load histories are such that any yield hinge, once formed, keeps yielding and never unloads. Under general, non-monotonic loading, the structure can fail by incremental collapse or plastic fatigue demonstrated in the previous chapter. Clearly, a more general concept of safety is needed to cover these situations.

A certain region in the load space represents all the load combinations that may occur during the service life of the structure. This region may be called the *service load domain*, $\mathcal{F}$. Safety analysis can be carried out on two different levels:

1. Check whether a given service load domain is safe.
2. Determine the safety factor.

On the first level, it must be checked that collapse cannot occur for any load history inside the given service load domain. On the second level, the load domains created by a radial scaling of the given service domain must be considered, and the safety factor is determined as the largest scaling factor for which the domain remains safe. Of course, on either level the safety against incremental collapse, as well as against plastic fatigue, must be checked.

10.2 STATIC APPROACH

10.2.1 Melan's Theorem

Let us restrict our attention for a while to the case of *cyclic loading histories*. Later we will show that the general case easily follows from the cyclic one. As already mentioned, a structure that adapts itself to a cyclic loading must, from some moment on, respond in a purely elastic manner. Let us denote by $\boldsymbol{f}(t)$ the periodic function defining the evolution of the external load vector in time, by $\boldsymbol{s}(t)$ the actual evolution of internal forces, and by $\boldsymbol{s}_e(t)$ the *elastic internal forces*, i.e. the internal forces that would be produced if the structure responded to the given loading only elastically, without any plastic deformation. The elastic internal forces are directly calculated from the applied loads as

$$\boldsymbol{s}_e(t) = \boldsymbol{D}_0\boldsymbol{B}\boldsymbol{K}_0^{-1}\boldsymbol{f}(t) \tag{10.1}$$

where the generalized material stiffness matrix $\boldsymbol{D}_0$ and the global (structural) stiffness matrix $\boldsymbol{K}_0$ are determined for the elastic material behavior. The difference

$$\bar{\boldsymbol{s}}_r(t) = \boldsymbol{s}(t) - \boldsymbol{s}_e(t) = \boldsymbol{s}(t) - \boldsymbol{D}_0\boldsymbol{B}\boldsymbol{K}_0^{-1}\boldsymbol{f}(t) \tag{10.2}$$

is the vector of *residual internal forces* (the bar in $\bar{\boldsymbol{s}}_r(t)$ helps to distinguish the time-dependent residual forces from the constant shakedown forces $\boldsymbol{s}_r$, to be introduced later). These are the forces that remain in the structure when it is elastically unloaded from the current state. As both the actual internal forces and the elastic internal forces are in equilibrium with the same applied loads, the residual internal forces are self-equilibrated, i.e. they are in equilibrium with zero external loads. This obvious property can formally be proven as follows:

$$\begin{aligned}\boldsymbol{B}^T\bar{\boldsymbol{s}}_r(t) &= \boldsymbol{B}^T\boldsymbol{s}(t) - \boldsymbol{B}^T\boldsymbol{s}_e(t) = \boldsymbol{f}(t) - \boldsymbol{B}^T\boldsymbol{D}_0\boldsymbol{B}\boldsymbol{K}_0^{-1}\boldsymbol{f}(t)\\ &= \boldsymbol{f}(t) - \boldsymbol{K}_0\boldsymbol{K}_0^{-1}\boldsymbol{f}(t) = \boldsymbol{0}\end{aligned} \tag{10.3}$$

Furthermore, the actual internal forces must, at any instant, satisfy the conditions of plastic admissibility

$$-\boldsymbol{s}_0 \le \boldsymbol{s}(t) \le \boldsymbol{s}_0 \tag{10.4}$$

which can be rewritten in terms of the residual internal forces as

$$-\boldsymbol{s}_0 \le \bar{\boldsymbol{s}}_r(t) + \boldsymbol{s}_e(t) \le \boldsymbol{s}_0 \tag{10.5}$$

If the structure shakes down (i.e. adapts to the cyclic loading) and begins at some time responding in a purely elastic manner, the increments of the actual internal forces are the same as the increments of the elastic internal forces, and so the residual internal forces remain constant. In other words, if the structure shakes down, there exist time-independent internal forces $\boldsymbol{s}_r$ that are self-equilibrated,

$$\boldsymbol{B}^T\boldsymbol{s}_r = \boldsymbol{0} \tag{10.6}$$

and satisfy the conditions

$$-\boldsymbol{s}_0 \le \boldsymbol{s}_r + \boldsymbol{s}_e(t) \le \boldsymbol{s}_0 \tag{10.7}$$

at any time t. Any time-independent internal forces $\boldsymbol{s}_r$ satisfying (10.6) and (10.7) will be called the *shakedown forces*. Evaluation of the safety of structures under cyclic loading is greatly simplified by the following classical theorem, first stated by Bleich (1932) and proven by Melan (1936). The proof was later simplified by Symonds (1951) and Koiter (1955; 1960).

Melan's Shakedown Theorem
If, for a given load history, there exist shakedown forces $\boldsymbol{s}_r$ such that the conditions of plastic admissibility (10.7) *are satisfied as a* strict *inequality at any time t, then the structure shakes down.*

Proof: Consider the actual residual forces $\bar{\boldsymbol{s}}_r(t)$ that vary in time, and the time-independent shakedown forces $\boldsymbol{s}_r$. The difference between these two internal force vectors can be measured using a scalar quantity $A(t)$ defined as

$$A(t) = \frac{1}{2}[\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r]^T\boldsymbol{C}_0[\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r] \tag{10.8}$$

where $\boldsymbol{C}_0 = \boldsymbol{D}_0^{-1}$ is the generalized elastic material compliance matrix (the inverse of the generalized elastic material stiffness matrix). Note that $A(t)$ is always non-negative because $\boldsymbol{C}_0$ is a positive definite matrix. Let us look at the time derivative

$$\dot{A}(t) = [\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r]^T\boldsymbol{C}_0\dot{\bar{\boldsymbol{s}}}_r(t) \tag{10.9}$$

Substituting from (10.2) and using the fact that both $\bar{\boldsymbol{s}}_r(t)$ and $\boldsymbol{s}_r$ are self-equilibrated ($\boldsymbol{B}^T\bar{\boldsymbol{s}}_r(t) = \boldsymbol{0}$ and $\boldsymbol{B}^T\boldsymbol{s}_r = \boldsymbol{0}$, see (10.3) and (10.6)), we get

$$\begin{aligned}\dot{A}(t) &= [\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r]^T\boldsymbol{C}_0[\dot{\boldsymbol{s}}(t) - \boldsymbol{D}_0\boldsymbol{B}\boldsymbol{K}_0^{-1}\dot{\boldsymbol{f}}(t)]\\ &= [\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r]^T\boldsymbol{C}_0\dot{\boldsymbol{s}}(t) - [\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r]^T\boldsymbol{B}\boldsymbol{K}_0^{-1}\dot{\boldsymbol{f}}(t)\\ &= [\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r]^T\boldsymbol{C}_0\dot{\boldsymbol{s}}(t)\end{aligned} \tag{10.10}$$

Recalling the constitutive equations, $\boldsymbol{s} = \boldsymbol{D}_0(\boldsymbol{e} - \boldsymbol{e}_p) = \boldsymbol{C}_0^{-1}(\boldsymbol{e} - \boldsymbol{e}_p)$, and the strain-displacement equations, $\boldsymbol{e} = \boldsymbol{B}\boldsymbol{d}$, we can write

$$\boldsymbol{C}_0\dot{\boldsymbol{s}}(t) = \dot{\boldsymbol{e}}(t) - \dot{\boldsymbol{e}}_p(t) = \boldsymbol{B}\dot{\boldsymbol{d}}(t) - \dot{\boldsymbol{e}}_p(t) \tag{10.11}$$

so that the expression for $\dot{A}(t)$ further simplifies to

$$\dot{A}(t) = [\bar{\boldsymbol{s}}_r(t) - \boldsymbol{s}_r]^T[\boldsymbol{B}\dot{\boldsymbol{d}}(t) - \dot{\boldsymbol{e}}_p(t)] = [\boldsymbol{s}_r - \bar{\boldsymbol{s}}_r(t)]^T\dot{\boldsymbol{e}}_p(t) \tag{10.12}$$

Here we again used the fact that both $\bar{\boldsymbol{s}}_r(t)$ and $\boldsymbol{s}_r$ are self-equilibrated. Now we substitute once more from (10.2) and get

$$\dot{A}(t) = [\boldsymbol{s}_r + \boldsymbol{s}_e(t) - \boldsymbol{s}(t)]^T\dot{\boldsymbol{e}}_p(t) \tag{10.13}$$

Since the internal forces $\boldsymbol{s}_r + \boldsymbol{s}_e(t)$ are plastically admissible, the postulate of maximum plastic dissipation (5.15) implies that

$$[\boldsymbol{s}_r + \boldsymbol{s}_e(t)]^T\dot{\boldsymbol{e}}_p(t) \le D_{\text{int}}(\dot{\boldsymbol{e}}_p(t)) = \boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_p(t) \tag{10.14}$$

which proves that the right-hand side of (10.13) is nonpositive, i.e. $\dot{A}(t) \le 0$. Moreover, since the conditions of plastic admissibility are satisfied as strict inequalities, $|\boldsymbol{s}_r + \boldsymbol{s}_e(t)| < \boldsymbol{s}_0$, $\dot{A}(t)$ is negative whenever the plastic strain rate vector $\dot{\boldsymbol{e}}_p(t)$ is different from a zero vector, i.e. whenever plastic flow takes place.

Now we know that $A(t)$ is a non-negative and nonincreasing function of time. From this, some proofs published in the literature immediately infer that $A(t)$ must be constant from some instant t_0 on, and so $\dot{A}(t) = 0$ for $t > t_0$. As $\dot{A}(t)$ is different from zero whenever any yielding takes place, it follows that for $t > t_0$ the structure responds purely elastically, i.e. shakedown is achieved.

We must realize, however, that there are nonnegative and nonincreasing functions such as $1/t$ that decrease monotonically while their derivative never becomes zero. The mathematically rigorous proof (e.g. Kamenjarzh, 1996, p. 389) is therefore somewhat more involved. It makes use of the assumption that the condition of plastic admissibility ought to be satisfied by the sum of the shakedown and elastic internal forces as a *strict* inequality. As this sum, $\boldsymbol{s}_r + \boldsymbol{s}_e(t)$, is a continuous periodic function of time, and the number of internal forces is finite, we can always find a constant $\beta > 1$ such that

$$-\frac{1}{\beta}\boldsymbol{s}_0 \le \boldsymbol{s}_r + \boldsymbol{s}_e(t) \le \frac{1}{\beta}\boldsymbol{s}_0 \tag{10.15}$$

at any time t. The postulate of maximum plastic dissipation (5.15) now yields

$$\beta[\boldsymbol{s}_r + \boldsymbol{s}_e(t)]^T \dot{\boldsymbol{e}}_p(t) \leq D_{\text{int}}(\dot{\boldsymbol{e}}_p(t)) = \boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_p(t) \tag{10.16}$$

and so

$$\begin{aligned}(\beta-1)D_{\text{int}}(\dot{\boldsymbol{e}}_p(t)) &= (\beta-1)\boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_p(t)\\ &= \beta\boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_p(t) - \boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_p(t)\\ &\leq \beta\boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_p(t) - \beta[\boldsymbol{s}_r + \boldsymbol{s}_e(t)]^T\dot{\boldsymbol{e}}_p(t)\\ &= \beta[\boldsymbol{s}(t) - \boldsymbol{s}_r - \boldsymbol{s}_e(t)]^T\dot{\boldsymbol{e}}_p(t)\\ &= -\beta\dot{A}(t)\end{aligned} \tag{10.17}$$

Consequently,

$$D_{\text{int}}(\dot{\boldsymbol{e}}_p(t)) \leq -\frac{\beta}{\beta-1}\dot{A}(t) \tag{10.18}$$

and

$$\begin{aligned}\int_0^\infty D_{\text{int}}(\dot{\boldsymbol{e}}_p(t))\,\mathrm{d}t &\leq -\frac{\beta}{\beta-1}\int_0^\infty \dot{A}(t)\,\mathrm{d}t\\ &= -\frac{\beta}{\beta-1}[A(\infty) - A(0)]\\ &\leq \frac{\beta}{\beta-1}A(0) < \infty\end{aligned} \tag{10.19}$$

This means that the total dissipation is finite, which is equivalent to shakedown.

Note, however, that for beams and frames $D_{\text{int}}(\dot{\boldsymbol{e}}_p) = \boldsymbol{s}^T\dot{\boldsymbol{e}}_p$ is only the part of the dissipation power produced by rotation of fully formed plastic hinges. In reality, some energy is also dissipated during the formation of the hinge, i.e. when the cross section is in an elastoplastic state. This is going to be addressed in the next subsection.

10.2.2 Alternating Plasticity

Melan's shakedown theorem makes it possible to prove that a certain load domain is safe against incremental collapse. It would apply to plastic fatigue as well if the inequality $-\boldsymbol{s}_0 < \boldsymbol{s} < \boldsymbol{s}_0$ were sufficient to guarantee the behavior of all points of the structure to be purely elastic. This is clearly the case for trusses where the first point yields only after the axial force has reached its plastic value. However, the extreme points of a cross section subjected to a bending moment begin yielding already at the limit elastic moment, i.e. before the limit plastic moment is reached. On the structural level of analysis, the moment-curvature diagram is simplified as bilinear, and the yield condition for the entire cross section states that the plastic hinge develops only if $|M| = M_0$. In reality, the first yielding occurs at the limit elastic moment $M_{el} = M_0/\alpha$ where α is a shape factor characterizing the cross section geometry. Under monotonic loading, the earlier idealization is justified, since the total dissipation due to partial yielding of the cross section in the elastoplastic regime $M_{el} < |M| < M_0$ is finite (each hinge going through this stage only once). Under cyclic loading, though, one must make sure that the cross section could not experience alternating yielding of the

extreme points in flexure that would lead to plastic fatigue. Fortunately, the concepts of shakedown can be applied on the level of the cross section, too.

Suppose that, on the global level, the structure shakes down and the response becomes periodic. The bending moment at a fixed cross section can be written as

$$M(t) = M_r + M_e(t) \tag{10.20}$$

where M_r is the residual (shakedown) moment, and $M_e(t)$ is the elastic moment calculated for the given load variation $\boldsymbol{f}(t)$ on a purely elastic structure. The actual moment $M(t)$ varies between two extreme values, M^{min} and M^{max}, that are plastically admissible:

$$-M_0 \leq M^{min} \leq M(t) \leq M^{max} \leq M_0 \tag{10.21}$$

On the level of the cross section, we can apply the moment $M(t)$ as 'external loading' and look at the evolution of stresses in the cross section. It follows from a generalized version of Melan's theorem that, if there exists a self-equilibrated residual stress distribution such that its sum with the stresses calculated from the applied moments according to the elastic theory satisfies the conditions od plastic admissibility as strict inequalities, then the cross section will shake down to a purely elastic response and no alternating plasticity can occur.

The elastically calculated stress σ_e varies within the limits

$$\frac{M^{min}}{I}z \leq \sigma_e \leq \frac{M^{max}}{I}z \quad \text{if } z \geq 0 \tag{10.22}$$

$$\frac{M^{max}}{I}z \leq \sigma_e \leq \frac{M^{min}}{I}z \quad \text{if } z \leq 0 \tag{10.23}$$

where z is the coordinate of the point under consideration, measured from the (elastic) neutral axis. The shakedown stress distribution can obviously be found only if, at any point, the variation of σ_e is smaller than $2\sigma_0$. The extreme points of the cross section are the most critical. The condition

$$\frac{M^{max} - M^{min}}{I}|z| < 2\sigma_0 \tag{10.24}$$

is satisfied at all the points if

$$M^{max} - M^{min} < \frac{2\sigma_0 I}{|z|_{max}} = 2\sigma_0 W_{el} = 2M_{el} = \frac{2M_0}{\alpha} \tag{10.25}$$

It can be shown that this condition is also sufficient for the existence of a shakedown stress distribution.

Condition (10.25) is easy to check because the difference $M^{max} - M^{min}$ can be derived from the elastic moment variation $M_e(t)$ alone, without any influence of the residual moment M_r:

$$\begin{aligned}M^{max} - M^{min} &= \max_t[M_r + M_e(t)] - \min_t[M_r + M_e(t)]\\ &= M_r + \max_t M_e(t) - M_r - \min_t M_e(t)\\ &= M_e^{max} - M_e^{min}\end{aligned} \tag{10.26}$$

This means that the safety against plastic fatigue (on the cross sectional level) can be checked independently of the safety against incremental collapse. It is sufficient to evaluate the extreme values of the elastically calculated moments at the critical cross sections, and to compare the difference between the maximum and minimum values to twice the limit elastic moment (for each critical cross section).

10.2.3 Lower Bound Theorem

According to Melan's theorem, a certain load domain $\mathcal{F}$ is safe against incremental collapse if shakedown internal forces $\boldsymbol{s}_r$ exist such that

$$\boldsymbol{B}^T \boldsymbol{s}_r = \boldsymbol{0} \tag{10.27}$$

$$|\boldsymbol{s}_r + \boldsymbol{s}_e(t)| < \boldsymbol{s}_0 \tag{10.28}$$

The internal forces $\boldsymbol{s}_e(t) = \boldsymbol{D}_0 \boldsymbol{B} \boldsymbol{K}_0^{-1} \boldsymbol{f}(t)$ are derived from the applied loads. Condition (10.28) should be satisfied for any cyclic loading history $\boldsymbol{f}(t)$ that stays inside the given load domain. As there are infinitely many possible loading histories, it is impossible to check the condition for each of them separately. Instead of that, we must find the extreme values of the elastic internal forces that can be produced by any load configuration from the load domain. If we find $\boldsymbol{s}_e^{min}$ and $\boldsymbol{s}_e^{max}$ such that

$$\boldsymbol{s}_e^{min} \le \boldsymbol{D}_0 \boldsymbol{B} \boldsymbol{K}_0^{-1} \boldsymbol{f} \le \boldsymbol{s}_e^{max} \quad \text{for any } \boldsymbol{f} \in \mathcal{F} \tag{10.29}$$

then (10.28) can be replaced by

$$\boldsymbol{s}_r + \boldsymbol{s}_e^{max} < \boldsymbol{s}_0 \tag{10.30}$$

$$\boldsymbol{s}_r + \boldsymbol{s}_e^{min} > -\boldsymbol{s}_0 \tag{10.31}$$

Typically, the load domain is the convex hull of a finite set of vertices, e.g. a convex polygon in a two-dimensional load space, or a convex polyhedron in a three-dimensional load space. Since the transformation from $\boldsymbol{f}$ to $\boldsymbol{s}_e$ is a linear mapping, the extreme elastic internal forces $\boldsymbol{s}_e^{min}$ and $\boldsymbol{s}_e^{max}$ can be evaluated by looking only at the values of $\boldsymbol{s}_e$ corresponding to the vertices of the load domain. An important corollary, which will be exploited in Section 10.3.2, is that the load history that periodically visits all the vertices of the load domain is the most critical one (in the sense that if this history does not lead to collapse, then the entire domain is safe).

Let us now proceed to level 2 analysis, i.e. to the evaluation of the safety factor. From the discussion of alternating plasticity it is clear that the safety factor against plastic fatigue can be computed as

$$\mu_{fat} = \min_{ij} \frac{2M_{el,ij}}{M_{eij}^{max} - M_{eij}^{min}} = \min_{ij} \frac{2M_{0,ij}}{\alpha_{(ij)}(M_{eij}^{max} - M_{eij}^{min})} \tag{10.32}$$

where M_{eij}^{min} and M_{eij}^{max} are the extreme elastically computed moments at cross section ij for any load from the service load domain $\mathcal{F}$, and the minimum is taken over all the critical cross sections; $\alpha_{(ij)}$ is the shape factor of cross section i in member ij, and it should not be confused with the angle α_{ij} characterizing the orientation of the member with respect to the global coordinate system.

The safety factor against incremental collapse, μ_{inc}, is somewhat more difficult to determine. Suppose that $\mathcal{F}$ is the given service load domain, and $\boldsymbol{s}_e^{min}$ and $\boldsymbol{s}_e^{max}$ are the extreme elastic internal forces that can be produced by any $\boldsymbol{f} \in \mathcal{F}$. A positive number μ_s is called a *statically admissible multiplier* if there are shakedown internal forces $\boldsymbol{s}_r$ such that

$$\boldsymbol{B}^T \boldsymbol{s}_r = \boldsymbol{0} \tag{10.33}$$

$$\boldsymbol{s}_r + \mu_s \boldsymbol{s}_e^{max} \le \boldsymbol{s}_0 \tag{10.34}$$

$$\boldsymbol{s}_r + \mu_s \boldsymbol{s}_e^{min} \ge -\boldsymbol{s}_0 \tag{10.35}$$

It follows from Melan's theorem that if μ_s is a statically admissible multiplier, then the load domain $\mathcal{F}_\mu = \{\mu \boldsymbol{f} |\ \boldsymbol{f} \in \mathcal{F}\}$, obtained by radial scaling of the service load domain, is safe against incremental collapse for any nonnegative scaling factor $\mu < \mu_s$. This means that the safety factor μ_{inc} cannot be smaller than any statically admissible multiplier μ_s. On the other hand, the safety factor by definition has the property that, for any $\mu < \mu_{inc}$, the load domain $\mathcal{F}_\mu$ is safe, and so μ is statically admissible. In summary, we have established the following theorem.

Lower Bound Theorem of Shakedown Analysis
The safety factor is the largest statically admissible multiplier.

Note that this formulation of the theorem applies to shakedown analysis as well as to limit analysis. The difference is only in the meaning of the statically admissible multiplier. In limit analysis the load factor multiplies a fixed load vector, while in shakedown analysis it multiplies a fixed load domain. In fact, shakedown analysis can be regarded as a generalization of limit analysis. If we shrink the service load domain to a single point representing the reference load vector (or to the straight segment connecting this point with the origin of the load space), conditions (10.33)–(10.35) become equivalent with (5.6) and (5.3), and limit analysis is obtained as a special case of shakedown analysis.

10.3 KINEMATIC APPROACH

In the previous section we have built the theoretical foundations of the static approach to shakedown analysis, which works with the distribution of internal forces and provides a lower bound on the safety factor. From analogy with limit analysis (see Chapter 5), it is natural to expect that there should exist a kinematic approach, which uses some collapse mechanisms and provides the upper bound on the safety factor. Indeed, such an approach was intuitively developed by Symonds and Neal (1951), and later theoretically justified by a bound theorem derived for a general continuum by Koiter (1956) and simultaneously for beams and frames by Neal (1956).

10.3.1 Koiter's Theorem

Again, we consider a cyclic load described by a periodic function $\boldsymbol{f}(t)$ with period T. The kinematic approach is based on the notion of *admissible plastic deformation cycle* given by a vector of deformation rates $\dot{\boldsymbol{e}}_k(t)$. To be called admissible, the deformation

rates must produce a compatible deformation increment over one cycle, i.e. there must be a displacement increment $\boldsymbol{\Delta d}_k$ such that

$$\boldsymbol{\Delta e}_k \equiv \int_0^T \dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t = \boldsymbol{B}\boldsymbol{\Delta d}_k \tag{10.36}$$

Note that we require only compatibility after the entire cycle, i.e. the intermediate increments of $\boldsymbol{e}_k$ do not have to be compatible. Let us define a (totally fictitious) displacement rate vector

$$\dot{\boldsymbol{d}}_k(t) = \boldsymbol{K}_0^{-1}\boldsymbol{B}^T\boldsymbol{D}_0\dot{\boldsymbol{e}}_k(t) \tag{10.37}$$

The displacement rates $\dot{\boldsymbol{d}}_k(t)$ are in general not compatible with the plastic deformation rates $\dot{\boldsymbol{e}}_k(t)$ because, for statically indeterminate structures, $\boldsymbol{B}\dot{\boldsymbol{d}}_k(t) = \boldsymbol{B}\boldsymbol{K}_0^{-1}\boldsymbol{B}^T\boldsymbol{D}_0\dot{\boldsymbol{e}}_k(t)$ does not have to be equal to $\dot{\boldsymbol{e}}_k(t)$. However, the total increment over one cycle is exactly the vector $\boldsymbol{\Delta d}_k$ from (10.36) because

$$\begin{aligned}\int_0^T \dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t &= \int_0^T \boldsymbol{K}_0^{-1}\boldsymbol{B}^T\boldsymbol{D}_0\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t = \boldsymbol{K}_0^{-1}\boldsymbol{B}^T\boldsymbol{D}_0\int_0^T \dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t \\ &= \boldsymbol{K}_0^{-1}\boldsymbol{B}^T\boldsymbol{D}_0\boldsymbol{B}\boldsymbol{\Delta d}_k = \boldsymbol{K}_0^{-1}\boldsymbol{K}_0\boldsymbol{\Delta d}_k = \boldsymbol{\Delta d}_k\end{aligned} \tag{10.38}$$

It should be emphasized that the plastic deformation rates $\dot{\boldsymbol{e}}_k(t)$ are completely arbitrary (except for the condition of compatibility after the entire cycle) and have nothing to do with the actual plastic deformations of the structure under consideration.

Having specified the notion of an admissible plastic deformation cycle, we can formulate and prove the aforementioned fundamental theorem (Koiter, 1956).

Koiter's Theorem

If there is an admissible plastic deformation cycle $\dot{\boldsymbol{e}}_k(t)$ such that

$$\int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\;\mathrm{d}t > \int_0^T D_{int}\left(\dot{\boldsymbol{e}}_k(t)\right)\;\mathrm{d}t \tag{10.39}$$

then the structure cannot shake down under the cyclic loading history $\boldsymbol{f}(t)$.

In (10.39), $\dot{\boldsymbol{d}}_k(t)$ is the fictitious displacement rate derived from $\dot{\boldsymbol{e}}_k(t)$ according to (10.37), and $D_{\mathrm{int}}\left(\dot{\boldsymbol{e}}_k(t)\right)$ is the dissipation rate corresponding to $\dot{\boldsymbol{e}}_k(t)$.

Proof: We will prove the assertion of the theorem by contradiction. Assume that the structure subjected to the cyclic loading history $\boldsymbol{f}(t)$ shakes down. Then, as we already know from Section 10.2, the actual internal forces can be expressed as

$$\boldsymbol{s}(t) = \boldsymbol{s}_r + \boldsymbol{s}_e(t) = \boldsymbol{s}_r + \boldsymbol{D}_0\boldsymbol{B}\boldsymbol{K}_0^{-1}\boldsymbol{f}(t) \tag{10.40}$$

where $\boldsymbol{s}_r$ are the shakedown forces, and $\boldsymbol{s}_e(t)$ are the elastic internal forces, computed as if the structure were purely elastic. Let us look at the total work that would be done

over one cycle by the actual internal forces $\boldsymbol{s}(t)$ on the assumed plastic deformation rates $\dot{\boldsymbol{e}}_k(t)$:

$$\begin{aligned}\int_0^T \boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t &= \boldsymbol{s}_r^T\int_0^T \dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t + \int_0^T \boldsymbol{s}_e^T(t)\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t \\ &= \boldsymbol{s}_r^T\boldsymbol{B}\boldsymbol{\Delta d}_k + \int_0^T \boldsymbol{f}^T(t)\boldsymbol{K}_0^{-1}\boldsymbol{B}^T\boldsymbol{D}_0\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t \\ &= \int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t\end{aligned} \tag{10.41}$$

In the last step, we used the fact that the shakedown forces $\boldsymbol{s}_r$ are self-equilibrated, i.e. $\boldsymbol{B}^T\boldsymbol{s}_r = \boldsymbol{0}$. As the actual internal forces $\boldsymbol{s}(t)$ must by plastically admissible (at any time t), we have, according to the postulate of maximum plastic dissipation,

$$\boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_k(t) \le D_{\mathrm{int}}\left(\dot{\boldsymbol{e}}_k(t)\right) \tag{10.42}$$

Using the identity derived in (10.41) we get

$$\int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t = \int_0^T \boldsymbol{s}^T(t)\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t \le \int_0^T D_{\mathrm{int}}\left(\dot{\boldsymbol{e}}_k(t)\right)\,\mathrm{d}t \tag{10.43}$$

Recall that we assumed the structure to shake down, and we ended up with the inequality

$$\int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t \le \int_0^T D_{\mathrm{int}}\left(\dot{\boldsymbol{e}}_k(t)\right)\,\mathrm{d}t \tag{10.44}$$

valid for any admissible plastic deformation cycle. By contradiction, this proves Koiter's theorem. □

10.3.2 Upper Bound Theorem

Koiter's theorem provides a means of showing that a certain cyclic loading history leads to collapse. We already know from Section 10.2.3 that, for load domains generated as the convex hull of a certain finite set of load configurations, the most critical load history is that which visits all of the vertices. If such a history leads to collapse, the domain is of course unsafe. This can be exploited for constructing upper bounds on the safety factor.

We will again look at load domains derived by radial scaling of a given service load domain. A real number μ_k is called a *kinematically admissible multiplier* if there is a cyclic load history $\boldsymbol{f}(t)$ contained within the service domain $\mathcal{F}$ and an admissible plastic deformation cycle $\dot{\boldsymbol{e}}_k(t)$ such that

$$\mu_k\int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t = \int_0^T D_{\mathrm{int}}(\dot{\boldsymbol{e}}_k(t))\,\mathrm{d}t \tag{10.45}$$

where $\dot{\boldsymbol{d}}_k(t)$ is the fictitious displacement rate derived from the deformation rate $\dot{\boldsymbol{e}}_k(t)$ according to (10.37).

If μ_k is a kinematically admissible multiplier, then, for any multiplier $\mu > \mu_k$, the structure cannot shake down under the loading history $\mu\boldsymbol{f}(t)$. This follows directly

from Koiter's theorem and from the definition of a kinematically admissible multiplier. Consequently, the safety factor cannot be larger than the smallest kinematically admissible multiplier. Intuitively, the safety factor itself may be expected to be a kinematically admissible multiplier. This property is not as obvious as it was for monotonic loading, for which it was clear that the actual collapse mechanism can be used to show the collapse load to be kinematically admissible. In shakedown analysis, it is clear only that the structure will not shake down but it is not guaranteed that an admissible plastic deformation cycle will develop. Koiter, nevertheless, showed the safety factor to be indeed kinematically admissible, and so we can state the upper bound theorem in exactly the same form as in limit analysis.

Upper Bound Theorem of Shakedown Analysis
The safety factor is the smallest kinematically admissible multiplier.

We can see that, in similarity to limit analysis, any kinematically admissible multiplier is an upper bound on the safety factor. To make practical use of this property, we need an efficient way of specifying a collapse pattern (admissible plastic deformation cycle) and calculating the corresponding kinematically admissible multiplier. It is inconvenient that the expression

$$\mu_k = \frac{\int_0^T D_{\mathrm{int}}(\dot{\boldsymbol{e}}_k(t))\,\mathrm{d}t}{\int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t} \tag{10.46}$$

derived directly from (10.45) depends on the evolution of $\dot{\boldsymbol{e}}_k(t)$ in time, and also on the specific loading history $\boldsymbol{f}(t)$. To remove the dependence on $\boldsymbol{f}(t)$, we will try to find the history for which μ_k is as small as possible. In this way, we would get the best upper bound on the safety factor. So the first step will be to consider a fixed admissible plastic deformation cycle and maximize the denominator, $\int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t$, over all possible histories from the service load domain $\mathcal{F}$. Note that the numerator, $\int_0^T D_{\mathrm{int}}(\dot{\boldsymbol{e}}_k(t))\,\mathrm{d}t$, remains constant once $\dot{\boldsymbol{e}}_k(t)$ is fixed.

We have already seen that

$$\int_0^T \boldsymbol{f}^T(t)\dot{\boldsymbol{d}}_k(t)\,\mathrm{d}t = \int_0^T \boldsymbol{s}_e^T(t)\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t \tag{10.47}$$

(recall (10.41) from the proof of Koiter's theorem). The integral $\int_0^T \boldsymbol{s}_e^T(t)\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t$ has the largest possible value if each component of $\boldsymbol{s}_e$, say $(\boldsymbol{s}_e)_i$, is equal to $(\boldsymbol{s}_e^{max})_i$ during the time intervals when $(\dot{\boldsymbol{e}}_k)_i$ is positive, and is equal to $(\boldsymbol{s}_e^{min})_i$ during the time intervals when $(\dot{\boldsymbol{e}}_k)_i$ is negative. Let us decompose $\dot{\boldsymbol{e}}_k$ into the positive part, $\dot{\boldsymbol{e}}_k^+ = (|\dot{\boldsymbol{e}}_k| + \dot{\boldsymbol{e}}_k)/2$, and the negative part, $\dot{\boldsymbol{e}}_k^- = (|\dot{\boldsymbol{e}}_k| - \dot{\boldsymbol{e}}_k)/2$, so that $\dot{\boldsymbol{e}}_k = \dot{\boldsymbol{e}}_k^+ - \dot{\boldsymbol{e}}_k^-$. Then we can write

$$\begin{aligned}\max \int_0^T \boldsymbol{s}_e^T(t)\dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t &= \int_0^T \left[(\boldsymbol{s}_e^{max})^T\dot{\boldsymbol{e}}_k^+(t) - (\boldsymbol{s}_e^{min})^T\dot{\boldsymbol{e}}_k^-(t)\right]\,\mathrm{d}t \\ &= (\boldsymbol{s}_e^{max})^T\int_0^T \dot{\boldsymbol{e}}_k^+(t)\,\mathrm{d}t - (\boldsymbol{s}_e^{min})^T\int_0^T \dot{\boldsymbol{e}}_k^-(t)\,\mathrm{d}t \\ &= (\boldsymbol{s}_e^{max})^T\Delta\boldsymbol{e}_k^+ - (\boldsymbol{s}_e^{min})^T\Delta\boldsymbol{e}_k^- \end{aligned} \tag{10.48}$$

where the maximum is taken over all possible histories of the elastic internal forces $\boldsymbol{s}_e(t)$ generated by loading histories $\boldsymbol{f}(t)$ from $\mathcal{F}$. An extremely useful property of the resulting expression is that it is independent of the details of evolution of the assumed plastic deformation during the cycle. What matters is only the total positive increment,

$$\Delta\boldsymbol{e}_k^+ = \int_0^T \dot{\boldsymbol{e}}_k^+(t)\,\mathrm{d}t \tag{10.49}$$

and the total negative increment,

$$\Delta\boldsymbol{e}_k^- = \int_0^T \dot{\boldsymbol{e}}_k^-(t)\,\mathrm{d}t \tag{10.50}$$

of the generalized plastic strain over the entire cycle. Note that

$$\Delta\boldsymbol{e}_k = \int_0^T \dot{\boldsymbol{e}}_k(t)\,\mathrm{d}t = \int_0^T \left[\dot{\boldsymbol{e}}_k^+(t) - \dot{\boldsymbol{e}}_k^-(t)\right]\,\mathrm{d}t = \Delta\boldsymbol{e}_k^+ - \Delta\boldsymbol{e}_k^- \tag{10.51}$$

but $\Delta\boldsymbol{e}_k^+$ and $\Delta\boldsymbol{e}_k^-$ are different from the positive and negative parts of $\Delta\boldsymbol{e}_k$, unless each component of $\dot{\boldsymbol{e}}_k(t)$ keeps the same sign for any time t between 0 and T.

The numerator from (10.46) can be simplified in the same manner. Realizing that the dissipation rate is given as

$$D_{\mathrm{int}}(\dot{\boldsymbol{e}}_k) = \boldsymbol{s}_0^T|\dot{\boldsymbol{e}}_k| = \boldsymbol{s}_0^T(\dot{\boldsymbol{e}}_k^+ + \dot{\boldsymbol{e}}_k^-) \tag{10.52}$$

we can write

$$\begin{aligned}\int_0^T D_{\mathrm{int}}(\dot{\boldsymbol{e}}_k(t))\,\mathrm{d}t &= \int_0^T \boldsymbol{s}_0^T(\dot{\boldsymbol{e}}_k^+(t) + \dot{\boldsymbol{e}}_k^-(t))\,\mathrm{d}t \\ &= \boldsymbol{s}_0^T\left(\int_0^T \dot{\boldsymbol{e}}_k^+(t)\,\mathrm{d}t + \int_0^T \dot{\boldsymbol{e}}_k^-(t)\,\mathrm{d}t\right) = \boldsymbol{s}_0^T(\Delta\boldsymbol{e}_k^+ + \Delta\boldsymbol{e}_k^-)\end{aligned} \tag{10.53}$$

which brings us to the final expression for the kinematically admissible multiplier,

$$\mu_k = \frac{\boldsymbol{s}_0^T(\Delta\boldsymbol{e}_k^+ + \Delta\boldsymbol{e}_k^-)}{(\boldsymbol{s}_e^{max})^T\Delta\boldsymbol{e}_k^+ - (\boldsymbol{s}_e^{min})^T\Delta\boldsymbol{e}_k^-} \tag{10.54}$$

□

PROBLEMS

Reformulate the theorems and proofs given in this chapter, considering a more general case of structures with unequal plastic capacities in tension and in compression:

Problem 10.1: Generalize the definition of shakedown forces.

Problem 10.2: Prove Melan's theorem.

Problem 10.3: Generalize the condition for the existence of a shakedown stress distribution in a cross section subjected to a varying bending moment.

Problem 10.4: Generalize the definition of a statically admissible multiplier and show the lower bound theorem of shakedown analysis to remain valid.

Problem 10.5: Explain what would change in the kinematic approach to shakedown analysis.

11

Methods of Shakedown Analysis

Based on the theorems derived in the preceding chapter, the safety factor μ_{inc} can be determined either by minimizing the kinematically admissible multiplier μ_k, or by maximizing the statically admissible multiplier μ_s. If the service load domain is specified by linear constraints, both approaches lead to a problem of linear programming.

11.1 STATIC APPROACH

In shakedown analysis, a statically admissible solution is represented by a load multiplier μ_s and a vector of shakedown forces $\boldsymbol{s}_r$. The corresponding problem of linear programming can be stated as follows:

$$\text{Maximize } g(\boldsymbol{s}_r, \mu_s) \equiv \mu_s \tag{11.1}$$

subject to

$$\boldsymbol{B}^T \boldsymbol{s}_r = \boldsymbol{0} \tag{11.2}$$

$$\boldsymbol{s}_r + \mu_s \boldsymbol{s}_e^{max} \leq \boldsymbol{s}_0 \tag{11.3}$$

$$\boldsymbol{s}_r + \mu_s \boldsymbol{s}_e^{min} \geq -\boldsymbol{s}_0 \tag{11.4}$$

The following examples demonstrate the formulation and solution of this problem for a simple truss and frame.

Example 11.1: Consider the truss from Example 9.3 subjected to a horizontal force F_x and a vertical force F_z varying arbitrarily inside the service domain

$$0 \leq F_x \leq S_0 \tag{11.5}$$

$$0 \leq F_z \leq 2S_0 \tag{11.6}$$

where S_0 is the plastic axial force. Determine the factor of safety against incremental collapse, μ_{inc}.

Solution: The service domain $\mathcal{F}$ given by (11.5)–(11.6) is a rectangle in the load space. To determine the vectors $\boldsymbol{s}_e^{max}$ and $\boldsymbol{s}_e^{min}$, we have to calculate the bounds on the axial forces in an elastically responding structure subjected to loads arbitrarily

varying inside $\mathcal{F}$. As shown in Example 9.3, the axial forces in the elastic regime can be expressed as

$$S_1 = 0.667F_x + 0.333F_z \tag{11.7}$$
$$S_2 = -0.083F_x + 0.583F_z \tag{11.8}$$
$$S_3 = -0.75F_x + 0.25F_z \tag{11.9}$$

For F_x and F_z satisfying (11.5) and (11.6), we get the following bounds:

$$S_{e1}^{max} = 0.667 \times S_0 + 0.333 \times 2S_0 = 1.333S_0 \tag{11.10}$$
$$S_{e2}^{max} = -0.083 \times 0 + 0.583 \times 2S_0 = 1.167S_0 \tag{11.11}$$
$$S_{e3}^{max} = -0.75 \times 0 + 0.25 \times 2S_0 = 0.5S_0 \tag{11.12}$$
$$S_{e1}^{min} = 0.667 \times 0 + 0.333 \times 0 = 0 \tag{11.13}$$
$$S_{e2}^{min} = -0.083 \times S_0 + 0.583 \times 0 = -0.083S_0 \tag{11.14}$$
$$S_{e3}^{min} = -0.75 \times S_0 + 0.25 \times 0 = -0.75S_0 \tag{11.15}$$

A load multiplier μ_s is statically admissible if there exists a set of self-equilibrated residual forces S_{ri}, $i = 1, 2, 3$, such that

$$S_{r1} + \mu_s S_{e1}^{max} \le S_0 \tag{11.16}$$
$$S_{r2} + \mu_s S_{e2}^{max} \le S_0 \tag{11.17}$$
$$S_{r3} + \mu_s S_{e3}^{max} \le S_0 \tag{11.18}$$
$$S_{r1} + \mu_s S_{e1}^{min} \ge -S_0 \tag{11.19}$$
$$S_{r2} + \mu_s S_{e2}^{min} \ge -S_0 \tag{11.20}$$
$$S_{r3} + \mu_s S_{e3}^{min} \ge -S_0 \tag{11.21}$$

To be be self-equilibrated, the residual forces must satisfy the conditions

$$0.6S_{r1} - 0.8S_{r3} = 0 \tag{11.22}$$
$$0.8S_{r1} + S_{r2} + 0.6S_{r3} = 0 \tag{11.23}$$

It is, therefore, possible to select one of the residual forces, say S_{r1}, and express the other residual forces as

$$S_{r3} = 0.75S_{r1} \tag{11.24}$$
$$S_{r2} = -0.8S_{r1} - 0.6S_{r3} = -1.25S_{r1} \tag{11.25}$$

Substituting (11.10)–(11.15) and (11.24)–(11.25) into (11.16)–(11.21) and separating S_{r1}, we get the following six inequalities:

$$-1 \le S_{r1}/S_0 \le 1 - 1.333\mu_s \tag{11.26}$$
$$-0.8 + 0.933\mu_s \le S_{r1}/S_0 \le 0.8 - 0.067\mu_s \tag{11.27}$$
$$-1.333 + \mu_s \le S_{r1}/S_0 \le 1.333 - 0.667\mu_s \tag{11.28}$$

The set of values S_{r1}/S_0 and μ_s satisfying all these constraints is constructed graphically in Figure 11.1. The maximum possible value of μ_s is given by the

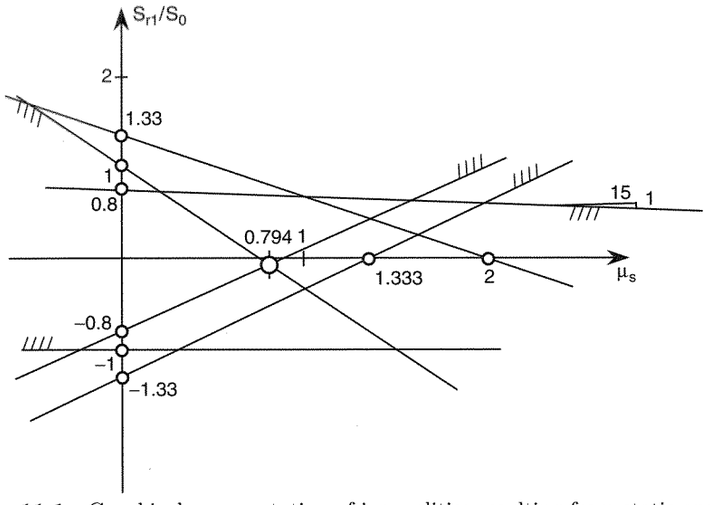


Figure 11.1 Graphical representation of inequalities resulting from static approach

intersection of the lines corresponding to $S_{r1}/S_0 = 1 - 1.333\mu_s$ and $S_{r1}/S_0 = -0.8 + 0.933\mu_s$, and so the safety factor is $\mu_{inc} = \max \mu_s = 0.794$.

Alternatively, it is possible to find max μ_s by comparing the left-hand sides and the right-hand sides in inequalities (11.26)–(11.28). For a solution to exist, each left-hand side must not be larger than any right-hand side. This provides us with nine inequalities for μ_s, which must be satisfied simultaneously. For example, comparing the left-hand side of the second inequality with all the right-hand sides, we get the following conditions:

$$-0.8 + 0.933\mu_s \le 1 - 1.333\mu_s \quad \ldots \quad \mu_s \le 0.794 \tag{11.29}$$
$$-0.8 + 0.933\mu_s \le 0.8 - 0.067\mu_s \quad \ldots \quad \mu_s \le 1.6 \tag{11.30}$$
$$-0.8 + 0.933\mu_s \le 1.333 - 0.667\mu_s \quad \ldots \quad \mu_s \le 1.333 \tag{11.31}$$

The strictest of all the nine conditions for μ_s determines the safety factor. Naturally, we again get $\mu_{inc} = \max \mu_s = 0.794$. Graphically, this so-called *method of inequalities* corresponds to calculating the intersections of all the pairs of ascending and descending lines in Figure 11.1 and taking the leftmost intersection as that specifying the safety factor.

Example 11.2: Consider the frame in Figure 11.2(a), with the service load domain given by the pentagon in Figure 11.2(b). Determine the shakedown safety factor using the static approach.

Solution: The shakedown safety factor is the largest number μ for which the μ-multiple of the service load domain $\mathcal{F}$ is safe against incremental collapse as well as plastic fatigue. The service load domain is the convex hull of five load combinations corresponding to the vertices of the pentagon in Figure 11.2(b).

Let us begin by examining the plastic fatigue. The elastic solution of the frame has already been given in Example 9.4. The dependence of the

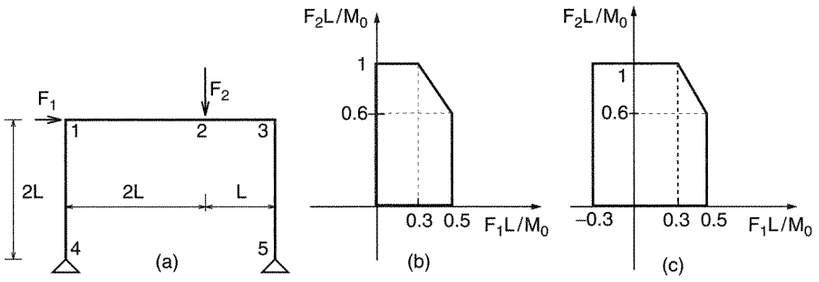


Figure 11.2 Shakedown analysis of a frame: a) geometry and loading, b)–c) service load domains

three critical bending moments on the applied forces is in the elastic range described by

$$M_{e14} = F_1 L - \frac{3}{13} F_2 L \tag{11.32}$$

$$M_{e21} = -\frac{1}{3} F_1 L + \frac{17}{39} F_2 L \tag{11.33}$$

$$M_{e32} = -F_1 L - \frac{3}{13} F_2 L \tag{11.34}$$

The moments generated by unit values of F_1 and F_2 are shown in the first two lines of Table 11.1. By combining these lines, we can easily calculate the moments for the five load combinations defining the service domain – see lines 3 to 7 of Table 11.1.

In line 8 we then write the maximum and in line 9 the minimum of the elastic moments that appear in lines 3–7, and in line 10 we write the difference between the maximum and minimum elastic moments. Noting the maximum magnitude among M_{eij}^{max} and M_{eij}^{min}, we see that, under a monotonic load increase, the smallest load factor μ for which a yield hinge forms is given by $-0.6385\mu_{hin} = -1$, i.e. $\mu_{hin} = 1.566$. Assuming the shape factors $\alpha = 1.2$, the elastic limit is reached for

Table 11.1 Shakedown analysis

Line	State	F_1L/M_0	F_2L/M_0	M_{e14}/M_0	M_{e21}/M_0	M_{e32}/M_0
1	unit F_1	1	0	1.0000	−0.3333	−1.0000
2	unit F_2	0	1	−0.2308	0.4359	−0.2308
3	combination 1	0.5	0	0.5000	−0.1667	−0.5000
4	combination 2	0.5	0.6	0.3615	0.0949	−0.6385
5	combination 3	0.3	1	0.0692	0.3359	−0.5308
6	combination 4	0	1	−0.2308	0.4359	−0.2308
7	combination 5	0	0	0.0000	0.0000	0.0000
8	M_{eij}^{max}			0.5000	0.4359	0.0000
9	M_{eij}^{min}			−0.2308	−0.1667	−0.6385
10	$M_{eij}^{max} - M_{ij}^{min}$			0.7308	0.6026	0.6385

$\mu_{el} = \mu_{hin}/\alpha = 1.566/1.2 = 1.305$. This would be the safety factor against the onset of yielding. The structure is safe against plastic fatigue if

$$\max_{ij} (\mu M_{eij}^{max} - \mu M_{eij}^{min}) < \frac{2M_0}{\alpha} \tag{11.35}$$

from which

$$\mu_{fat} \times 0.7308 M_0 = \frac{2M_0}{\alpha} \tag{11.36}$$

$$\mu_{fat} = \frac{2}{0.7308\alpha} = \frac{2}{0.7308 \times 1.2} = 2.281 \tag{11.37}$$

Let us now evaluate the safety factor against incremental collapse, μ_{inc}. We know that μ_{inc} is the largest statically admissible multiplier. For any statically admissible multiplier μ_s there must exist residual shakedown moments M_{r14}, M_{r21} and M_{r32} such that

$$M_{r14} + 0.5000\mu_s M_0 \le M_0 \tag{11.38}$$

$$M_{r21} + 0.4359\mu_s M_0 \le M_0 \tag{11.39}$$

$$M_{r32} + 0.0000\mu_s M_0 \le M_0 \tag{11.40}$$

$$M_{r14} - 0.2308\mu_s M_0 \ge -M_0 \tag{11.41}$$

$$M_{r21} - 0.1667\mu_s M_0 \ge -M_0 \tag{11.42}$$

$$M_{r32} - 0.6385\mu_s M_0 \ge -M_0 \tag{11.43}$$

The shakedown moments must also be self-equilibrated. As the structure is statically indeterminate to the first degree, only one of the moments can be regarded as an independent variable, and the other two must be expressed from equilibrium equations. Using the beam mechanism and the panel mechanism as virtual mechanisms for the principle of virtual work, we get the following conditions of equilibrium:

$$M_{r14} - M_{r32} = 0 \tag{11.44}$$

$$M_{r14} - 3M_{r21} + 2M_{r32} = 0 \tag{11.45}$$

We select M_{r14} as the independent moment and express $M_{r32} = M_{r14}$ and $M_{r21} = (M_{r14} + 2M_{r32})/3 = M_{r14}$. Substituting this back into the conditions of plastic admissibility, we get six inequalities for the shakedown moment M_{r14}. They can be rearranged as

$$-1 + 0.2308\mu_s \le M_{r14}/M_0 \le 1 - 0.5\mu_s \tag{11.46}$$

$$-1 + 0.1667\mu_s \le M_{r14}/M_0 \le 1 - 0.4359\mu_s \tag{11.47}$$

$$-1 + 0.6385\mu_s \le M_{r14}/M_0 \le 1 \tag{11.48}$$

Recall that a multiplier μ_s is statically admissible if there is at least one value of M_{r14} satisfying all the above inequalities. Obviously, the most severe restriction 'from the left' is $-1 + 0.6385\mu_s \le M_{r14}/M_0$, and the most severe restriction 'from the right' is $M_{r14}/M_0 \le 1 - 0.5\mu_s$. The multiplier μ_s is statically admissible if

$-1 + 0.6385\mu_s \leq 1 - 0.5\mu_s$, i.e. if $1.1385\mu_s \leq 2$. From this we get the safety factor against incremental collapse as the largest statically admissible multiplier, i.e. $\mu_{inc} = \max \mu_s = 2/1.1385 = 1.757$. Since μ_{inc} is smaller than $\mu_{fat} = 2.281$, the shakedown safety factor is in this case equal to $\mu_{inc} = 1.757$.

11.2 DESIGN IMPLICATIONS

In the previous example, we determined a number of different safety factors, each of them referring to a specific limit state. We can also easily compute the safety factor against collapse under monotonic loading along the paths contained in the service domain. The elastoplastic domain of the present frame was constructed in Example 9.2. By proportionally expanding the service load domain we find that one of the most dangerous monotonic paths would be to increase F_1 and keep $F_2 = 0$. Collapse would occur at $F_1 = M_0/L$, which is twice the extreme value lying still in the service domain, $F_1 = 0.5M_0/L$. The safety factor against collapse under monotonic loading is thus $\mu_0 = 2$. Geometrically, this corresponds to the fact that the service domain radially scaled by a factor of two hits the limit plastic surface (Figure 11.3).

The various safety factors are summarized in Table 11.2. The interpretation of these data depends on the specific design procedure. However, in general we can say that the safety factor in purely elastic design would be 1.305, and in plastic design it would be 2.0 if we consider only monotonic loading and 1.757 if we consider arbitrary load variations within the service load domain. Note that the shakedown safety factor is in this case governed by incremental collapse because the value for plastic fatigue is higher. This is not a general rule. For example, if we extend the service domain to that in Figure 11.2(c), we find the safety factors to be $\mu_{inc} = 1.703$ and $\mu_{fat} = 1.617$. In that case, plastic fatigue would be decisive.

Note also that the shakedown load for the service domain in Figure 11.2(b) is much larger than the load at the elastic limit. This is often the case when the given loads do not change their signs. However, if the major load alternates its sign between roughly equally large limits, the excess of the shakedown limit over the elastic limit is usually rather small. This can be observed in the results corresponding to the service domain in Figure 11.2(c), where the shakedown safety factor decreases to 1.617 while the load multiplier at the elastic limit remains at 1.305.

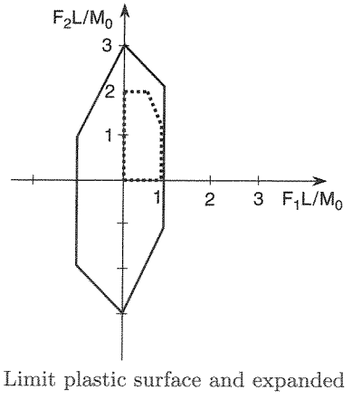


Figure 11.3 Limit plastic surface and expanded service domain

Table 11.2 Safety factors

Limit state	Factor	Value
onset of yielding	μ_{el}	1.305
formation of a plastic hinge	μ_{hin}	1.566
incremental collapse	μ_{inc}	1.757
plastic fatigue	μ_{fat}	2.281
monotonic collapse	μ_0	2.000

The joints in typical steel frame buildings exhibit a highly curved moment-rotation diagram. The calculations of shakedown can then hardly be carried out on the basis of the idealized rigid-plastic moment-rotation diagram consisting of a straight vertical line and a horizontal plateau. In such situations, a more sophisticated analysis is necessary. The alternate wind loads cause unloading at the ends of the beams in the frame and the shakedown residual moment almost cancels the elastic fixed end moment of the beams, causing that in typical high-rise building frames the bending moments at the end of the beams after many cycles of wind load nearly vanish. For this reason, the practice has for a long time been to determine the midspan moments in high-rise frames under the assumption of simple supports at the ends (Disque, 1973; Moncarz and Gerstle, 1981; Ackroyd and Gerstle, 1983).

11.3 KINEMATIC APPROACH

Let us now address the kinematic approach, which is easier to cast in the standard form solvable by the simplex method. A kinematically admissible cycle is represented by a positive deformation increment, $\Delta \boldsymbol{e}_k^+ \geq \boldsymbol{0}$, and a negative deformation increment, $-\Delta \boldsymbol{e}_k^- \leq \boldsymbol{0}$, such that the total increment $\Delta \boldsymbol{e}_k = \Delta \boldsymbol{e}_k^+ - \Delta \boldsymbol{e}_k^-$ would be compatible with some displacement increments $\Delta \boldsymbol{d}_k$. The corresponding load multiplier is given by

$$\mu_k = \frac{\boldsymbol{s}_0^T(\Delta \boldsymbol{e}_k^+ + \Delta \boldsymbol{e}_k^-)}{(\boldsymbol{s}_e^{max})^T \Delta \boldsymbol{e}_k^+ - (\boldsymbol{s}_e^{min})^T \Delta \boldsymbol{e}_k^-} \tag{11.49}$$

where $\boldsymbol{s}_0$ is the vector of plastic internal forces, and $\boldsymbol{s}_e^{max}$ and $\boldsymbol{s}_e^{min}$ are the previously defined vectors containing extreme internal forces in an elastically responding structure under loads from the reference loading domain $\mathcal{F}$. Similar to the kinematic method of limit analysis, the displacement increments $\Delta \boldsymbol{d}_k$ can be eliminated from the problem by transforming the kinematic equations into the equations of compatibility, and the kinematically admissible cycles can be normalized by setting the denominator in (11.49) equal to 1. The resulting problem of linear programming in standard form reads

$$\text{Minimize } \mu_k(\Delta \boldsymbol{e}_k^+, \Delta \boldsymbol{e}_k^-) \equiv \boldsymbol{s}_0^T \Delta \boldsymbol{e}_k^+ + \boldsymbol{s}_0^T \Delta \boldsymbol{e}_k^- \tag{11.50}$$

subject to

$$\boldsymbol{S}^T \Delta \boldsymbol{e}_k^+ - \boldsymbol{S}^T \Delta \boldsymbol{e}_k^- = \boldsymbol{0} \tag{11.51}$$

$$(\boldsymbol{s}_e^{max})^T \Delta \boldsymbol{e}_k^+ - (\boldsymbol{s}_e^{min})^T \Delta \boldsymbol{e}_k^- = 1 \tag{11.52}$$

$$\Delta \boldsymbol{e}_k^+ \geq \boldsymbol{0} \tag{11.53}$$

$$\Delta \boldsymbol{e}_k^- \geq \boldsymbol{0} \tag{11.54}$$

This problem is readily solved by the simplex method. Note the formal similarity to problem (7.34)–(7.37), resulting from the kinematic approach to limit analysis.

Example 11.3: Formulate the problem of linear programming describing shakedown analysis of the truss from the Example 11.1 using the kinematic approach, and solve it.

Solution: The compatibility equation

$$\Delta e_{k1} - 1.25\Delta e_{k2} + 0.75\Delta e_{k3} = 0 \tag{11.55}$$

is analogous to that derived in Example 7.1; see (7.27). The extreme elastic axial forces

$$\boldsymbol{s}_e^{max} = S_0 \begin{Bmatrix} 1.333 \\ 1.167 \\ 0.500 \end{Bmatrix}, \quad \boldsymbol{s}_e^{min} = S_0 \begin{Bmatrix} 0 \\ -0.083 \\ -0.750 \end{Bmatrix} \tag{11.56}$$

are known from Example 11.1. The corresponding problem of linear programming is described by the initial simplex tableau in Table 11.3(a). Note that the extreme elastic forces $\boldsymbol{s}_e^{max}$ and $\boldsymbol{s}_e^{min}$ and the plastic forces $\boldsymbol{s}_0$ have been scaled by $1/S_0$, which renders the coefficients in the simplex tableau dimensionless. Such scaling affects the numerator and the denominator in (11.49) in the same manner, and so it does not change the value of the kinematically admissible multiplier.

Table 11.3(b) shows the initial feasible solution corresponding to the basic variables Δe_{k2}^- and Δe_{k3}^-, for which the kinematically admissible multiplier is $\mu_k = 2$. In the next step, Δe_{k1}^+ enters the basis and Δe_{k2}^- leaves it (see Table 11.3(c)). The kinematically admissible multiplier decreases to $\mu_k = 1$, but the negative number in the third column of the last row indicates that the solution is not yet optimal. Another simplex transformation, during which Δe_{k2}^+ enters the basis and Δe_{k3}^- leaves it (see Table 11.3(d)), furnishes the correct safety factor $\mu_{inc} = 0.794$. The result is in perfect agreement with Example 11.1.

The final basic variables are Δe_{k1}^+ and Δe_{k2}^+, which means that the collapse is caused by the yielding of bars 1 and 2 in tension. Of course, during incremental collapse, the bars never yield simultaneously. This can again be compared with Example 9.3, in which we analyzed the response of the truss to a loading cycle that was contained in the 0.8-multiple of the service load domain considered in the present example, i.e. slightly beyond the limit $\mu_{inc} = 0.794$. We found that the plastic extension of bar 1 increased in each cycle by $0.12S_0L/EA$, and the plastic extension of bar 2 by $0.096S_0L/EA$. If we knew this collapse 'mechanism' in advance, we could calculate the safety factor directly as

$$\mu_k = \frac{\boldsymbol{s}_0^T\boldsymbol{\Delta e}_k^+ + \boldsymbol{s}_0^T\boldsymbol{\Delta e}_k^-}{(\boldsymbol{s}_e^{max})^T\boldsymbol{\Delta e}_k^+ - (\boldsymbol{s}_e^{min})^T\boldsymbol{\Delta e}_k^-} = \frac{S_0\Delta e_{k1}^+ + S_0\Delta e_{k2}^+}{S_{e1}^{max}\Delta e_{k1}^+ + S_{e2}^{max}\Delta e_{k2}^+}$$
$$= \frac{1 \times 0.12 + 1 \times 0.096}{1.333 \times 0.12 + 1.167 \times 0.096} = 0.794 \tag{11.57}$$

The simplex method evaluates the kinematically admissible multiplier for possible collapse 'mechanisms' in a systematic way, so that μ_k would always decrease, and after a finite number of steps would reach the minimum possible value. Note

Table 11.3 a) Initial simplex tableau (not feasible), b) initial feasible solution, c) improved feasible solution, d) optimal solution

(a)

$-\mu_k$	Δe_{k1}^+	Δe_{k2}^+	Δe_{k3}^+	Δe_{k1}^-	Δe_{k2}^-	Δe_{k3}^-	RHS
0	1.000	−1.250	0.750	−1.000	1.250	−0.750	0.000
0	1.333	1.167	0.500	0.000	0.083	0.750	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	0.000

(b)

$-\mu_k$	Δe_{k1}^+	Δe_{k2}^+	Δe_{k3}^+	Δe_{k1}^-	Δe_{k2}^-	Δe_{k3}^-	RHS
0	1.750	−0.062	0.938	−0.750	1.000	0.000	0.750
0	1.583	1.563	0.563	0.083	0.000	1.000	1.250
1	−2.333	−0.500	−0.500	1.667	0.000	0.000	−2.000

(c)

$-\mu_k$	Δe_{k1}^+	Δe_{k2}^+	Δe_{k3}^+	Δe_{k1}^-	Δe_{k2}^-	Δe_{k3}^-	RHS
0	1.000	−0.036	0.536	−0.429	0.571	0.000	0.429
0	0.000	1.619	−0.286	0.762	−0.905	1.000	0.571
1	0.000	−0.583	0.750	0.667	1.333	0.000	−1.000

(d)

$-\mu_k$	Δe_{k1}^+	Δe_{k2}^+	Δe_{k3}^+	Δe_{k1}^-	Δe_{k2}^-	Δe_{k3}^-	RHS
0	1.000	0.000	0.529	−0.412	0.551	0.022	0.441
0	0.000	1.000	−0.176	0.471	−0.559	0.618	0.353
1	0.000	0.000	0.647	0.941	1.007	0.360	−0.794

that the description of the collapse 'mechanism' can be read from the right-hand side of the optimal simplex tableau (Table 11.3(d)). The ratio of the basic variables, $\Delta e_{k1}^+/\Delta e_{k2}^+ = 0.441/0.353 = 1.25$, is the same as $0.12S_0L/EA$ divided by $0.096S_0L/EA$. □

The next two examples illustrate the kinematic approach to shakedown analysis of frames. We will first present the solution that checks all single degree-of-freedom mechanisms, and then formulate and solve the problem using the framework of linear programming.

Example 11.4: Using the kinematic approach, determine the safety factor against incremental collapse for the frame in Figure 11.4(a). The service load domain is given by the pentagon in Figure 11.4(b), same as Example 11.2.

Solution: During incremental collapse, yield hinges form and unload at various stages of the loading cycle. We actually do not have a collapse mechanism because the plastic rotations at these hinges do not occur simultaneously. However, the cumulative effect of the plastic hinge rotations over each cycle is the same as that of a deformation in the sense of a collapse mechanism. The elastic deformations have no effect over the cycle because the stress state returns to the same state after the cycle. We may, therefore, use an imaginary collapse mechanism to write the work equality for the entire cycle. Consider mechanism 1A in Figure 11.4. If we select θ_{14}, θ_{21} and θ_{32} as

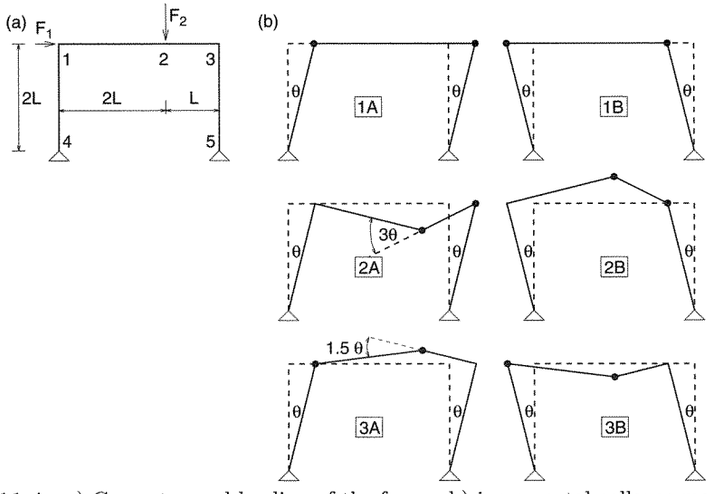


Figure 11.4 a) Geometry and loading of the frame, b) incremental collapse mechanisms

the generalized strains, mechanism 1A is characterized by

$$\boldsymbol{\Delta e}_k \equiv \begin{Bmatrix} \Delta\theta_{14} \\ \Delta\theta_{21} \\ \Delta\theta_{32} \end{Bmatrix} = \begin{Bmatrix} \Delta\theta \\ 0 \\ -\Delta\theta \end{Bmatrix} = \begin{Bmatrix} \Delta\theta \\ 0 \\ 0 \end{Bmatrix} - \begin{Bmatrix} 0 \\ 0 \\ \Delta\theta \end{Bmatrix} = \boldsymbol{\Delta e}_k^+ - \boldsymbol{\Delta e}_k^- \tag{11.58}$$

For the reader's convenience, the extreme elastic moments evaluated in Example 11.2 are reproduced in Table 11.4. Substituting them into (11.49), we obtain a kinematically admissible multiplier

$$\mu_k = \frac{\boldsymbol{s}_0^T \boldsymbol{\Delta e}_k^+ + \boldsymbol{s}_0^T \boldsymbol{\Delta e}_k^-}{(\boldsymbol{s}_e^{max})^T \boldsymbol{\Delta e}_k^+ - (\boldsymbol{s}_e^{min})^T \boldsymbol{\Delta e}_k^-} = \frac{M_0\Delta\theta + M_0\Delta\theta}{M_{e14}^{max}\Delta\theta - M_{e32}^{min}\Delta\theta}$$
$$= \frac{2M_0}{M_{e14}^{max} - M_{e32}^{min}} = \frac{2M_0}{0.5M_0 - (-0.6385M_0)} = 1.757 \tag{11.59}$$

The previous evaluation of μ_k can be described in physical terms as follows: Since $\Delta\theta_{14} = \Delta\theta$ is positive and $\Delta\theta_{32} = -\Delta\theta$ is negative, we must have $M_{14} = M_0$ during the interval in which hinge 1 yields, and $M_{32} = -M_0$ during the interval in which hinge 3 yields. Expressing the total moments M_{14} and M_{32} in terms of shakedown moments M_{r14} and M_{r32} and of the extreme elastic moments according to Table 11.4, we have the conditions

$$M_{14} = M_{r14} + \mu_k M_{e14}^{max} = M_{r14} + \mu_k \times 0.5M_0 = M_0 \tag{11.60}$$
$$M_{32} = M_{r32} + \mu_k M_{e32}^{min} = M_{r32} - \mu_k \times 0.6385M_0 = -M_0 \tag{11.61}$$

Table 11.4 Extreme elastic moments

ij	14	21	32
M_{eij}^{max}/M_0	0.5000	0.4359	0.0000
M_{eij}^{min}/M_0	−0.2308	−0.1667	−0.6385

from which we can express the shakedown moments in terms of the kinematically admissible multiplier as

$$M_{r14} = (1 - 0.5\mu_k)M_0 \tag{11.62}$$
$$M_{r32} = (-1 + 0.6385\mu_k)M_0 \tag{11.63}$$

Mechanism 1A in Figure 11.4 can at the same time be used as a virtual mechanism to set up an equilibrium equation containing only M_{r14} and M_{r32}. As the shakedown moments must be self-equilibrated, the external virtual work vanishes and we can write

$$M_{r14}\delta\theta - M_{r32}\delta\theta = 0 \tag{11.64}$$

from which

$$M_{r14} - M_{r32} = 0 \tag{11.65}$$

Substituting the previously derived expressions for the shakedown moments, we get a single equation for the load multiplier

$$(1 - 0.5\mu_k)M_0 - (-1 + 0.6385\mu_k)M_0 = 0 \tag{11.66}$$

from which

$$\mu_k = 1.757 \tag{11.67}$$

Now we consider the same mechanism but with the opposite sign (1B, Figure 11.4). In (11.59), we must replace M_{e14}^{max} and M_{e32}^{min} by the opposite extremes from

Table 11.5 Evaluation of kinematically admissible multipliers

ij		14	21	32	μ_k
M_{eij}^{max}/M_0		0.5000	0.4359	0.0000	
M_{eij}^{min}/M_0		−0.2308	−0.1667	−0.6385	
1A:	$\Delta\theta_{ij}^+$	1	0	0	
	$\Delta\theta_{ij}^-$	0	0	−1	2.0 / 1.138 = 1.757
1B:	$\Delta\theta_{ij}^+$	0	0	1	
	$\Delta\theta_{ij}^-$	−1	0	0	2.0 / 0.231 = 8.666
2A:	$\Delta\theta_{ij}^+$	0	3	0	
	$\Delta\theta_{ij}^-$	0	0	−3	6.0 / 3.223 = 1.862
2B:	$\Delta\theta_{ij}^+$	0	0	3	
	$\Delta\theta_{ij}^-$	0	−3	0	6.0 / 0.5 = 12.0
3A:	$\Delta\theta_{ij}^+$	1.5	0	0	
	$\Delta\theta_{ij}^-$	0	−1.5	0	3.0 / 1.0 = 3.0
3B:	$\Delta\theta_{ij}^+$	0	1.5	0	
	$\Delta\theta_{ij}^-$	−1.5	0	0	3.0 / 1.0 = 3.0

Table 11.4, which leads to

$$\mu_k = \frac{\boldsymbol{s}_0^T \Delta \boldsymbol{e}_k^+ + \boldsymbol{s}_0^T \Delta \boldsymbol{e}_k^-}{(\boldsymbol{s}_e^{max})^T \Delta \boldsymbol{e}_k^+ - (\boldsymbol{s}_e^{min})^T \Delta \boldsymbol{e}_k^-} = \frac{M_0 \Delta\theta + M_0 \Delta\theta}{M_{e14}^{max} \Delta\theta - M_{e32}^{min} \Delta\theta}$$
$$= \frac{2M_0}{M_{e14}^{max} - M_{e32}^{min}} = \frac{2M_0}{0 - (-0.2308 M_0)} = 8.666 \qquad (11.68)$$

For the other four mechanisms from Figure 11.4, the kinematically admissible multipliers are evaluated in Table 11.5. Among all the six kinematically admissible multipliers, the decisive limit for incremental collapse is the smallest one, $\mu_k = 1.757$, which is the same value as obtained in Example 11.2 using the static approach. The incremental collapse mechanism is the panel mechanism 1A. □

Example 11.5: Formulate the problem analyzed in the previous example as a problem of linear programming and solve it using the simplex method.

Solution: The general form of the kinematic approach to shakedown analysis as a linear programming problem is given by (11.50)–(11.54). In application to frames, the vector of unknown plastic deformation increments $\boldsymbol{\Delta e}_k$ contains the increments of

Table 11.6 a) Initial simplex tableau (not feasible), b) initial feasible solution, c) improved feasible solution, d) optimal solution

(a)

$-\mu_k$	$\Delta\theta_{14}^+$	$\Delta\theta_{21}^+$	$\Delta\theta_{32}^+$	$\Delta\theta_{14}^-$	$\Delta\theta_{21}^-$	$\Delta\theta_{32}^-$	RHS
0	1.000	1.000	1.000	−1.000	−1.000	−1.000	0.000
0	0.500	0.436	0.000	0.231	0.167	0.638	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	0.000

(b)

$-\mu_k$	$\Delta\theta_{14}^+$	$\Delta\theta_{21}^+$	$\Delta\theta_{32}^+$	$\Delta\theta_{14}^-$	$\Delta\theta_{21}^-$	$\Delta\theta_{32}^-$	RHS
0	4.000	3.616	1.000	0.385	0.000	2.831	6.000
0	3.000	2.616	0.000	1.385	1.000	3.831	6.000
1	−6.000	−5.231	0.000	−0.770	0.000	−5.662	−12.000

(c)

$-\mu_k$	$\Delta\theta_{14}^+$	$\Delta\theta_{21}^+$	$\Delta\theta_{32}^+$	$\Delta\theta_{14}^-$	$\Delta\theta_{21}^-$	$\Delta\theta_{32}^-$	RHS
0	1.000	0.904	0.250	0.096	0.000	0.708	1.500
0	0.000	−0.096	−0.750	1.096	1.000	1.708	1.500
1	0.000	0.192	1.500	−0.192	0.000	−1.416	−3.000

(d)

$-\mu_k$	$\Delta\theta_{14}^+$	$\Delta\theta_{21}^+$	$\Delta\theta_{32}^+$	$\Delta\theta_{14}^-$	$\Delta\theta_{21}^-$	$\Delta\theta_{32}^-$	RHS
0	1.000	0.944	0.561	−0.358	−0.414	0.000	0.878
0	0.000	−0.056	−0.439	0.642	0.586	1.000	0.878
1	0.000	0.113	0.878	0.716	0.829	0.000	−1.757

plastic rotations, $\Delta\theta_{14}$, $\Delta\theta_{21}$ and $\Delta\theta_{32}$. The vector of plastic internal forces contains the plastic moments of individual critical cross sections (assumed here to have the same value, M_0). The vectors $\boldsymbol{s}_e^{max}$ and $\boldsymbol{s}_e^{min}$ contain the extreme elastic moments that can be taken from Table 11.4. Finally, the compatibility condition

$$\Delta\theta_{14} + \Delta\theta_{21} + \Delta\theta_{32} = 0 \qquad (11.69)$$

has been constructed in Example 7.3. It is therefore easy to set up the initial simplex tableau in Table 11.6(a).

To obtain the initial feasible solution, we must select two variables as the basic ones. This choice corresponds to the choice of a specific 'collapse mechanism'. As we know from Example 11.4, the most critical mechanism is mechanism 1A from Figure 11.4, for which the rotation of hinge 1 exhibits a positive increment over the cycle while the rotation of hinge 3 exhibits a negative increment. So if we selected $\Delta\theta_{14}^+$ and $\Delta\theta_{32}^-$ as the basic variables, we would immediately get the optimal solution. However, even if we make the worst possible choice – the mechanism that gave the largest kinematically admissible multiplier – the optimal solution is obtained after two simplex transformations, i.e. we have to check only three mechanisms. This is documented in Tables 11.6(b)–(d).

The feasible solution in Table 11.6(b) corresponds to mechanism 2B from Figure 11.4 (positive rotation $\Delta\theta_{32}$ and negative rotation $\Delta\theta_{21}$), for which we get $\mu_k = 12$. During the first transformation, $\Delta\theta_{32}^+$ leaves the basis and $\Delta\theta_{14}^+$ enters it, which means that we go to mechanism 3B from Figure 11.4, and the kinematically admissible multiplier decreases to $\mu_k = 3$. Finally, $\Delta\theta_{21}^-$ is replaced in the basis by $\Delta\theta_{32}^-$, and we get the optimal solution with $\mu_s = 1.757$, which corresponds to mechanism 1A and is equal to the previously determined safety factor μ_{inc}.

PROBLEMS

Problem 11.1: Convert the linear programming problem (11.1)–(11.4) to the standard form and show how to find an initial basic feasible solution.
Hint: Use a technique similar to that applied in the static approach to limit analysis.

Problem 11.2: Generalize the linear programming problem (11.1)–(11.4) for structures with unequal plastic capacities in tension and in compression.

Problem 11.3: Generalize the linear programming problem (11.50)–(11.54) for structures with unequal plastic capacities in tension and in compression.

Problem 11.4: Solve Example 11.3 for the case $S_0^+ = 2S_0^-$.

12

Optimum Design

12.1 COST FUNCTION

The aim of this chapter is to expose the reader to the fundamental ideas of structural optimization. In practical applications to real problems, it is necessary to take into account many complicated aspects that we neglect here, e.g. the effect of normal and shear forces, the difference in labor costs and other costs of manufacturing different beams, the added cost of fabricating joints, or second-order effects (buckling), to name only the most important ones.

We consider a design problem in which the load is given, and the geometry of the structure has been selected and is fixed. We strive to find the optimum selection of cross sections for this structure. Various criteria can be introduced for the optimum. In most cases, the optimum criterion is that of minimum cost.

For the sake of simplicity, let us only consider structures made of a single material, e.g. steel. The simplest choice of the cost function z to be minimized in optimum design is the total weight of the structure or, equivalently, its volume. For a truss, the volume of each bar $V = \ell A = \ell S_0/\sigma_0$ is proportional to the product of length ℓ and plastic axial force S_0, and so we can minimize

$$z(\boldsymbol{s}_0) = \tilde{\boldsymbol{c}}^T \boldsymbol{s}_0 = \sum_i \ell_i S_{0,i} \tag{12.1}$$

where

$$\tilde{\boldsymbol{c}} = \{\ell_1, \ell_2, \ldots \ell_n\}^T \tag{12.2}$$

is the vector of given lengths of the bars, and

$$\boldsymbol{s}_0 = \{S_{0,1}, S_{0,2}, \ldots S_{0,n}\}^T \tag{12.3}$$

is the vector of plastic axial forces to be designed. It is often impractical to allow the cross sectional dimensions of each bar to vary independently. Additional design constraints can be introduced by setting

$$\boldsymbol{s}_0 = \boldsymbol{E}\boldsymbol{s}_d \tag{12.4}$$

where $\boldsymbol{s}_d$ is a vector of independent design variables and $\boldsymbol{E}$ is a given matrix expressing technological constraints and called the *technological matrix*. In the simplest case, $\boldsymbol{E}$ is a Boolean (incidence) matrix specifying for each bar which design variable determines

its cross section, but it can also be a more general matrix prescribing e.g. fixed ratios between certain cross sections. The objective function is then modified to

$$z(\boldsymbol{s}_d) = \tilde{\boldsymbol{c}}^T \boldsymbol{E} \boldsymbol{s}_d = \boldsymbol{c}^T \boldsymbol{s}_d \tag{12.5}$$

where $\boldsymbol{c} = \boldsymbol{E}^T \tilde{\boldsymbol{c}}$ is a vector containing (in the simplest case of a Boolean matrix $\boldsymbol{E}$) the total lengths of bars in each group of bars to be made with the same cross section.

For beams or frames, the problem is more complicated because the cost is proportional to the cross sectional area, A, but the yield moment is proportional to the plastic section modulus, W_0; see Section 2.2. The relation between A and W_0 is in general not linear, as illustrated by the first example.

Example 12.1: Derive the relation between the cross sectional area and the plastic section modulus for a rectangular cross section of depth h and width b, considering the following situations:

a) the depth remains constant while the width varies,
b) the width remains constant while the depth varies,
c) both depth and width vary but their ratio remains constant.

Solution: Given the cross sectional dimensions, the area and the plastic section modulus can be evaluated as

$$A = bh, \qquad W_0 = \frac{1}{4} b h^2 \tag{12.6}$$

In the first case, $h = \bar{h} = \text{const.}$, and we have

$$W_0 = \frac{1}{4} b \bar{h}^2 \quad \ldots \quad b = \frac{4W_0}{\bar{h}^2} \tag{12.7}$$

$$A = b\bar{h} = \frac{4W_0}{\bar{h}^2}\bar{h} = \text{const.} \times W_0 \tag{12.8}$$

In the second case, $b = \bar{b} = \text{const.}$, and we have

$$W_0 = \frac{1}{4} \bar{b} h^2 \quad \ldots \quad h = \sqrt{\frac{4W_0}{\bar{b}}} \tag{12.9}$$

$$A = \bar{b}h = \bar{b}\sqrt{\frac{4W_0}{\bar{b}}} = \text{const.} \times (W_0)^{1/2} \tag{12.10}$$

In the third case, let $b = \beta h$ where β is a constant. Then we have

$$W_0 = \frac{1}{4} \beta h^3 \quad \ldots \quad h = \left(\frac{4W_0}{\beta}\right)^{1/3} \tag{12.11}$$

$$A = bh = \beta h^2 = \beta \left(\frac{4W_0}{\beta}\right)^{2/3} = \text{const.} \times (W_0)^{2/3} \tag{12.12}$$

Only in the first case does the plastic section modulus grow proportionally to the area. If the depth of the cross section is not kept constant, the modulus grows faster than the area. The relation is in either case described by a power law. □

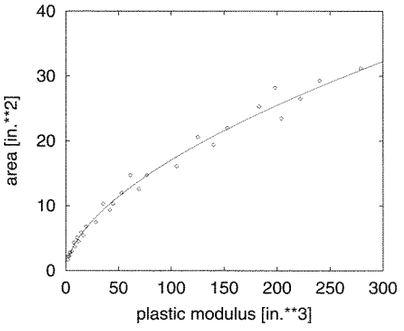


Figure 12.1 Dependence between the area and plastic section modulus for S-shapes of hot-rolled profiles

For standard hot-rolled steel profiles, the cross sectional area A may be plotted against the plastic section modulus, W_0. This provides a set of points that appear to lie close to a smooth curve, as shown in Figure 12.1. The dependence can often be well approximated by a power function

$$\frac{A}{\bar{A}} = \left(\frac{W_0}{\overline{W}_0}\right)^p \tag{12.13}$$

where $\bar{A}$ and $\overline{W}_0$ are, respectively, the area and plastic section modulus of a chosen reference section, and p is a constant exponent. For typical hot-rolled profiles, the depth grows somewhat faster than the width, and the value of p is between 1/2 and 2/3, which are the values corresponding to cases b) and c), respectively, from Example 12.1. The points in Figure 12.1 correspond to the S-shapes of hot-rolled profiles according to the AISC Manual of Steel Construction (AISC 1986), and the curve shows approximation (12.13) with $\bar{A} = 10$ in.^{2}, $\overline{W}_0 = 40$ in.^{3} and $p = 0.58$.

The nonlinearity caused by (12.13), and also by the fact that the true objective function should be based on the fabrication and material cost rather than on the total weight, would lead to a nonlinear programming problem. Designing a structure, we usually know a certain range in which the desired values of the plastic moment $M_0 = W_0\sigma_0$ for a given member i may be expected to lie. Within such a limited range, the nonlinear function

$$A = \bar{A}\left(\frac{M_0}{\sigma_0 \overline{W}_0}\right)^p \tag{12.14}$$

may usually be approximated by a linearized dependence,

$$A_i = a_i + b_i M_{0,i} \tag{12.15}$$

in which a_i and b_i are constants depending on the range.

Suppose that we know an estimate $\overline{M}_{0,i}$ of the plastic moment at member i. Linearizing the general relation $A = f(M_0)$ around $M_0 = \overline{M}_{0,i}$ we obtain

$$A_i = f(\overline{M}_{0,i}) + f'(\overline{M}_{0,i})(M_{0,i} - \overline{M}_{0,i}) \tag{12.16}$$

and so the constants a_i and b_i from (12.15) are given by

$$a_i = f(\overline{M}_{0,i}) - f'(\overline{M}_{0,i})\overline{M}_{0,i} \tag{12.17}$$

$$b_i = f'(\overline{M}_{0,i}) \tag{12.18}$$

In particular, for the power law (12.14) we get

$$a_i = (1-p)\bar{A}\left(\frac{\overline{M}_{0,i}}{\overline{M}_0}\right)^p \tag{12.19}$$

$$b_i = p\frac{\bar{A}}{\overline{M}_0}\left(\frac{\overline{M}_{0,i}}{\overline{M}_0}\right)^{p-1} \tag{12.20}$$

where $\overline{M}_0 = \overline{W}_0\sigma_0$ is the reference yield moment.

For members of a constant cross section, the total volume of the structure is

$$V = \sum_i A_i\ell_i = \sum_i a_i\ell_i + \sum_i b_i\ell_i M_{0,i} \tag{12.21}$$

Clearly, it is sufficient to minimize only the second term, $\sum_i b_i\ell_i M_{0,i}$, because the first term is constant. The cost function

$$z(\boldsymbol{s}_0) = \tilde{\boldsymbol{c}}^T\boldsymbol{s}_0 = \sum_i b_i\ell_i M_{0,i} \tag{12.22}$$

has the form (12.1), but this time with

$$\tilde{\boldsymbol{c}} = \{b_1\ell_1, b_2\ell_2, \dots b_n\ell_n\}^T \tag{12.23}$$

$$\boldsymbol{s}_0 = \{M_{0,1}, M_{0,2}, \dots M_{0,n}\}^T \tag{12.24}$$

After introducing the design constraints (12.4), we again end up with a cost function (12.5).

Setting up the problem, we can scale down the cost function by any positive factor without affecting the decision which design is optimal. This means that we need only the relative values of coefficients b_i for the individual cross sections, while a common factor can be eliminated by scaling down function z. A convenient property of the power law is that the ratio between two parameters b_i and b_j from (12.20),

$$\frac{b_i}{b_j} = \left(\frac{\overline{M}_{0,i}}{\overline{M}_{0,j}}\right)^{p-1} \tag{12.25}$$

depends only on one parameter of the law – the exponent, p. Even if no precise information on the specific sequence of hot-rolled steel profiles is available, we can approximately take $p = 0.6$.

It should be mentioned that, strictly speaking, the selection of the cross sectional profiles from a discrete catalog leads to a problem of *integer linear programming*, which is more difficult than the usual linear programming problem. This issue is, however, outside the scope of our book.

The optimum design can be based on the behavior of the structure under monotonic loading (*limit design*), or under cyclic loading (*shakedown design*). For each type of design, duality in linear programming permits four equivalent formulations, namely primal and dual based on the static theorem, and primal and dual based on the kiematic theorem. The choice of the most efficient formulation is not obvious; it depends on the nature of the problem (Maier, Srinivasan and Save, 1976). The

basic inconvenience of a true kinematic approach is that it is expressed in terms of deformation variables, and the objective function has the meaning of work, which is not a primary design objective. An alternative static approach using kinematically admissible constraints is possible, but of considerable practical difficulty in identifying all conceivable mechanisms or the active constraints at the specified limit load. For these reasons, we will focus attention on the standard static approach to optimum design.

12.2 LIMIT DESIGN

The aim of limit design is to find the design variables $\boldsymbol{s}_d$ such that the structure can safely resist a given reference loading $\bar{\boldsymbol{f}}$ with minimum cost, proportional to the objective function $z(\boldsymbol{s}_d)$.

According to the lower bound theorem of limit analysis, the design is safe if there is a vector of internal forces $\boldsymbol{s}$ that is plastically admissible and in equilibrium with the given load. This means that a safe design is characterized by a set of design variables $\boldsymbol{s}_d$ and internal forces $\boldsymbol{s}$ such that

$$\boldsymbol{B}^T\boldsymbol{s} = \bar{\boldsymbol{f}} \tag{12.26}$$

and

$$|\boldsymbol{s}| \le \boldsymbol{E}\boldsymbol{s}_d \tag{12.27}$$

The optimal design is obtained by looking for the safe design that minimizes the cost function $z(\boldsymbol{s}_d)$. This leads directly to the following formulation of the limit design as a linear programming problem:

$$\text{Minimize } z(\boldsymbol{s}_d, \boldsymbol{s}) \equiv \boldsymbol{c}^T\boldsymbol{s}_d \tag{12.28}$$

subject to

$$\boldsymbol{B}^T\boldsymbol{s} = \bar{\boldsymbol{f}} \tag{12.29}$$

$$\boldsymbol{E}\boldsymbol{s}_d - \boldsymbol{s} \ge \boldsymbol{0} \tag{12.30}$$

$$\boldsymbol{E}\boldsymbol{s}_d + \boldsymbol{s} \ge \boldsymbol{0} \tag{12.31}$$

Before we discuss the conversion of this problem to the standard form, let us illustrate the nature of limit design with an example solvable by simple means.

Example 12.2: Consider again the three-bar truss reproduced in Figure 12.2(a), this time subjected to given loads $F_x = \bar{F}$, $F_z = 2\bar{F}$. Find the optimum design satisfying the additional condition that bars 1 and 3 have the same cross section.

Solution: The design variables to be optimized are the plastic axial forces $S_{0,1}$ and $S_{0,2}$. The optimum design minimizes the cost function

$$\begin{aligned} z(S_{0,1}, S_{0,2}) &= (\ell_1 + \ell_3)S_{0,1} + \ell_2 S_{0,2} = (3L + 4L)S_{0,1} + 2.4LS_{0,2} \\ &= 7LS_{0,1} + 2.4LS_{0,2} \end{aligned} \tag{12.32}$$

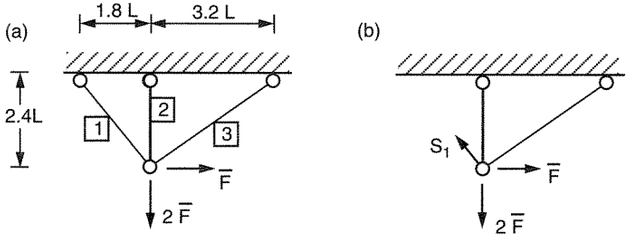


Figure 12.2 Truss to be optimized

The design is safe if there exists a set of axial forces S_1, S_2 and S_3 satisfying the equilibrium conditions

$$0.6S_1 - 0.8S_3 = \bar{F} \tag{12.33}$$

$$0.8S_1 + S_2 + 0.6S_3 = 2\bar{F} \tag{12.34}$$

and the conditions of plastic admissibility

$$|S_1| \le S_{1,0}, \qquad |S_2| \le S_{2,0}, \qquad |S_3| \le S_{1,0} \tag{12.35}$$

Obviously, if S_1, S_2 and S_3 are known, the most economical way of selecting the cross sections is to set

$$S_{1,0} = \max\left(|S_1|, |S_3|\right) \tag{12.36}$$

$$S_{2,0} = |S_2| \tag{12.37}$$

Consequently, the cost function can be rewritten in terms of the (yet unknown) axial forces as

$$z(S_1, S_2, S_3) = 7L \max\left(|S_1|, |S_3|\right) + 2.4L|S_2| \tag{12.38}$$

The structure is statically indeterminate to the first degree. So we select one of the axial forces as the redundant, and express the remaining forces from the equilibrium conditions (12.33)–(12.34). For example, moving the terms with S_1 to the right-hand side and solving for S_2 and S_3, we get

$$S_2 = 2.75\bar{F} - 1.25S_1 \tag{12.39}$$

$$S_3 = -1.25\bar{F} + 0.75S_1 \tag{12.40}$$

Of course, the same result could be obtained by solving the corresponding statically determinate structure (Figure 12.2(b)) subjected to the prescribed external loading and the statically redundant force S_1; this is the standard procedure of the *flexibility method* (force method).

Substituting (12.39) and (12.40) into (12.38), we get the cost function as a function of a single variable – the statically redundant force S_1:

$$z(S_1) = 7L \max\left(|S_1|, |-1.25\bar{F} + 0.75S_1|\right) + 2.4L\,|2.75\bar{F} - 1.25S_1| \tag{12.41}$$

This function is piecewise linear, and it is therefore easy to construct its graph; see Figure 12.3(b). Since $\bar{F}$ is a given force, we can plot the dimensionless cost function

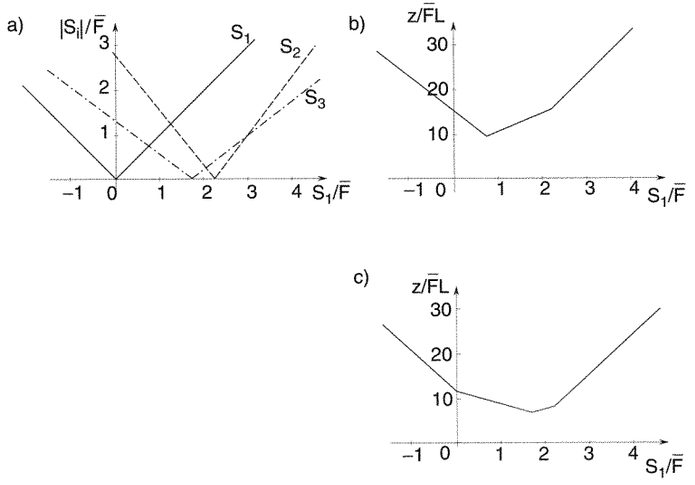


Figure 12.3 a) Absolute values of axial forces as functions of redundant force, b) cost function with design constraint $S_{0,1} = S_{0,3}$, c) cost function without design constraints

$z/\bar{F}L$ against the dimensionless variable $S_1/\bar{F}$. The cost function changes its slope only if one of the axial forces (the expressions inside the absolute value operators) changes its sign, or if the two expressions behind the maximum operator become equal. It is clear from the graph that the cost function is convex and assumes its minimum value at the point where $S_1 = -S_3$, i.e. at $S_1 = 0.714\bar{F}$. The optimum design is thus $S_{0,1} = S_{0,3} = 0.714\bar{F}$, $S_{0,2} = 1.857\bar{F}$ and $z = 7LS_{0,1} + 2.4LS_{0,2} = 9.457\bar{F}L$. □

Note that if we did not impose any design constraints, the cost function would change its slope only at points where one of the axial forces vanishes (Figure 12.3(c)), and the optimal design would consist of only two bars with a positive cross sectional area, i.e. the optimal structure would be statically determinate. A generalization of this result is left to the reader as an exercise (see Problems 12.1 and 12.2).

To convert the linear programming problem (12.28)–(12.31) to the standard form, we can use the technique that has already been explained in Chapter 7. The internal forces $\boldsymbol{s}$ are decomposed into the positive and negative parts. In terms of the new variables

$$\boldsymbol{s}^+ = (|\boldsymbol{s}| + \boldsymbol{s})/2, \qquad \boldsymbol{s}^- = (|\boldsymbol{s}| - \boldsymbol{s})/2 \tag{12.42}$$

we can express $\boldsymbol{s}$ and its absolute value as

$$\boldsymbol{s} = \boldsymbol{s}^+ - \boldsymbol{s}^-, \qquad |\boldsymbol{s}| = \boldsymbol{s}^+ + \boldsymbol{s}^- \tag{12.43}$$

and transform equilibrium equations (12.29) into

$$\boldsymbol{B}^T\boldsymbol{s}^+ - \boldsymbol{B}^T\boldsymbol{s}^- = \bar{\boldsymbol{f}} \tag{12.44}$$

The conditions of plastic admissibility (12.27), which are fully equivalent to (12.30)–(12.31), can be transformed into the conditions

$$\boldsymbol{s}^+ + \boldsymbol{s}^- \le \boldsymbol{E}\boldsymbol{s}_d \tag{12.45}$$

To transform the inequality constraints (12.45) into equality constraints as required by the standard form, we introduce the vector of slack variables

$$\boldsymbol{\Delta s} = \boldsymbol{E s}_d - (\boldsymbol{s}^+ + \boldsymbol{s}^-) \tag{12.46}$$

The basic unknowns of the problem now consist of the original design variables, $\boldsymbol{s}_d$, variables $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$ that have replaced $\boldsymbol{s}$, and the slack variables, $\boldsymbol{\Delta s}$. Note that conditions (12.45) are equivalent to $\boldsymbol{\Delta s} \geq \boldsymbol{0}$, and that the variables $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$ are also nonnegative due to their definition (12.42). The design variables $\boldsymbol{s}_d$ were not explicitly required to be nonnegative in the original formulation of the problem, but they can never become negative due to their physical meaning, and so the condition $\boldsymbol{s}_d \geq \boldsymbol{0}$ can be added without affecting the solution. In summary, we have the following linear programming problem in the standard form:

$$\text{Minimize } z(\boldsymbol{s}_d, \boldsymbol{s}^+, \boldsymbol{s}^-, \boldsymbol{\Delta s}) \equiv \boldsymbol{c}^T \boldsymbol{s}_d \tag{12.47}$$

subject to

$$\boldsymbol{B}^T \boldsymbol{s}^+ - \boldsymbol{B}^T \boldsymbol{s}^- = \bar{\boldsymbol{f}} \tag{12.48}$$

$$\boldsymbol{E s}_d - \boldsymbol{s}^+ - \boldsymbol{s}^- - \boldsymbol{\Delta s} = \boldsymbol{0} \tag{12.49}$$

$$\boldsymbol{s}_d \geq \boldsymbol{0}, \quad \boldsymbol{s}^+ \geq \boldsymbol{0}, \quad \boldsymbol{s}^- \geq \boldsymbol{0}, \quad \boldsymbol{\Delta s} \geq \boldsymbol{0} \tag{12.50}$$

Example 12.3: Present the limit design problem from Example 12.2 as a problem of linear programming, and convert it to the standard form.

Solution: We will start by rewriting the formulation (12.32)–(12.35) from Example 12.2 in the matrix form. The vector of design variables and the vector of internal forces are respectively given by

$$\boldsymbol{s}_d = \begin{Bmatrix} S_{0,1} \\ S_{0,2} \end{Bmatrix}, \quad \boldsymbol{s} = \begin{Bmatrix} S_1 \\ S_2 \\ S_3 \end{Bmatrix} \tag{12.51}$$

The cost function (12.32) can be written in the matrix form

$$z = \boldsymbol{c}^T \boldsymbol{s}_d = \{7L \;\; 2.4L\} \begin{Bmatrix} S_{0,1} \\ S_{0,2} \end{Bmatrix} \tag{12.52}$$

and the constraints (12.33)–(12.35) read

$$\begin{bmatrix} 0.6 & 0 & -0.8 \\ 0.8 & 1 & 0.6 \end{bmatrix} \begin{Bmatrix} S_1 \\ S_2 \\ S_3 \end{Bmatrix} = \begin{Bmatrix} \bar{F} \\ 2\bar{F} \end{Bmatrix} \tag{12.53}$$

$$\begin{Bmatrix} |S_1| \\ |S_2| \\ |S_3| \end{Bmatrix} \leq \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{Bmatrix} S_{0,1} \\ S_{0,2} \end{Bmatrix} \tag{12.54}$$

The constraints (12.35) containing absolute values could be rewritten in the form of linear inequalities (12.30)–(12.31) required by linear programming. However, we will bypass this step and proceed directly to the standard form (12.47)–(12.50). Following the general scheme, we decompose the axial forces into their positive and negative parts, $S_i = S_i^+ - S_i^-, i = 1, 2, 3$, which allows us to express their absolute values as $|S_i| = S_i^+ + S_i^-, i = 1, 2, 3$. Furthermore, we introduce the slack variables $\Delta S_i, i = 1, 2, 3$, corresponding to the difference of the right-hand side and the left-hand side in (12.54). For convenience, we can scale the coefficients in the cost function by $1/L$. As a result, we obtain the problem

$$\text{Minimize } z(S_{0,1}, \ldots \Delta S_3) \equiv \{7 \;\; 2.4\} \begin{Bmatrix} S_{0,1} \\ S_{0,2} \end{Bmatrix} \tag{12.55}$$

subject to

$$\begin{bmatrix} 0.6 & 0 & -0.8 \\ 0.8 & 1 & 0.6 \end{bmatrix} \begin{Bmatrix} S_1^+ \\ S_2^+ \\ S_3^+ \end{Bmatrix} + \begin{bmatrix} -0.6 & 0 & 0.8 \\ -0.8 & -1 & -0.6 \end{bmatrix} \begin{Bmatrix} S_1^- \\ S_2^- \\ S_3^- \end{Bmatrix} = \begin{Bmatrix} \bar{F} \\ 2\bar{F} \end{Bmatrix} \tag{12.56}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{Bmatrix} S_{0,1} \\ S_{0,2} \end{Bmatrix} - \begin{Bmatrix} S_1^+ \\ S_2^+ \\ S_3^+ \end{Bmatrix} - \begin{Bmatrix} S_1^- \\ S_2^- \\ S_3^- \end{Bmatrix} - \begin{Bmatrix} \Delta S_1 \\ \Delta S_2 \\ \Delta S_3 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix} \tag{12.57}$$

where all the unknown variables are non-negative. □

Example 12.4: Solve the optimization problem from Example 12.3 by the simplex method.

Solution: Based on the description of the problem derived in the previous example, we easily set up the initial simplex tableau; see Table 12.1. The crucial step is to find a basic feasible solution to serve as a start. The technique explained in Section 7.5 in the context of limit analysis is easily adapted to the present problem of limit design. Recall that the aim is to select the basic variables such that, for zero values of the remaining (nonbasic) variables, the basic variables could be uniquely computed from the constraints, and that their values be non-negative. In our particular problem, we have 11 variables and five constraints, and therefore we need to set six variables to zero.

Table 12.1 Initial simplex tableau

$S_{0,1}$	$S_{0,2}$	S_1^+	S_2^+	S_3^+	S_1^-	S_2^-	S_3^-	ΔS_1	ΔS_2	ΔS_3	RHS/$\bar{F}$
0	0	0.6	0	−0.8	−0.6	0	0.8	0	0	0	1
0	0	0.8	1	0.6	−0.8	−1	−0.6	0	0	0	2
1	0	−1	0	0	−1	0	0	−1	0	0	0
0	1	0	−1	0	0	−1	0	0	−1	0	0
1	0	0	0	−1	0	0	−1	0	0	−1	0
7	2.4	0	0	0	0	0	0	0	0	0	0

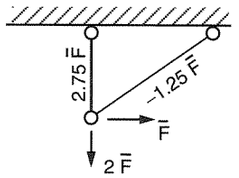


Figure 12.4 Primary structure

Let us begin by examining the first two constraints, which represent the equations of equilibrium. They can be satisfied by axial forces evaluated on any primary structure obtained from the actual structure by deleting a statically redundant bar. Consider for example the structure obtained by leaving out bar number 1; see Figure 12.4. The axial forces in the remaining bars are $S_2 = 2.75\bar{F}$ and $S_3 = -1.25\bar{F}$. To end up with positive values of the computed basic variables, we select S_2^+ and S_3^- as the first two members of the basis while the variables S_1^+, S_3^+, S_1^- and S_2^- stay out of the basis.

Now we have to specify additional three basic variables. For the axial forces just computed, the corresponding design variables $S_{0,1}$ and $S_{0,2}$ cannot be smaller than $\max(|S_1|, |S_3|) = |S_3| = 1.25\bar{F}$ and $|S_2| = 2.75\bar{F}$, respectively. This means that, being nonzero, both $S_{0,1}$ and $S_{0,2}$ must be basic variables. Recall that $\Delta S_1 = S_{0,1} - |S_1|$, $\Delta S_2 = S_{0,2} - |S_2|$, and $\Delta S_3 = S_{0,1} - |S_3|$. The variable ΔS_2 can be set to zero, and $S_{0,2}$ is then computed from the fourth constraint, which corresponds to the condition $\Delta S_2 = S_{0,2} - |S_2|$. On the other hand, as $|S_1| = 0$ and $|S_3| = 1.25\bar{F} > 0$, we cannot set $\Delta S_1 = 0$ because then $S_{0,1}$ would have to be zero as well and ΔS_3 would have to be negative. Therefore, ΔS_1 is the fifth basic variable and ΔS_3 stays out of the basis.

Indeed, selecting $S_{0,1}$, $S_{0,2}$, S_2^+, S_3^- and ΔS_1 as the basic variables, we obtain a feasible solution presented in Table 12.2. A negative value in the third column of the last row indicates that the solution is not optimal. The standard procedure detects the element in the fifth row of the third column as the key element. Consequently, ΔS_1 will leave the basis while S_1^+ will enter it. After the corresponding simplex transformation we obtain the solution in Table 12.3, which turns out to be optimal. The values of the basic variables $S_{0,1} = 0.714\bar{F}$, $S_{0,2} = 1.857\bar{F}$, $S_1^+ = 0.714\bar{F}$, $S_2^+ = 1.857\bar{F}$ and $S_3^- = 0.714\bar{F}$ are in perfect agreement with the results obtained in Example 12.2 using a simple approach. □

The strategy leading to the basic feasible solution in the preceding example is quite universal. In general, there are $3m + d$ variables and $2m - s$ constraints, where m

Table 12.2 Initial feasible solution

$S_{0,1}$	$S_{0,2}$	S_1^+	S_2^+	S_3^+	S_1^-	S_2^-	S_3^-	ΔS_1	ΔS_2	ΔS_3	RHS/$\bar{F}$
0.0	0.0	0.75	0.0	−1.0	−0.75	0.0	1.0	0.0	0.0	0.0	1.25
0.0	0.0	1.25	1.0	0.0	−1.25	−1.0	0.0	0.0	0.0	0.0	2.75
1.0	0.0	0.75	0.0	−2.0	−0.75	0.0	0.0	0.0	0.0	−1.0	1.25
0.0	1.0	1.25	0.0	0.0	−1.25	−2.0	0.0	0.0	−1.0	0.0	2.75
0.0	0.0	1.75	0.0	−2.0	0.25	0.0	0.0	1.0	0.0	−1.0	1.25
0.0	0.0	−8.25	0.0	14.0	8.25	4.8	0.0	0.0	2.4	7.0	−15.35

Table 12.3 Optimal solution

$S_{0,1}$	$S_{0,2}$	S_1^+	S_2^+	S_3^+	S_1^-	S_2^-	S_3^-	ΔS_1	ΔS_2	ΔS_3	RHS/$\bar{F}$
0.0	0.0	0.0	0.0	−0.143	−0.857	0.0	1.0	−0.429	0.0	0.429	0.714
0.0	0.0	0.0	1.0	1.429	−1.429	−1.0	0.0	−0.714	0.0	0.714	1.857
1.0	0.0	0.0	0.0	−1.143	−0.857	0.0	0.0	−0.429	0.0	−0.571	0.714
0.0	1.0	0.0	0.0	1.429	−1.429	−2.0	0.0	−0.714	−1.0	0.714	1.857
0.0	0.0	1.0	0.0	−1.143	0.143	0.0	0.0	0.571	0.0	−0.571	0.714
0.0	0.0	0.0	0.0	4.571	9.429	4.8	0.0	4.714	2.4	2.286	−9.457

is the number of independent internal forces (for a truss, m is the number of bars, while for a frame, m is the number of critical cross sections), d is the number of design variables, and s is the degree of static redundancy. We can begin by selecting s internal forces as the redundants. The values of the remaining internal forces follow from the equations of equilibrium on the corresponding primary structure. Depending on the sign of each force, we select $m-s$ basic variables as the positive or negative parts of the computed internal forces. Then we evaluate the minimum possible values of the design variables for which the solution is plastically admissible. For each design variable, at least one of the yield conditions is satisfied (otherwise the value of this design variable could be decreased without violating the conditions of plastic admissibility). The slack variables corresponding to these tight conditions are set to zero while the remaining slack variables become members of the basis. Hence, the basis consists of $m - s$ variables representing the nonzero internal forces, d design variables, and $m - d$ slack variables, which add up to a total of exactly $2m - s$ basic variables as required. Their values computed from the constraints are guaranteed to be nonnegative.

The approach that has been demonstrated can be applied to limit design of beams and frames without any difficulty, and it can be fully computerized. The next example will show that the number of variables and constraints resulting from a straightforward application of the general algorithm can be reduced by examining the structure more carefully. The algorithmic approach, however, allows easy implementation into a computer program. On the other hand, the approach making use of special knowledge is advantageous for a designer trying to solve a simple problem by hand. The reduction of variables and constraints could, of course, be described algorithmically and introduced into the program, but it is questionable whether the savings would be so dramatic as to justify the increased program complexity. As will be seen, one can save three variables and two constraints per joint that connects two members with the same properties, and two variables and one constraint per joint that connects two members with different properties. In both cases, it is assumed that the joint is not directly loaded by an external moment.

Example 12.5: Find the optimum design of the frame in Figure 12.5a. Each of the three straight segments (two columns and a beam) should have a constant profile along its entire length.

Solution: Let us divide the frame into four members as shown in Figure 12.5(a). Note that, in addition to the physical joints of the structure, a computational node

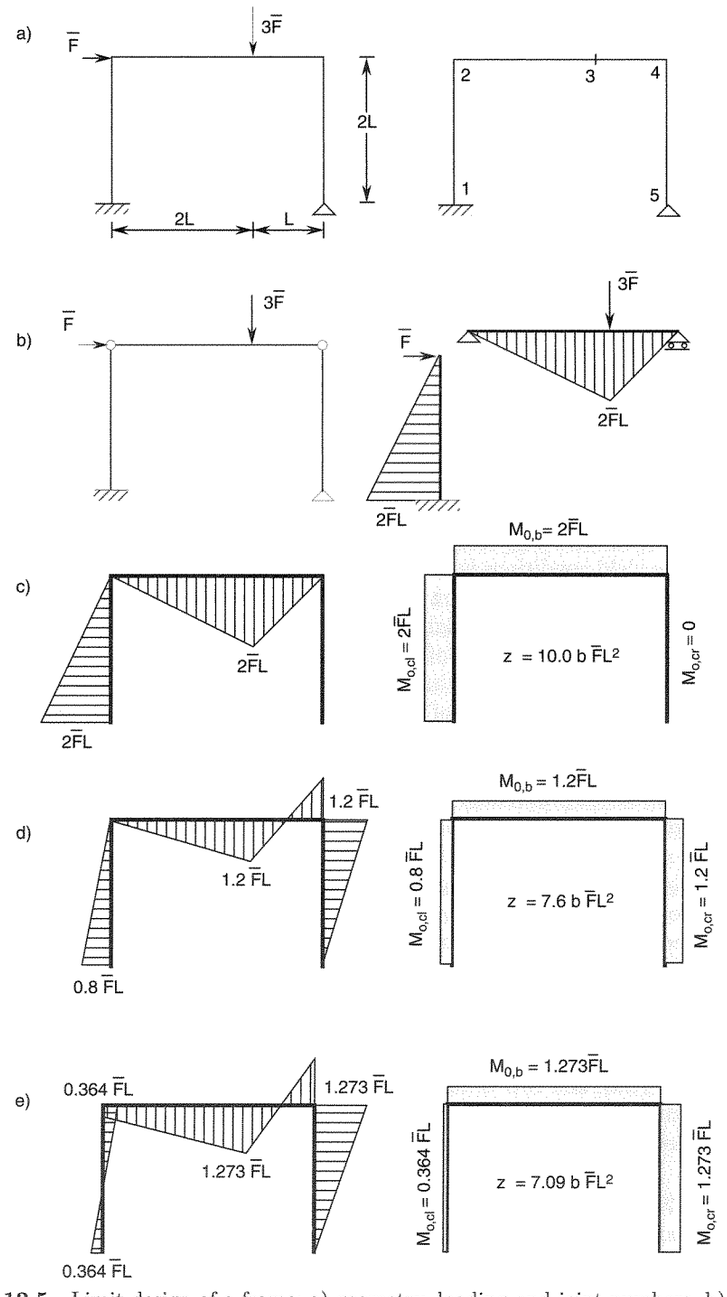


Figure 12.5 Limit design of a frame: a) geometry, loading and joint numbers, b) primary structure, c) initial design, d) improved design, e) optimal design

is formally introduced under the applied vertical load. A fully automatic procedure would treat the end moments at all end sections of the members as unknown internal forces, with the exception of the moment M_{54} at the hinge-supported section 5, which is a priori known to be zero. Hence, the vector of internal forces

$$\boldsymbol{s} = \{M_{12}, M_{21}, M_{23}, M_{32}, M_{34}, M_{43}, M_{45}\}^T \tag{12.58}$$

would contain seven unknown moments. As the structure is statically indeterminate to the second degree, these moments are linked by five independent equilibrium equations, which could be constructed, e.g. by applying the principle of virtual work to three joint mechanisms, one beam mechanism, and one panel mechanism. From the specification of the design requirements it follows that three design variables need to be introduced – the plastic moment of the beam, M_{0b}, and the plastic moments of the columns, M_{0cl} and M_{0cr}, where subscripts cl and cr stand for the left column and right column, respectively. Converting the problem to the standard form leads to $3 \times 7 + 3 = 24$ fundamental variables and $5 + 7 = 12$ equality constraints.

A structural analyst looking at the structure would no doubt introduce only one unknown moment at each of the joints, except for the hinge-supported one. Formally this follows from the moment equilibrium conditions at joints 2, 3 and 4 which, in terms of the end moments, read $M_{21} + M_{23} = 0$, $M_{32} + M_{34} = 0$ and $M_{43} + M_{45} = 0$. This means that, making use of the three equations that correspond to the joint mechanisms, we can, for example, reduce the vector of internal forces to

$$\boldsymbol{s} = \{M_{12}, M_{21}, M_{32}, M_{45}\}^T \tag{12.59}$$

The number of equilibrium conditions that remain to be satisfied decreases to two. However, it is important to realize that not all the conditions of plastic admissibility associated with the eliminated moments can be left out. Without knowing in advance which of the members will be stronger, we must make sure that the moments at the corners are plastically admissible both on the beam side and the column side. Only at joint 3, which connects two members with the same properties, it suffices to check only one moment. Hence the reduced formulation deals with $2 \times 4 + 6 + 3 = 17$ variables and $2 + 6 = 8$ constraints.

It is now quite easy to set up the initial simplex tableau. The first two rows of Table 12.4 represent the equilibrium conditions

$$0.5M_{12} + 0.5M_{21} + 0.5M_{45} = \bar{F}L \tag{12.60}$$

$$-0.5M_{21} + 1.5M_{32} + M_{45} = 3\bar{F}L \tag{12.61}$$

derived by applying the principle of virtual work, respectively, to the panel mechanism and the beam mechanism. The next six rows represent the conditions of plastic admissibility at sections 12, 21, 23, 3, 43 and 45. The last row contains the coefficients of the cost function. As we have no information on the expected values of the plastic moments, we assume that they do not differ substantially, and therefore we take the same coefficient $b_i = b$ for each member; cf. (12.18).

The selection of the initial basis can follow the scheme developed in Example 12.4 and later generalized. First, we select an appropriate number of internal forces as the redundants, and determine the values of the remaining internal forces from equilibrium. In our case, we can for example set $M_{21} = 0$ and $M_{45} = 0$ and then

solve the primary structure in Figure 12.5(b), which can be decomposed into a simply supported beam and a cantilever. The moments $M_{12} = 2\bar{F}L$ and $M_{32} = 2\bar{F}L$ could of course be solved directly from (12.60)–(12.61). As both moments are positive, we will select M_{12}^+ and M_{32}^+ as two of the basic variables. The basis will also include all the three design variables, M_{0cl}, M_{0cr} and M_{0b}, and the slack variables corresponding to the conditions of plastic admissibility with the exclusion of the tight ones (i.e. those satisfied as yield conditions with an equality sign). In our case, we have $M_{0cl} = \max(|M_{12}|, |M_{21}|) = |M_{12}|$, $M_{0b} = \max(|M_{23}|, |M_{32}|, |M_{43}|) = |M_{32}|$ and $M_{0cr} = |M_{45}|$, and so the slack variables that stay out of the basis are ΔM_{12}, ΔM_{32} and ΔM_{45}.

Since each simplex tableau for this problem is quite large, we leave out the intermediate ones and present only the optimal solution in Table 12.5. Rather, it is more instructive to look at the physical interpretation of each tableau. Figures 12.5(c)–(e) show the bending moment diagrams for the individual stages of the optimization process. The initial basic feasible solution corresponds to the design in Figure 12.5(c). Note that this design would not require any bending resistance of the right column. In reality, the cross section cannot totally disappear because the column must resist some axial force that is not accounted for in the present structural model. The value of the cost function for the initial design is

$$\begin{aligned} z &= b\sum l_i M_{0,i} = b(2LM_{0cl} + 3LM_{0b} + 2LM_{0cr}) \\ &= b(2L \times 2\bar{F}L + 3L \times 2\bar{F}L + 2L \times 0) = 10b\bar{F}L^2 \end{aligned} \tag{12.62}$$

The optimization process requires two cycles of the simplex method. In the first cycle, M_{45}^+ replaces ΔM_{43} in the basis. Physically this means that the moment at the upper right corner is increased so as to make it equal to the moment below the vertical load, and thus better exploit the plastic capacity of the beam; see Figure 12.5(d). This allows us to reduce the cross section of the beam. At the same time, we have to increase the cross section of the right column in order to guarantee plastic admissibility at the column section of the right corner, but this is compensated for by a decrease of the cross section of the left column where the maximum moment drops substantially. The value of the cost function is reduced to

$$z = b(2L \times 0.8\bar{F}L + 3L \times 1.2\bar{F}L + 2L \times 1.2\bar{F}L) = 7.6b\bar{F}L^2 \tag{12.63}$$

In the next cycle, M_{21}^+ replaces ΔM_{21} in the basis, which means that the moment at the upper left corner is increased to become equal to the moment at the fixed support; see Figure 12.5(e). This allows us to make better use of the plastic capacity of the left column, whose cross section can further be decreased while the other cross sections slightly increase. The value of the cost function attains its minimum possible value

$$z = b(2L \times 0.364\bar{F}L + 3L \times 1.273\bar{F}L + 2L \times 1.273\bar{F}L) = 7.09b\bar{F}L^2 \tag{12.64}$$

□

In the previous example, the optimum design was constructed under the assumption that the coefficients b_i from (12.15) are the same for all the members. Recall that these coefficients come from the linear approximation of the relationship between the cross sectional area and the plastic bending moments. We could now improve the accuracy of this approximation by linearizing the relationship separately around each plastic moment predicted by the first design.

Table 12.4 Initial simplex tableau

M_{0cl}	M_{0b}	M_{0cr}	M_{12}^+	M_{21}^+	M_{32}^+	M_{45}^+	M_{12}^-	M_{21}^-	M_{32}^-	M_{45}^-	ΔM_{12}	ΔM_{21}	ΔM_{23}	ΔM_{32}	ΔM_{43}	ΔM_{45}	RHS/$\bar{F}L$
0	0	0	0.5	0.5	0.0	0.5	−0.5	−0.5	0.0	−0.5	0	0	0	0	0	0	1
0	0	0	0.0	−0.5	1.5	1.0	0.0	0.5	−1.5	−1.0	0	0	0	0	0	0	3
1	0	0	−1.0	0.0	0.0	0.0	−1.0	0.0	0.0	0.0	−1	0	0	0	0	0	0
1	0	0	0.0	−1.0	0.0	0.0	0.0	−1.0	0.0	0.0	0	−1	0	0	0	0	0
0	1	0	0.0	−1.0	0.0	0.0	0.0	−1.0	0.0	0.0	0	0	−1	0	0	0	0
0	1	0	0.0	0.0	−1.0	0.0	0.0	0.0	−1.0	0.0	0	0	0	−1	0	0	0
0	1	0	0.0	0.0	0.0	−1.0	0.0	0.0	0.0	−1.0	0	0	0	0	−1	0	0
0	0	1	0.0	0.0	0.0	−1.0	0.0	0.0	0.0	−1.0	0	0	0	0	0	−1	0
2	3	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0	0	0	0	0	0

Table 12.5 Optimal solution

M_{0cl}	M_{0b}	M_{0cr}	M_{12}^{+}	M_{21}^{+}	M_{32}^{+}	M_{45}^{+}	M_{12}^{-}	M_{21}^{-}	M_{32}^{-}	M_{45}^{-}	ΔM_{12}	ΔM_{21}	ΔM_{23}	ΔM_{32}	ΔM_{43}	ΔM_{45}	RHS/$\bar{F}L$
0	0	0	0	0	1	0	−0.182	0.182	−0.091	−0.909	−0.091	0.091	0	0.455	−0.455	0	1.273
0	0	0	1	0	0	0	0.091	−1.091	0.545	−0.545	0.545	−0.545	0	0.273	−0.273	0	0.364
1	0	0	0	0	0	0	−0.909	−1.091	0.545	−0.545	−0.455	−0.545	0	0.273	−0.273	0	0.364
0	0	0	0	1	0	0	−0.909	−0.091	0.545	−0.545	−0.455	0.455	0	0.273	−0.273	0	0.364
0	1	0	0	0	0	0	−0.182	0.182	−1.091	−0.909	−0.091	0.091	0	−0.545	−0.455	0	1.273
0	0	0	0	0	0	0	0.727	1.273	−1.636	−0.364	0.364	−0.364	1	−0.818	−0.182	0	0.909
0	0	0	0	0	0	1	−0.182	0.182	−1.091	0.091	−0.091	0.091	0	−0.545	0.545	0	1.273
0	0	1	0	0	0	0	−0.182	0.182	−1.091	−0.909	−0.091	0.091	0	−0.545	0.545	−1	1.273
0	0	0	0	0	0	0	2.727	1.273	4.364	5.636	1.364	0.636	0	2.182	0.818	2	−7.091

Example 12.6: Increase the accuracy of the optimum design from Example 12.5 by using the results of that example as initial estimates for the determination of the sensitivity parameters. Assume that the relation between the area and the plastic moment is described by (12.14) with $p = 0.6$.

Solution: Recall that in Example 12.5 we ended up with the design values $M_{0cl} = 0.364\bar{F}L$ and $M_{0b} = M_{0cr} = 1.273\bar{F}L$. Since, according to (12.20),

$$\frac{b_{cl}}{b_{cr}} = \left(\frac{M_{0cl}}{M_{0cr}}\right)^{p-1} = \left(\frac{0.364}{1.273}\right)^{-0.4} = 1.65 \tag{12.65}$$

$$\frac{b_b}{b_{cr}} = \left(\frac{M_{0b}}{M_{0cr}}\right)^{p-1} = \left(\frac{1.273}{1.273}\right)^{-0.4} = 1 \tag{12.66}$$

the new cost function is given by

$$\begin{aligned}\tilde{z} &= \sum_{i=1}^{3} b_i l_i M_{0,i} = b_{cr}(1.65 \times 2LM_{0cl} + 1 \times 3LM_{0b} + 2LM_{0cr}) \\ &= b_{cr}L(3.3M_{0cl} + 3M_{0b} + 2M_{0cr})\end{aligned} \tag{12.67}$$

We have denoted the new cost function by $\tilde{z}$ to distinguish it from the original cost function

$$z = \sum_{i=1}^{3} bl_i M_{0,i} = bL(2M_{0cl} + 3M_{0b} + 2M_{0cr}) \tag{12.68}$$

It is important to realize that there is no need to rerun the complete analysis with the new cost function. First, the basis that provided the optimal solution for the original cost function will give at least a feasible (maybe even the optimal) solution for the new one. So we could start from the initial tableau and go directly to the basic solution corresponding to the optimal basis from the first analysis. However, we do not even have to perform this transformation. Realizing that the only difference between the original and the new problem resides in the definition of the cost function, we can use the final tableau from the original analysis and modify the bottom row that represents the cost function. Note that all the rows except for the bottom one are linear combinations of the constraints, and as such they are the same for both problems. The bottom row in Table 12.5 describes the original cost function as

$$\begin{aligned}z = bL(&7.091\bar{F}L + 2.727M_{12}^{-} + 1.273M_{21}^{-} + 4.364M_{32}^{-} + 5.636M_{45}^{-} \\ &+1.364\Delta M_{12} + 0.636\Delta M_{21} + 2.182\Delta M_{32} + 0.818\Delta M_{43} + 2\Delta M_{45})\end{aligned} \tag{12.69}$$

The value of the coefficient b in the original cost function was unspecified, and so we can set it equal to the coefficient b_{cr} (alternatively, we could drop the factors bL and $b_{cr}L$ from the cost functions and minimize only the linear combination of design moments in the parentheses). Then, subtracting (12.68) from (12.67), we get the difference

$$\tilde{z} - z = b_{cr}L \times 1.3M_{0cl} \tag{12.70}$$

The third row of Table 12.5 contains the expression of M_{0cl} in terms of the nonbasic variables,

$$M_{0cl} = 0.364\bar{F}L + 0.909M_{12}^- + 1.091M_{21}^- - 0.545M_{32}^- + 0.545M_{45}^- \\ +0.455\Delta M_{12} + 0.545\Delta M_{21} - 0.273\Delta M_{32} + 0.273\Delta M_{43} \qquad (12.71)$$

Combining (12.69)–(12.71) we arrive at an expression for the new cost function in terms of the nonbasic variables,

$$\tilde{z} = b_{cr}L(7.564\bar{F}L + 3.909M_{12}^- + 2.691M_{21}^- + 3.656M_{32}^- + 6.345M_{45}^- \\ +1.956\Delta M_{12} + 1.345\Delta M_{21} + 1.827\Delta M_{32} + 1.173\Delta M_{43} + 2\Delta M_{45}) \qquad (12.72)$$

The fact that all the coefficients in this expression are positive indicates that the solution is optimal even with the new cost function. □

The manipulations that we have done here in several steps in order to show their physical meaning can easily be described algorithmically. We simply re-evaluate the coefficients of the cost function, and add the differences between the new coefficients and the old ones to those entries of the bottom row in the final simplex tableau that correspond to the design variables. Then we eliminate the nonzero coefficients at basic variables using the other rows of the tableau. If the coefficients at the nonbasic variables remain non-negative, the solution is optimal even with the new cost function. If this is not the case, we proceed with the usual simplex transformations. Should the resulting optimal design differ substantially from the previous one, we could re-evaluate the coefficients of the cost function again, and repeat the whole procedure.

12.3 SHAKEDOWN DESIGN

The limit design expounded in the preceding section optimizes the properties of the structure with respect to a single loading case. In practice, the loads acting on a structure usually vary within a certain load domain. This aspect has already been discussed in Chapter 9, where the concept of shakedown has been introduced. We will now show that the structure can be tuned so that it is optimal with respect to arbitrary loading histories within a given domain in the load space.

Similar to limit design, we will adopt the static approach. Recall that, according to Melan's (lower bound) theorem stated in Chapter 10, the structure is safe against incremental collapse (and against plastic fatigue on the structural level) if there is a shakedown distribution of internal forces, i.e. a residual internal force vector $\boldsymbol{s}_r$ that is self-equilibrated,

$$\boldsymbol{B}^T\boldsymbol{s}_r = \boldsymbol{0} \qquad (12.73)$$

and satisfies the conditions of plastic admissibility

$$-\boldsymbol{E}\boldsymbol{s}_d \le \boldsymbol{s}_r + \boldsymbol{s}_e^{min}, \quad \boldsymbol{s}_r + \boldsymbol{s}_e^{max} \le \boldsymbol{E}\boldsymbol{s}_d \qquad (12.74)$$

where $\boldsymbol{s}_e^{min}$ and $\boldsymbol{s}_e^{max}$ are the extreme values of internal forces calculated for any load from the given load domain under the assumption that the response remains purely elastic. Again, we have imposed additional design constraints by expressing the plastic

internal forces as $\boldsymbol{s}_0 = \boldsymbol{E}\boldsymbol{s}_d$, where $\boldsymbol{s}_d$ is a vector of design variables and $\boldsymbol{E}$ is a given technological matrix.

The aim of shakedown design is to find, among all the safe designs, that for which the total weight of the structure is minimized. Similar to limit design, we assume the weight to be proportional to the cost function $z(\boldsymbol{s}_d) = \boldsymbol{c}^T\boldsymbol{s}_d$. The design variables $\boldsymbol{s}_d$ are fundamental unknowns of the problem, along with the shakedown forces $\boldsymbol{s}_r$. The resulting problem can be stated as

$$\text{Minimize } z(\boldsymbol{s}_d, \boldsymbol{s}_r) \equiv \boldsymbol{c}^T\boldsymbol{s}_d \qquad (12.75)$$

subject to (12.73)–(12.74). This looks again like a problem of linear programming. However, the elastically computed internal forces $\boldsymbol{s}_e^{min}$ and $\boldsymbol{s}_e^{max}$ that appear in (12.74) depend on the relative stiffnesses of individual members, and so they are functions of the design variables. To facilitate the solution, we will consider them as constants, bearing in mind that this is a gross simplification and that the actual shakedown design problem is intrinsically nonlinear. Further comments will be given at the end of this chapter.

Comparing (12.73)–(12.75) to (12.28)–(12.31), we realize that the limit design is covered by the present problem as a special case. Indeed, for limit design, the load domain shrinks to a single point in the load space, and the extreme elastically computed internal forces $\boldsymbol{s}_e^{min}$ and $\boldsymbol{s}_e^{max}$ are the same, $\boldsymbol{s}_e^{min} = \boldsymbol{s}_e^{max} = \boldsymbol{D}_0\boldsymbol{B}\boldsymbol{K}_0^{-1}\bar{\boldsymbol{f}}$. Denoting $\boldsymbol{s} = \boldsymbol{s}_r + \boldsymbol{s}_e^{min} = \boldsymbol{s}_r + \boldsymbol{s}_e^{max}$ and substituting into (12.73)–(12.74), we recover the constraints (12.29)–(12.31) of the limit design problem.

The present shakedown design problem can be converted to the standard form using the same principal strategy as in converting the limit design problem. The only tricky point is the definition of the new variables $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$. It is not a good idea to set $\boldsymbol{s}^+ = (|\boldsymbol{s}_r| + \boldsymbol{s}_r)/2$ and $\boldsymbol{s}^- = (|\boldsymbol{s}_r| - \boldsymbol{s}_r)/2$, in direct analogy to (12.42), because then the conditions of plastic admissibility (12.74) would require two different sets of slack variables, and the size of the problem would increase more than necessary. A better approach starts from replacing the extreme elastic moments, $\boldsymbol{s}_e^{max}$ and $\boldsymbol{s}_e^{min}$, by their average,

$$\boldsymbol{s}_e^{ave} = \frac{1}{2}(\boldsymbol{s}_e^{max} + \boldsymbol{s}_e^{min}) \qquad (12.76)$$

and one half of their difference,

$$\boldsymbol{s}_e^{amp} = \frac{1}{2}(\boldsymbol{s}_e^{max} - \boldsymbol{s}_e^{min}) \qquad (12.77)$$

Note that $\boldsymbol{s}_e^{amp} = \boldsymbol{s}_e^{max} - \boldsymbol{s}_e^{ave} = \boldsymbol{s}_e^{ave} - \boldsymbol{s}_e^{min}$ has the meaning of 'amplitudes' of elastic moments, i.e. of the differences between their mean and extreme values. The actual internal forces vary in the limits $\boldsymbol{s}_r + \boldsymbol{s}_e^{ave} \pm \boldsymbol{s}_e^{amp}$, and so the conditions of plastic admissibility (12.74) are equivalent to

$$|\boldsymbol{s}_r + \boldsymbol{s}_e^{ave}| \le \boldsymbol{E}\boldsymbol{s}_d - \boldsymbol{s}_e^{amp} \qquad (12.78)$$

If we decompose $\boldsymbol{s}_r + \boldsymbol{s}_e^{ave}$ into the positive and negative parts,

$$\boldsymbol{s}^+ = \frac{1}{2}\left(|\boldsymbol{s}_r + \boldsymbol{s}_e^{ave}| + \boldsymbol{s}_r + \boldsymbol{s}_e^{ave}\right) \qquad (12.79)$$

$$\boldsymbol{s}^- = \frac{1}{2}\left(|\boldsymbol{s}_r + \boldsymbol{s}_e^{ave}| - \boldsymbol{s}_r - \boldsymbol{s}_e^{ave}\right) \qquad (12.80)$$

and introduce the slack variables

$$\boldsymbol{\Delta s} = \boldsymbol{E s}_d - \boldsymbol{s}_e^{amp} - |\boldsymbol{s}_r + \boldsymbol{s}_e^{ave}| = \boldsymbol{E s}_d - \boldsymbol{s}_e^{amp} - \boldsymbol{s}^+ - \boldsymbol{s}^- \tag{12.81}$$

we can replace conditions (12.78) by

$$\boldsymbol{\Delta s} \geq \boldsymbol{0} \tag{12.82}$$

The residual forces $\boldsymbol{s}_r$ can be expressed in terms of the new variables as

$$\boldsymbol{s}_r = \boldsymbol{s}^+ - \boldsymbol{s}^- - \boldsymbol{s}_e^{ave} \tag{12.83}$$

and substituted into the equilibrium conditions (12.73) to yield

$$\boldsymbol{B}^T \boldsymbol{s}^+ - \boldsymbol{B}^T \boldsymbol{s}^- = \boldsymbol{B}^T \boldsymbol{s}_e^{ave} \tag{12.84}$$

In summary, the resulting problem of linear programming in the standard form reads

$$\text{Minimize } z(\boldsymbol{s}_d, \boldsymbol{s}^+, \boldsymbol{s}^-, \boldsymbol{\Delta s}) \equiv \boldsymbol{c}^T \boldsymbol{s}_d \tag{12.85}$$

subject to

$$\boldsymbol{B}^T \boldsymbol{s}^+ - \boldsymbol{B}^T \boldsymbol{s}^- = \boldsymbol{B}^T \boldsymbol{s}_e^{ave} \tag{12.86}$$

$$\boldsymbol{E s}_d - \boldsymbol{s}^+ - \boldsymbol{s}^- - \boldsymbol{\Delta s} = \boldsymbol{s}_e^{amp} \tag{12.87}$$

$$\boldsymbol{s}_d \geq \boldsymbol{0}, \quad \boldsymbol{s}^+ \geq \boldsymbol{0}, \quad \boldsymbol{s}^- \geq \boldsymbol{0}, \quad \boldsymbol{\Delta s} \geq \boldsymbol{0} \tag{12.88}$$

Again, this problem is very similar to the standard form of the limit design problem (12.47)–(12.50). The only difference is in the terms on the right-hand sides of the equality constraints.

So far we have considered only the conditions of incremental collapse. To be safe against plastic fatigue, the design must satisfy additional conditions

$$\boldsymbol{E s}_d \geq \boldsymbol{I}_\alpha \boldsymbol{s}_e^{amp} \tag{12.89}$$

where $\boldsymbol{I}_\alpha$ is a diagonal matrix with shape factors $\alpha_{ij} = M_{0,ij}/M_{el,ij}$ in its diagonal. In the simplest case, in which all the cross sections are expected to have the same shape factor, α, we have $\boldsymbol{I}_\alpha = \alpha \boldsymbol{I}$.

Conditions (12.89) can easily be added to problem (12.85)–(12.88) without any increase of the number of variables or constraints. Often, the technological matrix $\boldsymbol{E}$ is such that each inequality from (12.89) contains only one of the design variables, and then we can equivalently write

$$\boldsymbol{s}_d \geq \boldsymbol{s}_d^{min} \tag{12.90}$$

where the minimum required values of the design variables, $\boldsymbol{s}_d^{min}$, are deduced from $\boldsymbol{s}_e^{amp}$, $\boldsymbol{E}$ and $\boldsymbol{I}_\alpha$; see the next example. Now we introduce new variables

$$\boldsymbol{\Delta s}_d = \boldsymbol{s}_d - \boldsymbol{s}_d^{min} \tag{12.91}$$

and replace $\boldsymbol{s}_d$ in (12.87) by $\boldsymbol{s}_d^{min} + \boldsymbol{\Delta s}_d$. Condition (12.90) is equivalent to $\boldsymbol{\Delta s}_d \geq \boldsymbol{0}$, which replaces $\boldsymbol{s}_d \geq \boldsymbol{0}$ in (12.88). Note that $\boldsymbol{s}_d^{min} \geq \boldsymbol{0}$, and so condition $\boldsymbol{\Delta s}_d \geq \boldsymbol{0}$ is stronger than $\boldsymbol{s}_d \geq \boldsymbol{0}$. Needless to say, instead of $\boldsymbol{c}^T \boldsymbol{s}_d$ we can minimize just $\boldsymbol{c}^T \boldsymbol{\Delta s}_d$, because $\boldsymbol{c}^T \boldsymbol{s}_d = \boldsymbol{c}^T \boldsymbol{s}_d^{min} + \boldsymbol{c}^T \boldsymbol{\Delta s}_d = \text{const.} + \boldsymbol{c}^T \boldsymbol{\Delta s}_d$. The complete problem of shakedown design, comprising both the incremental collapse and plastic fatigue, reads

$$\text{Minimize } z(\boldsymbol{\Delta s}_d, \boldsymbol{s}^+, \boldsymbol{s}^-, \boldsymbol{\Delta s}) \equiv \boldsymbol{c}^T \boldsymbol{\Delta s}_d \tag{12.92}$$

subject to

$$\boldsymbol{B}^T \boldsymbol{s}^+ - \boldsymbol{B}^T \boldsymbol{s}^- = \boldsymbol{B}^T \boldsymbol{s}_e^{ave} \tag{12.93}$$

$$\boldsymbol{E}\, \boldsymbol{\Delta s}_d - \boldsymbol{s}^+ - \boldsymbol{s}^- - \boldsymbol{\Delta s} = \boldsymbol{s}_e^{amp} - \boldsymbol{E s}_d^{min} \tag{12.94}$$

$$\boldsymbol{\Delta s}_d \geq \boldsymbol{0}, \quad \boldsymbol{s}^+ \geq \boldsymbol{0}, \quad \boldsymbol{s}^- \geq \boldsymbol{0}, \quad \boldsymbol{\Delta s} \geq \boldsymbol{0} \tag{12.95}$$

Example 12.7: Optimize the shakedown design of the frame from Example 12.5, considering the load domain to be a triangle in the plane (F_{2x}, F_{3z}) with vertices at $(-\overline{F}, 0)$, $(2\overline{F}, 0)$, and $(0, 4\overline{F})$; see Figure 12.6.

Solution: First we have to find the extreme values of the moments at critical cross sections calculated as if the structure were elastic. Without elaborating on this trivial step, we present directly the results in Table 12.6. Each row corresponds to a specific load combination defined by the values of the horizontal force, F_{2x}, and the vertical force, F_{3z}. The first two rows represent the unit loading cases solved by the standard procedures of elastic frame analysis. Since we do not have any information on the expected cross sectional dimensions, we take the same cross sectional bending stiffness EI for all the members of the structure and neglect the effect of the axial compliance. For evaluating the extreme moments, it suffices to consider the three load combinations corresponding to the vertices of the triangular load domain; they are evaluated in the third to fifth row of the table and are graphically presented in Figure 12.7(a)–(c). Rows six and seven respectively show the minimum and maximum value of each critical moment; see Figure 12.7(d). The last two rows are respectively the average of the maximum and minimum value (Figure 12.9(a)) and one half of their difference (Figure 12.8(a)).

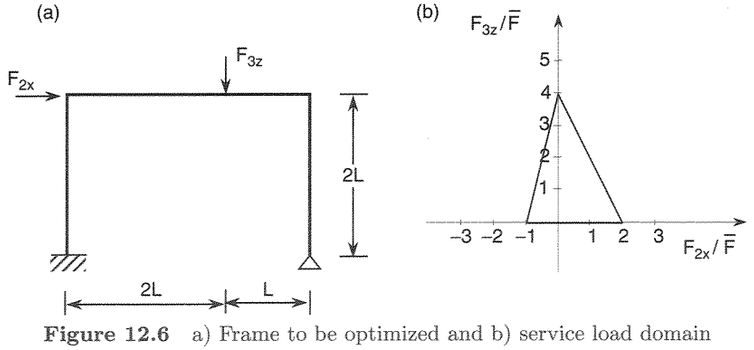


Figure 12.6 a) Frame to be optimized and b) service load domain

Table 12.6 Evaluation of extreme elastic moments

Row	State	$\frac{F_{2x}}{\overline{F}}$	$\frac{F_{3z}}{\overline{F}}$	$\frac{M_{12}}{\overline{F}L}$	$\frac{M_{21}}{\overline{F}L}$	$\frac{M_{32}}{\overline{F}L}$	$\frac{M_{45}}{\overline{F}L}$
1	unit F_{2x}	1	0	0.9634	0.5925	−0.0986	0.4441
2	unit F_{3z}	0	1	−0.0848	−0.1949	0.4153	0.2797
3	combination 1	−1	0	−0.9634	−0.5925	0.0986	−0.4441
4	combination 2	2	0	1.9268	1.1850	−0.1972	0.8882
5	combination 3	0	4	−0.3392	−0.7795	1.6612	1.1187
6	minimum		M_{ij}^{min}	−0.9634	−0.7795	−0.1972	−0.4441
7	maximum		M_{ij}^{max}	1.9268	1.1850	1.6612	1.1187
8	average		M_{ij}^{ave}	0.4817	0.2028	0.7320	0.3373
9	amplitude		M_{ij}^{amp}	1.4451	0.9822	0.9292	0.7814

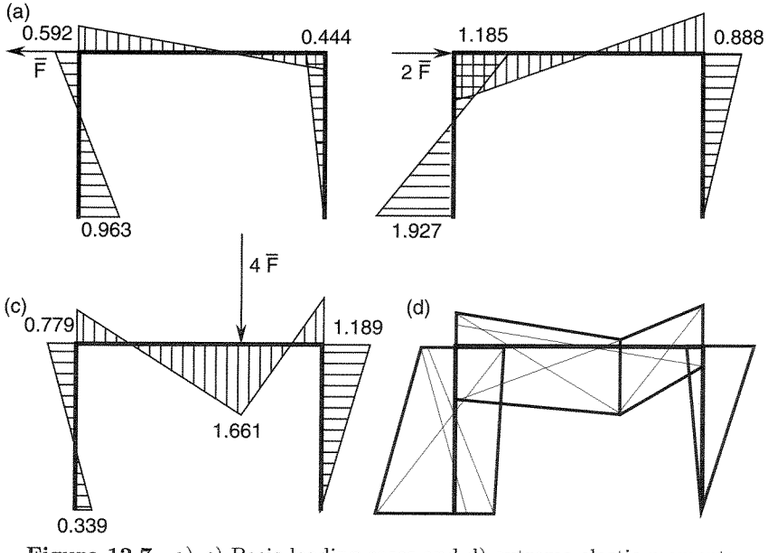


Figure 12.7 a)–c) Basic loading cases and d) extreme elastic moments

Based on the amplitudes of elastic moments, it is easy to evaluate the minimum resistances (plastic moments) required by safety against plastic fatigue. Let us assume that $\alpha = 1.2$ for all the cross sections. Conditions (12.89) then read

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} M_{0cl} \\ M_{0b} \\ M_{0cr} \end{Bmatrix} \geq \alpha \begin{Bmatrix} M_{e12}^{amp} \\ M_{e21}^{amp} \\ M_{e23}^{amp} \\ M_{e32}^{amp} \\ M_{e43}^{amp} \\ M_{e45}^{amp} \end{Bmatrix} = 1.2 \times \begin{Bmatrix} 1.4451 \\ 0.9822 \\ 0.9822 \\ 0.9292 \\ 0.7814 \\ 0.7814 \end{Bmatrix} \overline{F}L = \begin{Bmatrix} 1.7341 \\ 1.1786 \\ 1.1786 \\ 1.1150 \\ 0.9377 \\ 0.9377 \end{Bmatrix} \overline{F}L \quad (12.96)$$

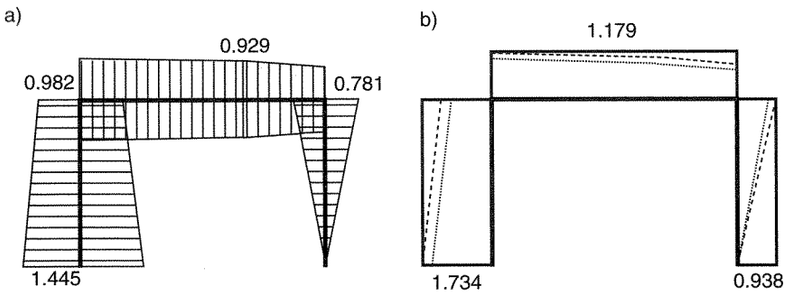


Figure 12.8 a) Amplitudes of elastic moments, b) minimum plastic capacities required by safety against plastic fatigue

The minimum allowed values of the design variables dictated by safety against plastic fatigue are

$$M_{0cl}^{min} = \max(1.7341,\ 1.1786)\,\overline{F}L = 1.7341\overline{F}L \quad (12.97)$$

$$M_{0b}^{min} = \max(1.1786,\ 1.1150,\ 0.9377)\,\overline{F}L = 1.1786\overline{F}L \quad (12.98)$$

$$M_{0cr}^{min} = 0.9377\overline{F}L \quad (12.99)$$

They can be constructed graphically by the procedure illustrated in Figure 12.8(b). The amplitudes of elastic moments indicated by the dotted lines are magnified by the shape factor $\alpha = 1.2$ to yield the distribution plotted by the dashed lines. The largest magnitudes on each structural member give the minimum plastic capacities shown by the solid lines.

Now we can construct the initial simplex tableau. In view of similarity to limit design, Table 12.7 closely resembles Table 12.4, except for the column with the right-hand side. The given loads $\bar{\boldsymbol{f}}$ on the right-hand side of the equilibrium constraints are replaced by fictitious loads

$$\boldsymbol{B}^T \boldsymbol{s}_e^{ave} = \begin{bmatrix} 0.5 & 0.5 & 0.0 & 0.5 \\ 0.0 & -0.5 & 1.5 & 1.0 \end{bmatrix} \begin{Bmatrix} M_{e12}^{ave} \\ M_{e21}^{ave} \\ M_{e32}^{ave} \\ M_{e45}^{ave} \end{Bmatrix} \quad (12.100)$$

$$= \begin{bmatrix} 0.5 & 0.5 & 0.0 & 0.5 \\ 0.0 & -0.5 & 1.5 & 1.0 \end{bmatrix} \begin{Bmatrix} 0.4817 \\ 0.2028 \\ 0.7320 \\ 0.3373 \end{Bmatrix} \overline{F}L = \begin{Bmatrix} 0.5109 \\ 1.3339 \end{Bmatrix} \overline{F}L$$

which have the physical meaning of external forces that would be in equilibrium with the average elastic moments; see Figure 12.9(a). The zeros on the right-hand side of the plastic admissibility constraints are replaced by internal forces

$$\boldsymbol{s}_e^{amp} - \boldsymbol{E}\boldsymbol{s}_d^{min} = \begin{Bmatrix} M_{e12}^{amp} \\ M_{e21}^{amp} \\ M_{e23}^{amp} \\ M_{e32}^{amp} \\ M_{e43}^{amp} \\ M_{e45}^{amp} \end{Bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} M_{0cl}^{min} \\ M_{0b}^{min} \\ M_{0cr}^{min} \end{Bmatrix} \quad (12.101)$$

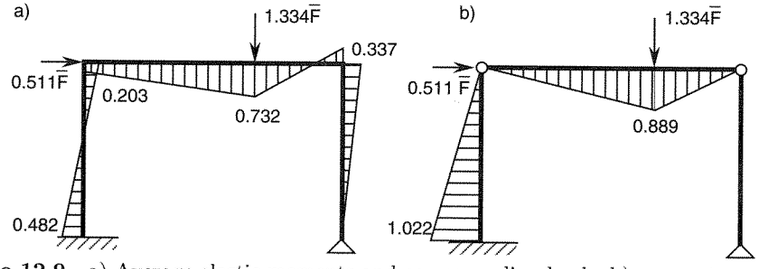


Figure 12.9 a) Average elastic moments and corresponding loads, b) moments caused by 'average loads' on a primary structure

$$= \begin{Bmatrix} 1.4451 \\ 0.9822 \\ 0.9822 \\ 0.9292 \\ 0.7814 \\ 0.7814 \end{Bmatrix} \overline{F}L - \begin{Bmatrix} 1.7341 \\ 1.7341 \\ 1.1786 \\ 1.1786 \\ 1.1786 \\ 0.9377 \end{Bmatrix} \overline{F}L = \begin{Bmatrix} -0.2890 \\ -0.7519 \\ -0.1964 \\ -0.2494 \\ -0.3972 \\ -0.1563 \end{Bmatrix} \overline{F}L$$

which represent the difference between the amplitudes of elastic moments and the minimum resistances obtained in the optimum design against plastic fatigue only. The magnitudes of these differences correspond to the distances between the solid and dotted lines in Figure 12.8(b).

An initial basic feasible solution for the present problem can be constructed using a similar strategy as in limit design. First we select two moments as the redundants, and we compute the bending moments on the corresponding primary structure subjected to the 'average loads', $F_{2x}^{ave} = 0.5109\overline{F}$ and $F_{3z}^{ave} = 1.3339\overline{F}$. Using the same primary structure as in Example 12.5, we obtain $M_{12} = 1.0218\overline{F}L$ and $M_{32} = 0.8893\overline{F}L$; see Figure 12.9(b). Both moments are positive, and so we select M_{12}^{+} and M_{32}^{+} as members of the initial basis. All the other variables M_{ij}^{+} and M_{ij}^{-} are set to zero.

Having determined the initial values of $\boldsymbol{s}^{+}$ and $\boldsymbol{s}^{-}$, we can rewrite conditions (12.94) as

$$\boldsymbol{E}\,\boldsymbol{\Delta s}_d - \boldsymbol{\Delta s} = \left(\boldsymbol{s}^{+} + \boldsymbol{s}^{-}\right) + \left(\boldsymbol{s}_e^{amp} - \boldsymbol{E}\boldsymbol{s}_d^{min}\right) \tag{12.102}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} \Delta M_{0cl} \\ \Delta M_{0b} \\ \Delta M_{0cr} \end{Bmatrix} - \begin{Bmatrix} \Delta M_{12} \\ \Delta M_{21} \\ \Delta M_{23} \\ \Delta M_{32} \\ \Delta M_{43} \\ \Delta M_{45} \end{Bmatrix} = \begin{Bmatrix} 1.0218 \\ 0 \\ 0 \\ 0.8893 \\ 0 \\ 0 \end{Bmatrix} \overline{F}L + \begin{Bmatrix} -0.2890 \\ -0.7519 \\ -0.1964 \\ -0.2494 \\ -0.3972 \\ -0.1563 \end{Bmatrix} \overline{F}L$$

$$= \begin{Bmatrix} 0.7328 \\ -0.7519 \\ -0.1964 \\ 0.6399 \\ -0.3972 \\ -0.1563 \end{Bmatrix} \overline{F}L \tag{12.103}$$

Table 12.7 Initial simplex tableau

ΔM_{0cl}	ΔM_{0b}	ΔM_{0cr}	M_{12}^{+}	M_{21}^{+}	M_{32}^{+}	M_{45}^{+}	M_{12}^{-}	M_{21}^{-}	M_{32}^{-}	M_{45}^{-}	ΔM_{12}	ΔM_{21}	ΔM_{23}	ΔM_{32}	ΔM_{43}	ΔM_{45}	RHS/$\overline{F}L$
0	0	0	0.5	0.5	0.0	0.5	−0.5	−0.5	0.0	−0.5	0	0	0	0	0	0	0.511
0	0	0	0.0	−0.5	1.5	1.0	0.0	0.5	−1.5	−1.0	0	0	0	0	0	0	1.334
1	0	0	−1.0	0.0	0.0	0.0	−1.0	0.0	0.0	0.0	−1	0	0	0	0	0	−0.289
1	0	0	0.0	−1.0	0.0	0.0	0.0	−1.0	0.0	0.0	0	−1	0	0	0	0	−0.752
0	1	0	0.0	−1.0	0.0	0.0	0.0	−1.0	0.0	0.0	0	0	−1	0	0	0	−0.196
0	1	0	0.0	0.0	−1.0	0.0	0.0	0.0	−1.0	0.0	0	0	0	−1	0	0	−0.249
0	1	0	0.0	0.0	0.0	−1.0	0.0	0.0	0.0	−1.0	0	0	0	0	−1	0	−0.397
0	0	1	0.0	0.0	0.0	−1.0	0.0	0.0	0.0	−1.0	0	0	0	0	0	−1	−0.156
2	3	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0	0	0	0	0	0.000

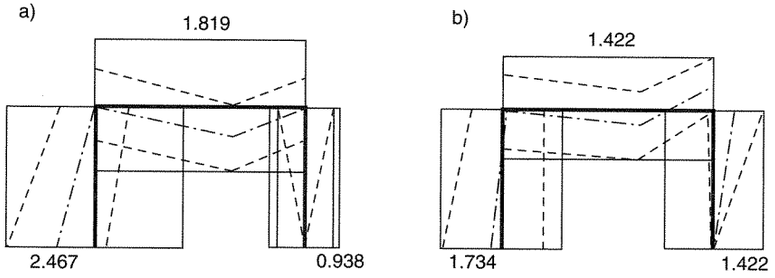


Figure 12.10 a) Initial design and b) optimum shakedown design

or, in scalar notation, as

$$\Delta M_{0cl} - \Delta M_{12} = 0.7328\overline{F}L \tag{12.104}$$
$$\Delta M_{0cl} - \Delta M_{21} = -0.7519\overline{F}L \tag{12.105}$$
$$\Delta M_{0b} - \Delta M_{23} = -0.1964\overline{F}L \tag{12.106}$$
$$\Delta M_{0b} - \Delta M_{32} = 0.6399\overline{F}L \tag{12.107}$$
$$\Delta M_{0b} - \Delta M_{43} = -0.3972\overline{F}L \tag{12.108}$$
$$\Delta M_{0cr} - \Delta M_{45} = -0.1563\overline{F}L \tag{12.109}$$

According to (12.95), all design variables ΔM_{0i} and all slack variables ΔM_{ij} must be nonnegative. If all the values on the right-hand side of the above equations were nonpositive, we could set $\Delta M_{0cl} = \Delta M_{0b} = \Delta M_{0cr} = 0$, and the minimum design against plastic fatigue would satisfy the conditions of incremental collapse, i.e. it would be the optimum shakedown design. However, the presence of some positive values indicates that this is not the case here. The minimum admissible values of ΔM_{0cl} and ΔM_{0b} are dictated by conditions (12.104) and (12.107), respectively. To satisfy the other conditions, the corresponding slack variables must be nonzero. Consequently, the initial basis consists of ΔM_{0cl}, ΔM_{0b}, M_{12}^+, M_{32}^+, ΔM_{21}, ΔM_{23}, ΔM_{43} and ΔM_{45}. We have constructed the basis such that the corresponding basic solution is guaranteed to be feasible. This is confirmed by the simplex tableau in Table 12.8.

Recall that the actual internal forces vary in the limits $\boldsymbol{s}_r + \boldsymbol{s}_e^{ave} \pm \boldsymbol{s}_e^{amp}$. As $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$ are the positive and negative parts of $\boldsymbol{s}_r + \boldsymbol{s}_e^{ave}$ (see definitions (12.79)–(12.80)), these limits are easily evaluated as $\boldsymbol{s}^+ - \boldsymbol{s}^- \pm \boldsymbol{s}_e^{amp}$. For example, moment M_{12} varies in the limits $M_{12}^+ - M_{12}^- \pm M_{e12}^{amp} = (1.022 - 0 \pm 1.445)\overline{F}L$, i.e. between $-0.423\overline{F}L$ and $2.467\overline{F}L$. Now it is easy to give a physical interpretation of the initial design. Moments $M_{ij}^+ - M_{ij}^-$ represent the mean values around which the actual moments can vary with amplitudes M_{eij}^{amp}; see Figure 12.10(a). The dash-dotted lines correspond to the mean moments $M_{ij}^+ - M_{ij}^-$ (with linear interpolation between the critical cross sections), and the dashed lines show the limits within which the actual moments can vary. The maximum magnitudes of moments in each structural member determine the minimum plastic resistance of that member required for safety against incremental collapse. At the same time, the plastic resistances must not be smaller than the minimum values needed for safety against plastic fatigue. The aim of the optimization process is to adjust the mean moments $M_{ij}^+ - M_{ij}^-$ such that the required plastic resistances weighted by the lengths of the respective members be minimized, while at the same time the mean bending moments must be in equilibrium with the average

Table 12.8 Initial feasible solution

ΔM_{0cl}	ΔM_{0b}	ΔM_{0cr}	M_{12}^+	M_{21}^+	M_{32}^+	M_{45}^+	M_{12}^-	M_{21}^-	M_{32}^-	M_{45}^-	ΔM_{12}	ΔM_{21}	ΔM_{23}	ΔM_{32}	ΔM_{43}	ΔM_{45}	RHS/$\overline{F}L$
0	0	0	0	−0.333	1	0.667	0	0.333	−1	−0.667	0	0	0	0	0	0	0.889
0	0	0	1	1.000	0	1.000	−1	−1.000	0	−1.000	0	0	0	0	0	0	1.022
1	0	0	0	1.000	0	1.000	−2	−1.000	0	−1.000	−1	0	0	0	0	0	0.733
0	0	0	0	2.000	0	1.000	−2	0.000	0	−1.000	−1	1	0	0	0	0	1.485
0	1	0	0	−0.333	0	0.667	0	0.333	−2	−0.667	0	0	0	−1	0	0	0.640
0	0	0	0	0.667	0	0.667	0	1.333	−2	−0.667	0	0	1	−1	0	0	0.836
0	0	0	0	−0.333	0	1.667	0	0.333	−2	0.333	0	0	0	−1	1	0	1.037
0	0	−1	0	0.000	0	1.000	0	0.000	0	1.000	0	0	0	0	0	1	0.156
0	0	2	0	−1.000	0	−4.000	4	1.000	6	4.000	2	0	0	3	0	0	−3.385

Table 12.9 Optimal solution

ΔM_{0cl}	ΔM_{0b}	ΔM_{0cr}	M_{12}^{+}	M_{21}^{+}	M_{32}^{+}	M_{45}^{+}	M_{12}^{-}	M_{21}^{-}	M_{32}^{-}	M_{45}^{-}	ΔM_{12}	ΔM_{21}	ΔM_{23}	ΔM_{32}	ΔM_{43}	ΔM_{45}	RHS/$\overline{F}L$
0.167	0	0	0	0	1	0	−0.333	0	0	−1	−0.167	0	0	0.5	−0.5	0	0.493
−1.000	0	0	1	0	0	0	1.000	0	0	0	1.000	0	0	0.0	0.0	0	0.289
0.833	0	0	0	1	0	0	−1.667	−1	1	−1	−0.833	0	0	0.5	−0.5	0	0.092
−1.833	0	0	0	0	0	0	1.667	2	−1	1	0.833	1	0	−0.5	0.5	0	0.660
0.167	1	0	0	0	0	0	−0.333	0	−1	−1	−0.167	0	0	−0.5	−0.5	0	0.243
−0.667	0	0	0	0	0	0	1.333	2	−2	0	0.667	0	1	−1.0	0.0	0	0.348
0.167	0	1	0	0	0	0	−0.333	0	−1	−1	−0.167	0	0	−0.5	0.5	−1	0.484
0.167	0	0	0	0	0	1	−0.333	0	−1	0	−0.167	0	0	−0.5	0.5	0	0.641
1.167	0	0	0	0	0	0	1.667	0	5	5	0.833	0	0	2.5	0.5	2	−1.699

loads F_{2x}^{ave} and F_{3z}^{ave}. Note that the value in the right bottom corner of the simplex tableau corresponds to the modified cost function (12.92). The total required plastic capacities are

$$M_{0cl} = M_{0cl}^{min} + \Delta M_{0cl} = 1.734\overline{F}L + 0.733\overline{F}L = 2.467\overline{F}L \tag{12.110}$$
$$M_{0b} = M_{0b}^{min} + \Delta M_{0b} = 1.179\overline{F}L + 0.640\overline{F}L = 1.819\overline{F}L \tag{12.111}$$
$$M_{0cr} = M_{0cr}^{min} + \Delta M_{0cr} = 0.938\overline{F}L + 0 = 0.938\overline{F}L \tag{12.112}$$

and the original cost function (12.75) attains the value

$$z = b(M_{0cl} \times 2L + M_{0b} \times 3L + M_{0cr} \times 2L) = 12.27b\overline{F}L^2 \tag{12.113}$$

We will skip the intermediate stages of the optimization process and present the optimum solution in Table 12.9. The corresponding design is shown in Figure 12.10(b). The plastic capacities required by the optimum design are

$$M_{0cl} = M_{0cl}^{min} + \Delta M_{0cl} = 1.734\overline{F}L + 0 = 1.734\overline{F}L \tag{12.114}$$
$$M_{0b} = M_{0b}^{min} + \Delta M_{0b} = 1.179\overline{F}L + 0.243\overline{F}L = 1.422\overline{F}L \tag{12.115}$$
$$M_{0cr} = M_{0cr}^{min} + \Delta M_{0cr} = 0.938\overline{F}L + 0.484\overline{F}L = 1.422\overline{F}L \tag{12.116}$$

and the minimum value of the cost function is

$$z = b(M_{0cl} \times 2L + M_{0b} \times 3L + M_{0cr} \times 2L) = 10.58b\overline{F}L^2 \tag{12.117}$$

After the selection of appropriate profiles, a careful designer would rerun the analysis with the actual values of moments of inertia, I_{ij}, of shape factors, α_{ij}, and of coefficients b_i from definition (12.22) of the cost function. Recall that the present design has been constructed under the simplified assumptions that $I_{ij} = I = \text{const}$, $\alpha_{ij} = 1.2$ and $b_i = b = \text{const}$. □

As already mentioned, the shakedown design problem is intrinsically nonlinear. We have presented a solution technique leading to a sequence of linear programming problems. In some cases, such an *iterated linear programming approach* may not converge. Then it is necessary to use more sophisticated techniques of *nonlinear programming*, e.g. see Cohn, Maier and Grierson (1979), Save and Prager (1990), or Cottle, Pang and Stone (1992). Principles of optimum design and applications to various types of structures are covered in a collection of papers edited by Save and Prager (1985). Mathematical optimization tools and general applications in engineering are treated, for example, by Papalambros and Wilde (2000). Modern optimization methodologies frequently employ stochastic search procedures such as *genetic algorithms* (Holland, 1975; Goldberg, 1989) or *simulated annealing* (Kirkpatrick, Gelatt and Vecchi, 1983).

PROBLEMS

Problem 12.1*: Prove that the limit design of trusses without any design constraints always leads to a statically determinate structure. Can this statement be generalized to shakedown design?

Problem 12.2*: What can you say about the degree of static redundancy of a truss resulting from limit design with d design variables?

Problem 12.3: Find the optimum limit design of the three-span continuous beam in Figure 12.11. The cross section should be constant within each span.

Problem 12.4: Find the optimum shakedown design of the three-bar truss from Figure 12.2(a), considering a service load domain described by inequalities $0 \leq F_x \leq \overline{F}$ and $0 \leq F_z \leq 2\overline{F}$. Bars 1 and 3 should have the same cross section.
Hint: Note that for trusses it suffices to check the safety against incremental collapse, and so formulation (12.85)–(12.88) can be used.

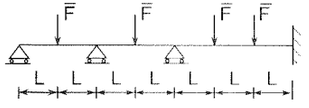


Figure 12.11 Three-span continuous beam

13

Combined Plastic Bending and Compression or Tension

13.1 GENERALIZED PLASTIC HINGE

So far we have assumed that a plastic hinge forms if the bending moment in a critical cross section reaches the plastic limit value, M_0. We have tacitly neglected the effect of the other internal forces on the formation of the yield hinge. This effect, however, can be appreciable, for example in multi-storey frames or frames with a large horizontal thrust where the axial forces are large. So, we will now refine our analysis and consider the interaction of bending moment M and normal force N. The resulting fully plasticized cross section will be referred to as a *generalized plastic hinge*. The effect of the shear force is usually less important, and is also more difficult to incorporate. We will postpone this problem to Part III, because it is beyond the scope of analysis under uniaxial stress.

A typical evolution of the strain and stress profiles during the formation of a generalized plastic hinge is depicted in Figure 13.1. As usual, it may be assumed that the cross sections remain planar at all stages of the loading process. Consequently, the variation of normal strains over the cross section is linear. The corresponding stress distribution is linear only in the elastic range (Figure 13.1(a)). As the loading process continues, yielding starts at the top or bottom fibers, and the plastic zone propagates into the interior of the cross section (Figure 13.1(b)). During the elastoplastic stage, the cross section has an elastic core with linear stress variation, and one or two plastic zones with constant stress equal to the positive or negative yield stress. For very large curvatures, the elastic core becomes negligibly small, and the stress distribution approaches a piecewise constant distribution, with one part of the cross section yielding in tension and the remaining part yielding in compression.

The plasticization process of course takes place in the neighboring cross sections as well, and the plastic zone occupies a certain volume (Figure 13.2(b)). However, for the purpose of modeling, we can lump the plastic hinge into one single cross section, the same as we did while analyzing the hinge under pure bending. The total plastic deformation in the idealized hinge is replaced by a rotation, θ, and longitudinal displacement, e_p (Figure 13.2(c)). From kinematic considerations, it follows that the plastic extension at an arbitrary point of the plasticized cross section can be expressed by a linear function

$$\bar{e}_p(z) = e_p + \theta z \tag{13.1}$$

where z is the centroidal coordinate perpendicular to the bending axis. The plastic extension at the centroid, $\bar{e}_p(0) = e_p$, corresponds to the 'gap' between the centroids

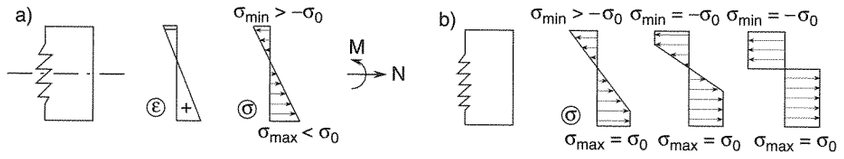


Figure 13.1 Strain and stress distributions

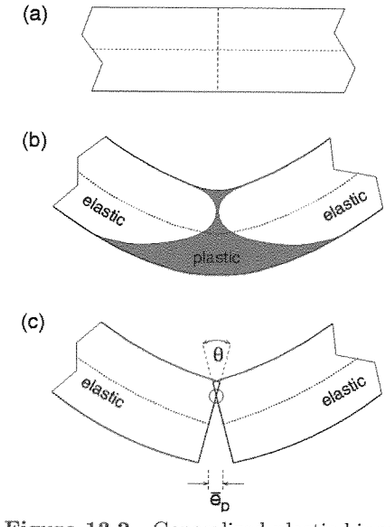


Figure 13.2 Generalized plastic hinge

of the end cross sections of the two elastic parts connected by the generalized yield hinge in Figure 13.2(c). The bar over e helps to distinguish the distribution of the plastic extension over the cross section, $\bar{e}_p(z)$, which is a function of the distance from the centroidal axis, from the variable e_p, which characterizes the deformation of the generalized plastic hinge and is equal to the value of $\bar{e}_p$ at $z = 0$.

Consider a given distribution of plastic extension, represented by the solid line in Figure 13.3(a). It might seem natural to assume that the fibers for which $\bar{e}_p(z) > 0$ yield in tension, and those for which $\bar{e}_p(z) < 0$ yield in compression; see Figure 13.4(b). This assumption corresponds to the so-called *deformation theory of plasticity*, in which

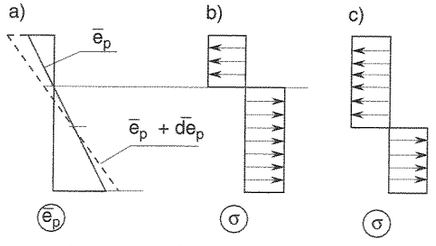


Figure 13.3 Distribution of a) strain, b) stress according to the deformation theory, c) stress according to the flow theory of plasticity

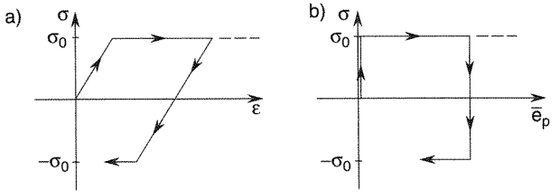


Figure 13.4 Stress-strain and stress-extension diagrams

the stress state is entirely determined by the current strain state, independently of the previous history of strain evolution (Hencky, 1924). Such a model is conceptually simple and convenient, but not always sufficiently realistic. The actual stress in an elastoplastic material is uniquely determined by the strain only as long as the loading process is monotonic. If the strain increment changes its sign, the material unloads elastically and eventually starts yielding in the opposite direction (Figure 13.4(a)). The deformation of a plastic hinge does not have any elastic component, and so the stress jumps to the opposite yield stress immediately after the plastic extension increment has changed its sign (Figure 13.4(b)). Clearly, the stress can be negative even at a positive extension. If we strictly adhere to the *flow theory of plasticity* we have to determine the sign of stress from the sign of plastic extension rate (increment) rather than from the sign of the plastic extension itself. This means that the position of the neutral axis separating the tensile zone from the compressive zone is determined by the condition $\dot{\bar{e}}_p(z) = 0$ rather than $\bar{e}_p(z) = 0$; see Figure 13.3(c).

Depending on the position of the neutral axis, a generalized plastic hinge can yield different combinations of internal forces M and N. The set of all such combinations is the so-called *plastic limit envelope* in the space of internal forces (also called the yield surface, yield locus, failure envelope, or interaction curve). States outside this envelope can never be reached. Similarly to the plastic hinge under pure bending, it is assumed that states inside the plastic limit envelope are entirely elastic, although this is not exactly true. One could define the *elastic limit envelope* that corresponds to states at which the first point of a cross section begins yielding (similar to the limit elastic moment in the case of pure bending). The states between the elastic and plastic limit envelopes are elastoplastic, but the corresponding plastic deformations that precede the formation of a full plastic hinge are usually neglected, especially when looking for the final collapse mechanism rather than for the details of the response in the elastoplastic range.

Example 13.1: Construct the elastic limit envelope and the plastic limit envelope for a rectangular cross section of width b and depth h.

Solution: 1) We start with the plastic limit envelope. Let us specify the position of the neutral axis by the coordinate $z_n = \eta h/2$ measured from the cross section centroid (Figure 13.5). Instead of integrating over the cross section, we can replace the positive stress by its resultant

$$N^+ = A^+\sigma_0 = \frac{1}{2}(1-\eta)bh\sigma_0 \tag{13.2}$$

and the negative stress by its resultant

$$N^- = A^-\sigma_0 = \frac{1}{2}(1+\eta)bh\sigma_0 \tag{13.3}$$

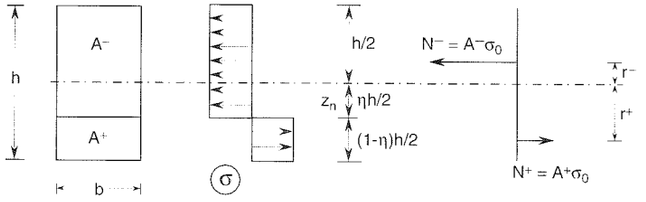


Figure 13.5 Stress distributions in a rectangular cross section

as indicated in Figure 13.5. The arms of the resultants of positive and negative stress, i.e. the distances of the centroids of the areas under tension and under compression from the centroid of the entire cross section, are given by

$$r^+ = \frac{\eta h}{2} + \frac{1}{2}(1-\eta)\frac{h}{2} = \frac{1}{4}(1+\eta)h \tag{13.4}$$

$$r^- = \frac{1}{2}(1+\eta)\frac{h}{2} + \frac{\eta h}{2} = \frac{1}{4}(1-\eta)h \tag{13.5}$$

The force and moment resultants at a plastic limit state are now obtained as

$$N = \int_A \sigma \, dA = N^+ + N^- = -\sigma_0 bh\eta \tag{13.6}$$

$$\begin{aligned} M &= \int_A \sigma z \, dA = N^+ r^+ + N^- r^- \\ &= \frac{1}{2}(1-\eta)bh\sigma_0 \times \frac{1}{4}(1+\eta)h + \frac{1}{2}(1+\eta)bh\sigma_0 \times \frac{1}{4}(1-\eta)h \\ &= \frac{1}{4}\sigma_0 bh^2(1-\eta^2) \end{aligned} \tag{13.7}$$

The parameter η describing the position of the neutral axis can vary within the limits $[-1, 1]$. The values $\eta = -1$, $\eta = 0$ and $\eta = 1$ correspond to the axial tension, pure bending and axial compression, respectively. For $\eta = -1$, the normal force assumes its maximum possible value, $N = \sigma_0 bh \equiv N_0$ = plastic normal force, while the bending moment is zero. For $\eta = 0$, the bending moment assumes its maximum possible value, $M = \sigma_0 bh^2/4 \equiv M_0$ = plastic moment, while the normal force is zero. Finally, for $\eta = 1$, the normal force assumes its minimum possible value, $N = -\sigma_0 bh = -N_0$, while the bending moment is zero. Note that the plastic normal force N_0 is identical to the plastic axial force S_0 we introduced in truss analysis; likewise, N and S are only two alternative symbols for the normal, or axial, force.

Relations (13.6) and (13.7), rewritten as

$$\left.\begin{aligned} N &= -N_0\eta \\ M &= M_0(1-\eta^2) \end{aligned}\right\} \quad -1 \le \eta \le 1 \tag{13.8}$$

parametrically describe the plastic limit envelope. Eliminating the parameter η, we get an explicit equation for the envelope,

$$\frac{M}{M_0} = 1 - \left(\frac{N}{N_0}\right)^2, \quad -1 \le \frac{N}{N_0} \le 1 \tag{13.9}$$

which obviously defines a parabolic arc.

Deriving the previous results, we have tacitly assumed that the bottom face of cross section is under tension and the top face under compression. In other words, equation (13.9) only applies if the bending moment is positive. To obtain a complete description of the plastic limit envelope, we have to consider the opposite case of tension at the top face and compression at the bottom. Reverting the signs of stresses and, consequently, of the internal forces, we get

$$\left.\begin{aligned} N &= N_0\eta \\ M &= -M_0(1-\eta^2) \end{aligned}\right\} \quad -1 \le \eta \le 1 \tag{13.10}$$

from which

$$\frac{M}{M_0} = -1 + \left(\frac{N}{N_0}\right)^2, \quad -1 \le \frac{N}{N_0} \le 1 \tag{13.11}$$

This equation describes a parabolic arc that is a mirror image of the previous one. The complete plastic limit envelope is shown as the solid curve in Figure 13.6.

Relations (13.9) and (13.11), separately describing the upper and lower branches of the plastic limit envelope, can be replaced by a single equation

$$f(N, M) \equiv \left(\frac{N}{N_0}\right)^2 + \left|\frac{M}{M_0}\right| - 1 = 0 \tag{13.12}$$

Note that the interior of the plastic limit envelope is characterized by $f(N, M) < 0$, and its exterior by $f(N, M) > 0$. A function f with this property is called the *yield function*.

2) The elastic limit envelope can easily be determined using the well-known expressions for stresses in the extreme points of the cross section,

$$\sigma_{max} = \frac{N}{A} + \frac{M}{W_{el}} \tag{13.13}$$

$$\sigma_{min} = \frac{N}{A} - \frac{M}{W_{el}} \tag{13.14}$$

where $A = bh$ is the cross sectional area, and $W_{el} = bh^2/6$ is the elastic cross sectional modulus. The section remains elastic as long as $|\sigma| \le \sigma_0$, i.e. the values of internal forces in the elastic domain satisfy the condition

$$\left|\frac{N}{\sigma_0 A} \pm \frac{M}{\sigma_0 W_{el}}\right| \le 1 \tag{13.15}$$

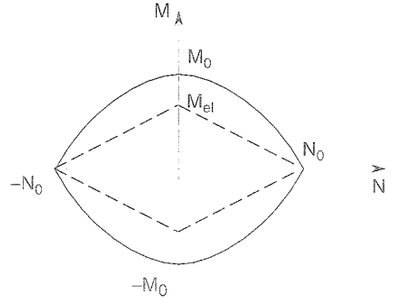


Figure 13.6 Elastic and plastic limit envelopes for a rectangular cross section

Note that (13.15) represents four inequalities that must be satisfied simultaneously. Using the relations $\sigma_0 A = N_0$ and $\sigma_0 W_{el} = M_{el}$ we can rewrite the condition as

$$\left| \frac{N}{N_0} \pm \frac{M}{M_{el}} \right| \leq 1 \tag{13.16}$$

The previous description of the elastic domain applies to cross sections of an arbitrary shape, not only to rectangular ones. According to (13.16), the elastic domain is a parallelogram whose vertices correspond to the elastic limits under axial tension or compression, and under pure bending. Under axial tension or compression, the elastic limit coincides with the plastic limit. Under pure bending, the ratio $M_0/M_{el} = \alpha$ depends on the shape of the cross section. For rectangular cross sections, $\alpha = 1.5$, and the corresponding elastic limit envelope is marked in Figure 13.6 by the dashed lines. □

Let us now explore the relation between the plastic deformations and the changes of internal forces. A generalized plastic hinge forms at a certain cross section when the internal forces reach the plastic limit envelope. In contrast to the plastic hinge under pure bending, the values of the internal forces do not have to remain constant as the hinge deforms. During plastic flow, the neutral axis can change its location while the cross section remains entirely plasticized. Consequently, the internal forces transmitted by the hinge move along the plastic limit envelope.

As explained before, the location of the neutral axis is determined, according to the deformation theory of plasticity, by the current values of the generalized strains e and θ or, according to the flow theory of plasticity, by their rates $\dot{e}$ and $\dot{\theta}$. The neutral axis divides the cross section into the part yielding in tension, A^+, and the part yielding in compression, A^-; see Figure 13.7. For a general cross section, the internal forces can be expressed as

$$N \quad \int_{A^+} \sigma_0 \, \mathrm{d}A + \int_{A^-} (-\sigma_0) \, \mathrm{d}A = \sigma_0 (A^+ - A^-) \tag{13.17}$$

$$M \quad \int_{A^+} \sigma_0 z \, \mathrm{d}A + \int_{A^-} (-\sigma_0) z \, \mathrm{d}A = \sigma_0 (S^+ - S^-) \tag{13.18}$$

where

$$S^+ = \int_{A^+} z \, \mathrm{d}A \tag{13.19}$$

is the static moment of the part under tension with respect to the centroidal axis, and

$$S^- = \int_{A^-} z \, \mathrm{d}A \tag{13.20}$$

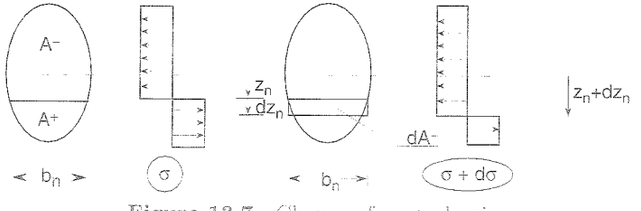


Figure 13.7 Change of neutral axis

is the static moment of the part under compression. Let z_n denote the current position of the neutral axis. If the neutral axis currently located at z_n moves by an infinitesimal distance $\mathrm{d}z_n$, the corresponding increments of the areas and static moments can be expressed as

$$\mathrm{d}A^+ = -b_n \mathrm{d}z_n \tag{13.21}$$

$$\mathrm{d}A^- = b_n \mathrm{d}z_n \tag{13.22}$$

$$\mathrm{d}S^+ = z_n \mathrm{d}A^+ = -b_n z_n \mathrm{d}z_n \tag{13.23}$$

$$\mathrm{d}S^- = z_n \mathrm{d}A^- = b_n z_n \mathrm{d}z_n \tag{13.24}$$

where b_n is the width of the cross section at the current level of the neutral axis; see Figure 13.7. Consequently, the internal forces change by

$$\mathrm{d}N = \sigma_0 (\mathrm{d}A^+ - \mathrm{d}A^-) = -2\sigma_0 b_n \mathrm{d}z_n \tag{13.25}$$

$$\mathrm{d}M = \sigma_0 (\mathrm{d}S^+ - \mathrm{d}S^-) = -2\sigma_0 b_n z_n \mathrm{d}z_n = z_n \mathrm{d}N \tag{13.26}$$

The fact that the slope of the plastic limit envelope is equal to the distance of the neutral axis from the centroid, i.e.

$$\frac{\mathrm{d}M}{\mathrm{d}N} = z_n \tag{13.27}$$

has very important implications:

1. **Convexity:** As the position of the neutral axis changes from the bottom of the cross section ($z_n = z_{max} > 0$) to the top ($z_n = z_{min} < 0$), the corresponding point on the plastic limit envelope travels along the upper branch from the left corner to the right corner, and along the lower branch from the right to the left; see Figure 13.8. During this process, the slope $\mathrm{d}M/\mathrm{d}N = z_n$ monotonically decreases. Consequently, for the upper branch we have $\mathrm{d}^2M/\mathrm{d}N^2 < 0$, and for the lower branch $\mathrm{d}^2M/\mathrm{d}N^2 > 0$. This implies that the interior of the plastic limit envelope is a *convex domain*.

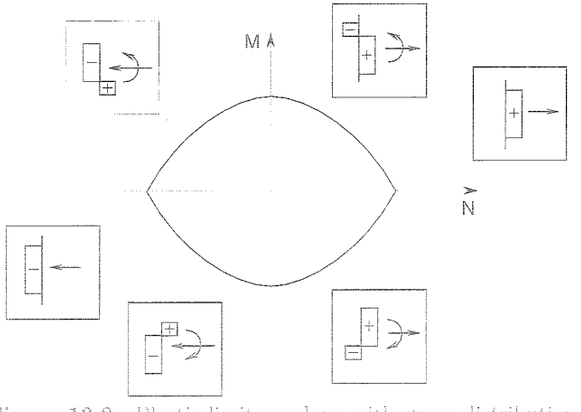


Figure 13.8 Plastic limit envelope with stress distributions

2. **Normality:** Recall that, according to the flow theory of plasticity, the position of the neutral axis is determined by the condition

$$\dot{\bar{e}}_p(z_n) \equiv \dot{e}_p + \dot{\theta} z_n = 0 \tag{13.28}$$

Combining this with the rate form of (13.27),

$$\frac{\dot{M}}{\dot{N}} = z_n \tag{13.29}$$

we obtain

$$\dot{N}\dot{e}_p + \dot{M}\dot{\theta} = 0 \tag{13.30}$$

or

$$\dot{\boldsymbol{s}}^T \dot{\boldsymbol{e}}_p = 0 \tag{13.31}$$

where

$$\boldsymbol{s} = \begin{Bmatrix} N \\ M \end{Bmatrix} \tag{13.32}$$

is the vector of internal forces and

$$\boldsymbol{e}_p = \begin{Bmatrix} e_p \\ \theta \end{Bmatrix} \tag{13.33}$$

is the vector of plastic generalized strains that is work-conjugate to $\boldsymbol{s}$.

During plastic flow, full plasticity of the cross section is preserved. Consequently, vector $\boldsymbol{s}$ travels along the yield surface (plastic limit envelope), and so vector $\dot{\boldsymbol{s}}$ represents the direction of the tangent to the yield surface; see Figure 13.9. Now we plot the plastic deformations in the same figure; we superimpose the axial extension e_p on axis N, and rotation θ on axis M. The plastic deformation rate vector, $\dot{\boldsymbol{e}}_p$, may now be displayed in the plane of internal forces. Equation (13.31) means that the scalar product of the tangent vector $\dot{\boldsymbol{s}}$ and the plastic deformation rate vector $\dot{\boldsymbol{e}}_p$ is zero, and so these vectors are orthogonal.[1]
We have thus deduced a very important property, which is typical not just for the interaction of normal force and bending moment, but for many types of plastic behavior, in general called *associated plasticity.*

Normality rule: *The increment of the plastic deformation vector has the direction of the normal to the yield surface (plastic limit envelope) at the point representing the current internal forces.*

At the corner points A and E in Figure 13.9, the preceding conclusion is ambiguous and does not apply literally. For those points the stress is uniformly distributed and equal to σ_0 or $-\sigma_0$. Such a uniform stress distribution occurs not only for pure extension ($\dot{\theta} = 0$), but also for all rotation rates for which the neutral axis location $z_n = -\dot{e}_p/\dot{\theta}$ lies outside the interval (z_{min}, z_{max}). All such z_n locations

[1] Strictly speaking, the vectors are orthogonal in the standard geometric sense only if the scales on the axes are chosen such that the 'work product' $\boldsymbol{s}^T\boldsymbol{e}_p$ be identical with the standard scalar product of geometric vectors.

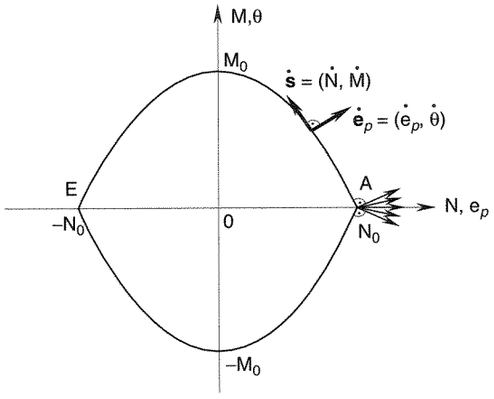


Figure 13.9 Normality rule

represent the case of uniform tension, or of uniform compression, depending on the sign of $\dot{e}_p$. A uniform stress distribution is obtained for any z_n between z_{max} and ∞, or between $-\infty$ and z_{min}. For the limit cases $z_n = z_{max}$ and $z_n = z_{min}$ the preceding derivation applies, and so the vector $\dot{\boldsymbol{e}}_p$ is normal to the upper branch of the yield surface at point A (if $\dot{\theta} = -\dot{e}_p/z_{min} > 0$), or to its lower branch (if $\dot{\theta} = -\dot{e}_p/z_{max} < 0$); see Figure 13.9. For any intermediate value of $\dot{\theta}$, the plastic deformation rate vector $\dot{\boldsymbol{e}}$ lies between the directions for the two limiting cases. Thus, we reach the following important conclusion for a corner on the yield surface (also known as a *vertex*).

Normality rule at a corner: *The increment of the plastic deformation vector at a corner of the yield surface lies within a fan of directions limited by the normals to the yield surface on both sides of the corner.*

This rule has a broad validity, not limited to the interaction of normal forces and bending moments. In the deformation theory of plasticity, similar statements hold when we replace the increments of plastic generalized strains by their total values.

13.2 BASIC THEOREMS

Recall that the dissipation power (under uniaxial stress) is defined as

$$D_{\text{int}} = \int_V \sigma \dot{\varepsilon}_p \, \mathrm{d}V \tag{13.34}$$

where $\dot{\varepsilon}_p$ is the plastic strain rate. Plastic strains in a generalized plastic hinge are lumped into the plastic extensions $\bar{e}_p(z)$ that are concentrated in a single cross section. Consequently, the rate of dissipation in a single hinge can be expressed as

$$D_{\text{int}} = \int_A \sigma(z) \dot{\bar{e}}_p(z) \, \mathrm{d}A \tag{13.35}$$

because the extension $\bar{e}_p(z)$ corresponds to the plastic strain integrated over the length of the plastic hinge. Substituting expression (13.1) for the plastic extension, we can

transform (13.35) into

$$D_{\text{int}} = \int_A \sigma(z)(\dot{e}_p + \dot{\theta} z)\,\mathrm{d}A = \int_A \sigma(z)\,\mathrm{d}A\,\dot{e}_p + \int_A \sigma(z)\,z\,\mathrm{d}A\,\dot{\theta} = N\dot{e}_p + M\dot{\theta} \quad (13.36)$$

The dissipative work is done by the normal force on the plastic extension of the beam axis, and by the bending moment on the rotation of the plastic hinge. Setting $N = 0$, we obtain as a special case formula (2.22) for the dissipation power in a plastic hinge under pure bending. Setting $M = 0$ and identifying the normal force N with the axial force S, we obtain formula (2.2) for the dissipation power in a bar under axial tension or compression. The total dissipation rate in a structure is given by the sum of dissipation rates in all plastic hinges.

The basic properties of the plastic flow in a generalized plastic hinge (convexity and normality) are closely related to the postulate of maximum plastic dissipation. In the context of multiaxial plasticity, they have to be postulated as the hypotheses of the theory; cf. Chapter 15. The actual behavior of some materials might violate these assumptions, and then a number of basic theorems that rely on the aforementioned postulate cease to hold. However, in the present context of plasticity under uniaxial stress, the postulate is almost automatic. Recall that, for one material point, we have

$$\sigma^* \dot{\varepsilon}_p - \mathcal{D}(\dot{\varepsilon}_p) \le \sigma \dot{\varepsilon}_p \quad (13.37)$$

where σ^* is any plastically admissible stress, $\dot{\varepsilon}_p$ is a given plastic strain rate, and $\mathcal{D}(\dot{\varepsilon}_p)$ is the corresponding plastic dissipation rate per unit volume.

Suppose that we are given the rates of generalized strains $\dot{e}_p$ and $\dot{\theta}$ at a certain plastic hinge, and some values of internal forces N^* and M^* that are completely arbitrary (independent of $\dot{e}_p$ and $\dot{\theta}$), but satisfy the condition of plastic admissibility, $f(N^*, M^*) \le 0$. Let us denote by N and M the actual internal forces generated by the given deformation rates, and by $\sigma(z)$ the corresponding stress distribution. As the forces N^* and M^* are inside the plastic limit envelope, they can be represented as the resultants of a normal stress distribution $\sigma^*(z)$ that is plastically admissible at any point of the cross section. Integrating (13.37) along each longitudinal 'fiber' of the generalized plastic hinge (i.e. along x, at constant y and z), we have

$$\sigma^*(z)\dot{\varepsilon}_p(z) \le \sigma(z)\dot{\varepsilon}_p(z) \quad (13.38)$$

Now we integrate this inequality over the entire cross section. Expressing the extension rate $\dot{\varepsilon}_p(z)$ in terms of the generalized strain rates according to (13.1) we can write the integral of the left-hand side as

$$\int_A \sigma^*(z)\dot{\varepsilon}_p(z)\,\mathrm{d}A = \int_A \sigma^*(z)(\dot{e}_p + \dot{\theta} z)\,\mathrm{d}A = N^*\dot{e}_p + M^*\dot{\theta} = \boldsymbol{s}^{*T}\dot{\boldsymbol{e}}_p \quad (13.39)$$

and the integral of the right-hand side as

$$\int_A \sigma(z)\dot{\varepsilon}_p(z)\,\mathrm{d}A = \int_A \sigma(z)(\dot{e}_p + \dot{\theta} z)\,\mathrm{d}A = N\dot{e}_p + M\dot{\theta} = \boldsymbol{s}^T\dot{\boldsymbol{e}}_p = D_{\text{int}}(\dot{\boldsymbol{e}}_p) \quad (13.40)$$

We have proven that

$$\boldsymbol{s}^{*T}\dot{\boldsymbol{e}}_p \le D_{\text{int}}(\dot{\boldsymbol{e}}_p) \equiv \boldsymbol{s}^T\dot{\boldsymbol{e}}_p \quad (13.41)$$

for any plastically admissible internal forces $\boldsymbol{s}^*$. Summing the contributions of all generalized plastic hinges, we obtain a similar inequality for the entire structure. Now, vector $\boldsymbol{s}$ collects the moments and normal forces at critical sections, and $\boldsymbol{e}_p$ is the work-conjugate vector of plastic generalized strains.

Postulate of maximum plastic dissipation
The actual internal forces $\boldsymbol{s}$ corresponding to the given rates of plastic generalized strains $\dot{\boldsymbol{e}}_p$ maximize the plastic dissipation rate among all the plastically admissible vectors of internal forces $\boldsymbol{s}^*$.

The previous postulate makes it possible to extend the basic theorems of perfect elastoplasticity to frames with generalized plastic hinges. The proofs remain formally unchanged, but the internal force vector $\boldsymbol{s}$ now contains both the normal forces and the bending moments, and the corresponding generalized strains $\boldsymbol{e}_p$ include the plastic axial extensions and the rotations of the generalized plastic hinges. The reader is advised to review the proofs of the bound theorems of limit analysis and of the shakedown theorems in order to verify their general validity.

Example 13.2: Calculate the collapse load for the frame in Figure 13.10(a) taking into account the effect of axial forces. The columns and the beam have the same rectangular cross section with a depth-to-span ratio $h/2L = 1 : 10$.

Solution: We will follow the static approach. The structure is statically indeterminate to the third degree, and so we select three forces as the redundants. Our choice will be the internal forces at the cross section just to the left of the axis of symmetry, which are identical to the end forces X_{32}, Z_{32} and M_{32}: see Figure 13.10(b). Given the redundants plus the load F, we can calculate all the internal forces from equilibrium. If the internal forces are plastically admissible at all the critical cross sections, the state is statically admissible. Our goal is to find the statically admissible state with the largest possible value of the load F.

A simplified analysis, neglecting the effect of axial forces, leads to the beam failure mechanism that corresponds to the collapse load $\hat{F}_0 = 4M_0/L$. When the axial force is taken into account, we must distinguish whether the hinge in a corner of a frame forms on the column side or on the beam side of the corner. This is because the axial forces on either side of the corner are different, although the bending moment is the same. Due to the type of loading, it is reasonable to expect that the hinges form on

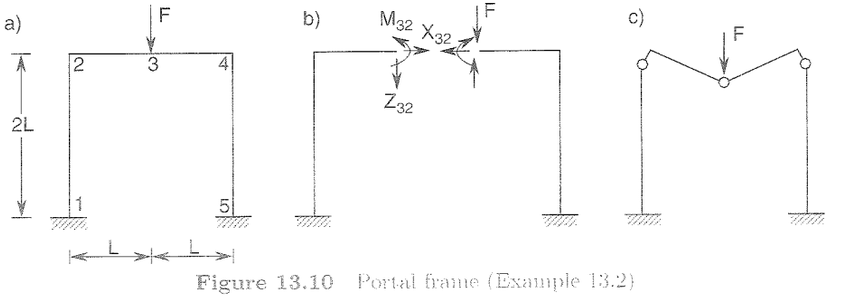


Figure 13.10 Portal frame (Example 13.2)

the column side of the corner; see Figure 13.10(c). Of course, in a complete analysis we must check the validity of this assumption once we have obtained the corresponding solution.

It immediately follows from symmetry and from equilibrium of joint 3 that $Z_{32} = F/2$, and so we can reduce the number of redundants to two. Based on equilibrium considerations, it is easy to compute the bending moment and normal force at cross section 21,

$$N_{21} = -\frac{1}{2}F \tag{13.42}$$

$$M_{21} = M_{32} - \frac{1}{2}FL \tag{13.43}$$

The internal forces at cross section 3 are directly the redundants $N_{32} = X_{32}$ and M_{32}, and due to symmetry, the conditions of plastic admissibility at cross section 45 are equivalent to those at cross section 21. We are looking for a triplet (F, N_{32}, M_{32}) satisfying the conditions of plastic admissibility

$$f(N_{21}, M_{21}) \le 0 \quad \dots \quad \left(\frac{F}{2N_0}\right)^2 + \frac{|M_{32} - FL/2|}{M_0} \le 1 \tag{13.44}$$

$$f(N_{32}, M_{32}) \le 0 \quad \dots \quad \left(\frac{N_{32}}{N_0}\right)^2 + \frac{|M_{32}|}{M_0} \le 1 \tag{13.45}$$

and maximizing the value of F. If the plastic normal force N_0 is large compared to M_0/L we can neglect the quadratic terms in (13.44)–(13.45). The corresponding set of statically admissible solutions is graphically presented in terms of dimensionless parameters FL/M_0 and M_{32}/M_0 in Figure 13.11(a). The largest statically admissible load, $\hat{F}_0 = 4M_0/L$, is the collapse load according to the simplified analysis with axial forces neglected. When the quadratic terms are taken into account, the constraint on M_{32} dictated by (13.45) does not change, because the redundant force N_{32} does not appear in any other condition and can be set to zero (the least restrictive case). Condition (13.44) now becomes quadratic, and the corresponding straight lines are replaced by the parabolas shown in Figure 13.11(b) (where the parabolas

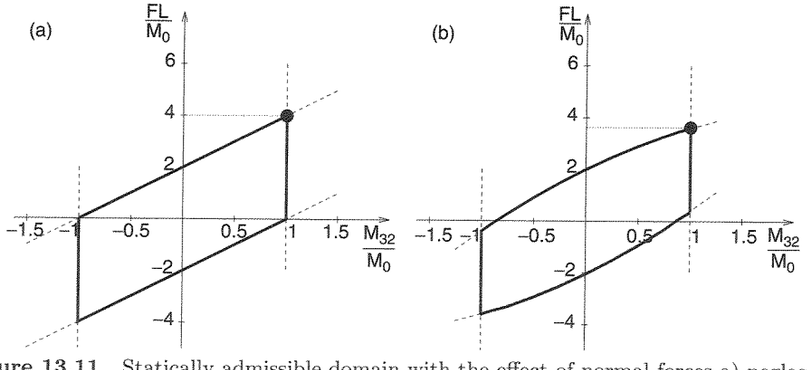


Figure 13.11 Statically admissible domain with the effect of normal forces a) neglected, b) taken into account

are deliberately plotted for a depth-to-span ratio much higher than the prescribed one, because for $h/2L = 1:10$ they would be almost identical with the straight lines from Figure 13.11(a)). For realistic depth-to-span ratios, the apex of the upper parabola is outside the interval $-1 \le M_{32}/M_0 \le 1$, and the maximum statically admissible value of F is determined by the intersection of this parabola with the straight line $M_{32}/M_0 = 1$. This leads to a quadratic equation for the collapse load F_0,

$$\left(\frac{M_0}{LN_0}\right)^2 \left(\frac{F_0 L}{M_0}\right)^2 + 2\frac{F_0 L}{M_0} - 8 = 0 \tag{13.46}$$

Since $M_0 = W_0\sigma_0 = bh^2\sigma_0/4$ and $N_0 = A\sigma_0 = bh\sigma_0$, the factor $(M_0/LN_0)^2$ can be expressed as $(h/4L)^2$. For the given depth-to-span ratio $h/2L = 1:10$ we have $(h/4L)^2 = (1/20)^2 = 1/400$, which means that the quadratic term in (13.46) is indeed very small compared to the linear term. The collapse load $F_0 = 3.98M_0/L$ differs from $\hat{F}_0$ only by 0.5%, which is negligible from the practical point of view. Even for $h/2L = 1:5$ we would still get a reduction of only 2%. The simplified solution neglecting the effect of axial forces is therefore fully justified in the present situation. □

13.3 SIMPLE ESTIMATES OF COLLAPSE LOAD

As we have seen, an exact analysis of the collapse load of frames based on the curvilinear yield surface is possible but complicated, since it leads to nonlinear equations. This is because the limit bending moment depends nonlinearly on the limit axial force. Example 13.2 has shown that in some cases, the effect of normal forces on the collapse load is negligible, and the standard analysis taking into account only the bending moments is fully sufficient. It is therefore important to develop simple methods permitting easy detection of such situations.

The relative importance of normal forces strongly depends on the slenderness of the frame members (beams and columns). If the shape of the cross section is kept constant, the plastic limit moment grows faster than the plastic normal force when the size of the cross section increases. This means that for bulky members the effect of normal forces on the plasticization process becomes more important. It is useful to introduce the parameter

$$h_0 = \frac{M_0}{N_0} = \frac{W_0\sigma_0}{A\sigma_0} = \frac{W_0}{A} \tag{13.47}$$

which is a geometrical characteristic of the cross section and has the dimension of length. In fact, h_0 is one half of the 'internal arm' in the plastic limit state, i.e. of the distance between the resultant of the tensile stresses and the resultant of the compressive stresses. For example, for a rectangular cross section we obtain

$$h_0 = \frac{W_0}{A} = \frac{bh^2/4}{bh} = \frac{h}{4} \tag{13.48}$$

The slenderness of a member is then characterized by the ratio ℓ/h_0 where ℓ is the member length. Here we define slenderness for the purpose of plastic analysis. Note that in the analysis of elastic stability, slenderness has a different meaning – the ratio

of the effective length to the cross sectional radius of inertia; e.g. see Bažant and Cedolin (1991).

For an *ideal sandwich cross section*, which is an idealized I-section with thin flanges and a negligible web, the internal arm is equal to the full depth of the cross section, h, and so $h_0 = h/2$. Note also another interesting property of the ideal sandwich section: the elastic domain coincides with the elastoplastic one, i.e. the plastic limit state is reached immediately after the onset of yielding. This is related to the fact that the web is neglected and the flanges are assumed to be thin compared to the depth of the section. In particular, the plastic limit moment M_0 is equal to the elastic limit moment M_{el}.

Let us denote the safety factors determined with and without taking into account the effect of normal forces by μ_0 and $\hat{\mu}_0$, respectively. Suppose that we have analyzed the structure by the standard technique (i.e. neglecting the effect of normal forces on the formation of plastic hinges), and we know the corresponding safety factor, $\hat{\mu}_0$. A rough lower bound on μ_0 can be constructed quite easily. The internal forces at collapse determined by the standard technique satisfy the equations of equilibrium but, due to the presence of nonzero normal forces, they are not plastically admissible in the generalized plastic hinges. A statically admissible state is obtained if we multiply the internal forces as well as the load multiplier by a common scalar factor such that the resulting forces become plastically admissible. This corresponds to a radial return in the space of internal forces such that the point representing the reduced state is just on the yield surface. For each critical cross section ij we determine a scaling factor ξ_{ij} from the condition

$$f(\xi_{ij}N_{ij}, \xi_{ij}M_{ij}) = 0 \tag{13.49}$$

where N_{ij} and M_{ij} are the internal forces at collapse evaluated by the standard analysis.

Multiplying the internal forces and the safety factor $\hat{\mu}_0$ by the smallest scaling factor

$$\xi_{min} = \min_{ij} \xi_{ij} \tag{13.50}$$

we obtain a statically admissible state, and so $\xi_{min}\hat{\mu}_0$ is a lower bound on μ_0. If ξ_{min} is sufficiently close to 1, the effect of normal forces is negligible, and we can avoid the laborious analysis of the model with generalized plastic hinges.

Example 13.3: Evaluate simple bounds on the safety factor for the frame from Examples 7.3 and 7.5, reproduced in Figure 13.12(a), taking into account the effect of normal forces. Assume that the cross section is close to an ideal sandwich cross section.

Solution: For an ideal sandwich, the plastic limit envelope coincides with the elastic limit envelope; see Problem 13.2. From the general expression for the elastic limit envelope (13.16) and from the relation $M_{el} = M_0$, it follows that the condition of plastic admissibility can be written as

$$\frac{|N|}{N_0} + \frac{|M|}{M_0} \le 1 \tag{13.51}$$

The bending moments at collapse calculated in Example 7.5 are reproduced in Figure 13.12(b), and the corresponding normal forces determined from equilibrium

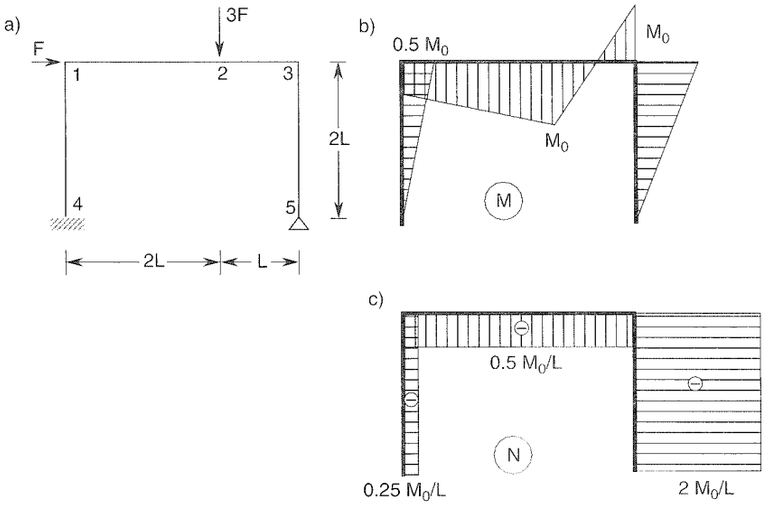


Figure 13.12 Frame from Example 13.3

considerations are given in Figure 13.12(c). According to the theory neglecting the effect of normal forces, the structure collapses when $\hat{P}_0 = 0.75M_0/L$. For each critical cross section ij we determine a scaling factor ξ_{ij} such that the internal forces N_{ij} and M_{ij} multiplied by ξ_{ij} are plastically admissible; cf. (13.49). From

$$\frac{\xi_{ij}|N_{ij}|}{N_0} + \frac{\xi_{ij}|M_{ij}|}{M_0} = 1 \tag{13.52}$$

we get

$$\xi_{ij} = \frac{1}{n_{ij} + m_{ij}} \tag{13.53}$$

where

$$n_{ij} = |N_{ij}|/N_0 \tag{13.54}$$
$$m_{ij} = |M_{ij}|/M_0 \tag{13.55}$$

For example, for cross section 12 we have

$$n_{12} = |N_{12}|/N_0 = 0.5M_0/LN_0 = 0.5h_0/L \tag{13.56}$$
$$m_{12} = |M_{12}|/M_0 = 0.5M_0/M_0 = 0.5 \tag{13.57}$$
$$\xi_{12} = \frac{1}{0.5h_0/L + 0.5} = \frac{2}{1 + h_0/L} \tag{13.58}$$

For an ideal sandwich, the parameter $h_0 = M_0/N_0 = W_0/A$ is equal to $h/2$ where h is the depth of the cross section. For physically meaningful (i.e. positive) values of h_0, the scaling factor ξ_{12} is clearly always greater than one, which means that the internal forces at this cross section are plastically admissible even without rescaling. The other cross sections can be processed similarly; see Table 13.1.

Table 13.1 Normalized internal forces at critical cross sections

ij	$n_{ij}L/h_0$	m_{ij}
14	0.25	0.5
12	0.5	0.5
23	0.5	1
32	0.5	1
35	2	1

The smallest value of the scaling factor is obviously obtained at cross section 35, i.e. at the top of the right column. This is true for any ratio h_0/L. The resulting safety reduction factor is

$$\xi_{min} = \min_{ij} \xi_{ij} \quad \xi_{35} \quad \frac{1}{1 + 2h_0/L} \tag{13.59}$$

As might have been expected, for $h_0/L \to 0$ we obtain $\xi_{min} \to 1$, which is the limit case in which the effect of axial forces vanishes. This is, of course, true only in the context of a geometrically linear theory that neglects the second-order effects of deflections on the distribution of internal forces. For slender columns, these effects must be taken into account, and then the role of the normal force becomes important.

For the ratio $h/2L = 1:10$, the safety reduction factor is

$$\xi_{min} \quad \frac{1}{1 + 2h_0/L} = \frac{1}{1 + h/L} = \frac{1}{1 + 0.2} = 0.833 \tag{13.60}$$

This means that the collapse load is no smaller than

$$P_c \quad \xi_{min}\hat{P}_0 = 0.833 \times 0.75\frac{M_0}{L} = 0.625\frac{M_0}{L} \tag{13.61}$$

The reduction of the ultimate load by 17% is already appreciable. However, this is only a conservative, crude estimate. The exact safety factor will be evaluated in Example 13.5. □

The ideal sandwich is an extreme case in the sense that, for given N_0 and M_0, its plastic limit envelope is the smallest possible plastic limit envelope. This readily follows from the property of convexity. The parallelogram may thus be used as a safe approximation of the plastic limit envelope for a cross section of any shape. As we always have $m_{ij} \leq 1$ (otherwise the condition of plastic admissibility under pure bending would be violated), we can write

$$\xi_{ij} = \frac{1}{n_{ij} + m_{ij}} \geq \frac{1}{n_{ij} + 1} \geq 1 - n_{ij} \tag{13.62}$$

and

$$\mu_0 \geq \xi_{min}\hat{\mu}_0 \geq \hat{\mu}_0 \left(1 - \max_{ij} \frac{|N_{ij}|}{N_0}\right) \tag{13.63}$$

Albeit very crude, this is a universally applicable safe estimate of the safety factor (i.e. an estimate preserving a lower bound). For example, if the normal forces calculated

for the standard model (i.e. the model with plastic hinges affected only by bending moments) nowhere exceed 10% of the plastic axial force, we can be sure that the reduction of the safety factor due to the effect of normal forces is not greater than 10%.

The behavior of frames with members of a rectangular cross section is sensitive to the effect of normal forces much less than is the behavior of an ideal sandwich. The shape of the plastic limit envelope is parabolic, and therefore the bending moment carried by the cross section decreases more slowly as the normal force increases. The yield condition

$$\left(\frac{\xi N}{N_0}\right)^2 + \frac{\xi|M|}{M_0} = 1 \tag{13.64}$$

leads to a quadratic equation

$$n^2\xi^2 + m\xi - 1 = 0 \tag{13.65}$$

for the scaling factor ξ. The solution

$$\xi = \frac{\sqrt{m^2 + 4n^2} - m}{2n^2} \tag{13.66}$$

can safely be estimated as

$$\xi \geq 1 - n^2 \tag{13.67}$$

This means that, for a rectangular section, we have

$$\mu_0 \geq \hat{\mu}_0 \left(1 - \max_{ij} \frac{N_{ij}^2}{N_0^2}\right) \tag{13.68}$$

For example, if the normal forces do not exceed 20% of the plastic axial force, the reduction of the safety factor is at most 4%. It should be emphasized that estimate (13.68) is valid only if all critical sections are rectangular, while estimate (13.63) is universally applicable. The behavior of massive cross sections is closer to the rectangular section while thin-walled cross sections is closer to the ideal sandwich. Thus, the latter are more sensitive to the effect of normal forces.

13.4 APPLICATION OF LINEAR PROGRAMMING

As already mentioned, the yield conditions derived by taking into account the interaction of bending moments and normal forces are typically nonlinear. Linear programming techniques can only be applied to such models if the yield conditions are linearized.

For this purpose we may replace the curvilinear yield surface by a polygon. The equations describing the sides of the polygon are obviously linear with respect to N and M. To preserve the applicability of the bound theorems, it is necessary to use interior and exterior approximations, i.e. linearized yield surfaces that lie totally inside or outside the exact yield surface. Examples of the interior and exterior approximations are shown in Figure 13.13. With regard to the approximate yield surface consisting of straight line segments, it is interesting to note that it represents

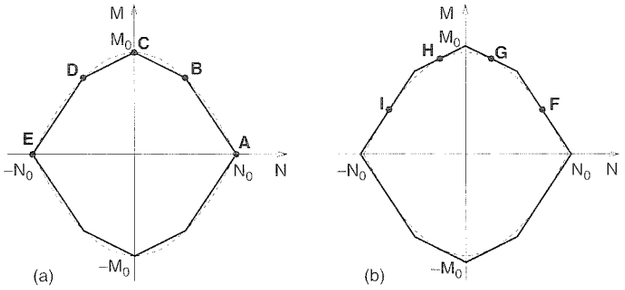


Figure 13.13 Linearization of the plastic limit envelope: a) interior approximation, b) exterior approximation

the exact yield surface for an idealized cross section that has no web and a finite number of thin horizontal flanges.

Consider first the interior approximation (Figure 13.13(a)). Let μ_0 denote the exact safety factor for the actual yield surface, and let μ_{int} denote the safety factor based on approximate yield surfaces that lie in the interior of the actual ones (for each critical cross section). Every stress state statically admissible with respect to the interior approximation is also statically admissible for the actual yield surface. This means that any lower bound $\mu_{\text{int,s}}$ on the approximate safety factor μ_{int} is at the same time a lower bound on the exact safety factor, μ_0.

Next, consider the exterior approximation (Figure 13.13(b)). Let μ_{ext} denote the safety factor for the approximate yield surface lying outside the actual one. Now the situation is the opposite – the actual internal forces at collapse are statically admissible for the exterior approximation, and so $\mu_0 \leq \mu_{\text{ext}}$. Furthermore, any kinematically admissible multiplier $\mu_{\text{ext,k}}$ based on the exterior approximation is greater than or equal to the exact safety factor μ_{ext} for the exterior approximation, according to the upper bound theorem applied to the exterior approximation. Hence, the upper bound is preserved.

The previous conclusions are summarized in the inequality

$$\mu_{\text{int,s}} \leq \mu_{\text{int}} \leq \mu_0 \leq \mu_{\text{ext}} \leq \mu_{\text{ext,k}} \tag{13.69}$$

Of course, this does not mean that the interior approximation must be used exclusively with the static approach, nor that the exterior approximation must be used exclusively with the kinematic approach. We can, for example, compute μ_{int} by the kinematic approach, however, keeping in mind that the intermediate values $\mu_{\text{int,k}}$ do not provide any bounds on the exact safety factor μ_0. Only the optimal value $\mu_{\text{int}} = \min \mu_{\text{int,k}}$ is a *lower* bound on μ_0.

After a slight generalization, the linear programming formulations developed in previous chapters are applicable to frames with generalized plastic hinges. In the *static formulation*, it is no longer possible to write the conditions of plastic admissibility simply as $|\boldsymbol{s}| \leq \boldsymbol{s}_0$. Every nonlinear condition of plastic admissibility at a critical cross section, $f(N, M) \leq 0$, will be replaced by several linearized conditions, each of which involves two internal forces – the bending moment and the normal force. Without any loss of generality, we can describe such conditions by a linear inequality

$$a\frac{N}{N_0} + b\frac{M}{M_0} \leq 1 \tag{13.70}$$

or

$$ah_0 N + bM \leq M_0 \tag{13.71}$$

where a and b are dimensionless coefficients and $h_0 = M_0/N_0$. We can always write the condition of plastic admissibility with a positive right-hand side and with the 'less than or equal to' sign because the stress-free state, $N = 0$ and $M = 0$, must always be inside the elastic domain. The complete system of linearized conditions of plastic admissibility can be written as

$$\boldsymbol{Y}\boldsymbol{s} \leq \tilde{\boldsymbol{s}}_0 \tag{13.72}$$

where matrix $\boldsymbol{Y}$ contains the coefficients ah_0 and b multiplying individual internal forces in the linearized conditions of plastic admissibility, and column matrix $\tilde{\boldsymbol{s}}_0$ contains the corresponding terms M_0 on the right-hand side. Note that $\tilde{\boldsymbol{s}}_0$ now contains one entry per linearized condition of plastic admissibility, and so the value $M_{0,ij}$ is repeated many times for every critical cross section ij. The notation with a superposed tilda should emphasize the difference compared to the column matrix $\boldsymbol{s}_0$, which contains only one entry per cross section.

It is useful to derive general expressions describing the segments of the linearized yield surface. The *interior approximation* is usually constructed by connecting selected points located on the exact yield surface. Due to convexity, the resulting polygon is guaranteed to lie inside the exact yield surface. Let $(\bar{N}_1, \bar{M}_1)$ and $(\bar{N}_2, \bar{M}_2)$ be two points on the exact yield surface. Denoting $\bar{n}_i = \bar{N}_i/N_0$, $\bar{m}_i = \bar{M}_i/M_0$, $i = 1, 2$, and solving the system of linear equations

$$a\bar{n}_1 + b\bar{m}_1 = 1 \tag{13.73}$$

$$a\bar{n}_2 + b\bar{m}_2 = 1 \tag{13.74}$$

for unknowns a and b, we obtain

$$a = \frac{\bar{m}_2 - \bar{m}_1}{c}, \qquad b = \frac{\bar{n}_1 - \bar{n}_2}{c} \qquad \text{where} \quad c = \bar{n}_1\bar{m}_2 - \bar{n}_2\bar{m}_1 \tag{13.75}$$

The *exterior approximation* usually consists of segments that are tangent to the exact yield surface. Again, convexity implies that the exact yield surface is fully contained in the resulting polygon. Suppose that we want to describe a segment that is tangential to the nonlinear yield surface at a certain point $(\bar{N}, \bar{M})$. Let us denote by

$$\bar{f}_{,N} \equiv \frac{\partial f}{\partial N}(\bar{N}, \bar{M}), \qquad \bar{f}_{,M} \equiv \frac{\partial f}{\partial M}(\bar{N}, \bar{M}) \tag{13.76}$$

the components of the gradient of the yield function at the given point. The tangent line must be orthogonal to the gradient vector and pass through point $(\bar{N}, \bar{M})$, and so it is described by

$$\bar{f}_{,N}(N - \bar{N}) + \bar{f}_{,M}(M - \bar{M}) = 0 \tag{13.77}$$

After assembling the constant terms at the right-hand side and rescaling, we can cast (13.77) in the format of (13.70) with

$$a = \frac{N_0\bar{f}_{,N}}{c}, \qquad b = \frac{M_0\bar{f}_{,M}}{c}, \qquad \text{where } c = \bar{f}_{,N}\bar{N} + \bar{f}_{,M}\bar{M} \tag{13.78}$$

It can be proven that c is always positive, and so the denominator in the expressions for a and b cannot vanish; see Problem 13.3.

Example 13.4: Set up the linearized conditions of plastic admissibility for a rectangular cross section corresponding to a) the interior approximation in Figures 13.13(a),(b) the exterior approximation in Figure 13.13(b).

Solution: a) The normalized coordinates of points A to E from Figure 13.13(a) are listed in Table 13.2(a). Consider for example the condition corresponding to segment CB. Substituting $\bar{n}_1 = 0$, $\bar{m}_1 = 1$, $\bar{n}_2 = 0.5$ and $\bar{m}_2 = 0.75$ into (13.75), we obtain $a = 0.5$ and $b = 1$. Consequently, the linearized condition of plastic admissibility (13.71) is

$$0.5h_0N + M \le M_0 \tag{13.79}$$

The remaining segments of the octagon approximating the plastic limit envelope can be processed in a similar manner. The resulting set of inequalities (13.72) reads

$$\begin{bmatrix} 0.5h_0 & 1 \\ h_0 & 0.667 \\ 0.5h_0 & -1 \\ h_0 & -0.667 \\ -0.5h_0 & -1 \\ -h_0 & -0.667 \\ -0.5h_0 & 1 \\ -h_0 & 0.667 \end{bmatrix} \begin{Bmatrix} N \\ M \end{Bmatrix} \le \begin{Bmatrix} M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \end{Bmatrix} \tag{13.80}$$

b) Differentiation of the yield function $f(N, M) = (N/N_0)^2 + |M|/M_0 - 1$ that corresponds to a rectangular section yields

$$f_{,N} \equiv \frac{\partial f}{\partial N} = \frac{2N}{N_0^2}, \qquad f_{,M} \equiv \frac{\partial f}{\partial M} = \frac{\operatorname{sgn} M}{M_0} \tag{13.81}$$

The normalized coordinates of points F to I from Figure 13.13(b) and the corresponding gradient components are listed in Table 13.2(b). Consider for example the condition corresponding to the tangent at point F. Substituting $\bar{N} = 0.75N_0$, $\bar{M} = 0.4375M_0$, $\bar{f}_{,N} = 1.5/N_0$ and $\bar{f}_{,M} = 1/M_0$ into (13.78), we obtain $c = 1.5625$, $a = 0.96$ and $b = 0.64$. Consequently, the linearized condition of plastic admissibility (13.71) is

$$0.96h_0N + 0.64M \le M_0 \tag{13.82}$$

The remaining segments of the octagon approximating the plastic limit envelope can be processed in a similar manner. The resulting set of inequalities (13.72) reads

$$\begin{bmatrix} 0.4706h_0 & 0.9412 \\ 0.96h_0 & 0.64 \\ 0.4706h_0 & -0.9412 \\ 0.96h_0 & -0.64 \\ -0.4706h_0 & -0.9412 \\ -0.96h_0 & -0.64 \\ -0.4706h_0 & 0.9412 \\ -0.96h_0 & 0.64 \end{bmatrix} \begin{Bmatrix} N \\ M \end{Bmatrix} \le \begin{Bmatrix} M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \\ M_0 \end{Bmatrix} \tag{13.83}$$

□

Table 13.2 Points on the exact yield surface serving as a basis of the a) interior approximation, b) exterior approximation

point	$n = \frac{N}{N_0}$	$m = \frac{M}{M_0}$
A	1	0
B	0.5	0.75
C	0	1
D	−0.5	0.75
E	−1	0

(a)

point	$n = \frac{N}{N_0}$	$m = \frac{M}{M_0}$	$N_0\bar{f}_{,N}$	$M_0\bar{f}_{,M}$
F	0.75	0.4375	1.5	1
G	0.25	0.9375	0.5	1
H	−0.25	0.9375	−0.5	1
I	−0.75	0.4375	−1.5	1

(b)

In the *kinematic formulation*, it is not convenient to express the dissipation rate directly in terms of the generalized strain rates. Suppose that we know on which segment of the linearized yield surface plastic flow takes place. Vector $\{ah_0,\ b\}^T$ is the gradient of the linear yield function $f(N, M) = ah_0N + bM - M_0$ associated with condition (13.71), and so it is an outward normal to the corresponding segment. Due to normality, the plastic flow can be described by

$$\begin{Bmatrix} \dot{e}_p \\ \dot{\theta} \end{Bmatrix} = \dot{\lambda} \begin{Bmatrix} ah_0 \\ b \end{Bmatrix} \tag{13.84}$$

where $\dot{\lambda} \ge 0$ is the rate of a scalar variable called the *plastic multiplier*. The dissipation rate is now easily evaluated as

$$D_{\text{int}} = N\dot{e}_p + M\dot{\theta} = (ah_0N + bM)\dot{\lambda} = M_0\dot{\lambda} \tag{13.85}$$

In the last step, we have exploited the fact that during plastic flow the yield condition must be satisfied, and so $ah_0N + bM = M_0$. If the active segment of the yield surface is not known *a priori*, we express the plastic deformation rates by a linear combination

$$\dot{\boldsymbol{e}}_p = \sum_i \boldsymbol{n}_i\dot{\lambda}_i = \boldsymbol{N}\dot{\boldsymbol{\lambda}} \tag{13.86}$$

where the columns of $\boldsymbol{N}$ (vectors $\boldsymbol{n}_i$) are outward normal vectors to the yield surface constructed at individual straight segments, and $\dot{\boldsymbol{\lambda}} \ge \boldsymbol{0}$ is the vector of rates of the corresponding plastic multipliers.

As we have seen, the coefficients of the linearized condition of plastic admissibility (13.71) associated with a certain segment are at the same time components of a vector normal to this segment. Recall that the coefficients of plastic admissibility conditions form the rows of matrix $\boldsymbol{Y}$, and that the components of normals to the yield surface form the columns of matrix $\boldsymbol{N}$. With a proper ordering, we have $\boldsymbol{Y} = \boldsymbol{N}^T$, and then the linearized conditions of plastic admissibility (13.72) read

$$\boldsymbol{N}^T\boldsymbol{s} \le \tilde{\boldsymbol{s}}_0 \tag{13.87}$$

The rate of a certain plastic multiplier can be nonzero only if the corresponding yield condition is satisfied. This statement, mathematically written as

$$(\boldsymbol{N}^T\boldsymbol{s} - \tilde{\boldsymbol{s}}_0)^T\dot{\boldsymbol{\lambda}} = 0 \tag{13.88}$$

is one of the Karush–Kuhn–Tucker conditions, to be discussed in Section 15.2.4. The dissipation rate can now be expressed as a function of the plastic multiplier rates,

$$D_{\text{int}} = \boldsymbol{s}^T\dot{\boldsymbol{e}}_p = \boldsymbol{s}^T\boldsymbol{N}\dot{\boldsymbol{\lambda}} = (\boldsymbol{N}^T\boldsymbol{s})^T\dot{\boldsymbol{\lambda}} = \tilde{\boldsymbol{s}}_0^T\dot{\boldsymbol{\lambda}} \tag{13.89}$$

This is obviously a generalization of (13.85). It is important to realize that the *independent* variables characterizing the plastic flow are now $\dot{\boldsymbol{\lambda}}$ and not $\dot{\boldsymbol{e}}_p$. The dissipation rate is uniquely determined by $\dot{\boldsymbol{\lambda}}$ because $\tilde{\boldsymbol{s}}_0$ is a given constant vector. By contrast to that, the rates of plastic generalized strains $\dot{\boldsymbol{e}}_p$ in the expression $D_{\text{int}} = \boldsymbol{s}^T\dot{\boldsymbol{e}}_p$ are multiplied by the internal forces $\boldsymbol{s}$ that are not known *a priori* and not always uniquely determined by $\dot{\boldsymbol{e}}_p$. Moreover, the elastic domain can be unbounded for certain cases (e.g. for an idealized material that yields only in tension, but not in compression), and then some combinations of plastic generalized strains may be inadmissible.

We have prepared all the prerequisites for the formulation of limit analysis of frames with generalized plastic hinges in the format of linear programming. Now we briefly present the two fundamental approaches.

Static approach: A statically admissible state is defined by plastically admissible internal forces that are in equlibrium with a certain multiple of the reference loading. According to the lower bound theorem of limit analysis, the safety factor is obtained by maximizing the statically admissible load multiplier. This can be stated as a problem of linear programming

$$\text{Maximize } g(\boldsymbol{s}, \mu_s) \equiv \mu_s \tag{13.90}$$

subject to

$$\boldsymbol{B}^T\boldsymbol{s} - \bar{\boldsymbol{f}}\mu_s = \boldsymbol{0} \tag{13.91}$$

$$\boldsymbol{N}^T\boldsymbol{s} \le \tilde{\boldsymbol{s}}_0 \tag{13.92}$$

Kinematic approach: A kinematically admissible state is defined by a collapse mechanism that satisfies the conditions of compatibility and leads to a positive external power supplied by the reference loading. According to the upper bound theorem of limit analysis, the safety factor is obtained by minimizing the kinematically admissible load multiplier. This can be stated as a problem of linear programming

$$\text{Minimize } f(\dot{\boldsymbol{\lambda}}, \dot{\boldsymbol{d}}) \equiv \tilde{\boldsymbol{s}}_0^T\dot{\boldsymbol{\lambda}} \tag{13.93}$$

subject to

$$\boldsymbol{N}\dot{\boldsymbol{\lambda}} - \boldsymbol{B}\dot{\boldsymbol{d}} = \boldsymbol{0} \tag{13.94}$$

$$\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}} = 1 \tag{13.95}$$

$$\dot{\boldsymbol{\lambda}} \ge \boldsymbol{0} \tag{13.96}$$

The reader can verify that the static formulation (13.90)–(13.92) is the dual of the kinematic formulation (13.93)–(13.96), and vice versa; see Problem 13.4.

The linear programming problem in (13.93)–(13.96) is a generalization of problem (7.7)–(7.11), which ensued from the simplified method. We will convert it to the standard form using a similar technique as in Section 7.2. The displacement rates can be eliminated by applying appropriate Gauss transformations to the kinematic relations (13.94). This process can formally be described using matrices $\boldsymbol{S}$ and $\boldsymbol{P}$ introduced in Figure 7.2, which have the following properties:

$$\boldsymbol{S}^T\boldsymbol{B} = \boldsymbol{O} \tag{13.97}$$

$$\boldsymbol{P}^T\boldsymbol{B} = \boldsymbol{I} \tag{13.98}$$

Multiplying (13.94) from the left respectively by $\boldsymbol{S}$ and $\boldsymbol{P}$ leads to

$$\boldsymbol{S}^T\boldsymbol{N}\dot{\boldsymbol{\lambda}} = \boldsymbol{0} \tag{13.99}$$

$$\boldsymbol{P}^T\boldsymbol{N}\dot{\boldsymbol{\lambda}} = \dot{\boldsymbol{d}} \tag{13.100}$$

Physically, (13.99) are the compatibility conditions, and (13.100) are the expressions for the displacement rates in terms of the plastic multiplier rates. Substituting (13.100) into (13.95) we eliminate the displacement rates from the problem and arrive at the standard form

$$\text{Minimize } f(\dot{\boldsymbol{\lambda}}) \equiv \tilde{\boldsymbol{s}}_0^T\dot{\boldsymbol{\lambda}} \tag{13.101}$$

subject to

$$\boldsymbol{S}^T\boldsymbol{N}\dot{\boldsymbol{\lambda}} = \boldsymbol{0} \tag{13.102}$$

$$\bar{\boldsymbol{f}}^T\boldsymbol{P}^T\boldsymbol{N}\dot{\boldsymbol{\lambda}} = 1 \tag{13.103}$$

$$\dot{\boldsymbol{\lambda}} \ge \boldsymbol{0} \tag{13.104}$$

Of course, when the problem is solved by hand we do not have to perform all the matrix manipulations because the conditions of compatibility and the expressions for the displacement rates can be constructed directly using the principle of virtual work.

Example 13.5: Evaluate the safety factor for the structure in Figure 13.14(a). Take into account the effect of normal forces. The cross section is a rectangle of depth $h = L/8$.

Solution: The structure is twice statically indeterminate, and so the failure mechanism involves three plastic hinges. First let us neglect the effect of axial forces. Elementary limit analysis easily reveals the failure mechanism in Figure 13.14(b); the corresponding distribution of the bending moments and normal forces is plotted in Figures 13.14(c), (d). According to the simplified theory taking into account only the effect of bending moments, the structure would fail at $\hat{F} = 2.5M_0/L$.

Note that the simplified condition of plastic admissibility, $|M| \le M_0$, can be regarded as an exterior approximation of the 'exact' condition, $(N/N_0)^2 + |M|/M_0 \le 1$; see Figure 13.15(a). Consequently, the safety factor evaluated from the simplified theory is an upper bound on the actual safety factor. We can construct a crude lower bound using the technique explained in Section 13.3. The internal forces corresponding

Table 13.3 Internal forces and scaling factors at critical cross sections (simplified solution)

ij	M_{ij}/M_0	N_{ij}/N_0	ξ_{ij}	$\xi_{ij}M_{ij}/M_0$	$\xi_{ij}N_{ij}/N_0$
21	−1.0	0.375	0.889	−0.889	0.333
32	0.5	0.375	> 1		
34	−0.5	−0.250	> 1		
45	−1.0	−0.250	0.944	−0.944	−0.236
54	−1.0	−0.250	0.944	−0.944	−0.236

to the simplified solution and the scaling factors that bring them back to the plastic limit envelope are listed in Table 13.3; see also Figure 13.15(b). We consider cross sections 32 and 34 separately because they carry different normal forces. On the other hand, at each of nodes 2, 4 and 5 we have to inspect only one critical cross section. The smallest reduction factor $\xi_{min} = \min \xi_{ij} = \xi_{21} = 0.889$ multiplied by $\hat{F}$ gives a lower bound on the safety factor F_0. Now we can bracket the safety factor by

$$0.889\hat{F} \le F_0 \le \hat{F}, \qquad \text{i.e.} \quad 2.222\frac{M_0}{L} \le F_0 \le 2.5\frac{M_0}{L} \tag{13.105}$$

In order to improve this estimate, we will solve the problem using a more accurate approximation of the nonlinear yield surface. In general, we should replace the yield surface by a polygon that matches it as closely as possible. Each side of the polygon at each critical cross section is associated with one plastic multiplier, and so the number of unknowns in the linear programming problem (13.101)–(13.104) rapidly increases as the polygon approaches the nonlinear interaction diagram. Typically, however, only one plastic multiplier per active yield hinge is nonzero (or at most two

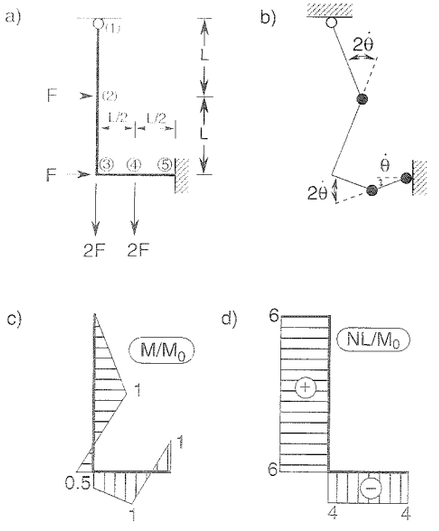


Figure 13.14 Simplified solution of Example 13.5

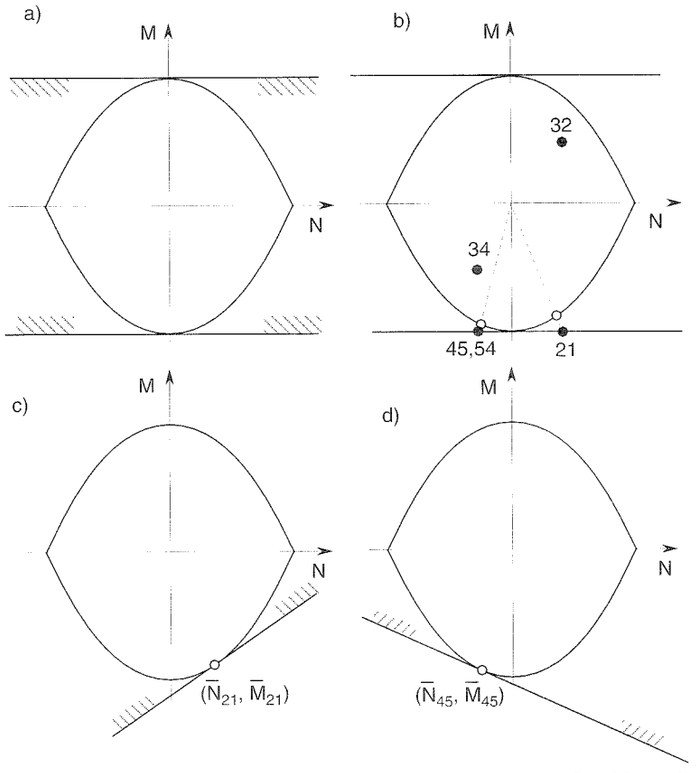


Figure 13.15 Scaling to limit envelope and approximation of yield conditions

if the corresponding point in the interaction diagram is located in a corner of the polygon). The simplified solution, neglecting the effect of normal forces, gives us a clue regarding the segments of the polygons on which plastic flow is expected to take place. We can assume that the actual values of internal forces at collapse will not be too far from the scaled values computed by the simplified method. Therefore, we can dramatically reduce the number of unknowns in our problem by taking into account only one plastic multiplier per critical cross section. This multiplier corresponds to plastic flow taking place in the direction determined by the normal to the interaction curve at the point representing the scaled internal forces from the simplified solution. In other words, we approximate the nonlinear elastic domain of internal forces by a halfplane bounded by a tangent to the actual domain; see Figures 13.15(c),(d). Such an approximation is in general very crude, but the results are excellent if the actual and simplified solutions are not too far apart. We can even assume that plastic hinges will form at the same cross sections as for the simplified solution, and then the number of unknowns further decreases. For the present problem, we introduce only three plastic multipliers corresponding to cross sections 21, 45 and 54. The ratio between the rate of plastic rotation and plastic extension at each plastic hinge is determined by the flow rule. For example, for cross section 21 the internal forces estimated by the simplified method and scaled to the yield surface are $\bar{N}_{21} = \xi_{21}N_{21} = 0.333N_0$ and $\bar{M}_{21} = \xi_{21}M_{21} = -0.889M_0$. Substituting this into expressions (13.81) for the

gradient components of the yield function we get

$$\bar{f}_{,N} = \frac{2\bar{N}}{N_0^2} = \frac{2 \times 0.333 N_0}{N_0^2} = \frac{0.667}{N_0} \tag{13.106}$$

$$\bar{f}_{,M} = \frac{\operatorname{sgn} \bar{M}}{M_0} = \frac{\operatorname{sgn}(-0.889)}{M_0} = -\frac{1}{M_0} \tag{13.107}$$

and then formulae (13.78) furnish $c = 1.111$, $a = 0.6$ and $b = -0.9$. The linearized condition of plastic admissibility (13.71) at cross section 21 is thus

$$0.6 h_0 N_{21} - 0.9 M_{21} \le M_0 \tag{13.108}$$

and the flow rule (13.84) gives

$$\dot{e}_{p,21} = 0.6 h_0 \dot{\lambda}_{21} \tag{13.109}$$

$$\dot{\theta}_{21} = -0.9 \dot{\lambda}_{21} \tag{13.110}$$

In a completely analogous manner, we get the expressions for the deformation rates at cross sections 45 and 54.

$$\dot{e}_{p,45} = -0.4472 h_0 \dot{\lambda}_{45} \tag{13.111}$$

$$\dot{\theta}_{45} = -0.9472 \dot{\lambda}_{45} \tag{13.112}$$

$$\dot{e}_{p,54} = -0.4472 h_0 \dot{\lambda}_{54} \tag{13.113}$$

$$\dot{\theta}_{54} = -0.9472 \dot{\lambda}_{54} \tag{13.114}$$

Now we could easily construct the matrix $\boldsymbol{N}$ and set up the linear programming problem (13.93)–(13.96). The kinematic matrix $\boldsymbol{B}$ would contain only the rows corresponding to the generalized strains of interest, i.e. $\dot{e}_{p,21}$, $\dot{\theta}_{21}$, $\dot{e}_{p,45}$, $\dot{\theta}_{45}$, $\dot{e}_{p,54}$ and $\dot{\theta}_{54}$. Instead of setting up and transforming the kinematic relations (13.94), we will directly construct the conditions of compatibility (13.99) and expressions (13.100), which lead to the standard form (13.101) (13.104).

The structure is twice statically indeterminate, and so we need to find two independent compatibility conditions. Figure 13.16 shows a primary structure subjected to two loading cases in which we apply unit values of the redundants. This provides us with two independent self-equilibrated states of internal forces. According to the priciple of virtual work, the work of these forces on the actual deformations must vanish (because no virtual external forces are applied on the original structure). In this way we obtain the compatibility conditions

$$-\dot{e}_{p,21}/L + 0.5\dot{\theta}_{21} - 0.5\dot{e}_{p,45}/L - 0.5\dot{\theta}_{45} - 0.5\dot{e}_{p,54}/L = 0 \tag{13.115}$$

$$\dot{e}_{p,21}/L - 0.5\dot{\theta}_{45} + \dot{\theta}_{54} = 0 \tag{13.116}$$

Note that it was essential to include the virtual work of the normal forces, since otherwise we would have obtained the compatibility conditions of the simplified method, which neglects the axial plastic deformations. Multiplying by $2L$ and substituting expressions (13.109)–(13.114), we express the conditions of compatibility in terms of the plastic multiplier rates:

$$(-0.9L - 1.2h_0)\dot{\lambda}_{21} + (0.9472L + 0.4472h_0)\dot{\lambda}_{45} + 0.4472 h_0 \dot{\lambda}_{54} = 0 \tag{13.117}$$

$$1.2 h_0 \dot{\lambda}_{21} + 0.9472 L \dot{\lambda}_{45} - 1.8944 L \dot{\lambda}_{54} = 0 \tag{13.118}$$

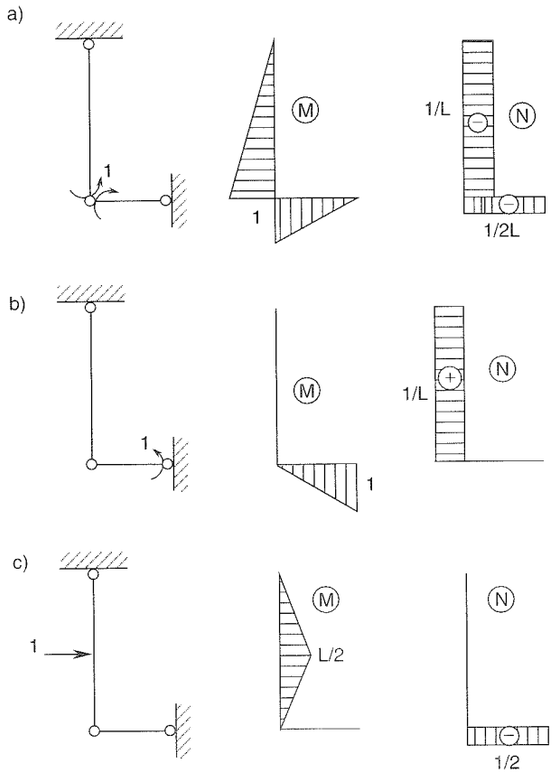


Figure 13.16 Virtual states considered in Example 13.5

We have two conditions for three unknown plastic multiplier rates, which is correct because the assumed mechanism should have one degree of freedom. We select one unknown, say $\dot{\lambda}_{54}$, as an independent one, and express the two remaining unknowns by standard elimination. This would be too tedious in the general symbolic form with arbitrary L and h_0. Therefore, we continue the calculation with the given ratio $h_0 : L = h : 2L = 1 : 16$. After elimination we get

$$\dot{\lambda}_{21} = 1.8801 \dot{\lambda}_{54} \tag{13.119}$$

$$\dot{\lambda}_{45} = 1.8511 \dot{\lambda}_{54} \tag{13.120}$$

Note that both coefficients in the resulting expressions are positive, for if they were not, it would be impossible to satisfy the condition that all plastic multiplier rates be nonnegative, and the assumed mechanism would not be kinematically admissible. The dissipation rate can now be expressed in terms of the independent plastic multiplier rate as

$$D_{\text{int}} = M_0(\dot{\lambda}_{21} + \dot{\lambda}_{45} + \dot{\lambda}_{54}) = 4.731 M_0 \dot{\lambda}_{54} \tag{13.121}$$

We also need to express the rates of displacements on which work is being done by external forces in terms of the plastic multiplier rates. First, we use the principle of virtual work to express them in terms of the nonzero deformation rates. For example,

the unit virtual state in Figure 13.16(c) leads to

$$\dot{u}_2 = -0.5L\dot{\theta}_{21} - 0.5\dot{e}_{45} - 0.5\dot{e}_{54} \tag{13.122}$$

In a similar fashion, we obtain expressions for the other displacement rates of interest,

$$\dot{u}_3 = -\dot{e}_{45} - \dot{e}_{54} \tag{13.123}$$
$$\dot{w}_3 = \dot{e}_{21} \tag{13.124}$$
$$\dot{w}_4 = -0.25L\dot{\theta}_{45} + 0.5\dot{e}_{21} \tag{13.125}$$

Using (13.109)–(13.114) and then (13.119) (13.120), we obtain

$$\dot{u}_2 = 0.8859L\dot{\lambda}_{54} \tag{13.126}$$
$$\dot{u}_3 = 0.0797L\dot{\lambda}_{54} \tag{13.127}$$
$$\dot{w}_3 = 0.0705L\dot{\lambda}_{54} \tag{13.128}$$
$$\dot{w}_4 = 0.1736L\dot{\lambda}_{54} \tag{13.129}$$

The external power done by the kinematically admissible external forces is

$$\dot{W}_{\text{ext}} = F_k\dot{u}_2 + F_k\dot{u}_3 + 2F_k\dot{w}_3 + 2F_k\dot{w}_4 = 2.054F_kL\dot{\lambda}_{54} \tag{13.130}$$

Finally, combining (13.121) and (13.130) into the power equality, we get an improved kinematically admissible load

$$F_k \quad \frac{1.731M_0}{2.054L} = 2.303\frac{M_0}{L} \tag{13.131}$$

To keep the derivations easy to follow, we were showing only four significant digits for all the intermediate results. When the calculation is performed with a high accuracy, the final value of the kinematically admissible load is $F_k = 2.303645M_0/L$. Now we can evaluate the corresponding internal forces from equilibrium and normality conditions. As expected, the bending moments and normal forces at the yield hinges are not too different from the scaled values from the simplified solution. The results are plotted in Figure 13.17 and listed in Table 13.4, which shows also the scaling factors needed to bring the internal forces back to the exact yield surface. Their values indicate that the improved upper bound is very accurate. Internal forces at sections 32 and 34 still remain inside the elastic domain, which means that the assumed locations of plastic hinges have been chosen correctly. The smallest scaling factor is $\xi_{min} = 0.999989$, and so the relative error of the upper bound is about 10^{-5}. The collapse load can be bracketed as

$$0.999989F_k \le F_0 \le F_k \tag{13.132}$$
$$2.303567\frac{M_0}{L} \le F_0 \le 2.303645\frac{M_0}{L} \tag{13.133}$$

There is no need for further improvement of this estimate. If the results were less accurate, we could repeat the entire procedure, starting from the construction of new tangent lines to the interaction diagram. Note that, in this example, the presence of

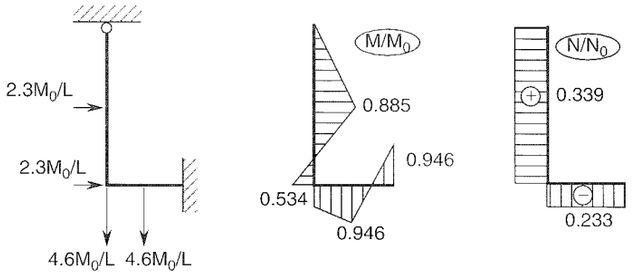


Figure 13.17 Internal forces at collapse

Table 13.4 Internal forces and scaling factors at critical sections (improved solution)

ij	M_{ij}/M_0	N_{ij}/N_0	ξ_{ij}
21	−0.884818	0.339440	0.999966
32	0.534001	0.339440	> 1
34	0.534001	−0.232654	> 1
45	−0.945883	−0.232654	0.999989
54	−0.945883	−0.232654	0.999989

normal forces has reduced the collapse load by 8% compared to the simplified solution that takes into account only the bending moments. □

To emphasize the physical meaning of individual variables, equations and transformations, we have presented the solution of the preceding example in simple terms and in a scalar notation. The reader is advised to think over the structure of the problem, the relation between the scalar and matrix notations, and the relation between the straightforward solution technique used in the example and the simplex method. Problems 13.7 and 13.8 should provide some useful guidelines regarding these considerations.

PROBLEMS

Problem 13.1: Construct the elastic and plastic limit envelopes for a) the I-section in Figures 13.18(a),(b) the T-section in Figure 13.18(b).

Problem 13.2: Construct the elastic and plastic limit envelopes for an idealized sandwich section (consisting of two thin flanges and a negligible web) and derive the expression for the parameter $h_0 = M_0/N_0$.

Problem 13.3*: Prove that the quantity c evaluated from (13.78) is always positive. Hint: Note that c is the scalar product of two vectors. Consider their physical meaning, and exploit the properties of the elastic domain.

Problem 13.4: Construct a table showing duality of the static and kinematic formulations derived in Section 13.3.

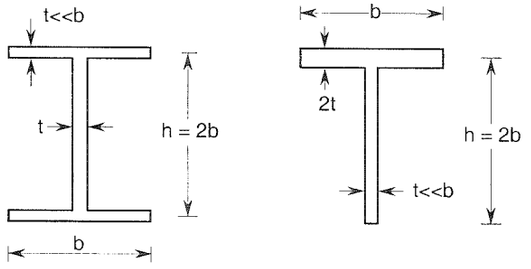


Figure 13.18 Cross-sections for Problem 13.1

Hint: Some equalities will have to be multiplied by -1 to match the signs in the primal and dual problem.

Problem 13.5: Show that the simplified kinematic formulation (7.7)–(7.11) can be regarded as a special case of the general kinematic formulation (13.93)–(13.96).

Problem 13.6*: Formulate precisely the conditions under which the simplified theory (neglecting the effect of axial forces) gives an upper bound on the actual safety factor. In other words: in which cases can the presence of normal forces increase the safety factor?
Hint: Note that the simplified condition of plastic admissibility $|M| \leq M_0$ does not always have to be an exterior approximation of the more exact condition that takes into account interaction with normal forces.

Problem 13.7: Present the solution of Example 13.5 in the matrix form. Link equations appearing in the example to the general matrix equations, write matrices $\boldsymbol{N}$, $\boldsymbol{S}$, $\boldsymbol{P}$, $\boldsymbol{S}^I\boldsymbol{N}$, $\boldsymbol{P}^I\boldsymbol{N}$ and $\boldsymbol{P}^T\boldsymbol{N}\bar{\boldsymbol{f}}$ explicitly, and show which matrix operations correspond to individual steps of the solution. Construct the kinematic matrix $\boldsymbol{B}$ and derive matrices $\boldsymbol{S}$ and $\boldsymbol{P}$ by Gauss transformations.

Problem 13.8*: It seems that, in Example 13.5, we solved a linear programming problem without applying the simplex technique. Formulate the complete problem (13.101)–(13.104) and show how the solution procedure applied in Example 13.5 relates to the simplex method.

14

Plasticity Aspects of Reinforced Concrete

During the 1970s it was fashionable to develop plasticity models for all kinds of failure of reinforced concrete structures. A decade later, though, it transpired that in most situations concrete fails as a quasibrittle material properly described by cohesive fracture mechanics or damage models with localization limiters (Bažant, 1986). The main practical consequence is that the nominal strength of concrete structures exhibits a size effect, which cannot be captured by plasticity, as well as shape effects different from plasticity. In beams and frames, the size effect gets manifested when the moment-rotation relation of the inelastic hinge exhibits post-peak softening. This aspect, quite important for proper design of large reinforced concrete frames, is briefly explained in Appendix F. In this chapter we will focus only on the behavior of reinforced concrete beams, slabs and frames of the kind that can be handled by plasticity.

An exception where plasticity does apply to concrete is the behavior under extreme confining pressures, such that the compressive principal stress of the smallest magnitude is several times larger than the uniaxial compression strength. Such conditions are attained at the tip of a missile (or explosively driven anchor) penetrating into a concrete wall, and in powerful explosions within the mass of concrete. In theory, they would also be attained in the failure of concrete columns reinforced by an unusually heavy steel spiral, of a volume at least 14% of the volume of concrete, or confined in a similarly heavy steel tube (Caner and Bažant, 2000b). Such behavior, however, is beyond the scope of this treatise.

Despite the innate brittleness of concrete, some kinds of failure of reinforced concrete nonetheless exhibit a plastic yielding plateau. They can, therefore, be tackled by the theory expounded in this book. A plastic yield plateau is exhibited by weakly reinforced concrete beams and frames without significant axial compressive force. The reason is that the bending failure of a beam with weak enough tensile reinforcement, called 'under-reinforced', is entirely dominated by the yielding of the tensile steel bars, and the strength of concrete does not matter for bending (although it does for shear and torsion). We will first focus attention on such behavior.

14.1 PLASTIC BENDING CAPACITY OF UNDER-REINFORCED BEAMS WITHOUT AXIAL FORCE

14.1.1 Singly Reinforced Rectangular Cross Section

Consider slender enough beams (approximately, length $\geq$ 8×depth), for which the Bernoulli–Navier hypothesis of cross sections remaining plane and normal holds. Then the longitudinal normal strain ε is distributed linearly throughout the cross section (Figure 14.1). Therefore, the distribution of longitudinal normal stress σ over the beam depth has the same shape as the uniaxial stress-strain diagram. In this chapter, we will use the notation of the American Concrete Institute (ACI) Standard 318 (ACI, 1999), according to which the maximum bending moment (called by ACI the *nominal bending moment*) is denoted as M_n. The stress distribution at M_n (Figure 14.1) is highly curved on the compression side and may involve some strain softening near the compression face, to the extent that it is stabilized against localization by the restraint of adjacent non-softening concrete. Distributed tensile cracks extend at maximum moment from the tensile face almost to the neutral axis (defined as the line where $\varepsilon_x = 0$).

As for tensile behavior, concrete in reinforced beams is generally assumed to resist no tension (i.e. represents a 'no-tension' material); see Figure 14.1. This simplification is on the side of safety and is not as poor as it might seem because the cohesive (crack-bridging) stresses of the tensile cracks are significant only near the neutral axis, where they can contribute to M_n only very little, and also because, in the case of any loading of longer duration, these stresses rapidly relax with time; see Section 29.2.3.

Based on extensive experimental studies in the middle of 20th century, the curved stress-strain distribution on the compression side may be replaced for practical purposes by an equivalent rectangular stress block having a uniform stress magnitude $0.85f'_c$, and a depth a that is less than the distance c from the compressive face to the neutral axis (Figure 14.1); f'_c is the *standard compression strength*, measured on cylinders of diameter 6 in. and length 12 in. aged 28 days.

For validating the code formula, f'_c needs to be taken as the statistical average of test results, $\bar{f}_c$. However, for design purposes, ACI Standard 318 (ACI, 1999, Art. 5.3.2.1), introduces a large safety margin by stipulating that $f'_c = \min(\bar{f}_c - 1.34s, \bar{f}_c - 2.33s + 500$ psi) where 1 psi = 6895 Pa and s = standard deviation of measured strength values. For the frequent case that the data does not assess the standard deviation reliably, ACI Standard specifies $\bar{f}_c = f'_c + 1000$ psi if $f'_c < 3000$ psi, $\bar{f}_c = f'_c + 1400$ psi if $f'_c > 5000$ psi, and $\bar{f}_c = f'_c + 1200$ psi in between; f'_c can be calculated for given $\bar{f}_c$ by inspection.

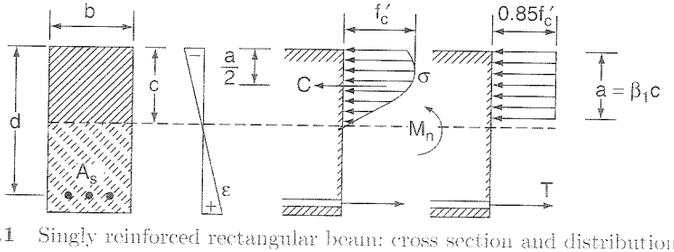


Figure 14.1 Singly reinforced rectangular beam: cross section and distributions of strain and stress

Consider first a singly reinforced rectangular cross section of width b and depth d measured from the compression face to the centroid to steel bars (Figure 14.1). Equilibrium of horizontal forces requires that $(0.85f'_c)ba = A_s f_y = T$, where T = tensile steel resultant, f_y (in ACI notation) = σ_0 = uniaxial yield strength of the bars (statistical average, in this case), and A_s = combined cross section area of all tensile steel bars. So,

$$a = \frac{A_s f_y}{0.85 f'_c b} \tag{14.1}$$

Since the distance of the resultant C of compressive stresses in concrete from the tensile resultant T is $d - \frac{1}{2}a$ (Figure 14.1), the nominal (maximum) bending moment, equivalent to M_0 in our earlier text, is

$$M_n = M_0 = A_s f_y \left(d - \tfrac{1}{2}a\right) \tag{14.2}$$

The design requirement is that $M_u \leq \phi M_n$, where M_u is called by ACI the *ultimate moment* and is obtained by applying specified load factors in various design load combinations, and ϕ is the *capacity reduction factor* (or briefly *understrength factor*), which is specified as $\phi = 0.9$ for the case of bending (and takes into account the additional uncertainty associated with the type of failure mechanism).

Equation (14.2) is valid only if the tensile steel yields before compressed concrete can get crushed. According to the ACI, it may be assumed that concrete gets crushed at strain -0.003 (which is a rather crude approximation, neglecting the size and shape effects). If strain -0.003 in concrete and strain f_y/E_s in steel are reached simultaneously, it follows from the similarity of triangles in Figure 14.1 (based on linearity of strain distribution) that

$$\frac{c}{d} = \frac{0.003}{0.003 + f_y/E_s} \tag{14.3}$$

where $E_s = 29 \times 10^6$ psi = 200 GPa = elastic modulus of steel. This case is called the *balanced condition* of the cross section. To interpret it, one must know the relation of c to a, which is specified by ACI empirically: $a = \beta_1 c$ where $\beta_1 = 0.85$ for $f'_c \leq 4000$ psi = 27.6 MPa and $\beta_1 = 0.65$ for $f'_c \geq 8000$ psi = 55.2 MPa, with linear interpolation in between. Expressing a from (14.1) and substituting $c = a/\beta_1$ into (14.3), one gets

$$\rho_b = \frac{0.85\beta_1 f'_c}{f_y} \frac{0.003}{0.003 + f_y/E_s} \tag{14.4}$$

where ρ_b is the value of the steel ratio $\rho = A_s/bd$ at the balanced condition. The case $\rho < \rho_b$ is called the *under-reinforced condition*, and $\rho > \rho_b$ the *over-reinforced condition*.

ACI prohibits steel ratios $\rho > 0.75\rho_b$ from being used in design, with the factor 0.75 representing a safety margin. In other words, ACI permits only significantly under-reinforced designs. The reason is that, for $\rho > \rho_b$, the failure, taking place as a result of concrete crushing, is brittle. This is manifested by the fact that the bending moment decreases with increasing curvature, which is undesirable as it leads to high sensitivity to minor disturbances and prevents plastic redistribution of internal forces in redundant beams and frames. On the other hand, for $\rho < \rho_b$ the failure is ductile.

or plastic, with the bending moment remaining constant at increasing curvature, as demonstrated by many tests.

Consequently, for under-reinforced concrete beams and frames designed according to the code, plastic limit analysis is applicable. However, it is not applicable for beams and frames with a large compressive force because it is impossible to avoid the over-reinforced condition (i.e. concrete failing before the steel can yield). This is the case for columns with a high axial compressive force, frames or arches with a large horizontal thrust, and generally prestressed concrete beams, columns and frames.

The cross section of a *reinforced concrete slab* (a term favored by concrete engineers instead of the term 'plate', which is standard in mechanics) failing along a hinge line is analyzed in the same way as a rectangular beam. The practical steel ratios are low enough for the slabs to always be under-reinforced (unless they are prestressed). Therefore, the yield hinge theory is valid, and no special discussion is needed. The yield criterion for yield lines in reinforced concrete plates, to be explained in Section 18.4, needs the value of the arm between compression resultant C and tension resultant T, which is $d - a/2$, where a is determined on the basis of (14.1).

14.1.2 Doubly Reinforced Cross Section

Often longitudinal reinforcing steel bars are placed near the compressed face of beams (Figure 14.2). If this compression steel, of combined cross section area A_s', yields, the cross section is analyzed as follows. The axial force from the yielding tensile steel, $A_s f_y$, is imagined to be subdivided into two parts: part $(A_s - A_s')f_y$ that must balance axial compression force $(0.85 f_c')ba$ in concrete; and part $A_s' f_y$ that balances the axial force in the yielding compression steel. So, the maximum bending moment (nominal moment) is

$$M_n \quad M_0 \quad 0.85 f_c' ba \left(d - \tfrac{1}{2}a\right) + A_s' f_y (d - d') \tag{14.5}$$

where d' is the distance of the compression steel centroid from the tensile face and

$$a = \frac{(A_s - A_s')f_y}{0.85 f_c' b} \tag{14.6}$$

To check whether the compression steel indeed starts yielding before concrete reaches the compressive strain of magnitude $|\varepsilon_c| = 0.003$, one must use the neutral axis depth $c = a/\beta_1$ and the linearity of strain distribution (similar triangles) to verify that when the strain at the centroid of compression steel is $-f_y/E_s$, the magnitude of the extreme strain in compressed concrete,

$$|\varepsilon_c| = \frac{f_y}{E_s} \frac{a/\beta_1}{a/\beta_1 - d'} \tag{14.7}$$

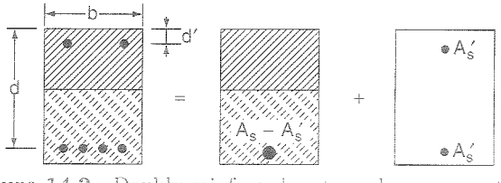


Figure 14.2 Doubly reinforced rectangular cross section

does not exceed 0.003. This yields the condition

$$\rho - \rho' \geq \frac{0.85\beta_1 f_c' d'}{f_y d} \frac{0.003}{0.003 - f_y/E_s} \tag{14.8}$$

where $\rho' = A_s'/bd$ = compression steel ratio.

If condition (14.8) is violated, the compression steel is still in an elastic state. One may then use an iterative procedure, in which the magnitude $|f_s'|$ of stress in compression steel is repeatedly evaluated from the condition

$$\frac{|f_s'|}{E_s} \frac{a/\beta_1}{a/\beta_1 - d'} = 0.003 \tag{14.9}$$

and then replaces f_y in (14.5), where also $a = (A_s f_y - A_s'|f_s'|)/0.85 f_c' b$, from equilibrium of axial forces. The balanced tensile steel ratio is in either case (yielding or nonyielding compression steel) expressed as

$$\rho_b = \bar{\rho}_b + \rho' |f_s'|/f_y \tag{14.10}$$

where $\bar{\rho}_b$ is the balanced steel ratio calculated for a singly reinforced beam with tensile steel area $A_s - A_s'$. The ACI design requirement for under-reinforced beam is $\rho \leq 0.75\bar{\rho}_b + \rho' |f_s'|/f_y$.

14.1.3 T and L Beams

In T beams (or L beams), two cases can arise. The first case occurs when the rectangular stress block of depth a fits within the depth h_f of the flange, $h \leq h_f$. The cross section may then be analyzed as a rectangular beam of width b (b now representing the width of the flange) because the excess area of concrete in such a beam is in extension and thus carries no stress (Figure 14.3(a)).

The second case occurs when $a > h_f$ (Figure 14.3(b)). The concrete cross section is imagined subdivided into two parts: the web part of width b_w and full beam depth,

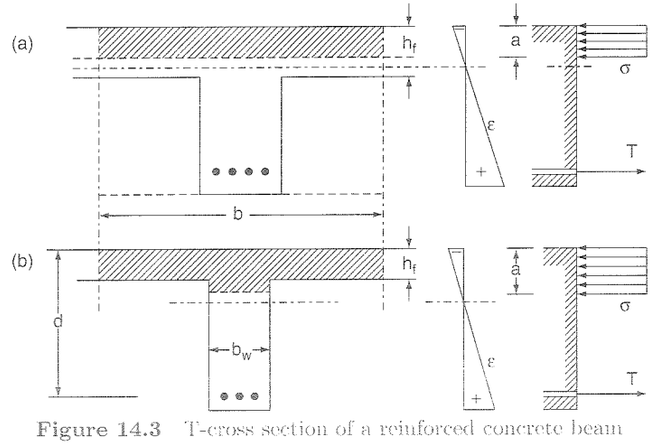


Figure 14.3 T-cross section of a reinforced concrete beam

and the flange part of width $b - b_w$, comprising the flanges on the sides of the web. First we determine the steel area required to balance the compression force resultant in the overhanging portions of the flanges:

$$A_{sf} = 0.85 f'_c h_f (b - b_w)/f_y \tag{14.11}$$

Then we consider the web part as a rectangular cross section containing the remaining reinforcement, and calculate a for this part from formula (14.1) with A_s replaced by $A_s - A_{sf}$ and b by b_w. Adding the bending moments of both steel parts about the centroids of the flange and of the block of depth a in the web, we obtain

$$M_n = A_{sf} f_y \left(d - \tfrac{1}{2} h_f\right) + (A_s - A_{sf}) f_y \left(d - \tfrac{1}{2} a\right) \tag{14.12}$$

To achieve the under-reinforced condition, and thus to ensure plastic response, it is now necessary that

$$\frac{A_s - A_{sf}}{b_w d} \le 0.75 \rho_b \tag{14.13}$$

where ρ_b is calculated from (14.4), same as before.

14.1.4 Moment-Curvature Diagram

A typical moment-curvature diagram of a reinforced concrete beam (unprestressed) is shown in Figure 14.4. At a relatively small load, typically 1/6 to 1/4 of the maximum service load, bending moment M_{cr} that initiates cracking at the tensile face is reached. This moment is calculated from the elastic bending formula for stress, by setting the stress at the tensile face equal to the so-called modulus of rupture (flexural strength) of concrete (allowed by ACI Standard, Eq. 9.9, to be estimated as $f_r \approx 7.5$ psi $\sqrt{f_c/\text{psi}}$, for design purposes).

A rational calculation of the subsequent moment-curvature diagram must take into account the distributed tensile cracks, their cohesive nature, and the so-called *tension stiffening*, i.e. the fact that concrete between two adjacent cracks still adheres to the bar (as a result of either bond or friction) and thus carries a part of the tensile force and restrains the bar. For more detail, in the context of creep, see Section 29.2.3.

To avoid the complexity of such rational calculation, ACI recommends a simplified formula (due to Branson, 1968) for the effective moment of inertia, I_e, giving an

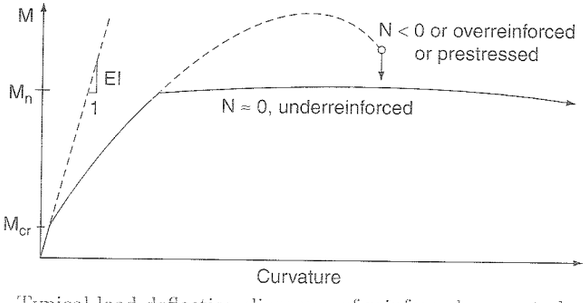


Figure 14.4 Typical load-deflection diagrams of reinforced concrete beams failing by yielding of tensile steel reinforcement or by crushing of concrete

approximate secant bending stiffness applicable throughout the service load range and, as a cruder approximation, up to the start of yielding of steel (if the cross section is under-reinforced):

$$I_e = I_{cr} + \left(\frac{M_{cr}}{M}\right)^n (I_g - I_{cr}) \qquad \text{for } M \ge M_{cr} \tag{14.14}$$

where I_g is the centroidal moment of inertia of the gross cross section (a section in which the steel bars are replaced with concrete, i.e. the cross section is taken as a plain concrete section without reinforcement); I_{cr} is the centroidal moment of inertia of the cracked cross section based on elastic no-tension analysis; and $n = 4$ for curvature at an individual cross section. This formula is also used for quasi-elastic deflection calculation of the whole beam considered to have a uniform bending stifness; in that case $n = 3$ is recommended and M is the maximum bending moment along the beam.

For estimating the elastic modulus at an age of 28 days, ACI specifies Pauw's formula (Pauw, 1960)

$$E_c = \left(\frac{w_c}{1 \text{ lb/ft}^3}\right)^{3/2} \sqrt{\frac{f'_c}{1 \text{ psi}}} \times 33 \text{ psi} \approx \sqrt{\frac{f'_c}{1 \text{ psi}}} \times 57000 \text{ psi} \tag{14.15}$$

where w_c is the specific weight of concrete (between 90 and 155 lb/ft^3) and the last expression is valid for normal weight concrete ($w_c \approx 144$ lb/ft^3). Note that (in contrast to the CEB formula (14.16) below) ACI specifies here the use of the reduced compression strength f'_c rather than the average strength $\bar{f}_c$. However, formula (14.15) was calibrated by Pauw (1960) using the mean strength data. This means that the strains, curvatures and deflections calculated from (14.15) are some increased values. To obtain the average values, f'_c must be replaced in (14.15) by $\bar{f}_c$.

14.1.5 Specifications of CEB Model Code

The design code specifications described so far have been those of ACI. The Euro-International Committee for Concrete (CEB) in collaboration with the International Federation for Prestressing (FIB) developed the CEB-FIP Model Code 1990 (CEB-FIP, 1991). The ACI and CEB codes are based on the same general ideas, but some detailed specifications are different. According to CEB, the design is based on the characteristic compression strength, f_{ck}, defined as the strength below which 5% of all strength measurements may be expected to fall. Testing is performed on cylinders 150 mm in diameter and 300 mm in height (practically the same dimensions as the ACI cylinders) at concrete age 28 days. The characteristic strength may be estimated as $f_{ck} = f_{cm} - 8$ MPa where f_{cm} = mean strength from experiments. We now use the CEB notation, but f_{cm} is obviously the same as $\bar{f}_c$ and f_{ck} is similar to f'_c. The design strength is obtained as $f_{cd} = f_{ck}/\gamma_c$ where γ_c is the partial safety factor, equal to 1.5 for persistent and transient design situations and 1.2 for accidental design situations. For the estimation of the conventional elastic modulus, CEB recommends

$$E_c = \left(\frac{f_{cm}}{10 \text{ MPa}}\right)^{1/3} \times 21.5 \text{ GPa} \tag{14.16}$$

For the design of cross sections under bending, CEB allows the use of an equivalent rectangular stress block. The design strength is reduced by the usual factor 0.85 and

by an additional factor $1 - f_{ck}/(250\ \text{MPa})$, which approximately takes into account the increased brittleness of higher-strength concretes. The extreme compressive strain must not exceed $\varepsilon^*_{cu} = 0.004 - 0.002(f_{ck}/100\ \text{MPa})$. As an alternative, the CEB code specifies a nonlinear uniaxial stress-strain diagram for compression, defined as

$$|\sigma(\varepsilon)| = \begin{cases} 0.85 f_{cd}\left(\dfrac{2|\varepsilon|}{\varepsilon_{c1}} - \dfrac{\varepsilon^2}{\varepsilon_{c1}^2}\right) & \text{for } |\varepsilon| < \varepsilon_{c1} \\ 0.85 f_{cd} & \text{for } \varepsilon_{c1} \le |\varepsilon| \le \varepsilon_{cu} \\ 0 & \text{for } \varepsilon_{cu} < |\varepsilon| \end{cases} \tag{14.17}$$

where $\varepsilon_{c1} = 0.002$ and

$$\varepsilon_{cu} = \begin{cases} 0.0035 & \text{for } f_{ck} \le 50\ \text{MPa} \\ 0.035 \times \dfrac{50\ \text{MPa}}{f_{ck}} & \text{for } 50\ \text{MPa} \le f_{ck} \le 80\ \text{MPa} \end{cases} \tag{14.18}$$

The design of the cross section requires finding the position of the neutral axis from the condition that the force resultant of stress be equal to the normal force (i.e. to zero under pure bending). If the nonlinear stress-strain law (14.17) is used, an iterative solution is required. In view of the recent understanding of localization effects and the sensitivity of post-peak stress-strain curves of concrete to specimen size, it is debatable whether the stress-strain relation leads to more realistic design than does the use of the classical equivalent rectangular stress block, which is simpler.

In the design of singly reinforced rectangular cross sections subject to bending moment, the requirement of underreinforced design is checked by the inequality

$$0.0035\,\frac{d-x}{d} > \frac{f_{yd}}{E_s} \tag{14.19}$$

where d = depth measured from the compression face to the centroid to steel bars, x = depth of compression zone (i.e. c in the ACI notation), f_{yd} = design yield strength of steel, and E_s = elastic modulus of steel. The design yield strength is defined as $f_{yd} = f_{yk}/\gamma_s$ where f_{yk} is the characteristic yield stress of steel and γ_s is the partial safety factor, equal to 1.15 for persistent and transient design situations and 1.0 for accidental design situations. Condition (14.19) is similar but not completely equivalent to that which we derived for the ACI approach; see the balanced condition (14.3).

14.2 LOAD CAPACITY OF REINFORCED CONCRETE COLUMNS AND PRESTRESSED BEAMS

14.2.1 Columns

The failure of reinforced concrete columns is often caused by compression failure of concrete. The behavior is then brittle and the load deflection diagram exhibits post-peak softening, with no horizontal yield plateau. So, strictly speaking, plastic limit analysis is not justified even though it is sanctioned by the current design codes. Therefore, our discussion of this broad topic will be quite terse. The buckling of slender concrete columns in the inelastic range (Bažant and Cedolin, 1991, Chapter 8), will be skipped altogether, being beyond the scope of this book.

Columns are generally loaded by axial compressive force P and bending moment M (P will be considered positive when compressive, and $P = -N$ = normal force). The ratio $e = M/P$ is the eccentricity of load or stress resultant P. When e is very large (much larger than the cross section dimension), the axial force plays almost no role and the column behaves as a beam. The failure states of column fill in the plane (M, P) the failure envelope, whose typical shape is shown in Figure 14.5 (in contrast to the usage in steel plasticity, it is customary to put M as the coordinate and P as the ordinate). The general points of the envelope are somewhat more tedious to calculate than three characteristic points, corresponding to: (1) the case of pure bending, $P = 0$ (or $e \to \infty$), which we already discussed; (2) the case of centric axial compression, $M = 0$ (or $e = 0$); and (3) the balanced condition, in which the compression failure of concrete is reached simultaneously with tensile yielding of steel bars at the tensile face.

In the case of axial compression, load capacity P is taken, according to codes, as the resultant of the yield stresses in the axial bars and the stresses in concrete assumed to be equal to $0.85 f'_c$ (because of uniform stress distribution, factor 0.85 is not justified as in bending, but simply provides an extra safety margin, which is beneficial for large columns as the deterministic size effect is still ignored by codes). In the case of balanced condition, the knowledge of the extreme compression strain in concrete, -0.003, and tensile strain in steel, f_y/E_s, along with the linearity of strain distribution, are used to locate the neutral axis, from which one obtains the depth a of the equivalent rectangular stress block, just like in bending. Taking the axial force and moment resultants, one gets the values of P and M that give the point b shown in Figure 14.5. The segments of the failure envelope from the axial compression point to cb and from b to the pure bending point m are generally curved (and convex), but it is customary to consider them as straight lines, which is on the safe side and avoids more tedious iterative calculation of the intermediate points.

Considering the case of pure tension, in which the load capacity is provided solely by the steel bars, the concrete contributing nothing, one can easily get a fourth point t lying on axis P below the origin (Figure 14.5). Connecting this point one gets a complete failure envelope. Note that this envelope resembles that for combined plastic bending and normal force, which we determined in Chapter 13 (Figure 13.8), but is more pointed and is shifted in the direction of axial compression. The shift is caused by the difference between the tensile and compressive strengths of concrete.

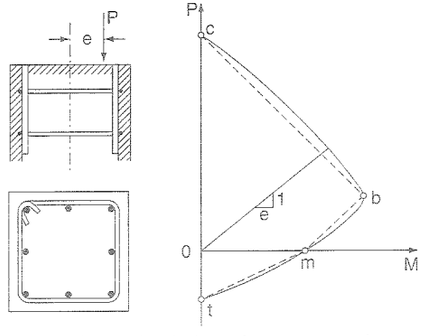


Figure 14.5 Cross-section of a tied reinforced concrete column and its failure envelope (interaction diagram)

It should be noted that, from the viewpoint of the latest research in quasibrittle compression failures, the failure envelope should depend on the size of column (Bažant and Planas, 1998). It should shrink as the size is increased.

14.2.2 Prestressed Concrete Beams

Like columns with not too large load eccentricities, prestressed concrete beams exhibit load-deflection diagrams with post-peak softening instead of a horizontal yield plateau. Therefore, plastic limit analysis is not really justified even though it is embodied in the codes. The main practical consequences of the lack of yield plateau are the lack of moment redistribution in statically indeterminate structures, and a deterministic (non-statistical) size effect (Appendix F). The latter is currently still ignored by the codes, while the former is recognized and handled empirically by limiting the degree to which plastic moment redistribution may be considered in design.

The bending moment capacity of prestressed beams is analyzed in a manner similar to unprestressed beams. Aside from the difference in strain between steel and adjacent concrete, caused by prestressing, the main difference stems from the fact that the prestressed reinforcement is continuously hardening and does not reach maximum stress before the concrete at the compression face gets crushed. This complicates the analysis slightly, but a simplified semiempirical approach is available.

14.2.3 Size Effect

Recently it became clear that the (non-statistical) size effect was implicated in some major catastrophes (e.g. Sleipner platform, bridge columns in earthquakes in Kobe and San Francisco, Malpasset Dam, or San Francis Dam). Nevertheless, in view of the previous comments about the inapplicability of plastic limit analysis and lack of attention to size effect, one might wonder: How is it that there have not been many more failures of large columns with small load eccentricities or large prestressed beams? Two explanations can be offered.

One explanation is that the true safety margins in design are much larger than those implied by the load factors and understrength factors in the code. For example, extra safety is gained by basing the design on concrete strength f'_c much lower than the mean strength, and further by reducing the concrete strength for the design of centrically loaded columns by factor 0.85.

Another explanation is that a large size effect, is hidden in prescribing an excessive dead load factor (Bažant and Frangopol, 2000). ACI code prescribes this factor as 1.4, while the own weight of a large-span or tall structure can practically never be underestimated by more than 5%. This penalizes heavy structures compared to light ones, and the heavy ones are generally those of large spans or heights. However, such a hidden size effect factor, which is about 1.3 for very large structures, is not a rational approach because it does not discriminate between different types of failure. For some types, such as the bending failure of an under-reinforced unprestressed beam, there is no size effect. For some others (e.g. some shear and torsional failures), a factor much larger than 1.3 would be appropriate. There is great need for further practical research combining fracture mechanics with reliability analysis.

PROBLEMS

Problem 14.1: Derive equation (14.10) for the balanced tensile steel ratio in the presence of compression steel.

Problem 14.2: Using the simplified CEB stress-strain law (14.17) for concrete under uniaxial compression and neglecting the stresses in concrete under tension, construct the moment-curvature diagram for a singly reinforced rectangular cross section of width $b = 200$ mm and effective depth $d = 400$ mm. Consider the material properties of concrete C-30 ($f_{ck} = 30$ MPa) and steel of grade 380 ($f_{yk} = 380$ MPa, $E_s = 210$ GPa) in a persistent design situation. Solve the problem for two different reinforcement ratios, 1% and 5%, and comment on the differences.

Problem 14.3: Consider a column of square cross section with side $b = 400$ mm, reinforced by 4 axial steel bars placed in corners, with 25 mm cover, having combined cross section area 4800 mm^2. Assuming the rectangular stress block, calculate e and P for the balanced condition.

Part III

Plastic Analysis of Structures Under Multiaxial Stress

PART III.
PLASTICITY ANALYSIS OF STRUCTURES
UNDER MULTIAXIAL STRESS

15

Simple Elastoplastic Constitutive Models

Before studying the plastic behavior of materials subjected to multiaxial stress, a reader who might not be acquainted with the basic theory of Cartesian tensors and their invariants is advised to consult the brief review in Appendix D. In the derivations to come, we preferentially use the compact tensor notation to take advantage of its conciseness. A reader not comfortable with this approach can rewrite all the formulae in the indicial notation, as explained in Appendix D.

Many plasticity models rely on the volumetric-deviatoric split of stress and strain. Recall that the stress tensor $\boldsymbol{\sigma}$ can be decomposed into

$$\boldsymbol{\sigma} = \sigma_V \boldsymbol{\delta} + \boldsymbol{s} \tag{15.1}$$

where σ_V is the mean (volumetric) stress, $\boldsymbol{\delta}$ is the unit second-order tensor (Kronecker delta), and $\boldsymbol{s}$ is the stress deviator. Similarly, the strain tensor $\boldsymbol{\varepsilon}$ can be decomposed into

$$\boldsymbol{\varepsilon} = \varepsilon_V \boldsymbol{\delta} + \boldsymbol{e} \tag{15.2}$$

where ε_V is the mean (volumetric) strain, and $\boldsymbol{e}$ is the strain deviator. One advantage is that the constitutive equations for an isotropic linearly elastic material,

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} \tag{15.3}$$

can be written separately for the volumetric and deviatoric components:

$$\sigma_V = 3K\varepsilon_V \tag{15.4}$$

$$\boldsymbol{s} = 2G\boldsymbol{e} \tag{15.5}$$

where K is the bulk modulus and G is the shear modulus of elasticity.

15.1 YIELD CONDITIONS FOR ISOTROPIC MATERIALS

The *yield condition*, in general written as $f(\boldsymbol{\sigma}) = 0$, defines the stress states for which the material exhibits plastic flow. The set of stress states that satisfy the yield condition forms the so-called *yield surface* in the stress space. The sign of the expression defining the yield function can always be selected such that $f(\boldsymbol{\sigma}) < 0$ corresponds to *elastic* stress states, $f(\boldsymbol{\sigma}) = 0$ corresponds to *plastic* stress states, and

stress states for which $f(\boldsymbol{\sigma}) > 0$ cannot be supported by the material. All the stress states for which $f(\boldsymbol{\sigma}) \leq 0$ are called *plastically admissible.*

The decision as to whether the material is yielding or not must be independent of the particular coordinate system in which we perform our calculations. Since, for isotropic materials, we can have no orientation effect, a yield condition must depend only on the invariants of the stress tensor. A convenient set of invariants consists of the trace of the stress tensor,

$$I_1 = \boldsymbol{\sigma} : \boldsymbol{\delta} = \sigma_{kk} = \sigma_1 + \sigma_2 + \sigma_3 = 3\sigma_{\text{oct}} = 3\sigma_V \tag{15.6}$$

and the deviator invariants,

$$J_2 = \tfrac{1}{2}\boldsymbol{s} : \boldsymbol{s} = \tfrac{1}{2}s_{ij}s_{ij} = \tfrac{1}{6}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] = \tfrac{3}{2}\tau_{\text{oct}}^2 \tag{15.7}$$

and

$$J_3 = \det \boldsymbol{s} = \det(s_{ij}) = s_1 s_2 s_3 = (\sigma_1 - \sigma_V)(\sigma_2 - \sigma_V)(\sigma_3 - \sigma_V) \tag{15.8}$$

where σ_i $(i = 1, 2, 3)$ are principal stresses, s_i $(i = 1, 2, 3)$ are principal deviatoric stresses, and σ_{oct} and τ_{oct} are the octahedral stresses defined in Appendix D.

In terms of these invariants, the yield condition can be written as

$$f(I_1, J_2, J_3) = 0 \tag{15.9}$$

Alternatively, we may express the yield condition in terms of other combinations of invariants, e.g. the principal stresses σ_1, σ_2 and σ_3, or the Haigh–Westergaard coordinates ξ, ρ and θ; see Appendix D.

15.1.1 Von Mises Criterion

In ductile materials such as metals, inelastic deformation typically takes place by plastic slip along crystallographic planes. The onset of yielding in this case does not depend on the volumetric part of the stress tensor (i.e. on the pressure, or the mean stress). Such materials are described by yield conditions that are independent of the invariant I_1. The simplest and most useful example is the yield condition of von Mises (1913),

$$f(J_2) \equiv \sqrt{J_2} - \tau_0 = 0 \tag{15.10}$$

in which τ_0 is a material parameter. Note that the second deviatoric invariant J_2 is always nonnegative, and so its square root is a real number. The theory based on von Mises yield condition is often referred to as *J_2-plasticity.*

In the case of pure shear stress ($\sigma_{12} = \sigma_{21} \neq 0$ while all other σ_{ij} are zero), we have $J_2 = \sigma_{12}^2$, and the yield condition reduces to $|\sigma_{12}| - \tau_0 = 0$, which indicates that the constant τ_0 is the magnitude of the shear stress at yielding in pure shear. We may thus define the *stress intensity* $\bar{\tau} = \sqrt{J_2}$, which coincides with the shear stress in the case of pure shear, and rewrite the von Mises condition as $\bar{\tau} = \tau_0$. According to (D.47) from Appendix D, the invariant J_2 is proportional to the distortional strain energy, i.e. to the energy due to the changes of shape. Consequently, the von Mises criterion is equivalent to the condition that the distortional strain energy at the onset of yielding reaches a critical value.

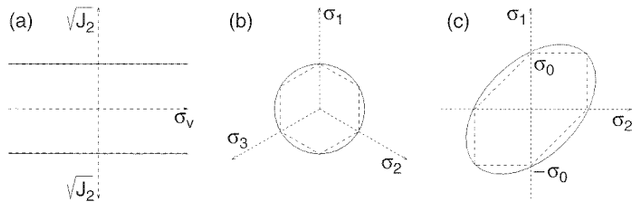


Figure 15.1 Von Mises (solid) and Tresca (dashed) yield surfaces: a) tensile and compressive meridians, b) deviatoric sections (or π-plane projections), c) plane stress representation

It is customary to present the yield conditions graphically in the principal stress space (Westergaard space) with axes σ_1, σ_2 and σ_3. The line on which all the three principal stresses are equal is the hydrostatic axis, and planes perpendicular to this line are the deviatoric planes (also called the π-planes). The intersections of the yield surface with the deviatoric planes are deviatoric sections (or π-plane projections), and intersections with the planes that contain the hydrostatic axis are called meridians.

The geometrical interpretation of the von Mises yield surface in terms of Westergaard coordinates (see Appendix D.4) is straightforward. Owing to (D.70), the von Mises yield condition means that the material yields when the distance of the corresponding stress point from the hydrostatic axis in the principal stress space reaches a certain limit value, $\rho_0 = \tau_0\sqrt{2}$. In the deviatoric plane, the points satisfying the yield condition fill a circle; see Figure 15.1(b). Since the hydrostatic component of stress has no effect on the J_2 invariant, the yield surface corresponding to the von Mises criterion is a cylinder, rotationally symmetric with respect to the hydrostatic axis. Its meridians are straight lines parallel to the hydrostatic axis (Figure 15.1(a)). The meridian plotted above the hydrostatic axis corresponds to Lode angle $\theta = 0$. It is called the *tensile meridian*, because it also contains the point corresponding to yielding under uniaxial tension. The meridian plotted below the hydrostatic axis corresponds to Lode angle $\theta = 60°$, and is called the *compressive meridian.*

Under plane stress ($\sigma_3 = 0$), expression (15.7) for the second deviatoric invariant simplifies to $J_2 = (\sigma_1^2 - \sigma_1\sigma_2 + \sigma_2^2)/3$. Therefore, the intersection of the Mises cylinder with the coordinate plane (σ_1, σ_2) is an ellipse (Figure 15.1(c)) described by the equation $\sigma_1^2 - \sigma_1\sigma_2 + \sigma_2^2 = 3\tau_0^2$. For uniaxial stress ($\sigma_2 = \sigma_3 = 0$), we obtain $J_2 = \sigma_1^2/3$, and so the yield condition reads $\sigma_1^2 = 3\tau_0^2$. This means that the uniaxial yield stress σ_0 and the pure shear yield stress τ_0 are related as

$$\sigma_0 = \tau_0\sqrt{3} \tag{15.11}$$

and the von Mises yield condition for plane stress can be written as

$$\sigma_1^2 - \sigma_1\sigma_2 + \sigma_2^2 = \sigma_0^2 \tag{15.12}$$

15.1.2 Tresca Criterion

Another simple and useful yield criterion, historically the oldest one, is due to Tresca (1868). He proposed that the material starts to yield when the maximum shear stress reaches a certain limit value τ_0. Since, among planes of all orientations, there are six

local maxima of the shear stress magnitude, the Tresca yield condition requires that one of the following six equalities be satisfied:

$$\frac{\sigma_1-\sigma_2}{2}=\pm\tau_0 \quad \text{or} \quad \frac{\sigma_2-\sigma_3}{2}=\pm\tau_0 \quad \text{or} \quad \frac{\sigma_3-\sigma_1}{2}=\pm\tau_0 \tag{15.13}$$

Thus, the yield condition of Tresca may be written as

$$f(\sigma_1,\sigma_2,\sigma_3)\equiv\left[(\sigma_1-\sigma_2)^2-4\tau_0^2\right]\left[(\sigma_2-\sigma_3)^2-4\tau_0^2\right]\left[(\sigma_3-\sigma_1)^2-4\tau_0^2\right]=0 \tag{15.14}$$

Of course, if the principal stresses are ordered such that $\sigma_1\geq\sigma_2\geq\sigma_3$, then the Tresca condition can be written simply as $\sigma_1-\sigma_3=2\tau_0$, because the largest shear stress is in this case $(\sigma_1-\sigma_3)/2$. In terms of invariants I_1, J_2 and J_3, the Tresca condition can be expressed as (Reuß, 1933)

$$f(I_1,J_2,J_3)\equiv 4J_2^3-27J_3^2-36\tau_0^2J_2^2+96\tau_0^4J_2-64\tau_0^6=0 \tag{15.15}$$

Expressing the shear yield stress in terms of the uniaxial yield stress, we obtain the relation

$$\sigma_0=2\tau_0 \tag{15.16}$$

which is different from relation (15.11) derived for the von Mises condition. This offers the possibility of easily checking the yield condition experimentally. Experiments on most metals lie between the two values, usually closer to the condition of von Mises. An intermediate criterion formulated in terms of J_2 and J_3 and providing a transition between the criteria of Tresca and von Mises has been proposed by Edelman and Drucker; see Lemaitre and Chaboche (1990, p. 182).

Under plane stress, it may readily be checked that the six equations in (15.13) represent straight lines that pass through the points $\pm\sigma_0$ on the stress axes in Figure 15.1(c). Therefore, the Tresca condition is represented by a hexagon that is inscribed in the von Mises ellipse (under the assumption that the uniaxial yield stress σ_0 is the same for both yield conditions). In the deviatoric plane, the Tresca yield surface appears as a regular hexagon inscribed into the von Mises circle (Figure 15.1(b)). The six sides of the hexagon correspond to the six planes of maximum shear stress. In the principal stress space, the Tresca condition is represented by an infinite regular hexahedral prism inscribed into the von Mises cylinder. The meridians are straight lines parallel to the hydrostatic axis. In contrast to the von Mises criterion, the distance of the meridians from the axis varies between $\sigma_0/\sqrt{2}$ and $\sigma_0\sqrt{2/3}$, depending on the Lode angle. This is not apparent from Figure 15.1(a), which shows only the tensile and compressive meridians. One can also define the *shear meridian* corresponding to $\theta=30°$, which, for the Tresca criterion, is the closest to the hydrostatic axis.

15.1.3 Gurson Criterion

Yielding of metals and similar ductile materials is usually considered as pressure-independent, and is described by the von Mises or Tresca criteria. However, large positive hydrostatic stresses can lead to the nucleation and propagation of voids, which is on the macroscopic scale perceived as a plastic increase of volume. The most popular model taking into account the effect of porosity on the yield surface is due

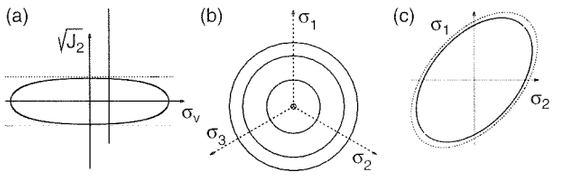


Figure 15.2 Gurson yield surface (solid curves) compared to von Mises surface (dotted curves): a) meridians, b) deviatoric sections, c) plane stress representation

to Gurson (1977). Based on a micromechanical analysis, he originally proposed the yield function

$$f(I_1,J_2)\equiv\frac{3J_2}{\sigma_0^2}+2p\cosh\left(\frac{I_1}{2\sigma_0}\right)-1-p^2=0 \tag{15.17}$$

where σ_0 is the uniaxial yield stress of the basic material (without pores) and p is the volume fraction of voids (porosity). For the case of no porosity, $p=0$, the Gurson criterion is equivalent to the von Mises condition (15.10), with $\tau_0\equiv\sigma_0/\sqrt{3}$. The Gurson yield surface for porosity $p=0.06$ is plotted in Figure 15.2.

Based on micromechanical studies of materials with periodic distributions of voids, Tvergaard (1981) took into account void interaction and proposed the porosity p in (15.17) to be multiplied by a correction factor, $q=1.5$. To improve the prediction of final failure due to void coalescence, Tvergaard and Needleman (1984) replaced the porosity p by

$$p^*=\begin{cases}p & \text{if } p\leq p_C\\ p_C+\frac{p_U^*-p_C}{p_F-p_C}(p-p_C) & \text{if } p\geq p_C\end{cases} \tag{15.18}$$

where p_C is the porosity at which the effect of coalescence begins (e.g. $p_C=0.15$), p_F is the porosity at which the stress-carrying capacity completely vanishes (e.g. $p_F=0.25$), and $p_U^*=1/q=2/3$. The resulting yield condition

$$f(I_1,J_2)\equiv\frac{3J_2}{\sigma_0^2}+3p^*\cosh\left(\frac{I_1}{2\sigma_0}\right)-1-(1.5p^*)^2=0 \tag{15.19}$$

with p^* given by (15.18), is a part of the *modified Gurson's model*.

Alternative yield conditions for porous ductile solids have been developed for the purpose of modeling powder compacted metals, e.g. by Shima and Oyane (1976).

15.1.4 Burzyński Criterion and Drucker–Prager Criterion

In materials with internal friction, such as sands, soils, rocks and concrete, the slip surfaces are rough and the shear stress needed to activate slip is affected by the stress normal to the slip plane. For such materials, often called *pressure-sensitive*, it is not possible to neglect the effect of the first invariant I_1 on the yield condition. The influence of the first invariant was already considered by Schleicher (1926), but the first effective yield criterion written in terms of I_1 and J_2 was proposed by Burzyński (1929) in the form

$$c_1I_1+c_2I_1^2+c_3J_2-1=0 \tag{15.20}$$

where c_1, c_2 and c_3 are material parameters. The corresponding yield surface in the principal stress space is rotationally symmetric about the hydrostatic axis, and usually

has the shape of an ellipsoid or a hyperboloid. As degenerated cases, one may obtain a paraboloid or a cone. The conical yield surface was subsequently used by several Polish and Russian researchers. In Western literature, it was made popular by Drucker and Prager (1952). The Drucker–Prager criterion is usually stated as

$$f(I_1, J_2) \equiv \alpha I_1 + \sqrt{J_2} - \tau_0 = 0 \tag{15.21}$$

in which α and τ_0 are material parameters. This condition may also be written as $\bar{\tau} = \tau'$, in which $\bar{\tau} = \sqrt{J_2}$ is the previously defined stress intensity, and $\tau' = \tau_0 - 3\alpha\sigma_V$ is a pressure-dependent shear yield stress. This indicates that the Drucker–Prager yield condition is the same as the von Mises condition, except that the value of the shear yield stress is adjusted according to the hydrostatic (volumetric) stress, $\sigma_V = I_1/3 = \sigma_{kk}/3$.

The effect of the hydrostatic stress on the shear yield stress is obviously a frictional effect. At a constant value of hydrostatic stress σ_V, the yield surface is the same as the von Mises yield surface, i.e. a circle in the deviatoric plane. However, the radius of the circle varies as σ_V changes, i.e. while moving along the hydrostatic axis; see Figure 15.2(b). Therefore, the Drucker–Prager yield surface represents a circular cone in the principal stress space. The generating lines of the cone (meridians) are straight if α, the friction coefficient, is constant; see Figure 15.3(a). In the principal plane σ_1–σ_2, the Drucker–Prager condition is represented by an ellipse, but unlike the von Mises ellipse, this one is not symmetric with respect to the origin (Figure 15.3(c)). The parameters α and τ_0 can be identified from the stress values in two different plastic states; see Problem 15.1.

15.1.5 Mohr–Coulomb Criterion

A pressure-sensitive generalization of the Tresca criterion, often used for soils, is due to Mohr (1900). The basic idea is that the shear stress necessary to induce yielding on a certain plane increases with increasing compressive stress normal to that plane. In the simplest case, the critical value of shear stress is given by a linear equation suggested as early as 1776 by Coulomb,

$$|\tau| = c - \sigma \tan\phi \tag{15.22}$$

where τ and σ are, respectively, the shear and normal stresses, and c and ϕ are material constants. Plastic flow is initiated as soon as there is at least one plane on which (15.22) holds. Parameter c represents the shear resistance of the material at zero normal stress, and is called the *cohesion*. Parameter ϕ is the *angle of internal*

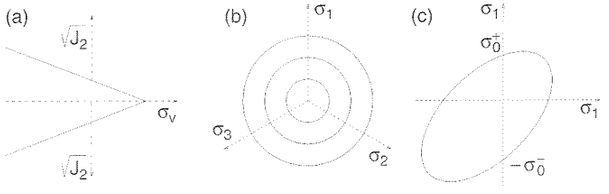


Figure 15.3 Drucker–Prager yield surface: a) meridians, b) deviatoric sections, c) plane stress representation

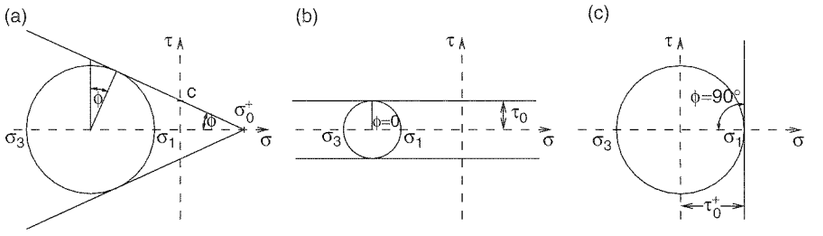


Figure 15.4 Yield surfaces σ–τ plane: a) Mohr–Coulomb criterion, b) Tresca criterion, c) Rankine criterion

friction. Equation (15.22) can equivalently be written as

$$|\tau| = (\sigma_0^+ - \sigma)\tan\phi \tag{15.23}$$

where $\sigma_0^+ = c/\tan\phi$ is the yield stress in tension.

The Mohr–Coulomb criterion can be most easily visualized in the σ–τ plane. Yielding starts when the Mohr circle representing the current stress state touches the Mohr failure envelope that can be a straight line corresponding to (15.22), or a more general curve described by a nonlinear function $\tau(\sigma)$. From Figure 15.4(a) it is clear that, at the onset of yielding,

$$\frac{\sigma_1 - \sigma_3}{2} = \left(c - \frac{\sigma_1 + \sigma_3}{2}\tan\phi\right)\cos\phi \tag{15.24}$$

if (15.21) is adopted with ϕ constant. From this,

$$(1 + \sin\phi)\sigma_1 - (1 - \sin\phi)\sigma_3 = 2c\cos\phi \tag{15.25}$$

Note that the shear yield stress in (15.22) does not depend on the volumetric stress, but on the stress normal to the slip plane. So the resulting criterion (15.25) is insensitive to the intermediate principal stress, σ_2, which is its weakness in some cases.

If the left-hand side of (15.25) is smaller than the right-hand side, the material is in the elastic state. Of course, this condition is sufficient only if the principal stresses are ordered such that $\sigma_1 \geq \sigma_2 \geq \sigma_3$. In general it must be tested for any combination of two principal stresses, and so the Mohr–Coulomb criterion of plastic admissibility is described by a set of six linear inequalities. The intersection of the corresponding six half-spaces is a hexagonal pyramid in the principal stress space. Its axis is the hydrostatic axis, and its intersection with the deviatoric plane is a hexagon. However, unlike the Tresca criterion, this hexagon is not regular, i.e. its vertices are not located on a circle; see Figure 15.5(b). Consequently, the slopes of the compressive and tensile meridians are different; see Figure 15.5(a). The intersection of the Mohr–Coulomb yield surface with the principal plane $\tau_1 - \tau_2$ is a hexagon with only one axis of symmetry (Figure 15.5(c)); the yield stress in uniaxial tension and in uniaxial compression is different, same as for the Drucker–Prager yield surface.

Many materials can be described better by a curved Mohr envelope. The formulation is not complicated if the envelope is a parabola or hyperbola.

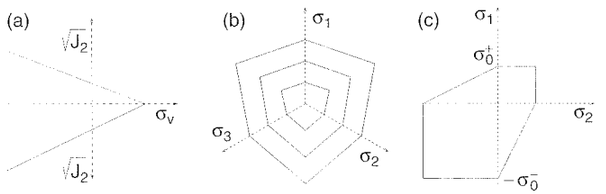


Figure 15.5 Mohr–Coulomb yield surface: a) meridians, b) deviatoric sections, c) plane stress representation

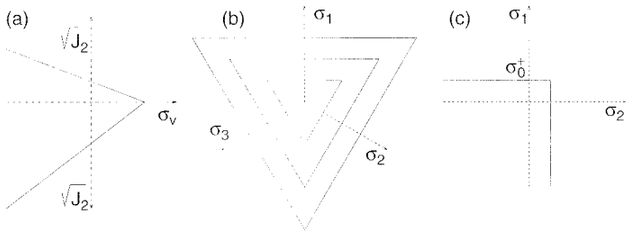


Figure 15.6 Rankine yield surface: a) meridians, b) deviatoric sections, c) plane stress representation

15.1.6 Rankine Criterion

The Mohr–Coulomb criterion is a generalization of the Tresca criterion, to which it reduces for $\phi = 0$; see Figure 15.4(b). The other extreme, $\phi = 90°$, leads to the Rankine (1858) criterion, which provides the simplest (albeit often oversimplified) basis for plasticity modeling of distributed cracking of concrete. For the sake of generality, we perform the limit transition $\phi \to 90°$ at constant yield stress σ_0^+ rather than at constant cohesion c; see Figure 15.4(c). In the limit, equation (15.25) rewritten as

$$(1 + \sin\phi)\sigma_1 - (1 - \sin\phi)\sigma_3 = 2\sigma_0^+ \sin\phi \tag{15.26}$$

gives

$$\sigma_1 = \sigma_0^+ \tag{15.27}$$

According to the Rankine criterion, yielding starts when the maximum principal stress reaches the tensile yield stress, σ_0^+. The hexagonal Mohr–Coulomb pyramid degenerates into a triangular Rankine pyramid; see Figure 15.6(b).

15.2 FLOW THEORY OF PLASTICITY

15.2.1 Flow Rule

Similar to one-dimensional plasticity, we will decompose the total strain tensor

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_p \tag{15.28}$$

into a sum of two symmetric tensors — the elastic (reversible) strain $\boldsymbol{\varepsilon}_e$ and the plastic (irreversible) strain $\boldsymbol{\varepsilon}_p$. We assume the material to exhibit plastic flow, but suffer no damage. In that case, the elastic stiffness of the material remains unchanged (in other words, the unloading follows the initial slope). This means that the relationship between the elastic strain and the stress,

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon}_e \tag{15.29}$$

is the same as in linear elasticity; see (15.3).

As long as the stress remains inside the elastic domain, the deformation process is purely elastic and the plastic strain does not change. On the other hand, when the stress state reaches the yield surface, plastic flow is initiated. During plastic flow, the stress state remains on the yield surface, and so the yield condition $f(\boldsymbol{\sigma}) = 0$ must be satisfied. The yield condition only provides one additional equation for the evaluation of the plastic strain, which has six independent components. Therefore, an additional rule governing the evolution of plastic flow must be postulated. It is customary to supplement the yield condition by a *flow rule* that defines the direction of plastic flow. Of course, the flow rule should be deduced from experimental observations. From the theoretical point of view, it is convenient to work with a rule that preserves the validity of the postulate of maximum plastic dissipation (von Mises, 1928), which plays a key role in the proofs of the fundamental theorems of limit and shakedown analyses.

Postulate of maximum plastic dissipation
Let $\dot{\boldsymbol{\varepsilon}}_p$ be a given plastic strain rate. Among all plastically admissable stress states $\boldsymbol{\sigma}^*$, the power $\boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p$ is maximized by the actual stress, $\boldsymbol{\sigma}$. This statement can be mathematically written as

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p) \equiv \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p = \max_{f(\boldsymbol{\sigma}^*)\le 0} (\boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p) \tag{15.30}$$

where $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p)$ is the dissipation rate per unit volume.

Recalling that the double contraction of second-order tensors has the same meaning as the scalar product of nine-dimensional vectors formed from the nine tensor components, we may interpret equation (15.30) as the condition that the projection of the admissible stress states onto the direction of the plastic strain rate be maximized by the actual stress state. This condition holds if

1. the elastic domain is *convex*, and
2. the direction of the plastic strain rate is *normal* to the yield surface.

These two requirements will be referred to as *convexity* and *normality*. They characterize the so-called *standard materials*. This expression was first used by Radenković (1961), and later became quite common, especially in the French literature. An extension of the postulate of maximum dissipation beyond perfect elastoplasticity lead to the concept of *generalized standard materials*, to be described in Section 23.5.

Figure 15.7(a) illustrates the violation of the postulate of maximum plastic dissipation due to the lack of normality (the flow direction is not normal to the yield surface). Clearly, the vector $\boldsymbol{\sigma}^*$ in this figure has a larger projection onto the direction of plastic flow than does the actual stress $\boldsymbol{\sigma}$. Similarly, Figure 15.7(b) illustrates the implications of non-convexity of the elastic domain. Convexity and normality are

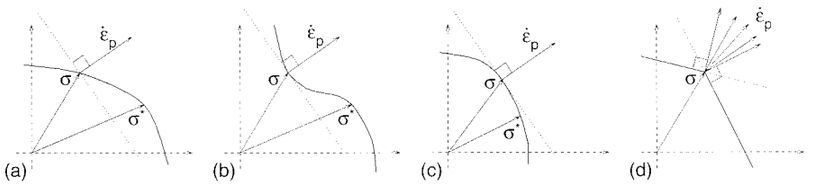


Figure 15.7 Illustration of convexity and normality

simultaneously satisfied in Figure 15.7(c), in which no vector with its end point in the closure of the elastic domain has a larger projection onto the flow direction than does $\boldsymbol{\sigma}$.

As shown in Figure 15.7(d), the direction of plastic flow is not uniquely defined by the normality condition at a sharp corner (vertex) of the yield surface. In this case, the postulate of maximum plastic dissipation is satisfied for any flow direction lying in the fan between the normals to the 'tangent from the left' and 'tangent from the right' to the yield surface. Such a fan is called the *normal cone*.

Since the yield surface is a graphical representation of $f(\boldsymbol{\sigma}) = 0$, the direction normal to the yield surface (at a point where the surface is smooth) is determined by the gradient of f. So, we arrive at a specific form of the flow rule,

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda} \frac{\partial f}{\partial \boldsymbol{\sigma}} \tag{15.31}$$

which is called the *normality rule* or the *associated flow rule*. The symbol λ stands for a scalar multiplier (called the *plastic multiplier*) that controls the magnitude of the plastic strain, and the superior dot marks explicitly that $\dot{\lambda}$ has the meaning of rate (it will later be shown to be proportional to the strain rate). To simplify notation, we will denote the gradient $\partial f/\partial \boldsymbol{\sigma}$ by $\boldsymbol{f}_\sigma$. As usual, the rate form of the flow rule can be replaced by the incremental form,

$$\mathrm{d}\boldsymbol{\varepsilon}_p = \mathrm{d}\lambda\, \boldsymbol{f}_\sigma \tag{15.32}$$

Note that, for a given stress state, the gradient $\boldsymbol{f}_\sigma$ is a fixed tensor specifying the direction of the plastic strain rate. The magnitude of this rate remains unspecified by the flow rule, and it must be determined from the condition that the stress during plastic flow remains on the yield surface.

The idea of a flow rule associated with the yield condition was first advanced by von Mises (1928), who verified its validity for a simple model of plastic slip in crystals. The response of most metals can be described well by an associated flow rule, except when the loading path is highly nonproportional. On the other hand, for pressure-sensitive materials the associated flow rule is often unrealistic and must be replaced by a more general nonassociated flow rule

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda} \frac{\partial g}{\partial \boldsymbol{\sigma}} \tag{15.33}$$

where $g(\boldsymbol{\sigma})$ is a new function called the *plastic potential*. So we can say that the flow rule is associated if the yield function is used at the same time as the plastic potential. A general plastic potential defines a set of equipotential surfaces such that the plastic strain always grows in the direction normal to the surface on which the current stress state is located (Figure 15.8).

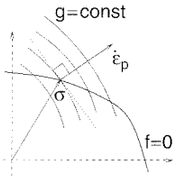


Figure 15.8 Nonassociated flow

A material with a nonassociated flow rule does not satisfy the postulate of maximum plastic dissipation, and so the standard upper and lower bound theorems cease to be applicable. In Part III of this book (Chapters 15–19), we will restrict our attention to materials with an associated flow rule. The nonassociated case will be discussed in Section 20.3.

15.2.2 Fundamental Equations

Having introduced the flow rule, we have a sufficient number of equations for the three-dimensional constitutive description of an elastoplastic material. In the standard displacement-based finite element analysis, the strain evolution is determined by the displacement increments computed on the structural level (see Sections 22.1–22.2). The basic task on the level of a material point is to evaluate the stress evolution generated by a given strain history. Compared to the elastic case, the problem involves additional unknowns – the components of plastic strain and the plastic multiplier. The governing equations consist of the elastic constitutive law

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) \tag{15.34}$$

and the associated flow rule

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda} \boldsymbol{f}_\sigma \tag{15.35}$$

Furthermore, we know that in the elastic regime the yield function must remain negative and the rate of the plastic multiplier is zero (plastic strain remains constant), while in the plastic regime the yield function must be equal to zero (stress remains on the yield surface) and the rate of the plastic multiplier is positive. Both cases can be simultaneously covered by the *loading-unloading conditions*

$$\dot{\lambda} \geq 0, \quad f(\boldsymbol{\sigma}) \leq 0, \quad \dot{\lambda} f(\boldsymbol{\sigma}) = 0 \tag{15.36}$$

Indeed, the last condition means that if $\dot{\lambda} > 0$ (plastic flow) then $f(\boldsymbol{\sigma}) = 0$ (stress on the yield surface), and if $f(\boldsymbol{\sigma}) < 0$ (stress inside the elastic domain) then $\dot{\lambda} = 0$ (no plastic flow). From the mathematical point of view, (15.35)–(15.36) are the so-called *Karush–Kuhn–Tucker conditions* (Karush, 1939; Kuhn and Tucker, 1951); see Section 15.2.4 for more details.

During plastic flow the stress remains on the yield surface, and so the yield function remains equal to zero for a certain period of time. Consequently, the time derivative of the yield function $\dot{f}$ vanishes whenever the rate of the plastic multiplier $\dot{\lambda}$ is nonzero. So we may state the *consistency condition*

$$\dot{\lambda}\, \dot{f} = 0 \tag{15.37}$$

In this form, the condition has a general validity. Sometimes it is stated that the consistency condition reads $\dot{f} = 0$, but then it must be borne in mind that this is true only during plastic yielding and not during an elastic process.

15.2.3 Loading Criteria and Elastoplastic Stiffness

Relations (15.34)–(15.36) represent the mathematical description of associated perfect elastoplasticity. Numerical aspects of this formulation will be discussed in Chapter 22. Here we will only establish the criteria for loading and unloading and derive the incremental stress-strain relations. At the same time, we will show that the problem is well-posed in the sense that the solution is uniquely determined by a given strain tensor history.

1. First, suppose that the current *stress state* $\boldsymbol{\sigma}$ *is inside the elastic domain*, i.e. $f(\boldsymbol{\sigma}) < 0$. Then the third condition from (15.36), $\dot{\lambda} f(\boldsymbol{\sigma}) = 0$, implies that $\dot{\lambda} = 0$, and so the plastic strain remains constant; see the flow rule (15.35). The strain increment is then entirely elastic, and the rate (increment) of stress can be computed from the rate (increment) of strain using the standard linear elastic law (Hooke's law)
$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \tag{15.38}$$
2. Now suppose that the current *stress state* $\boldsymbol{\sigma}$ *is on the yield surface*, i.e. $f(\boldsymbol{\sigma}) = 0$. In this case, the material can either experience further plastic flow, or start unloading elastically.
   (a) During plastic flow the *stress remains on the yield surface*, and so we may use the consistency condition (15.37) in the simpler form
   $$\dot{f} = 0 \tag{15.39}$$
   Applying the chain rule of differentiation we get
   $$\dot{f} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} = 0 \tag{15.40}$$
   and, substituting for $\dot{\boldsymbol{\sigma}}$ the rate form of the elastic law (15.34) and then for $\dot{\boldsymbol{\varepsilon}}_p$ the flow rule (15.35), we can transform this condition into
   $$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\lambda} \boldsymbol{f}_\sigma) = 0 \tag{15.41}$$
   If the strain evolution is prescribed, the foregoing equation contains a single scalar unknown $\dot{\lambda}$. The elastic stiffness tensor $\boldsymbol{D}_e$ is positive definite, and so
   $$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma > 0 \tag{15.42}$$
   for any nonzero tensor $\boldsymbol{f}_\sigma$. Equation (15.41) is therefore uniquely solvable for the rate of the plastic multiplier
   $$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma} \tag{15.43}$$

However, due to the first Kuhn–Tucker condition in (15.36), the solution is admissible only if $\dot{\lambda} \geq 0$. In view of (15.42), this is equivalent to the *loading criterion*
$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \geq 0 \tag{15.44}$$
The case when the loading criterion is satisfied as an equality, $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} = 0$, is just at the limit between loading and unloading (stress remains on the yield surface but the plastic strain rate is zero), and so it is referred to as *neutral loading*.

If the loading criterion is satisfied (as a strict inequality), plastic flow takes place and the rate of plastic strain can be evaluated by substituting the expression derived for $\dot{\lambda}$ into the flow rule:
$$\dot{\boldsymbol{\varepsilon}}_p = \boldsymbol{f}_\sigma \dot{\lambda} = \frac{\boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma} : \dot{\boldsymbol{\varepsilon}} \tag{15.45}$$
Note that we switched the order of $\dot{\lambda}$ and $\boldsymbol{f}_\sigma$ in the flow rule, which is allowed because λ is a scalar. The reason for this is that we want to end up with an expression in which the total strain rate $\dot{\boldsymbol{\varepsilon}}$ is multiplied from the left by a fourth-order tensor. Using (15.45), the plastic strain rate can be eliminated from the rate form of the constitutive law (15.34). We arrive at an expression for the stress rate in terms of the total strain rate,
$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} - \boldsymbol{D}_e : \frac{\boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma} : \dot{\boldsymbol{\varepsilon}} \tag{15.46}$$
which can be rewritten as
$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_{ep} : \dot{\boldsymbol{\varepsilon}} \tag{15.47}$$
where
$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma} \tag{15.48}$$
is the *elastoplastic material stiffness tensor* (or tangential moduli tensor). This tensor is used in place of the standard elastic material stiffness tensor $\boldsymbol{D}_e$ for calculating the tangential stiffness matrix in finite element calculations. Section 22.4 will explain that, in order to preserve the quadratic rate of convergence of the Newton–Raphson iteration process, the tangential stiffness matrix should be computed from the so-called algorithmic stiffness tensor, which is somewhat different from the elastoplastic material stiffness tensor.

(b) Finally, let us check under which condition the *stress returns from the yield surface into the elastic domain*. In this case, the yield function must decrease (from zero to negative values), and so its time derivative $\dot{f}(\boldsymbol{\sigma})$ is negative. The plastic strain rate must be zero, which means that the linear elastic law (15.38) applies to the stress rate and the (total) strain rate. Using again the chain rule and the elastic law, we get for this case
$$\dot{f}(\boldsymbol{\sigma}) = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} = \boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \tag{15.49}$$
This means that the *unloading criterion* reads
$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} < 0 \tag{15.50}$$

Table 15.1 Basic processes in perfect elastoplasticity

Initial state	Elastic	Plastic		
Process		Unloading	Neutral	Loading
$f(\boldsymbol{\sigma})$	negative	zero	zero	zero
$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$	arbitrary	negative	zero	positive
$\dot{f}$	arbitrary	negative	zero	zero
$\dot{\lambda}$	zero	zero	zero	positive
stiffness	$\boldsymbol{D}_e$	$\boldsymbol{D}_e$	$\boldsymbol{D}_e$ or $\boldsymbol{D}_{ep}$	$\boldsymbol{D}_{ep}$

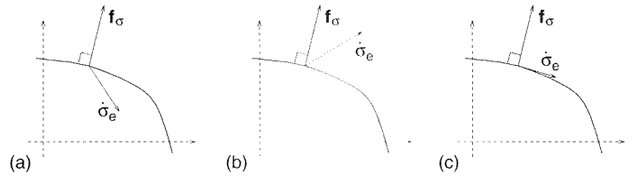


Figure 15.9 Position of the elastic stress rate with respect to the yield surface: a) loading, b) unloading, c) neutral loading

Since the loading criterion (15.44) and the unloading criterion (15.50) are complementary, the problem always has a unique solution. We have tacitly assumed that the yield function is smooth, and so the gradient $\boldsymbol{f}_\sigma$ is uniquely defined. The treatment of singularities (corners) of the yield surface is more complicated, and is postponed to Section 20.4.

The principal results of the previous analysis are summarized in Table 15.1. The third and fourth line represent the criteria for identifying the type of process. In the case of neutral loading, it is not clear whether to use the elastic stiffness or the elastoplastic stiffness. The rate equations $\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$ and $\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_{ep} : \dot{\boldsymbol{\varepsilon}}$ are satisfied with either stiffness (see Problem 15.7), but the corresponding stiffness matrices lead to different displacement corrections in the next iteration. Fortunately this is not a problem, because neutral loading is a degenerated situation that can appear only very exceptionally.

The loading-unloading criterion has a clear geometric interpretation in terms of the so-called elastic stress rate vector, $\dot{\boldsymbol{\sigma}}_e = \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$. This is the stress rate produced by the given strain rate $\dot{\boldsymbol{\varepsilon}}$ if no plastic flow takes place. The product $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$ can be rewritten as $\boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}}_e$ and interpreted as the scalar product of the elastic stress rate vector[1] and the outward normal to the yield surface:

- The unloading criterion $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \equiv \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}}_e < 0$ means that the elastic stress rate vector points to the interior of the elastic domain $f \leq 0$; see Figure 15.9(a). Such a stress evolution is plastically admissible. The process is purely elastic, and the actual stress rate is equal to the elastically computed rate.
- The loading criterion $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \equiv \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}}_e > 0$ means that the elastic stress rate vector points to the exterior of the elastic domain $f \leq 0$; see Figure 15.9(b). Such a stress evolution would not be plastically admissible. Plastic flow must occur, and the actual stress rate differs from the elastically computed one. The difference is equal to the product of the elastic stiffness and the plastic strain rate.
- If $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \equiv \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}}_e = 0$ then the elastic stress rate vector is tangent to the yield surface; see Figure 15.9(c). For infinitesimally small increments (or even for finite increments if the yield surface consists of straight segments) the stress point remains on the yield surface without any correction by plastic flow. However, a small change of the strain rate direction can lead to the onset of plastic flow.

[1] Vector is meant here as an element of the nine-dimensional stress-space

Example 15.1: Derive the specific form of the associated flow rule for J_2-plasticity.

Solution: The general form of an associated flow rule is given by (15.31). We will substitute the von Mises yield function (15.10) for f and evaluate its gradient

$$\boldsymbol{f}_\sigma = \frac{\partial f}{\partial \boldsymbol{\sigma}} = \frac{\partial}{\partial \boldsymbol{\sigma}} \left(\sqrt{J_2} - \tau_0 \right) = \frac{1}{2\sqrt{J_2}} \frac{\partial J_2}{\partial \boldsymbol{s}} : \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{\sigma}} \tag{15.51}$$

According to equations (D.51) and (D.52) from Appendix D we have

$$\frac{\partial J_2}{\partial \boldsymbol{s}} = \frac{\partial}{\partial \boldsymbol{s}} \left(\frac{1}{2} \boldsymbol{s} : \boldsymbol{s} \right) = \boldsymbol{s} \tag{15.52}$$

$$\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{\sigma}} = \frac{\partial}{\partial \boldsymbol{\sigma}} \left(\boldsymbol{\sigma} - \frac{1}{3} \boldsymbol{\sigma} : \boldsymbol{\delta} \otimes \boldsymbol{\delta} \right) = \boldsymbol{I} - \frac{1}{3} \boldsymbol{\delta} \otimes \boldsymbol{\delta} \tag{15.53}$$

where $\boldsymbol{I}$ is the fourth-order unit tensor and $\boldsymbol{\delta}$ is the second-order unit tensor (Kronecker delta). Substituting (15.52) and (15.53) back into (15.51), we get

$$\boldsymbol{f}_\sigma = \frac{1}{2\sqrt{J_2}} \frac{\partial J_2}{\partial \boldsymbol{s}} : \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{\sigma}} = \frac{1}{2\sqrt{J_2}} \boldsymbol{s} : (\boldsymbol{I} - \frac{1}{3} \boldsymbol{\delta} \otimes \boldsymbol{\delta}) = \frac{1}{2\sqrt{J_2}} \boldsymbol{s} \tag{15.54}$$

because $\boldsymbol{s} : \boldsymbol{I} = \boldsymbol{s}$ and $\boldsymbol{s} : \boldsymbol{\delta}$ = first invariant of the stress deviator = 0; see (D.19) in Appendix D.

During plastic flow, the yield condition must be satisfied, and so $J_2 = \tau_0^2$. So, the flow rule associated with the von Mises yield condition is

$$\dot{\boldsymbol{\varepsilon}}_p = \frac{\dot{\lambda}}{2\tau_0} \boldsymbol{s} \tag{15.55}$$

Introducing a rescaled plastic multiplier $\bar{\lambda} = \lambda / 2\tau_0$, we arrive at the famous *Prandtl-Reuß relations* (Prandtl, 1924; Reuß, 1930),

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\bar{\lambda}} \boldsymbol{s} \tag{15.56}$$

□

Since $\boldsymbol{s}$ is a purely deviatoric tensor ($\mathrm{tr}(\boldsymbol{s}) = \boldsymbol{\delta} : \boldsymbol{s} = 0$), the plastic flow described by the Prandtl-Reuß relations is purely deviatoric, i.e. the volumetric part of the plastic strain vanishes. This property, called the *plastic incompressibility*, is characteristic of metals at the usual levels of the mean stress. In contrast to that, materials with internal friction (such as concrete, rocks and soils), as well as porous materials, change their volume during plastic flow.

Example 15.2: Derive the specific form of the loading criterion for J_2-plasticity.

Solution: The loading criterion checks the sign of the product $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$. Evaluating $\boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$, we can take advantage of the volumetric-deviatoric split and write

$$\dot{\boldsymbol{\varepsilon}} = \dot{\varepsilon}_V \boldsymbol{\delta} + \dot{\boldsymbol{e}} \tag{15.57}$$

$$\boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} = 3K\dot{\varepsilon}_V \boldsymbol{\delta} + 2G\dot{\boldsymbol{e}} \tag{15.58}$$

The gradient of the von Mises yield function $\boldsymbol{f}_\sigma = \boldsymbol{s}/2\tau_0$ has been computed in the previous example, and so it is easy to evaluate

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} = \frac{3K}{2\tau_0}\dot{\varepsilon}_V \boldsymbol{s} : \boldsymbol{\delta} + \frac{2G}{2\tau_0}\boldsymbol{s} : \dot{\boldsymbol{e}} = \frac{G}{\tau_0}\boldsymbol{s} : \dot{\boldsymbol{e}} \tag{15.59}$$

The factor G/τ_0 is a positive constant and therefore may be omitted when checking the sign of the expression. Consequently, the von Mises material exhibits plastic flow if the yield condition is satisfied ($J_2 = \tau_0^2$) and

$$\boldsymbol{s} : \dot{\boldsymbol{e}} > 0 \tag{15.60}$$

This means that the deviatoric component of the strain increment must have a positive projection onto the direction of the current deviatoric stress. □

Example 15.3: Derive the elastoplastic material stiffness tensor for associated J_2-plasticity.

Solution: We have to substitute $\boldsymbol{f}_\sigma = \boldsymbol{s}/2\tau_0$ into the general expression (15.48). As both the numerator and the denominator of the second term (plastic correction) contain the symbol $\boldsymbol{f}_\sigma$ twice in a product-like expression, we can omit the scalar factor $1/2\tau_0$ and write

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{s} \otimes \boldsymbol{s} : \boldsymbol{D}_e}{\boldsymbol{s} : \boldsymbol{D}_e : \boldsymbol{s}} \tag{15.61}$$

Realizing that $\boldsymbol{s}$ is a deviatoric tensor (its trace is zero), we can again exploit the volumetric-deviatoric split of the elastic stiffness tensor and write

$$\boldsymbol{D}_e : \boldsymbol{s} = 2G\boldsymbol{s} \tag{15.62}$$

A formal derivation of (15.62) can be based on formula (D.25) for the elastic stiffness tensor. Due to the major symmetry of the elastoplastic stiffness tensor, we have $D^e_{ijkl}s_{kl} = s_{kl}D^e_{klij}$, i.e. $\boldsymbol{D}_e : \boldsymbol{s} = \boldsymbol{s} : \boldsymbol{D}_e$. Substituting this into (15.61), and realizing that $\boldsymbol{s} : \boldsymbol{s} = 2J_2 = 2\tau_0^2$, we get

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{(2G\boldsymbol{s}) \otimes (2G\boldsymbol{s})}{(2G\boldsymbol{s}) : \boldsymbol{s}} = \boldsymbol{D}_e - \frac{4G^2\boldsymbol{s} \otimes \boldsymbol{s}}{4G\tau_0^2} = \boldsymbol{D}_e - \frac{G}{\tau_0^2}\boldsymbol{s} \otimes \boldsymbol{s} \tag{15.63}$$

This means that the plastic correction of the stiffness tensor is a material constant G/τ_0^2 times the direct product of the stress deviator with itself. To check this result, consider two elementary situations – uniaxial tension and pure shear.

1. Under uniaxial tension in direction x_1 we have

$$\boldsymbol{\sigma} = \sigma \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{s} = \frac{\sigma}{3} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{15.64}$$

and, at the onset of yielding, $\sigma = \sigma_0 = \tau_0\sqrt{3}$. The stiffness coefficients associated with the main stress σ_{11} are evaluated as follows:

$$D^{ep}_{1111} = D^e_{1111} - \frac{G}{\tau_0^2}s_{11}^2 = K + \frac{4}{3}G - \frac{G}{\tau_0^2}\frac{4\tau_0^2}{3} = K \tag{15.65}$$

$$D^{ep}_{1122} = D^{ep}_{1133} = D^e_{1122} - \frac{G}{\tau_0^2}s_{11}s_{22}$$

$$= K - \frac{2}{3}G + \frac{G}{\tau_0^2}\frac{2\tau_0^2}{3} = K \tag{15.66}$$

The elastic stiffness coefficients were taken from Table D.1. Clearly, the coefficients D^{ep}_{1112}, D^{ep}_{1123} and D^{ep}_{1131} vanish, the same as in the elastic state. So the rate equation for the stress in the main direction reads

$$\dot{\sigma}_{11} = K(\dot{\varepsilon}_{11} + \dot{\varepsilon}_{22} + \dot{\varepsilon}_{33}) = 3K\dot{\varepsilon}_V \tag{15.67}$$

It is interesting to note that the stiffness terms associated with the main direction did not completely disappear, but the increments of σ_{11} will be zero anyway because, in the present case, plastic flow takes place at constant volume ($\dot{\varepsilon}_V = 0$); see also Problem 15.3.

2. Under pure shear in the (x_1, x_2) plane we have

$$\boldsymbol{\sigma} = \boldsymbol{s} = \tau \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{15.68}$$

and, at the onset of yielding, $\tau = \tau_0$. The only affected stiffness coefficients are

$$D^{ep}_{1212} = D^{ep}_{1221} = D^{ep}_{2112} = D^{ep}_{2121} = G - \frac{G}{\tau_0^2}\tau_0^2 = 0 \tag{15.69}$$

This means that, when the plastic flow begins, the corresponding shear stiffness vanishes.

15.2.4 Karush–Kuhn–Tucker Conditions

For materials obeying the postulate of maximum plastic dissipation, the problem of finding the stress corresponding to a given plastic strain rate is, from the mathematical point of view, a problem of *convex programming*. This branch of optimization theory grew out from the seminal paper by Kuhn and Tucker (1951), which was preceded by an unnoticed Master's thesis of Karush (1939). In the 1980s it became fashionable in mathematical plasticity to use for conditions such as (15.36) the Kuhn–Tucker label.

However, since this disconnects the terminology from the classical plasticity literature, where the criterion appears to have an older history, it is preferred in this book to use the classical term 'loading-unloading conditions' even when written in Kuhn–Tucker form, which might more properly be called the Karush–Kuhn–Tucker form.

The general problem of convex programming reads

$$\text{Minimize } g(\boldsymbol{x}) \tag{15.70}$$

subject to

$$\boldsymbol{f}(\boldsymbol{x}) \leq \boldsymbol{0}, \qquad \boldsymbol{x} \geq \boldsymbol{0} \tag{15.71}$$

where g and $\boldsymbol{f}$ are convex functions of $\boldsymbol{x}$. Any vector $\boldsymbol{x}$ satisfying (15.71) is a *feasible solution*, and if $\bar{\boldsymbol{x}}$ minimizes $g(\boldsymbol{x})$ among all feasible solutions, it is called the *optimal solution*. The optimal solution can be characterized in terms of the Lagrangian function

$$L(\boldsymbol{x}, \boldsymbol{u}) \equiv g(\boldsymbol{x}) + \boldsymbol{u}^T \boldsymbol{f}(\boldsymbol{x}) \tag{15.72}$$

where $\boldsymbol{u}$ is an auxiliary vector of Lagrange multipliers.

Kuhn–Tucker Theorem
A vector $\bar{\boldsymbol{x}} \geq \boldsymbol{0}$ is the optimal solution of the convex programming problem (15.70)–(15.71) if and only if there is a vector $\bar{\boldsymbol{u}} \geq \boldsymbol{0}$ such that

$$L(\bar{\boldsymbol{x}}, \boldsymbol{u}) \leq L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) \leq L(\boldsymbol{x}, \bar{\boldsymbol{u}}) \qquad \forall \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{u} \geq \boldsymbol{0} \tag{15.73}$$

If g and $\boldsymbol{f}$ are continuously differentiable functions, the optimality conditions can be written in the form

$$\frac{\partial L}{\partial \boldsymbol{x}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) \geq \boldsymbol{0} \qquad \frac{\partial L}{\partial \boldsymbol{u}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) \leq \boldsymbol{0} \tag{15.74}$$

$$\bar{\boldsymbol{x}}^T \frac{\partial L}{\partial \boldsymbol{x}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) = 0 \qquad \bar{\boldsymbol{u}}^T \frac{\partial L}{\partial \boldsymbol{u}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) = 0 \tag{15.75}$$

$$\bar{\boldsymbol{x}} \geq \boldsymbol{0} \qquad \bar{\boldsymbol{u}} \geq \boldsymbol{0} \tag{15.76}$$

For problems in which the variable $\boldsymbol{x}$ is not restricted to nonnegative values, it is sufficient to set $\boldsymbol{x} = \boldsymbol{x}^+ - \boldsymbol{x}^-$ with $\boldsymbol{x}^+ \geq \boldsymbol{0}$ and $\boldsymbol{x}^- \geq \boldsymbol{0}$. Since $\partial L/\partial \boldsymbol{x}^+ = -\partial L/\partial \boldsymbol{x}^-$, the Karush–Kuhn–Tucker conditions (15.74)–(15.76) in this case reduce to

$$\frac{\partial L}{\partial \boldsymbol{x}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) = \boldsymbol{0} \qquad \frac{\partial L}{\partial \boldsymbol{u}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) \leq \boldsymbol{0} \tag{15.77}$$

$$\bar{\boldsymbol{u}}^T \frac{\partial L}{\partial \boldsymbol{u}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) = 0 \tag{15.78}$$

$$\bar{\boldsymbol{u}} \geq \boldsymbol{0} \tag{15.79}$$

After this brief excursion into the mathematical optimization theory, we can return to plasticity. According to the postulate of maximum plastic dissipation, the stress $\boldsymbol{\sigma}$ corresponding to a given plastic strain rate $\dot{\boldsymbol{\varepsilon}}_p$ maximizes the linear function $\boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p$ among all the stress states satisfying the constraint $f(\boldsymbol{\sigma}^*) \leq 0$. This means that $\boldsymbol{\sigma}$ is the optimal solution of the convex programming problem

$$\text{Minimize } g(\boldsymbol{\sigma}^*) \equiv -\boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p \tag{15.80}$$

subject to

$$f(\boldsymbol{\sigma}^*) \leq 0 \tag{15.81}$$

The objective function $g(\boldsymbol{\sigma}^*)$ is linear in terms of $\boldsymbol{\sigma}^*$ and is therefore convex. The yield function $f(\boldsymbol{\sigma}^*)$ is also convex (since convexity is implied by the postulate of maximum plastic dissipation), and so (15.80)–(15.81) is indeed a problem of convex programming (without the nonnegativity constraints on $\boldsymbol{\sigma}^*$). Constructing the Lagrangian function

$$L(\boldsymbol{\sigma}^*, \dot{\lambda}^*) \equiv -\boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p + \dot{\lambda}^* f(\boldsymbol{\sigma}^*) \tag{15.82}$$

and evaluating the derivatives $\partial L/\partial \boldsymbol{\sigma}^* = -\dot{\boldsymbol{\varepsilon}}_p + \dot{\lambda}^* \partial \boldsymbol{f}/\partial \boldsymbol{\sigma}^*$ and $\partial L/\partial \dot{\lambda}^* = f$, we can rewrite the Karush–Kuhn–Tucker optimality conditions (15.77)–(15.79) as

$$-\dot{\boldsymbol{\varepsilon}}_p + \dot{\lambda} \boldsymbol{f}_{,\sigma}(\boldsymbol{\sigma}) = 0 \qquad f(\boldsymbol{\sigma}) \leq 0 \tag{15.83}$$

$$\dot{\lambda} f(\boldsymbol{\sigma}) = 0 \tag{15.84}$$

$$\dot{\lambda} \geq 0 \tag{15.85}$$

This shows that the associated flow rule (15.35) and the loading-unloading conditions (15.36) are the Karush–Kuhn–Tucker conditions of the convex programming problem (15.80)–(15.81) resulting from the postulate of maximum plastic dissipation.

PROBLEMS

Problem 15.1: Derive the expressions for parameters α and τ_0 of the Drucker–Prager criterion (15.21) in terms of the yield stress under uniaxial tension, σ_0^+, and the yield stress under uniaxial compression, σ_0^-. Plot the line corresponding to the uniaxial stress states in Figure 15.2 and mark the plastic states under uniaxial tension and compression.

Problem 15.2: Derive the expressions for parameters c and ϕ of the Mohr–Coulomb criterion (15.22) in terms of the yield stress under uniaxial tension, σ_0^+, and the yield stress under uniaxial compression, σ_0^-. Find the slope of the tensile meridian and of the compressive meridian.

Problem 15.3: Using the results of Example 15.3, write the rate form of the stress-strain relations for von Mises material after the onset of yielding under uniaxial tension. Consider that the lateral stresses σ_{22} and σ_{33} vanish, and express the lateral strain rates $\dot{\varepsilon}_{22}$ and $\dot{\varepsilon}_{33}$ in terms of the applied strain rate $\dot{\varepsilon}_{11}$. Demonstrate that the main stress σ_{11} remains constant.

Problem 15.4: Derive the specific form of the flow rule associated with the Drucker–Prager yield condition. Does plastic flow in this case take place at constant volume?

Problem 15.5: Derive the loading-unloading criteria for Drucker–Prager elasto-plasticity.

Problem 15.6: Consider von Mises material subjected to equibiaxial tension ($\sigma_{11} = \sigma_{22}$) at plain strain conditions ($\varepsilon_{33} = 0$). Set up the rate form of the stress-strain relations after the onset of yielding. Do the principal stresses remain constant after yielding has started? Explain why. How will the results change if we consider plane stress ($\sigma_{33} = 0$)? How will the results change if we consider Drucker–Prager material?

Problem 15.7: Consider the case of neutral loading. Show that the stress rate given by $\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$ is the same as that given by $\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_{ep} : \dot{\boldsymbol{\varepsilon}}$.

Problem 15.8: Derive the general expression for the elastoplastic stiffness matrix in nonassociated plasticity. Compared to the associated case, which fundamental property affecting the storage scheme for the global stiffness matrix is different?

Problem 15.9: The parabolic Leon (1935) criterion is a simple generalization of the Mohr–Coulomb criterion. It uses a parabolic envelope in the σ–τ plane, described by $\sigma/\sigma_0^+ + a(\tau/\sigma_0^+)^2 = 1$, where σ_0^+ is the yield stress under uniaxial tension and a is a dimensionless parameter. Express the yield function in terms of the principal stresses.

16

Theorems of Plastic Analysis in Multiaxial Case

The upper and lower bound theorems that were derived in Chapters 5 and 10 for structures under uniaxial stress may be generalized to structures under multiaxial stress, provided that the flow rule is associated and that the yield surface is convex.

16.1 THEOREMS OF LIMIT ANALYSIS

First, we must generalize the definitions of statically and kinematically admissible states. We will use the notation introduced in Appendix D. Let us specify the reference loading by given body forces, $\bar{\boldsymbol{b}}$, and given surface tractions, $\bar{\boldsymbol{t}}$. A *statically admissible state* is described by a stress field $\boldsymbol{\sigma}_s$ and a load multiplier μ_s such that

$$-\boldsymbol{\sigma}_s \cdot \boldsymbol{\nabla} = \mu_s \bar{\boldsymbol{b}} \qquad \text{in } V \tag{16.1}$$

$$\boldsymbol{\sigma}_s \cdot \boldsymbol{n} = \mu_s \bar{\boldsymbol{t}} \qquad \text{on } S_t \tag{16.2}$$

$$f(\boldsymbol{\sigma}_s) \leq 0 \qquad \text{in } V \tag{16.3}$$

Note that s is not a tensorial subscript; it only labels statically admissible quantities. The above equations are the Cauchy equations of equilibrium, static boundary conditions and conditions of plastic admissibility. In the indicial notation, they could be written as

$$-\frac{\partial \sigma_{ij}^s}{\partial x_j} = \mu_s \bar{b}_i \qquad \text{in } V \tag{16.4}$$

$$\sigma_{ij}^s n_j = \mu_s \bar{t}_i \qquad \text{on } S_t \tag{16.5}$$

$$f(\sigma_{ij}^s) \leq 0 \qquad \text{in } V \tag{16.6}$$

For easy reference, let us denote the set of all the plastically admissible stress fields by

$$\bar{\mathcal{E}}_V = \{\boldsymbol{\sigma} \mid f(\boldsymbol{\sigma}(\boldsymbol{x})) \leq 0 \ \ \forall\, \boldsymbol{x} \in V\} \tag{16.7}$$

Note that (16.3) can be rewritten as $\boldsymbol{\sigma}_s \in \bar{\mathcal{E}}_V$.

A *kinematically admissible state* is described by a displacement rate field $\dot{\boldsymbol{u}}_k$ and a plastic strain rate field $\dot{\boldsymbol{\varepsilon}}_k$ such that

$$\dot{\boldsymbol{\varepsilon}}_k = (\boldsymbol{\nabla} \dot{\boldsymbol{u}}_k)_{\text{sym}} \qquad \text{in } V \tag{16.8}$$

$$\dot{\boldsymbol{u}}_k = \boldsymbol{0} \qquad \text{on } S_u \tag{16.9}$$

$$\int_V \bar{\boldsymbol{b}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}V + \int_{S_t} \bar{\boldsymbol{t}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}S > 0 \tag{16.10}$$

In the present context, k is not a tensorial subscript, but it refers to the kinematical admissibility. The previous equations are the strain-displacement relations, kinematic boundary conditions, and condition of positive external power. In the indicial notation, they could be written as

$$\dot{\varepsilon}_{ij}^k = \frac{1}{2}\left(\frac{\partial \dot{u}_i^k}{\partial x_j} + \frac{\partial \dot{u}_j^k}{\partial x_i}\right) \qquad \text{in } V \tag{16.11}$$

$$\dot{u}_i^k = 0 \qquad \text{on } S_u \tag{16.12}$$

$$\int_V \bar{b}_i \dot{u}_i^k \, \mathrm{d}V + \int_{S_t} \bar{t}_i \dot{u}_i^k \, \mathrm{d}S > 0 \tag{16.13}$$

The corresponding kinematically admissible multiplier

$$\mu_k = \frac{\int_V \mathcal{D}(\dot{\boldsymbol{\varepsilon}}_k) \, \mathrm{d}V}{\int_V \bar{\boldsymbol{b}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}V + \int_{S_t} \bar{\boldsymbol{t}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}S} \tag{16.14}$$

is evaluated from the power equality.

As we already know, an associated elastoplastic model with a convex yield surface obeys (at each material point) the postulate of maximum plastic dissipation, which states that

$$\boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p \le \mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p) \text{ for any } \boldsymbol{\sigma}^* \text{ satisfying } f(\boldsymbol{\sigma}^*) \le 0 \tag{16.15}$$

Integrating over V we can extend this postulate to the entire structure.

Postulate of Maximum Plastic Dissipation
Consider an elastoplastic material with a convex yield surface and an associated flow rule. Let $\dot{\boldsymbol{\varepsilon}}_p$ be a given plastic strain rate field describing the flow induced by a certain stress field $\boldsymbol{\sigma}$. Then, the power that would be produced by an arbitrary plastically admissible stress field $\boldsymbol{\sigma}^* \in \bar{\mathcal{E}}_V$ cannot exceed the actual dissipation rate, i.e.

$$D_{\text{int}} \equiv \int_V \mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p) \, dV \equiv \int_V \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p \, dV = \max_{\boldsymbol{\sigma}^* \in \bar{\mathcal{E}}_V} \int_V \boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p \, dV \tag{16.16}$$

To extend the theorems originally proven for trusses and frames to bodies under multiaxial stress, it is very helpful to make use of the direct analogy between the fundamental variables and equations describing a discrete structural model and those describing a continuum; see Tables 16.1 and 16.2. It is often possible to prove a theorem for the three-dimensional continuum by formally rewriting the individual steps of the proof for a discrete model.

Let us demonstrate the derivation of the fundamental theorem of limit analysis, analogous to that given in Chapter 5 for trusses and frames. We consider an arbitrary statically admissible stress field $\boldsymbol{\sigma}_s$, and an arbitrary kinematically admissible displacement rate field $\dot{\boldsymbol{u}}_k$ (completely independent of each other). Since $\boldsymbol{\sigma}_s$ is

Table 16.1 Fundamental variables for a discrete structural model and a continuum

Discrete model		Continuum model	
$\boldsymbol{d}$	nodal displacements	$\boldsymbol{u}(\boldsymbol{x})$	displacement field
$\boldsymbol{e}$	generalized strains	$\boldsymbol{\varepsilon}(\boldsymbol{x})$	strain field
$\boldsymbol{s}$	internal forces	$\boldsymbol{\sigma}(\boldsymbol{x})$	stress field
$\bar{\boldsymbol{f}}$	external nodal forces	$\bar{\boldsymbol{b}}(\boldsymbol{x})$, $\bar{\boldsymbol{t}}(\boldsymbol{x})$	body and surface forces

plastically admissible, the postulate of maximum plastic dissipation (in which we set $\boldsymbol{\sigma}^* = \boldsymbol{\sigma}_s$ and $\dot{\boldsymbol{\varepsilon}}_p = \dot{\boldsymbol{\varepsilon}}_k$) yields

$$\int_V \boldsymbol{\sigma}_s : \dot{\boldsymbol{\varepsilon}}_k \, \mathrm{d}V \le \int_V \mathcal{D}(\dot{\boldsymbol{\varepsilon}}_k) \, \mathrm{d}V \tag{16.17}$$

As $\boldsymbol{\sigma}_s$ is in equilibrium with body forces $\mu_s \bar{\boldsymbol{b}}$ and surface tractions $\mu_s \bar{\boldsymbol{t}}$, and $\dot{\boldsymbol{\varepsilon}}_k$ is compatible with $\dot{\boldsymbol{u}}_k$, we can use the Clapeyron theorem (see equation (D.88) in Appendix D) and transform the left-hand side of (16.17) into

$$\begin{aligned} \int_V \boldsymbol{\sigma}_s : \dot{\boldsymbol{\varepsilon}}_k \, \mathrm{d}V &= \int_V \mu_s \bar{\boldsymbol{b}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}V + \int_{S_t} \mu_s \bar{\boldsymbol{t}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}S \\ &= \mu_s \left(\int_V \bar{\boldsymbol{b}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}V + \int_{S_t} \bar{\boldsymbol{t}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}S \right) \end{aligned} \tag{16.18}$$

The power represented by the sum of integrals in the parentheses is positive according to assumption (16.10). Combining (16.18) with (16.17) and using the definition of the kinematically admissible multiplier (16.14), we obtain the inequality

$$\mu_s \le \frac{\int_V \mathcal{D}(\dot{\boldsymbol{\varepsilon}}_k) \, \mathrm{d}V}{\int_V \bar{\boldsymbol{b}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}V + \int_{S_t} \bar{\boldsymbol{t}} \cdot \dot{\boldsymbol{u}}_k \, \mathrm{d}S} = \mu_k \tag{16.19}$$

which shows that the theorems of limit analysis from Chapter 5 can be extended to bodies under general multiaxial stress.

Fundamental Theorem of Limit Analysis
No statically admissible multiplier is larger than any kinematically admissible multiplier.

Lower Bound Theorem
The safety factor is the largest statically admissible multiplier.

Upper Bound Theorem
The safety factor is the smallest kinematically admissible multiplier.

As already mentioned in Chapter 5, the above theorems were stated by Gvozdev (1938), and their formal proof for plane strain was later given by Drucker, Greenberg and Prager (1951) and for general three-dimensional stress by Drucker, Prager and Greenberg (1952). For rigid-plastic materials, similar theorems were derived by Hill (1951, 1952).

Table 16.2 Fundamental equations for a discrete structural model and a continuum

Discrete model		Continuum model
$\boldsymbol{e} = \boldsymbol{B}\boldsymbol{d}$	kinematic equations	$\boldsymbol{\varepsilon} = (\boldsymbol{\nabla}\boldsymbol{u})_{\text{sym}}$ in V, $\boldsymbol{u} = \bar{\boldsymbol{u}}$ on S_u
$\boldsymbol{s} = \boldsymbol{D}_e\boldsymbol{e}_e$	elastic constitutive equations	$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon}_e$
$\mu\bar{\boldsymbol{f}} = \boldsymbol{B}^T\boldsymbol{s}$	static equations	$\mu\bar{\boldsymbol{b}} = \boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$ in V, $\mu\bar{\boldsymbol{t}} = \boldsymbol{\sigma}\cdot\boldsymbol{n}$ on S_t
$\boldsymbol{s}_s^T\boldsymbol{e}_k = \boldsymbol{f}_s^T\boldsymbol{d}_k$	work equality (Clapeyron theorem)	$\int_V \boldsymbol{\sigma}_s : \boldsymbol{\varepsilon}_k \,\mathrm{d}V = \int_V \boldsymbol{b}_s\cdot\boldsymbol{u}_k \,\mathrm{d}V + \int_S \boldsymbol{t}_s\cdot\boldsymbol{u}_k \,\mathrm{d}S$
$\boldsymbol{f}(\boldsymbol{s}) \le \boldsymbol{0}$	plastic admissibility	$f(\boldsymbol{\sigma}) \le 0$ in V
$\mathcal{E} = \{\boldsymbol{s} \mid \boldsymbol{f}(\boldsymbol{s}) < \boldsymbol{0}\}$	elastic domain	$\mathcal{E}_V = \{\boldsymbol{\sigma} \mid f(\boldsymbol{\sigma}(x)) < 0 \text{ for all } x \in V\}$
$D_{\text{int}} = \boldsymbol{s}^T\dot{\boldsymbol{e}}_p$	dissipation power	$D_{\text{int}} = \int_V \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p \,\mathrm{d}V$
$\dot{W}_{\text{ext}} = \mu\bar{\boldsymbol{f}}^T\dot{\boldsymbol{d}}$	external power	$\dot{W}_{\text{ext}} = \int_V \mu\bar{\boldsymbol{b}}\cdot\dot{\boldsymbol{u}} \,\mathrm{d}V + \int_{S_t} \mu\bar{\boldsymbol{t}}\cdot\dot{\boldsymbol{u}} \,\mathrm{d}S$
$D_{\text{int}} = \max_{\boldsymbol{s}^*\in\mathcal{E}} \boldsymbol{s}^{*T}\dot{\boldsymbol{e}}_p$	postulate of maximum plastic dissipation	$D_{\text{int}} = \max_{\boldsymbol{\sigma}^*\in\mathcal{E}_V} \int_V \boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p \,\mathrm{d}V$
$\boldsymbol{s}_s^T\dot{\boldsymbol{e}}_k = \boldsymbol{f}_s^T\dot{\boldsymbol{d}}_k$	power equality	$\int_V \boldsymbol{\sigma}_s : \dot{\boldsymbol{\varepsilon}}_k \,\mathrm{d}V = \int_V \boldsymbol{b}_s\cdot\dot{\boldsymbol{u}}_k \,\mathrm{d}V + \int_{S_t} \boldsymbol{t}_s\cdot\dot{\boldsymbol{u}}_k \,\mathrm{d}S$

Let us emphasize again the essential role played by the assumptions of normality and convexity. If either of them is violated, the foregoing theorems cease to hold, as demonstrated by the examples given in Koiter (1953a), Drucker (1956) and Salençon (1974). Radenković (1961) derived a theorem that provides a lower bound on the collapse load for a certain class of nonassociated models. This extended lower bound theorem applies, e.g. to models using the Drucker–Prager yield condition and a flow rule derived from a plastic potential that has the same general form as the yield condition, but the friction coefficient is replaced by the dilatancy coefficient. However, if the two coefficients are substantially different, the best lower bound obtained from the Radenković theorem is much lower than the actual collapse load.

16.2 SHAKEDOWN THEOREMS

In the previous section, we have easily generalized the theorems of limit analysis to the multiaxial stress state. Let us now demonstrate the analogy between the discrete structural models and the continuum models by generalizing Melan's shakedown theorem. We consider a body subjected to a cyclic loading history described by body forces $\bar{\boldsymbol{b}}(\boldsymbol{x}, t)$ and surface tractions $\bar{\boldsymbol{t}}(\boldsymbol{x}, t)$ periodic in time. The actual history of the stress field will be denoted by $\boldsymbol{\sigma}(\boldsymbol{x}, t)$, the elastic stress field that would arise in a structure responding purely elastically by $\boldsymbol{\sigma}_E(\boldsymbol{x}, t)$, and the residual stress field by $\bar{\boldsymbol{\sigma}}_r(\boldsymbol{x}, t) = \boldsymbol{\sigma}(\boldsymbol{x}, t) - \boldsymbol{\sigma}_E(\boldsymbol{x}, t)$.

Suppose there is a time-independent distribution of shakedown stresses $\boldsymbol{\sigma}_r(\boldsymbol{x})$ such that $\boldsymbol{\sigma}_r(\boldsymbol{x}) + \boldsymbol{\sigma}_E(\boldsymbol{x}, t)$ remains inside the elastic domain for any point $\boldsymbol{x} \in V$ and at any time t. From now on, the dependence on the spatial coordinate vector $\boldsymbol{x}$ will not be marked explicitly in order to simplify the notation. We define

$$A(t) = \frac{1}{2}\int_V [\bar{\boldsymbol{\sigma}}_r(t) - \boldsymbol{\sigma}_r] : \boldsymbol{C}_e : [\bar{\boldsymbol{\sigma}}_r(t) - \boldsymbol{\sigma}_r] \,\mathrm{d}V \tag{16.20}$$

where $\boldsymbol{C}_e$ is the elastic compliance tensor. We will make use of the fact that the work of any self-equilibrated stress field on any compatible strain field is zero, which is a direct consequence of the Clapeyron theorem; see (D.90). Note that the residual stress field $\bar{\boldsymbol{\sigma}}_r(t)$, as well as $\boldsymbol{\sigma}_r$, is self-equilibrated, and that the actual strain rate field $\dot{\boldsymbol{\varepsilon}}(t)$ and the strain rate field $\dot{\boldsymbol{\varepsilon}}_E(t) = \boldsymbol{C}_e : \dot{\boldsymbol{\sigma}}_e(t)$ that would be produced in a purely elastic body are compatible. Also, note that $\dot{\boldsymbol{\varepsilon}}_E(t)$ is different from the actual elastic strain field, $\dot{\boldsymbol{\varepsilon}}_e(t) = \dot{\boldsymbol{\varepsilon}}(t) - \dot{\boldsymbol{\varepsilon}}_p(t)$, which is in general incompatible. Differentiating (16.20) with respect to time we obtain

$$\begin{aligned}
\dot{A}(t) &= \int_V [\bar{\boldsymbol{\sigma}}_r(t) - \boldsymbol{\sigma}_r] : \boldsymbol{C}_e : \dot{\bar{\boldsymbol{\tau}}}_r(t) \,\mathrm{d}V \\
&= \int_V [\bar{\boldsymbol{\sigma}}_r(t) - \boldsymbol{\sigma}_r] : [\dot{\boldsymbol{\varepsilon}}(t) - \dot{\boldsymbol{\varepsilon}}_p(t)] \,\mathrm{d}V - \int_V [\bar{\boldsymbol{\sigma}}_r(t) - \boldsymbol{\sigma}_r] : \dot{\boldsymbol{\varepsilon}}_E(t) \,\mathrm{d}V \\
&= \int_V [\boldsymbol{\sigma}_r - \bar{\boldsymbol{\sigma}}_r(t)] : \dot{\boldsymbol{\varepsilon}}_p(t) \,\mathrm{d}V \\
&= \int_V [\boldsymbol{\sigma}_r + \boldsymbol{\sigma}_E(t)] : \dot{\boldsymbol{\varepsilon}}_p(t) \,\mathrm{d}V - \int_V \mathcal{D}\left(\dot{\boldsymbol{\varepsilon}}_p(t)\right) \,\mathrm{d}V \\
&= \int_V [\boldsymbol{\sigma}_r + \boldsymbol{\sigma}_E(t)] : \dot{\boldsymbol{\varepsilon}}_p(t) \,\mathrm{d}V - D_{\text{int}}(t)
\end{aligned} \tag{16.21}$$

To complete the proof, we need to know that there is a constant $\beta > 1$ such that

$$f\left(\beta\left[\boldsymbol{\sigma}_r(\boldsymbol{x}) + \boldsymbol{\sigma}_E(\boldsymbol{x}, t)\right]\right) \leq 0 \tag{16.22}$$

for any $\boldsymbol{x} \in V$ and at any time t (cf. equation (10.15)). However, since we now work in a space of tensor fields that has an infinite dimension (in contrast to the space of internal forces, which had a finite dimension), the property cannot be deduced from the assumption that $f\left(\boldsymbol{\sigma}_r(\boldsymbol{x}) + \boldsymbol{\sigma}_E(\boldsymbol{x}, t)\right) < 0$, and it must itself become a hypothesis of the theorem, which is stronger than the original one. Based on (16.22), we can apply the postulate of maximum plastic dissipation, and conclude that

$$\beta \int_V \left[\boldsymbol{\sigma}_r + \boldsymbol{\sigma}_E(t)\right] : \dot{\boldsymbol{\varepsilon}}_p(t) \leq \int_V \boldsymbol{\sigma}(t) : \dot{\boldsymbol{\varepsilon}}_p(t)\, \mathrm{d}V = D_{\mathrm{int}}(t) \tag{16.23}$$

Consequently,

$$(\beta - 1) D_{\mathrm{int}}(t) \leq \beta D_{\mathrm{int}}(t) - \beta \int_V \left[\boldsymbol{\sigma}_r + \boldsymbol{\sigma}_E(t)\right] : \dot{\boldsymbol{\varepsilon}}_p(t)\, \mathrm{d}V = -\beta \dot{A}(t) \tag{16.24}$$

from which

$$\int_0^\infty D_{\mathrm{int}}(t)\, \mathrm{d}t \leq -\frac{\beta}{\beta - 1} \int_0^\infty \dot{A}(t)\, \mathrm{d}t \leq \frac{\beta}{\beta - 1} A(0) < \infty \tag{16.25}$$

This means that the total dissipation is finite, which is equivalent to shakedown.

Melan's Shakedown Theorem
If, for a given load history, there is a self-equilibrated time-independent stress field $\boldsymbol{\sigma}_r$ and a constant $\beta > 1$ such that condition (16.22) is satisfied at all times and at all the points of the body, then the structure will shake down.

Condition (16.22) can be interpreted as the condition of plastic admissibility for the stress amplified by a factor β. As the value of β is allowed to be arbitrarily close to 1, this is for all practical purposes equivalent to the standard condition of plastic admissibility. We have introduced the refined condition only to show that the theorem can be proven in a mathematically rigorous way.

It is interesting to note that Melan's shakedown theorem (Melan, 1938a) covers the lower bound theorem of limit analysis as a special case. Even though Melan had already derived it for three-dimensional continua in 1938, it seems that he never discussed the case of proportional loading separately. The reason was perhaps that he approached the problem from the practical engineering viewpoint, and considered the actual loading as highly fluctuating, far from the idealized proportional path.

The generalization of Koiter's shakedown theorem (Koiter, 1956, 1960) is left to the reader as an exercise. The theorem was extended to nonassociated flow by Maier (1969), and to dynamic loading by Corradi and Maier (1974).

Numerical methodologies that make it possible to efficiently use finite element commercial codes in limit and shakedown analysis were developed by Zarka, Frelat, Inglebert and Kasmai-Navidi (1988), Hamilton, Boyle, Shi and Mackenzie (1996), and Ponter and Carter (1997). Extension of the shakedown theory to second-order geometric effects was elaborated by Maier (1973). Important theoretical results for the general case of large displacements and large strains were obtained, e.g. by Weichert (1990), Stumpf (1993) and Polizzotto and Borino (1996). A novel theoretical framework for shakedown of structures with nonassociated flow was proposed by Saxcé and Bousshine (1998). Extensions of shakedown theory to hardening materials were studied among others by Comi and Corigliano (1991) and Stein and Huang (1994). A collection of papers on advanced topics in shakedown analysis was edited by Weichert and Maier (2000). For additional references, see the review paper by Maier, Carvelli and Cocchetti (2000).

16.3 GLOBAL UNIQUENESS THEOREM

The fundamental theorem of limit analysis implies that the collapse load is, for a given loading path, unique. An interesting question is whether the collapse mode is also unique. For perfectly plastic materials, the answer is negative. We have already noted that if several plastic hinges form simultaneously at failure, then the number of failure mechanisms is infinite. However, it can be at least shown that the stress distribution at failure is unique. This is a special case of the following general statement.

Theorem
For perfectly elastoplastic material models obeying the normality rule, the static response of the structure to a given loading history remains unique in terms of the stress history (but not necessarily in terms of the strain and displacement histories).

Before we proceed to the proof, let us derive an auxiliary lemma. Consider a material point in a given state, and let $\dot{\boldsymbol{\sigma}}^A$ and $\dot{\boldsymbol{\sigma}}^B$ be the stress rates generated by arbitrary strain rates $\dot{\boldsymbol{\varepsilon}}^A$ and $\dot{\boldsymbol{\varepsilon}}^B$, respectively (as usual, rates could be replaced by infinitesimal increments). Let us denote the difference between the rates of any quantities corresponding to the two processes by Δ, e.g. $\Delta\dot{\boldsymbol{\sigma}} = \dot{\boldsymbol{\sigma}}^A - \dot{\boldsymbol{\sigma}}^B$ and $\Delta\dot{\boldsymbol{\varepsilon}} = \dot{\boldsymbol{\varepsilon}}^A - \dot{\boldsymbol{\varepsilon}}^B$. We claim that the product $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}$ is always positive, unless the rates $\dot{\boldsymbol{\sigma}}^A$ and $\dot{\boldsymbol{\sigma}}^B$ are the same. To show that, we write

$$\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}} = \Delta\dot{\boldsymbol{\sigma}} : (\Delta\dot{\boldsymbol{\varepsilon}}_e + \Delta\dot{\boldsymbol{\varepsilon}}_p) = \Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{C}_e : \Delta\dot{\boldsymbol{\sigma}} + \Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p \tag{16.26}$$

The expression $\Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{C}_e : \Delta\dot{\boldsymbol{\sigma}}$ is positive for all $\Delta\dot{\boldsymbol{\sigma}} \neq \boldsymbol{0}$, because the elastic material compliance $\boldsymbol{C}_e$ is a positive definite tensor. So it is sufficient to show that $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p$ is always nonnegative. If the current state is elastic, both plastic rates $\dot{\boldsymbol{\varepsilon}}_p^A$ and $\dot{\boldsymbol{\varepsilon}}_p^B$ vanish, and so $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p = 0$. If the current state is plastic, we can use the associated flow rule (15.28) for both processes A and B, to obtain $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p = \Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{f}_\sigma\, \Delta\dot{\lambda}$. Since $\boldsymbol{f}_\sigma \equiv \partial f / \partial \boldsymbol{\sigma}$ is the gradient of the yield function, the product $\dot{\boldsymbol{\sigma}}^A : \boldsymbol{f}_\sigma$ is equal to the rate of yield function $\dot{f}^A$ for process A and, similar to that, $\dot{\boldsymbol{\sigma}}^B : \boldsymbol{f}_\sigma = \dot{f}^B$. Consequently,

$$\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p = \Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{f}_\sigma\, \Delta\dot{\lambda} = (\dot{f}^A - \dot{f}^B)(\dot{\lambda}^A - \dot{\lambda}^B) = \dot{f}^A\dot{\lambda}^A + \dot{f}^B\dot{\lambda}^B - \dot{f}^A\dot{\lambda}^B - \dot{f}^B\dot{\lambda}^A \tag{16.27}$$

By virtue of the consistency condition (15.34), the products $\dot{f}^A\dot{\lambda}^A$ and $\dot{f}^B\dot{\lambda}^B$ must vanish. The product $\dot{f}^A\dot{\lambda}^B$ is not tied by the consistency condition, because $\dot{f}^A$ and $\dot{\lambda}^B$

correspond to two different processes. Nevertheless, the rate of the plastic multiplier $\dot{\lambda}^B$ must be nonnegative in view of the loading-unloading conditions (15.33), and the rate of the yield function $\dot{f}^A$ must be nonpositive, because the initial value of $f^A = f = 0$ and the yield function must not become positive. This shows that $\dot{f}^A \dot{\lambda}^B \leq 0$, and by similar arguments we get $\dot{f}^B \dot{\lambda}^A \leq 0$. Hence, $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p \geq 0$, and from (16.26) it follows that

$$\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}} \geq 0 \tag{16.28}$$

where the equality sign applies only if $\Delta\dot{\boldsymbol{\sigma}} = \mathbf{0}$.

Having finished this preparatory step, we can proceed to the actual proof of uniqueness. The loading is specified by the prescribed histories of body forces $\bar{\boldsymbol{b}}$ inside the body V, surface tractions $\bar{\boldsymbol{t}}$ on the part of the surface with traction boundary conditions, S_t, and displacements $\bar{\boldsymbol{u}}$ on the part of the surface with kinematic boundary conditions, S_u. Suppose that the history of the stress field remains unique up to a certain time instant t_0, at which it bifurcates into two different solutions, denoted by superscripts ^{A} and ^{B}. For both solutions, the stress $\boldsymbol{\sigma}$ at t_0 is the same, but the stress rates $\dot{\boldsymbol{\sigma}}^A$ and $\dot{\boldsymbol{\sigma}}^B$ are different (at least in a part of the body with nonzero volume). The strain rates $\dot{\boldsymbol{\varepsilon}}^A$ and $\dot{\boldsymbol{\varepsilon}}^B$ associated with the two solutions are compatible with displacement rates $\dot{\boldsymbol{u}}^A$ and $\dot{\boldsymbol{u}}^B$, which both satisfy the same kinematic boundary conditions on S_u, and so the strain rate difference $\Delta\dot{\boldsymbol{\varepsilon}} = \dot{\boldsymbol{\varepsilon}}^A - \dot{\boldsymbol{\varepsilon}}^B$ is compatible with the displacement rate difference $\Delta\dot{\boldsymbol{u}} = \dot{\boldsymbol{u}}^A - \dot{\boldsymbol{u}}^B$, which satisfies homogeneous kinematic boundary conditions on S_u. The stress rates $\dot{\boldsymbol{\sigma}}^A$ and $\dot{\boldsymbol{\sigma}}^B$ are in equilibrium with the same rates of body forces and satisfy the same static boundary conditions on S_t, and so their difference $\Delta\dot{\boldsymbol{\sigma}} = \dot{\boldsymbol{\sigma}}^A - \dot{\boldsymbol{\sigma}}^B$ is self-equilibrated. Due to these special properties of $\Delta\dot{\boldsymbol{\varepsilon}}$ and $\Delta\dot{\boldsymbol{\sigma}}$, it follows from (D.90) that

$$\int_V \Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}} \, \mathrm{d}V = 0 \tag{16.29}$$

However, according to the previously derived inequality (16.28), the integrand $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}$ is nonnegative, and so (16.29) can hold only if the integrand vanishes at all points of V (with the possible exception of a set of measure zero). We have also seen that $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}$ can vanish only if $\Delta\dot{\boldsymbol{\sigma}} = \mathbf{0}$, which means that the rates $\dot{\boldsymbol{\sigma}}^A$ and $\dot{\boldsymbol{\sigma}}^B$ cannot be different. Since the bifurcation time t_0 was arbitrary, the bifurcation cannot take place and the history of stress must remain unique.

The uniqueness in terms of stress does not necessarily mean that the strain history is also unique. Only the elastic strain uniquely depends on the stress, but the plastic strain does not. The stress only determines the direction of the plastic strain rate but not its magnitude. The simplest example is provided by the uniaxial tensile test, in which one end of a homogeneous prismatic bar is kept fixed while the other end is subjected to a monotonically increasing displacement along the beam axis. After the onset of yielding, the stress remains constant at all points of the bar, but the distribution of the plastic strain along the bar is almost arbitrary (it is constrained only by the compatibility condition – the integral of strain along the bar must be equal to the applied displacement). In Section 20.2 we will show that the uniqueness of strains and displacements can be guaranteed for hardening materials.

PROBLEMS

Problem 16.1: Formulate and prove Koiter's shakedown theorem for bodies under general multiaxial stress.

Problem 16.2: Formulate and prove a uniqueness theorem for discrete structural models.

17

Plastic Torsion and Shear

17.1 YIELD CONDITION

As the first and simplest demonstration of a problem in which the plastic flow takes place under multiaxial stress, we turn our attention to the plastic torsion of beams. The usual assumption of the beam theory is that the stresses in the plane perpendicular to the beam axis, σ_y, σ_z and τ_{yz}, are negligible. The only nonzero stress components are the shear stresses $\tau_{yx} = \tau_{xy}$ and $\tau_{zx} = \tau_{xz}$, and the normal stress σ_x, now denoted briefly as σ; see Figure 17.1(a).

Consider first the von Mises yield condition, for the present purpose written as $J_2 = \sigma_0^2/3$. Substituting (D.40) for the value of the second deviator invariant J_2, we obtain the yield condition in terms of the three nonzero stresses:

$$\sigma^2 + 3(\tau_{xy}^2 + \tau_{xz}^2) = \sigma_0^2 \tag{17.1}$$

Next, consider the Tresca yield condition. The shear stresses τ_{xy} and τ_{xz} in the plane of the cross section give the resultant shear stress τ (Figure 17.1(b)), the magnitude of which is given by

$$\tau = \sqrt{\tau_{xy}^2 + \tau_{xz}^2} \tag{17.2}$$

The state of stress is (locally) plane stress in the plane given by the shear resultant τ and the normal stress σ, acting along the axis x. To find the maximum shear stress, which must act in this plane, we consider the Mohr's circle in Figure 17.2. The point representing the maximum shear stress is the top of the circle, and so

$$\max \tau = \sqrt{\left(\frac{\sigma}{2}\right)^2 + \tau^2} = \sqrt{\frac{\sigma^2}{4} + \tau_{xy}^2 + \tau_{xz}^2} \tag{17.3}$$

According to the Tresca yield criterion, this value must be equal to $\sigma_0/2$. This yields

$$\sigma^2 + 4(\tau_{xy}^2 + \tau_{xz}^2) = \sigma_0^2 \tag{17.4}$$

We see that, for the specific stress state considered here ($\sigma_y = \sigma_z = \tau_{yz} = 0$), both yield criteria (von Mises and Tresca) can be written in a unified form

$$\sigma^2 + \alpha_0^2(\tau_{xy}^2 + \tau_{xz}^2) = \sigma_0^2 \tag{17.5}$$

in which $\alpha_0 = \sigma_0/\tau_0$ is the ratio between the yield stresses in uniaxial tension and in pure shear: $\alpha_0^2 = 3$ for the von Mises criterion and $\alpha_0^2 = 4$ for the Tresca criterion.

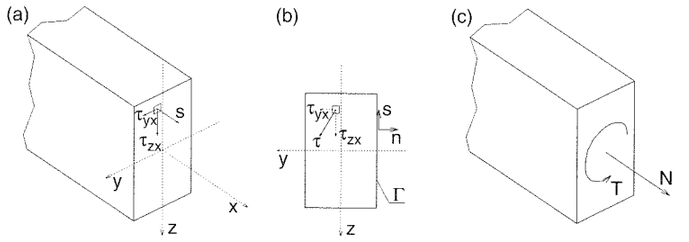


Figure 17.1 a) Beam coordinate system and nonzero stress components, b) shear stress resultant, c) normal force and torque

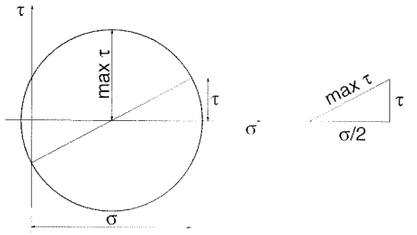


Figure 17.2 Mohr's circle characterizing the plane stress state

The actual plastic behavior of metals falls within the range

$$3 \leq \alpha_0^2 \leq 4 \tag{17.6}$$

usually somewhat closer to the value 3.

17.2 PURE TORSION

First let us analyze the case of *pure torsion*, defined as the torsion that produces no axial normal stresses, i.e. $\sigma = 0$. The pure torsion is obtained when the axial warping of the cross section is unopposed, which happens when the torque is distributed uniformly along the beam, the boundary conditions do not oppose the warping of the end cross sections, and the cross section is uniform. If these conditions are not satisfied, warping of the various cross sections of the beam is not identical, which produces axial stresses in the beam even when the loading is purely torsional. This case is known as the warping torsion. Consideration of the axial stresses due to warping torsion, however, is important only for thin-walled beams, usually of open cross sections. This problem will not be discussed here.

The first of the differential equations of equilibrium (D.74) reads

$$\frac{\partial \sigma}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} = 0 \tag{17.7}$$

In the case of pure torsion ($\sigma = 0$), this equation may be identically satisfied by introducing a stress function $\phi(y, z)$ and defining the stress components as (Prandtl, 1903)

$$\tau_{xy} = \frac{\partial \phi}{\partial z}, \qquad \tau_{xz} = -\frac{\partial \phi}{\partial y} \tag{17.8}$$

Indeed, substitution of these expressions into (17.7) leads to a left-hand side identically equal to zero for any stress function ϕ. Let us set up the boundary condition that must be satisfied by the stress function. For pure torsion it is reasonable to assume that no external surface tractions are applied along the bar (the torque being introduced by applying external forces at the end sections). From the general expression for tractions, $t_i = \sigma_{ij} n_j$, we obtain for the present stress state

$$t_x = \tau_{xy} n_y + \tau_{xz} n_z \tag{17.9}$$

$$t_y = \tau_{yx} n_x \tag{17.10}$$

$$t_z = \tau_{zx} n_x \tag{17.11}$$

where n_x, n_y and n_z are Cartesian components of the outer normal to the beam surface. For a prismatic beam we have $n_x = 0$ and so the surface traction components t_y and t_z always vanish. The surface traction component t_x can be rewritten in terms of the stress function as

$$t_x = \tau_{xy} n_y + \tau_{xz} n_z = \frac{\partial \phi}{\partial z} n_y - \frac{\partial \phi}{\partial y} n_z = \begin{Bmatrix} \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{Bmatrix} \begin{Bmatrix} -n_z \\ n_y \end{Bmatrix} \tag{17.12}$$

Since we will now deal with the plane (y, z), let us introduce the two-dimensional position vector $\boldsymbol{y} = \{y, z\}^T$. The right-hand side of (17.12) is the scalar product of the gradient of the stress function,

$$\frac{\partial \phi}{\partial \boldsymbol{y}} \equiv \begin{Bmatrix} \partial \phi / \partial y \\ \partial \phi / \partial z \end{Bmatrix} \tag{17.13}$$

with the unit vector $\boldsymbol{s} = \{-n_z, n_y\}^T$, which is perpendicular to $\{n_y, n_z\}^T$ and thus is tangential to the boundary of the cross section, Γ; see Figure 17.1(b). The scalar product

$$\left(\frac{\partial \phi}{\partial \boldsymbol{y}} \right)^T \boldsymbol{s} = \frac{\partial \phi}{\partial s} \tag{17.14}$$

represents the derivative of the stress function in the direction tangential to the boundary. According to (17.12), the traction-free boundary condition $t_x = 0$ implies that the derivative $\partial \phi / \partial s$ vanishes at all the points of the boundary, and so the value of the stress function must be constant along Γ. For the exterior boundary of the cross section we choose

$$\phi = 0 \tag{17.15}$$

In the case of multiply connected cross sections, e.g. cross sections with holes, the value of the stress function along the interior boundary (boundaries of the holes) is in general nonzero, but constant along each connected part of the boundary; see Figure 17.3.

The moment resultant of the shear stresses within the cross section is the torque, T; see Figure 17.1(c). It is clear from Figure 17.1(b) that the arms of stress components τ_{xy} and τ_{xz} are $-z$ and y, respectively. The torque is therefore given by

$$T = \int_A (y \tau_{xz} - z \tau_{xy}) \, \mathrm{d}A \tag{17.16}$$

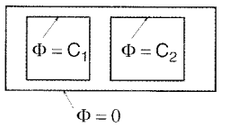


Figure 17.3 Multiply connected cross section

in which A is the area of the cross section. Expressing again the stresses in terms of the stress function, we have

$$T = -\int_A \left(y\frac{\partial \phi}{\partial y} + z\frac{\partial \phi}{\partial z} \right) \mathrm{d}A \tag{17.17}$$

In the case of a rectangular cross section, this integral may be simplified by integration by parts; for a general cross section we apply Green's integral theorem, which yields

$$\begin{aligned} T &= -\int_A \left(y\frac{\partial \phi}{\partial y} + z\frac{\partial \phi}{\partial z} \right) \mathrm{d}A \\ &= -\int_\Gamma (y\phi n_y + z\phi n_z)\, d\Gamma + \int_A \left(\frac{\partial y}{\partial y}\phi + \frac{\partial z}{\partial z}\phi \right) \mathrm{d}A = 2\int_A \phi \, \mathrm{d}A \end{aligned} \tag{17.18}$$

Here we have set the boundary integral equal to zero, which is justified by the fact that $\phi = 0$ on Γ if the cross section is simply connected.

For a multiply connected cross section (Figure 17.3), it is convenient to imagine the holes to be filled with a fictitious material that carries no stress. Defining the stress function in each hole by a constant value equal to the value of ϕ on the boundary of that hole, we do not change the torque because, for $\phi = \text{const.}$, we get $\tau_{xy} = \partial\phi/\partial z = 0$ and $\tau_{xz} = -\partial\phi/\partial y = 0$. Therefore, in (17.18), we can take A as the total area of the cross section, including the holes, and Γ only as the outer boundary curve. So we conclude that formula

$$T = 2\int_A \phi \, \mathrm{d}A \tag{17.19}$$

remains valid for the stress function extended over the holes.

So far we have established only the boundary conditions for the stress function and the relation between the stress function and the torque applied. We need to derive an additional equation governing the stress function inside the cross section. In elasticity, this additional equation is provided by the compatibility condition combined with the elastic constitutive law (Prandtl, 1903); see Problem 17.1. For perfect plasticity, the stress function in the plastic limit state can be derived from the yield condition. Substituting expressions (17.8) into the yield condition (17.5) rewritten as $\tau_{xy}^2 + \tau_{xz}^2 = (\sigma_0/\alpha_0)^2$, we obtain

$$\left(\frac{\partial \phi}{\partial y} \right)^2 + \left(\frac{\partial \phi}{\partial z} \right)^2 = \left(\frac{\sigma_0}{\alpha_0} \right)^2 \tag{17.20}$$

This is equivalent to

$$\left\| \frac{\partial \phi}{\partial \boldsymbol{y}} \right\| = \frac{\sigma_0}{\alpha_0} \tag{17.21}$$

in which the left-hand side represents the magnitude of the gradient vector of the stress function $\phi(y, z)$. Thus, we see that in a plastic state under pure torsion the

gradient vector of the stress function must have a constant magnitude related to the yield stress. This fact motivates a geometrical construction of the solution, based on the so-called *sand heap analogy* (Nádai, 1923). The function $\phi(y, z)$ has, according to (17.21), a constant slope at all points, which is a typical property of a heap of dry sand when we pour it over a base having the shape of our given cross section. The slope of the heap depends on the friction angle of the sand, and for the present analogy must be equal to σ_0/α_0. According to (17.19), the plastic torque is twice the volume of the sand heap. In the case of a multiply connected cross section, we must make an exception by considering the sand heap to be made flat over each hole. Another possible analogy is that of a roof. Roofs are often built with a constant slope, and recalling the shapes of roofs helps in constructing graphs of stress functions.

Example 17.1: Determine the plastic torque a) for a circular cross section of radius R, b) for a square cross section of side $2a$.

Solution: a) The sand heap has the shape of a circular cone of height $R\sigma_0/\alpha_0$, see Figure 17.4(a). The plastic torque is given by

$$T_0 = 2\int_A \phi \, \mathrm{d}A = 2 \times \frac{1}{3} \times \pi R^2 \times \frac{\sigma_0}{\alpha_0} R = \frac{2\pi}{3} \frac{\sigma_0 R^3}{\alpha_0} \tag{17.22}$$

b) The sand heap has the shape of a regular square pyramid of height $a\sigma_0/\alpha_0$, shown in Figure 17.4(b), for which we quickly calculate

$$T_0 = 2\int_A \phi \, \mathrm{d}A = 2 \times \frac{1}{3} \times (2a)^2 \times \frac{\sigma_0}{\alpha_0} a = \frac{8}{3} \frac{\sigma_0 a^3}{\alpha_0} \tag{17.23}$$

Example 17.2: Calculate the plastic torque for a thin-walled closed cross section of wall thickness t, which may in general be variable along the centerline of the section.

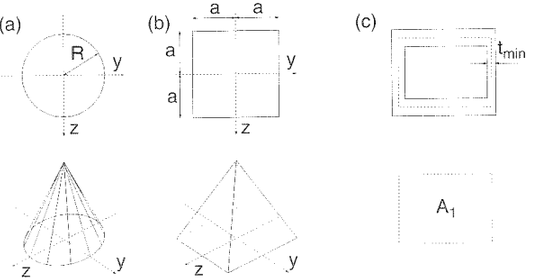


Figure 17.4 a) Sand heap over a circular base, b) sand heap over a square base, c) thin-walled section

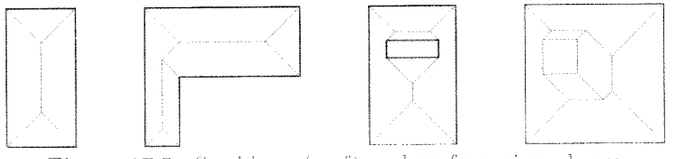


Figure 17.5 Sand heap (roof) analogy for various shapes

Solution: Within the hole of the cross section, the sand heap must be flat, and its height cannot exceed $t_{\min}\sigma_0/\alpha_0$, where $t_{\min}$ is the minimum thickness of the section wall; see Figure 17.4(c). This indicates that, for a small wall thickness, the plastic torque is

$$T_0 = 2\frac{\sigma_0}{\alpha_0} t_{\min} A_1 \tag{17.24}$$

in which A_1 represents the area closed by the centerline of the wall. Note that only the segments having the minimum thickness are fully plasticized. □

Further examples illustrating the shape of the sand heap (roof) for various other cross sections are sketched in Figure 17.5.

17.3 COMBINED TORSION AND BENDING

The exact solution of combined bending and torsion is difficult. However, for practical purposes it is usually sufficient to consider the lower and upper bounds, of which the former is especially easy to obtain. For the sake of simplicity, we restrict our attention to the case of bending about an axis of symmetry of the cross section. A generalization is left to the reader as an exercise; see Problem 17.3.

17.3.1 Lower Bound

When looking for an equilibrium stress distribution within the cross section subjected to torque T and bending moment M (Figure 17.6), we may pick the limit stress distributions for pure torsion and for pure bending (which are both in equilibrium), scale them down and superimpose them, with scaling factors such that the resulting stress distribution is plastically admissible. The pure plastic torque, T_0, generates the shear yield stress, $\tau_0 = \sigma_0/\alpha_0$. Scaling the shear stress to some value $\tau < \tau_0$, the torque is

$$T = T_0\frac{\tau}{\tau_0} = T_0\frac{\alpha_0\tau}{\sigma_0} \tag{17.25}$$

Similarly, for pure bending, the yield moment, M_0, corresponds to the uniaxial yield stress, σ_0, and so the bending moment M corresponding to normal stress σ is

$$M = M_0\frac{\sigma}{\sigma_0} \tag{17.26}$$

The relative values

$$t = \frac{T}{T_0} = \frac{\alpha_0\tau}{\sigma_0}, \qquad m = \frac{M}{M_0} = \frac{\sigma}{\sigma_0} \tag{17.27}$$

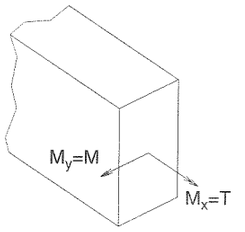


Figure 17.6 Bending moment M and torque T

may be called the non-dimensional torque and the non-dimensional bending moment. Expressing τ and σ from these equations and substituting them into the condition $\sigma^2 + \alpha_0^2\tau^2 \le \sigma_0^2$, we obtain the condition

$$m^2 + t^2 \le 1 \tag{17.28}$$

which ensures that the superposition of the pure torsion and pure bending fields leads to a plastically admissible stress field.

Equation (17.28) represents an interior approximation of the set of plastically admissible states bounded by the yield surface (yield locus, interaction curve) for combined bending and torsion. Geometrically, this curve represents a circle.

17.3.2 Upper Bound

Let $\dot\kappa_y$ denote the curvature rate associated with bending about the y axis, and let $\dot\kappa_x$ be the rate of specific twist of the beam. For symmetric cross sections, the strain rates are approximately given as

$$\dot\varepsilon_x = \dot\kappa_y z, \quad \dot\gamma_{xy} = -\dot\kappa_x z, \quad \dot\gamma_{xz} = \dot\kappa_x y \tag{17.29}$$

The first equation is exact only for pure bending. In the plastic limit state, the elastic strains remain constant and the plastic strain rates are equal to the total strain rates. The energy dissipation rate per unit length of the beam may be expressed as

$$\begin{aligned}\bar D_{\text{int}} &= \int_A (\sigma\dot\varepsilon_x + \tau_{xy}\dot\gamma_{xy} + \tau_{xz}\dot\gamma_{xz})\,\mathrm{d}A \\ &= \int_A \sigma z\,\mathrm{d}A\;\dot\kappa_y + \int_A (-\tau_{xy}z + \tau_{xz}y)\,\mathrm{d}A\;\dot\kappa_x = M\dot\kappa_y + T\dot\kappa_x\end{aligned} \tag{17.30}$$

Recall that the yield condition (17.5) is

$$f(\sigma, \tau_{xy}, \tau_{xz}) \equiv \sigma^2 + \alpha_0^2(\tau_{xy}^2 + \tau_{xz}^2) - \sigma_0^2 = 0 \tag{17.31}$$

Using the normality rule, we obtain

$$\dot\varepsilon_x = \dot\lambda\frac{\partial f}{\partial\sigma} = 2\dot\lambda\sigma \tag{17.32}$$

$$\dot\gamma_{xy} = \dot\lambda\frac{\partial f}{\partial\tau_{xy}} = 2\dot\lambda\alpha_0^2\tau_{xy} \tag{17.33}$$

$$\dot\gamma_{xz} = \dot\lambda\frac{\partial f}{\partial\tau_{xz}} = 2\dot\lambda\alpha_0^2\tau_{xz} \tag{17.34}$$

in which $\dot\lambda$ is an undetermined rate variable. Expressing the stresses σ, τ_{xy} and τ_{xz} from these equations and using (17.29), the yield function defined in (17.31) may be written as

$$f = \left(\frac{\dot\kappa_y z}{2\dot\lambda}\right)^2 + \alpha_0^2\left(\frac{\dot\kappa_x y}{2\dot\lambda\alpha_0^2}\right)^2 + \alpha_0^2\left(\frac{\dot\kappa_x z}{2\dot\lambda\alpha_0^2}\right)^2 - \sigma_0^2 \tag{17.35}$$

During plastic flow, the yield function must vanish, and the plastic multiplier is then given by

$$\dot\lambda = \frac{1}{2\sigma_0}\sqrt{z^2\dot\kappa_y^2 + \frac{y^2+z^2}{\alpha_0^2}\dot\kappa_x^2} \tag{17.36}$$

Note that the plastic multiplier $\dot\lambda$ varies over the cross section.

To simplify the notation, we introduce the radial coordinate $r = \sqrt{y^2+z^2}$ (distance from the centroid) and rewrite (17.36) as

$$\dot\lambda = \frac{1}{2\sigma_0}\sqrt{z^2\dot\kappa_y^2 + \alpha_0^{-2}r^2\dot\kappa_x^2} \tag{17.37}$$

Using (17.32)–(17.34) and (17.29), the bending moment and torque may be expressed as

$$M = \int_A z\sigma\,\mathrm{d}A = \int_A z\frac{\dot\varepsilon_x}{2\dot\lambda}\,\mathrm{d}A = \int_A \frac{z^2}{2\dot\lambda}\,\mathrm{d}A\,\dot\kappa_y \tag{17.38}$$

$$T = \int_A (y\tau_{xz} - z\tau_{xy})\,\mathrm{d}A = \int_A \left(y\frac{\dot\gamma_{xz}}{2\dot\lambda\alpha_0^2} - z\frac{\dot\gamma_{xy}}{2\dot\lambda\alpha_0^2}\right)\mathrm{d}A$$
$$= \int_A \frac{r^2}{2\dot\lambda\alpha_0^2}\,\mathrm{d}A\,\dot\kappa_x \tag{17.39}$$

Substituting (17.37) for $\dot\lambda$ we obtain relatively complicated expressions which describe the bending moment and the torque in the parametric form

$$M = \int_A \frac{z^2}{2\dot\lambda}\,\mathrm{d}A\,\dot\kappa_y = \sigma_0\int_A |z|\left[1+\left(\frac{r}{\alpha_0 z}\frac{\dot\kappa_x}{\dot\kappa_y}\right)^2\right]^{-1/2}\mathrm{d}A \equiv f_M\left(\frac{\dot\kappa_x}{\dot\kappa_y}\right) \tag{17.40}$$

$$T = \int_A \frac{r^2}{2\dot\lambda\alpha_0^2}\,\mathrm{d}A\,\dot\kappa_x = \frac{\sigma_0}{\alpha_0}\int_A r\left[1+\left(\frac{\alpha_0 z}{r}\frac{\dot\kappa_y}{\dot\kappa_x}\right)^2\right]^{-1/2}\mathrm{d}A \equiv f_T\left(\frac{\dot\kappa_x}{\dot\kappa_y}\right) \tag{17.41}$$

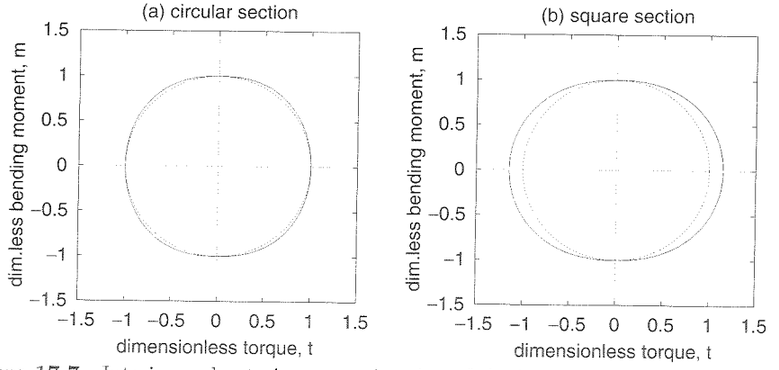


Figure 17.7 Interior and exterior approximation of the interaction curve for combined bending and torsion: a) circular cross section, b) square cross section. Dashed circles represent the lower bounds (17.28), solid curves correspond to the upper bounds (17.40)–(17.41)

These final equations allow us to determine the interaction curve. To this end, we choose a series of values for the ratio $\dot\kappa_x/\dot\kappa_y$. For each of them we calculate M and T and plot them as a point in the plane (m,t); see the solid curves in Figure 17.7. In this way, one can graphically construct an exterior approximation of the yield surface.

From the previous analysis, it follows that this approximation always lies outside the circle obtained for the lower bound. The shapes for a circular cross section and a square cross section are indicated in Figure 17.7. The maximum distance of these upper bound interaction curves from the lower bound interaction curve is relatively small and so the bounds are acceptable.

The use of the approximate interaction curves for the calculation of collapse loads of spatially loaded frames is similar to what we described for the combined bending and axial compression in Chapter 13.

17.4 COMBINED BENDING AND SHEAR

A plastic hinge formed under a distributed load between the supports of a beam is located at a cross section with an extreme value of the bending moment and a vanishing shear force. However, if the hinge forms under a concentrated load or at a support, the shear force does not vanish, and so the bending moment transmitted by a fully developed plastic hinge is reduced compared to the case of pure bending. The effect of the shear force is often negligible, but for deep beams it can play an important role. In this section we will derive estimates of the plastic limit envelope for cross sections under combined bending and shear. We will restrict attention to bending in a plane of symmetry of the cross section. The only nonzero stress components are σ_x and τ_{xz}, further denoted as σ and τ.

17.4.1 Lower Bound

Recall that, according to the usual elastic beam theory, the shear stress τ is computed from the differential equation

$$\frac{\partial}{\partial x}\left[b(z)\sigma(x,z)\right] + \frac{\partial}{\partial z}\left[b(z)\tau(x,z)\right] = 0 \tag{17.42}$$

where $b(z)$ is the width of the cross section at the level given by the coordinate z; see Figure 17.8(a). Equation (17.42) describes the equilibrium of a material layer of infinitesimal dimensions $\mathrm{d}x$ and $\mathrm{d}z$ and a finite dimension $b(z)$. For a constant width $b(z) = b$, (17.42) is equivalent to the Cauchy equation of equilibrium (17.7). If the surface of the beam is free of surface tractions parallel to the beam axis, the boundary conditions read

$$\tau(x, z_{\min}) = 0, \qquad \tau(x, z_{\max}) = 0 \tag{17.43}$$

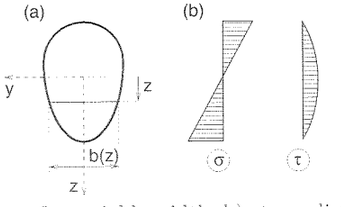


Figure 17.8 a) Cross-section of a variable width, b) stress distribution in an elastic state

can be substituted into (17.50). The resulting condition can be presented in the form

$$\frac{|M|}{M_0} + \frac{3Q^2}{4Q_0^2} \leq 1 \tag{17.52}$$

where

$$Q_0 = \frac{bh\sigma_0}{\alpha_0} = bh\tau_0 \tag{17.53}$$

is the shear force corresponding to a constant shear stress $\tau = \tau_0 = \sigma_0/\alpha_0$ across the entire cross section. Q_0 should not be understood as the maximum shear force that can be transmitted by the cross section (which is difficult to find), but rather as a convenient reference value.

Note that relation (17.51) is valid only for $M_{el} \leq |M| \leq M_0$, where $M_{el} = bh^2\sigma_0/6$ is the elastic limit moment. For bending moments below M_{el}, the depth of the elastic core evaluated from (17.51) would exceed the total depth of the cross section, and expression (17.48) for the shear stress would not be valid. Analysis of failure under large shear forces is not easy because the failure mechanism depends not only on the internal forces M and Q in one critical cross section, but also on their distribution in a finite neighborhood of that section. Since large shear forces typically occur below concentrated loads, the effect of normal stress σ_z is often non-negligible and the problem must be treated as two-dimensional. As a safe estimate, one can limit the shear force to the range in which (17.51) and (17.52) remain valid, i.e. to

$$|Q| \leq Q_0\sqrt{\frac{4}{3}\left(1 - \frac{M_{el}}{M_0}\right)} = Q_0\sqrt{\frac{4}{3}\left(1 - \frac{2}{3}\right)} = \frac{2}{3}Q_0 \tag{17.54}$$

The resulting interior approximation of the interaction curve is shown in Figure 17.10. □

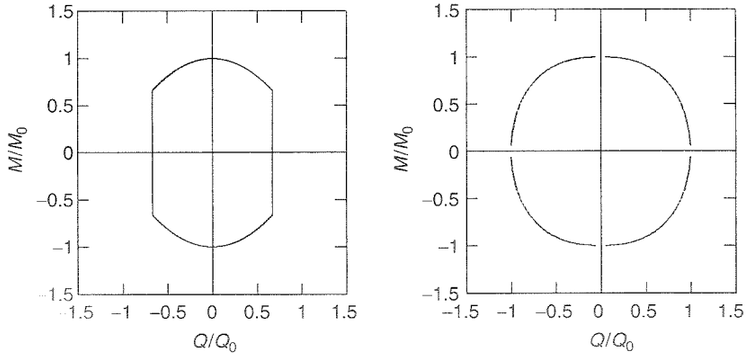


Figure 17.10 a) Interior and b) exterior approximation of the interaction curve for combined bending and shear

17.4.2 Upper Bound

Exterior approximations of the interaction curve are usually constructed by the kinematic approach. For combined bending and shear, a useful exterior approximation can be obtained by maximizing the internal forces corresponding to all plastically admissible distributions of the normal and shear stress (Rzhanitsyn, 1949; Hodge, 1957). At a first glance, it might seem strange why this 'static' approach provides an upper bound. The reason is that the actual state at collapse must be plastically admissible, and so the largest plastically admissible internal forces cannot be smaller than the actual internal forces at collapse.

More precisely, the states that will be considered satisfy the conditions of plastic admissibility and the relations between the stresses in the cross section and the internal forces, which can be regarded as one part of the equilibrium equations. Such states are in general not statically admissible, because the internal equilibrium conditions (17.42) and the boundary conditions for the shear stress (17.43) are not enforced. For one single state of this kind, it is impossible to say whether the internal forces are smaller or larger than the internal forces at collapse. However, if the internal forces are maximized over all plastically admissible states, the result provides an upper bound on the internal forces at failure. It is, of course, necessary to specify what exactly is maximized, because the bending moment and the shear force cannot attain their maximum values for the same stress distribution. In other words, we do not look for one single limit value but for the interaction curve.

To obtain a description of the entire envelope, we maximize the linear combination $M + kQ$, where k is a real parameter. For a fixed k, the equation $M + kQ = \text{const.}$ corresponds to a straight line in the (M, Q)-plane. Let $\mathcal{P}$ denote the set of all plastically admissible combinations of M and Q. If

$$\max_{(M,Q)\in\mathcal{P}} (M + kQ) \qquad M + kQ = \bar{\Phi} \tag{17.55}$$

then the line described by $M + kQ \quad \Phi$ is tangent to the boundary of $\mathcal{P}$ at its point $(\bar{M}, \bar{Q})$. By varying k we obtain a parametric description of the exterior approximation of the interaction curve.

Example 17.4: Derive a general formula for the exterior approximation of the plastic limit envelope under the combined effect of bending and shear, and plot this approximation for a rectangular cross section of width b and depth h.

Solution: Following the approach outlined above, we will maximize the expression

$$M + kQ = \int_A z\sigma(z)\,\mathrm{d}A + k\int_A \tau(z)\,\mathrm{d}A \tag{17.56}$$

over all plastically admissible stress distributions. It can be expected that the most favorable distribution satisfies the yield condition $\sigma^2(z) + \alpha_0^2\tau^2(z) = \sigma_0^2$ at all points of the cross section. Expressing the shear stress in terms of the normal stress and eliminating it from (17.56), we obtain the functional

$$\Phi(\sigma) = \int_A z\sigma(z)\,\mathrm{d}A + k\int_A \frac{1}{\alpha_0}\sqrt{\sigma_0^2 - \sigma^2(z)}\,\mathrm{d}A \tag{17.57}$$

where $z_{\min}$ and $z_{\max}$ are the extreme values of the z-coordinate at the top and bottom of the cross section, respectively.

The distribution of the normal stress is in the elastic beam theory given by the well-known formula

$$\sigma(x,z) = \frac{zM(x)}{I} \tag{17.44}$$

where $M(x)$ is the bending moment and I is the moment of inertia of the cross section with respect to its centroidal axis y. For the sake of simplicity, we assume that the beam is prismatic, and so the moment of inertia does not depend on the coordinate x. Since the distribution of the normal stress is known, the distribution of the shear stress can be obtained by integrating the differential equation (17.42) with boundary conditions (17.43). This leads to another well-known formula,

$$\tau(x,z) = \frac{Q(x)\bar{S}(z)}{Ib(z)} \tag{17.45}$$

where $Q(x) = M'(x)$ is the shear force, and $\bar{S}(z)$ is the static moment of the part of the cross section below the horizontal line at level z (shaded area in Figure 17.8(a)) with respect to the centroidal axis y.

The elastic solution, graphically illustrated in Figure 17.8(b), can be exploited for the construction of a statically admissible stress distribution providing a lower bound on the internal forces in the plastic limit state. Bezukhov (1936) proposed using the distribution shown in Figure 17.9(b). The cross section is divided into a core part, in which the elastic solution remains valid, and two plastic regions, in which the shear stress vanishes and the normal stress is equal in magnitude to the uniaxial yield stress, σ_0.

The plastic regions are also present in the neighboring cross sections, and so the normal stress in these regions is constant along the beam axis, at least in a small neighborhood of the critical cross section. Since $\partial\sigma/\partial x = 0$, vanishing shear stress in the plastic regions satisfies the equilibrium equation (17.42). The boundary conditions (17.43) are also satisfied.

The stresses in the elastic core are described by formulae similar to (17.44) and (17.45), however, with all the quantities refered to the core only. For example, $\bar{S}(z)$ is now the static moment of the shaded area in Figure 17.9(a) with respect to the centroidal axis of the core, and M is only the part of the bending moment transmitted by the core. If the stress at every point of the core is plastically admissible, the state

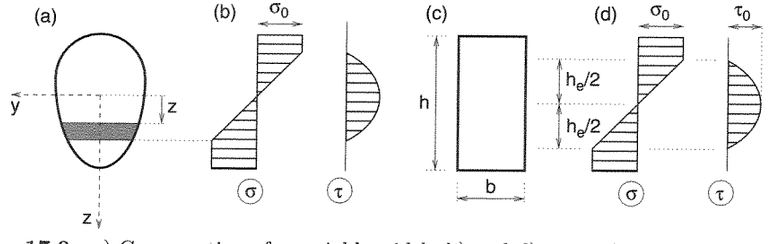


Figure 17.9 a) Cross-section of a variable width, b) and d) stress distributions according to Bezukhov, c) rectangular cross section

is statically admissible and the corresponding internal forces M and Q do not trespass on the outside of the plastic limit envelope.

Example 17.3: Construct the interior approximation of the plastic limit envelope for a rectangular cross section of width b and depth h under the combined effect of bending and shear.

Solution: Consider the distribution of stresses proposed by Bezukhov (1936), with the depth of the elastic core given by h_e (Figure 17.9(c)). The bending moment is easily calculated from the difference between the assumed normal stress distribution (Figure 17.9(d)) and the double rectangular distribution in the plastic limit case under pure bending:

$$|M| = M_0 - 2 \times \frac{1}{3}\frac{h_e}{2} \times \frac{1}{2} b \frac{h_e}{2}\sigma_0 = \frac{1}{4}bh^2\sigma_0 - \frac{1}{12}bh_e^2\sigma_0 = M_0\left(1 - \frac{h_e^2}{3h^2}\right) \tag{17.46}$$

Now we look for the maximum value of the shear force Q such that the combination of the normal stress and shear stress remains plastically admissible everywhere in the elastic core (outside the elastic core the yield condition is satisfied by construction). The moment of inertia of the elastic core is $I = bh_e^3/12$ and the variable static moment entering (17.45) is given by

$$\bar{S}(z) = b\left(\frac{h_e}{2} - z\right) \times \frac{1}{2}\left(\frac{h_e}{2} + z\right) = \frac{b}{2}\left(\frac{h_e^2}{4} - z^2\right), \qquad -\frac{h_e}{2} \le z \le \frac{h_e}{2} \tag{17.47}$$

The normal stress, $\sigma = 2\sigma_0 z/h_e$, and the shear stress,

$$\tau = \frac{Q\bar{S}(z)}{Ib} = \frac{3Q}{2bh_e^3}(h_e^2 - 4z^2) \tag{17.48}$$

must satisfy the plastic admissibility condition corresponding to the yield condition (17.5), i.e.

$$\frac{4\sigma_0^2 z^2}{h_e^2} + \alpha_0^2 \frac{9Q^2}{4b^2h_e^6}(h_e^2 - 4z^2)^2 \le \sigma_0^2 \tag{17.49}$$

The expression on the left-hand side is quadratic in terms of z^2, and it is easy to show that (17.49) holds for all $z \in [-h_e/2, h_e/2]$ if it holds for $z = 0$. In other words, if the maximum shear stress (which occurs at the centroid of the cross section) does not exceed the yield stress in shear, the entire stress distribution is plastically admissible. Consequently, the shear force Q must satisfy the condition

$$\alpha_0^2 \frac{9Q^2}{4b^2h_e^2} \le \sigma_0^2 \tag{17.50}$$

Since the depth of the elastic core, h_e, is related to the bending moment by (17.46), the value

$$h_e = h\sqrt{3\left(1 - \frac{|M|}{M_0}\right)} \tag{17.51}$$

that should be maximized over all functions $\sigma(z)$ for which the right-hand side makes sense (this replaces the condition of plastic admissibility). If the maximum of Φ is attained for certain normal stress distribution $\bar{\sigma}(z)$, the variation

$$\begin{aligned}\delta\Phi(\bar{\sigma},\delta\sigma) &= \int_A z\,\delta\sigma(z)\,\mathrm{d}A + \frac{k}{\alpha_0}\int_A \frac{-\bar{\sigma}(z)\,\delta\sigma(z)}{\sqrt{\sigma_0^2-\bar{\sigma}^2(z)}}\,\mathrm{d}A \\ &= \int_A \left(z - \frac{k\bar{\sigma}(z)}{\alpha_0\sqrt{\sigma_0^2-\bar{\sigma}^2(z)}}\right)\delta\sigma(z)\,\mathrm{d}A \end{aligned} \tag{17.58}$$

must vanish for an arbitrary variation $\delta\sigma(z)$. Consequently, the expression in parentheses must vanish at every point of the cross section. From this condition it is possible to find the most favorable distribution of normal stress,

$$\bar{\sigma}(z) = \frac{\alpha_0\sigma_0 z}{\sqrt{k^2+\alpha_0^2 z^2}} \tag{17.59}$$

and the corresponding internal forces

$$\bar{M} = \int_A z\bar{\sigma}(z)\,\mathrm{d}A = \alpha_0\sigma_0\int_A \frac{z^2}{\sqrt{k^2+\alpha_0^2z^2}}\,\mathrm{d}A \tag{17.60}$$

$$\bar{Q} = \int_A \bar{\tau}(z)\,\mathrm{d}A = \int_A \frac{1}{\alpha_0}\sqrt{\sigma_0^2-\bar{\sigma}^2(z)}\,\mathrm{d}A = \frac{k\sigma_0}{\alpha_0}\int_A \frac{\mathrm{d}A}{\sqrt{k^2+\alpha_0^2z^2}} \tag{17.61}$$

These are the general formulae providing a parametric description of the exterior approximation of the interaction curve. For $k=0$ we obtain

$$\bar{M} = \sigma_0\int_A |z|\,\mathrm{d}A \tag{17.62}$$

and $\bar{Q}=0$. For bisymmetric cross sections, (17.62) gives the plastic limit moment corresponding to pure bending, M_0, while for monosymmetric sections it gives a larger value. This is caused by the fact that we have ignored the condition $N=\int_A\sigma(z)\,\mathrm{d}A = 0$, which must be satisfied by the actual normal stress distribution in the absence of a normal force.

Let us now proceed to the special case of a rectangular cross section. Equations (17.60) (17.61) can be integrated in a closed form:

$$\begin{aligned}\bar{M} &= \alpha_0\sigma_0 b\int_{-h/2}^{h/2}\frac{z^2}{\sqrt{k^2+\alpha_0^2z^2}}\,\mathrm{d}z \\ &= \frac{M_0}{\eta}\sqrt{1+\eta^2} - \frac{M_0}{\eta^2}\ln\left(\eta+\sqrt{1+\eta^2}\right)\end{aligned} \tag{17.63}$$

$$\bar{Q} = \frac{k\sigma_0 b}{\alpha_0}\int_{-h/2}^{h/2}\frac{\mathrm{d}z}{\sqrt{k^2+\alpha_0^2z^2}} = \frac{Q_0}{\eta}\ln\left(\eta+\sqrt{1+\eta^2}\right) \tag{17.64}$$

To simplify the result, we have replaced the parameter k (which has the dimension of length) by a dimensionless parameter $\eta = h\alpha_0/2k$. The resulting approximation of the interaction curve is plotted in Figure 17.10(b). It turns out to be very close to a circle. □

PROBLEMS

Problem 17.1: Derive the differential equation governing the stress function inside the cross section for a linear elastic material.
Hint: Start from the compatibility equation

$$\frac{\partial\gamma_{xz}}{\partial y} - \frac{\partial\gamma_{xy}}{\partial z} = 2\kappa_x \tag{17.65}$$

where κ_x is the relative angle of twist (i.e. the curvature associated with the torque).

Problem 17.2: Derive the formula for the plastic torque of a rectangular cross section of sides $2a$ and $2b$. Plot the distribution of shear stresses in the plastic limit state.

Problem 17.3: Generalize the lower bound and upper bound approaches to combined bending and torsion to the case of bending about an arbitrary axis. For the sake of simplicity, assume the section to have two axes of symmetry.

Problem 17.4*: Improve the interior approximation of the interaction curve for combined bending and shear of a rectangular cross section.
Hint: Modify the Bezukhov distribution of normal and shear stresses in the core region such that the yield condition is satisfied everywhere in the cross section.

18

Limit Loads of Plates

18.1 BASIC RELATIONS

Complex problems of plasticity of plates can today be readily solved with good accuracy by finite element programs (Belytschko, Liu and Moran, 2000; Zienkiewicz and Taylor, 2000). Therefore, it makes sense to limit our attention to the basic principles and those solutions that are simple enough to provide insight and serve as benchmarks of finite element programs.

Similar to the elastic theory, the inelastic behavior of thin plates may be analyzed under Kirchhoff's assumption that the normals to the middle plane of the plate remain straight and normal to the deformed middle plane. This assumption yields

$$u = -z\frac{\partial w}{\partial x}, \qquad v = -z\frac{\partial w}{\partial y} \tag{18.1}$$

in which x, y = coordinates in the plane of the plate (Figure 18.1(a)), z = distance from the middle plane, w = deflection, and u, v = displacements in the x and y directions. By differentiation, the strains are obtained as

$$\varepsilon_x = \frac{\partial u}{\partial x} = -z\frac{\partial^2 w}{\partial x^2} = z\kappa_x, \quad \varepsilon_y = \frac{\partial v}{\partial y} = -z\frac{\partial^2 w}{\partial y^2} = z\kappa_y \tag{18.2}$$

$$\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = -2z\frac{\partial^2 w}{\partial x \partial y} = z\kappa_{xy} \tag{18.3}$$

in which

$$\kappa_x = -\frac{\partial^2 w}{\partial x^2}, \quad \kappa_y = -\frac{\partial^2 w}{\partial y^2}, \quad \kappa_{xy} = -2\frac{\partial^2 w}{\partial x \partial y} \tag{18.4}$$

are the curvatures. The internal virtual work can be expressed as

$$\begin{aligned} \delta W_{\text{int}} &= \int_V (\sigma_x\,\delta\varepsilon_x + \sigma_y\,\delta\varepsilon_y + \tau_{xy}\,\delta\gamma_{xy})\,\mathrm{d}V \\ &= \int_A \int_{-h/2}^{h/2} (\sigma_x z\,\delta\kappa_x + \sigma_y z\,\delta\kappa_y + \tau_{xy} z\,\delta\kappa_{xy})\,\mathrm{d}z\,\mathrm{d}A \\ &= \int_A (M_x\,\delta\kappa_x + M_y\,\delta\kappa_y + M_{xy}\,\delta\kappa_{xy})\,\mathrm{d}A \end{aligned} \tag{18.5}$$

where V is the volume of the plate, A is the area of the middle plane, h is the thickness of the plate, and

$$M_x = \int_{-h/2}^{h/2} z\sigma_x \,\mathrm{d}z, \quad M_y = \int_{-h/2}^{h/2} z\sigma_y \,\mathrm{d}z \tag{18.6}$$

$$M_{xy} = \int_{-h/2}^{h/2} z\tau_{xy} \,\mathrm{d}z = \int_{-h/2}^{h/2} z\tau_{yx} \,\mathrm{d}z = M_{yx} \tag{18.7}$$

are the bending moments and twisting moment in the plate (per unit length). The notation traditionally used in the plate theory is different from the general convention, according to which M_x would be the moment acting about the x-axis (recall the beam theory, in which x is the beam axis and M_x is the torque). In the plate theory, M_x is the moment produced by stress σ_x, but this moment acts about the y-axis; see Figure 18.1(c). Note also that the positive orientations of M_y and M_{yx} in cross sections with positive normals do not coincide with the positive orientation of the coordinate axis x. The reason is that we want to keep equations (18.6)–(18.7) simple and natural, but then we have to sacrifice consistency with the general notation. Since the internal moments are work-conjugate to curvatures κ_x, κ_y and $\kappa_{xy} = \kappa_{yx}$, one must be careful and interpret κ_x as the curvature related to the gradient of strain ε_x across the plate depth, and not as the curvature related to bending about the x-axis.

The principle of virtual work implies that

$$\int_A (M_x\,\delta\kappa_x + M_y\,\delta\kappa_y + M_{xy}\,\delta\kappa_{xy})\,\mathrm{d}A = \int_A p\,\delta w\,\mathrm{d}A \tag{18.8}$$

for any virtual deflection δw satisfying the homogeneous kinematic boundary conditions and for the corresponding virtual curvature fields satisfying the kinematic equations (18.4). The external work is done by a distributed load perpendicular to the middle plane of the plate, characterized by its intensity p (force per area). After substitution from (18.2)–(18.3) and twofold application of Green's theorem, we obtain the differential equation of equilibrium,

$$\frac{\partial^2 M_x}{\partial x^2} + 2\frac{\partial^2 M_{xy}}{\partial x \partial y} + \frac{\partial^2 M_y}{\partial y^2} = -p \tag{18.9}$$

For the sake of simplicity, we do not pay attention here to the static boundary conditions and we do not consider loading by concentrated external forces or by external moments.

In the plastic limit state, the elastic strains remain constant and the plastic strain rates are equal to the total strain rates. The rate of dissipation per unit area of the plate may then be expressed as

$$\overline{\overline{D}}_{\mathrm{int}} = \int_{-h/2}^{h/2} (\sigma_x\dot{\varepsilon}_x + \sigma_y\dot{\varepsilon}_y + \tau_{xy}\dot{\gamma}_{xy})\,\mathrm{d}z = M_x\dot{\kappa}_x + M_y\dot{\kappa}_y + M_{xy}\dot{\kappa}_{xy} \tag{18.10}$$

Similar to the bending of beams, the limit state of greatest capacity is obtained for a double rectangular distribution of the stresses across the depth of the plate (Figure 18.1(b)). Thus, the limit values of the bending moments in the x and y directions and of the twisting moment are

$$M_x = \sigma_x\frac{h^2}{4}, \quad M_y = \sigma_y\frac{h^2}{4}, \quad M_{xy} = \tau_{xy}\frac{h^2}{4} \tag{18.11}$$

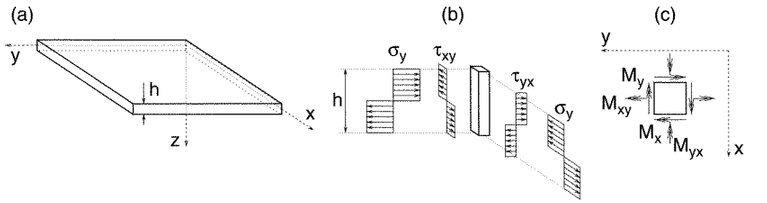


Figure 18.1 a) Plate coordinate system, b) distribution of normal and shear stresses in a fully plasticized infinitesimal plate element, c) bending and twisting moments on an infinitesimal plate element seen from the top

In contrast to the bending of beams, however, the normal stresses σ_x and σ_y are not equal to the yield limits in uniaxial tension; rather, they must satisfy a yield condition for plane stress, taking into account the in-plane shear stress τ_{xy}. The out-of-plane shear stresses, τ_{xz} and τ_{yz}, are usually neglected. Consider a general yield condition of the form

$$f(\sigma_x, \sigma_y, \tau_{xy}) = 0 \tag{18.12}$$

applicable to plane stress states. In a fully plasticized cross section, the stresses σ_x, σ_y and τ_{xy} are of constant magnitude. Expressing them from (18.11) in terms of the bending and twisting moments, we obtain the corresponding yield criterion for a plate,

$$f\left(\frac{4M_x}{h^2}, \frac{4M_y}{h^2}, \frac{4M_{xy}}{h^2}\right) = 0 \tag{18.13}$$

As an illustration, consider the *von Mises yield criterion* (15.10), specialized for plane stress ($\sigma_z = \tau_{xz} = \tau_{yz} = 0$) and rewritten in the form

$$f(\sigma_x, \sigma_y, \tau_{xy}) \equiv \sigma_x^2 - \sigma_x\sigma_y + \sigma_y^2 + 3\tau_{xy}^2 - \sigma_0^2 = 0 \tag{18.14}$$

Expressing from (18.11) stresses in terms of moments, the criterion takes the form

$$F(M_x, M_y, M_{xy}) \equiv M_x^2 - M_xM_y + M_y^2 + 3M_{xy}^2 - M_0^2 = 0 \tag{18.15}$$

where

$$M_0 = \sigma_0\frac{h^2}{4} \tag{18.16}$$

The relations among the curvature rates may be determined from the normality rule. By differentiating (18.15) it follows

$$\dot{\kappa}_x = \dot{\lambda}\frac{\partial F}{\partial M_x} = \dot{\lambda}(2M_x - M_y), \quad \dot{\kappa}_y = \dot{\lambda}(2M_y - M_x), \quad \dot{\kappa}_{xy} = 6\dot{\lambda}M_{xy} \tag{18.17}$$

in which $\dot{\lambda}$ is an arbitrary plastic multiplier rate.

If we consider the *Tresca criterion* (15.13), we have, for plane stress,

$$f(\sigma_1, \sigma_2) \equiv \max(|\sigma_1|,\ |\sigma_2|,\ |\sigma_1 - \sigma_2|) - \sigma_0 = 0 \tag{18.18}$$

in which σ_1 and σ_2 are the principal stresses in the plane of the plate. Expressing these stresses in terms of the principal bending moments M_1 and M_2, the Tresca yield criterion becomes

$$F(M_1, M_2) \equiv \max(|M_1|,\ |M_2|,\ |M_1 - M_2|) - M_0 = 0 \tag{18.19}$$

18.2 KINEMATIC APPROACH (YIELD LINE THEORY)

Experience indicates that good upper bound estimates of the collapse load of plates can usually be obtained assuming plastic deformation to be concentrated in narrow strips, called *yield lines* or yield hinges (or plastic hinges), in which the plastic curvature is cylindrical. Denoting the direction parallel to the yield line as x_2 and the normal direction as x_1 (Figure 18.2(a)), we have only one non-zero curvature rate, namely $\dot{\kappa}_1$. The rates of the curvature along the yield line and of the twisting curvature are zero, $\dot{\kappa}_2 = \dot{\kappa}_{12} = 0$. These conditions determine the direction of plastic flow, and so the corresponding moments can be deduced from the flow rule and the yield condition. For example, for the *von Mises yield condition* we get

$$\dot{\kappa}_2 = \dot{\lambda}(2M_2 - M_1) = 0 \tag{18.20}$$

$$\dot{\kappa}_{12} = 6\dot{\lambda}M_{12} = 0 \tag{18.21}$$

which implies that $M_2 = 0.5M_1$ and $M_{12} = 0$ since, during plastic flow, the rate of the plastic multiplier, $\dot{\lambda}$, cannot vanish. Substituting this into the yield condition (18.15) we obtain

$$M_1^2 - M_1 \times 0.5M_1 + (0.5M_1)^2 = M_0^2 \tag{18.22}$$

and the moment in the direction normal to the yield line, denoted in general as M_p, is in this case

$$M_p = M_1 = \frac{2}{\sqrt{3}} M_0 \tag{18.23}$$

Note that the corresponding stress

$$\sigma_p = \frac{4M_p}{h^2} = \frac{2}{\sqrt{3}} \frac{4M_0}{h^2} = \frac{2}{\sqrt{3}} \sigma_0 \tag{18.24}$$

exceeds the uniaxial yield stress by 15%. The reason is that the lateral stresses $\sigma_2 = \sigma_p/2$ along the yield line, which must develop for the bending to be cylindrical (i.e.

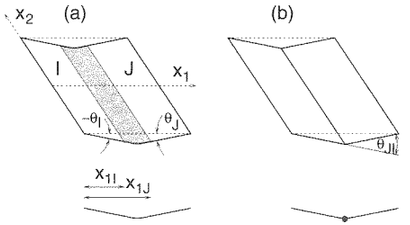


Figure 18.2 a) Narrow zone of high curvature, b) idealized yield line

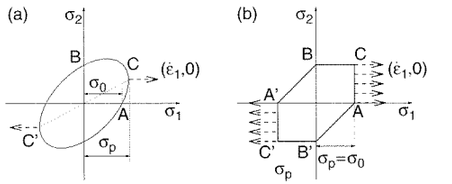


Figure 18.3 Plastic flow for cylindrical bending for a) von Mises condition, b) Tresca condition

$\dot{\varepsilon}_2 = 0$), decrease the value of the second deviatoric invariant compared to the uniaxial case. The situation is sketched in Figure 18.3(a), in which point A represents the uniaxial yield stress σ_0, while point C corresponds to the actual stress state in a yield line.

A similar sketch for the *Tresca criterion* is shown in Figure 18.3(b). There are now many points for which the strain rate vector is horizontal ($\dot{\varepsilon}_2 = 0$). They all occupy the sides AC and A'C' of the Tresca hexagon. For all these points, the normal stress across the yield line is the same, $\sigma_p = \sigma_0$, and the yield bending moment across the yield line,

$$M_p = \sigma_0 \frac{h^2}{4} = M_0 \tag{18.25}$$

is the same as for bending of a beam with zero lateral stress. This is because the lateral stress, σ_2, does not affect the maximum shear stress, which is equal to $(\sigma_1 - \sigma_3)/2 = \sigma_1/2$ as long as σ_2 remains between 0 and σ_1.

The yield zone is usually a narrow strip, and so we model it by a single line. Similar to the yield hinge of a beam, the rate of inelastic curvature across the yield zone is lumped into a sudden change of the rotation rate in the yield line, $\dot{\theta}_{JI} = \dot{\theta}_J - \dot{\theta}_I$; see Figure 18.2(b). The dissipation rate per unit length of the line is then given by

$$\overline{D}_{\text{int}} = \int_{x_{1I}}^{x_{1J}} M_1 \dot{\kappa}_1 \, dx_1 = M_1 \int_{x_{1I}}^{x_{1J}} \dot{\kappa}_1 \, dx_1 = M_1(\dot{\theta}_J - \dot{\theta}_I) = M_1 \dot{\theta}_{JI} = M_p |\dot{\theta}_{JI}| \tag{18.26}$$

This expression is completely analogous to formula (2.22) for the dissipation in a yield hinge.

During plastic collapse, the plate segments behave as rigid because, at constant stress, the elastic deformations remain unchanged. As is well known from elementary mechanics, small rotations of rigid bodies can be represented by vectors. The components of the rotation vector represent rotations about the individual coordinate axes. If the rotations of rigid segments I and J are described by vectors $\boldsymbol{\theta}_I = \{\theta_{Ix}, \theta_{Iy}\}^T$ and $\boldsymbol{\theta}_J = \{\theta_{Jx}, \theta_{Jy}\}^T$, respectively, the relative rotation of segment J with respect to segment I is given by

$$\boldsymbol{\theta}_{JI} = \boldsymbol{\theta}_J - \boldsymbol{\theta}_I = \begin{Bmatrix} \theta_{Jx} - \theta_{Ix} \\ \theta_{Jy} - \theta_{Iy} \end{Bmatrix} \tag{18.27}$$

We work with only two components of rotation vectors because, for the plate segments considered here, the third component – rotation about the vertical axis z – is always zero. The magnitude of the relative rotation vector,

$$\theta_{JI} = ||\boldsymbol{\theta}_J - \boldsymbol{\theta}_I|| = \sqrt{(\theta_{Jx} - \theta_{Ix})^2 + (\theta_{Jy} - \theta_{Iy})^2} \tag{18.28}$$

is the rotation in the yield line between the segments.

Now we have all that we need for the kinematic limit analysis of plates. As usual, we select a suitable, kinematically admissible, assumed collapse mechanism with one degree of freedom and then calculate an upper bound on the actual collapse load from the dissipation equality.

Example 18.1: Calculate the upper bound on the collapse load for a simply supported square plate of side $2a$ loaded by a uniform pressure.

Solution: We denote by w_0 the deflection at the center of the plate. From experience, the yield lines can be expected to form along the diagonals; see Figure 18.4(a), in which the thick dashed lines indicate positive yield hinges (tension in the bottom fibers) and the thin double lines represent simply supported edges. We must now relate the rotation rate in the yield lines to the deflection rate $\dot{w}_0$. To this end we consider a vertical cross section of the plate normal to the yield hinge; see section AC in Figure 18.4(a). From the figure, the rotation rate in yield line OB is $\dot{\theta}_{OB} = 2\dot{w}_0/(a\sqrt{2}) = \sqrt{2}\dot{w}_0/a$. Alternatively, we could obtain $\dot{\theta}_{OB}$ as the relative rotation rate between segments $1 \equiv$ OAB and $2 \equiv$ OBC. The rotation vector $\dot{\boldsymbol{\theta}}_1$ has components $\dot{\theta}_{1x} = -\dot{w}_0/a$ and $\dot{\theta}_{1y} = 0$, while $\dot{\boldsymbol{\theta}}_2$ has components $\dot{\theta}_{2x} = 0$ and $\dot{\theta}_{2y} = \dot{w}_0/a$. Substitution into (18.28) leads to

$$\dot{\theta}_{21} = \sqrt{\left(0 + \frac{\dot{w}_0}{a}\right)^2 + \left(\frac{\dot{w}_0}{a} - 0\right)^2} = \sqrt{2}\frac{\dot{w}_0}{a} \tag{18.29}$$

This rotation rate is the same in all the yield hinges, due to symmetry. Thus, the dissipation rate in the plate is

$$D_{\text{int}} \quad 4l_{OB}M_p|\dot{\theta}_{OB}| = 4\sqrt{2}aM_p\frac{\sqrt{2}\dot{w}_0}{a} = 8M_p\dot{w}_0 \tag{18.30}$$

where $l_{OB} = a\sqrt{2}$ is the length of segment OB. For a uniform load p, the external power is

$$\dot{W}_{\text{ext}} = \int_A p\dot{w}dA = p\int_A \dot{w}dA = p\dot{V} \tag{18.31}$$

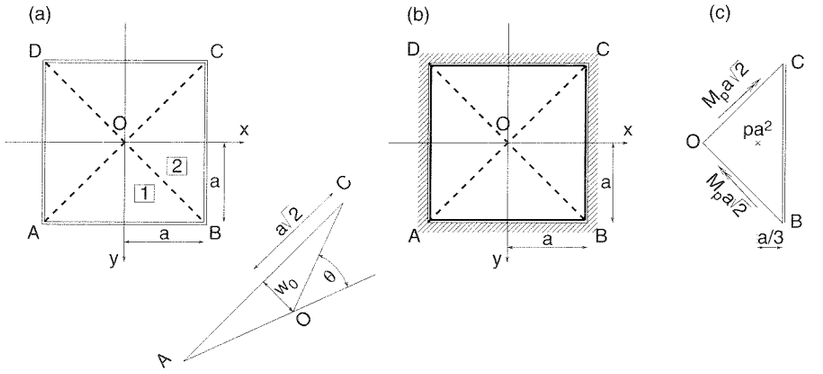


Figure 18.4 Plastic collapse of a square plate: a) simply supported plate with diagonal yield lines, b) clamped plate with diagonal yield lines and negative yield lines along the supports, c) moments and load resultant on segment OBC

in which V denotes the volume under the deflected surface of the plate. This surface is a square pyramid of height w_0 and base $2a \times 2a$ (Figure 18.4(a)), and this immediately yields

$$\dot{W}_{\text{ext}} = p\frac{1}{3}(2a)^2\dot{w}_0 = \frac{4}{3}pa^2\dot{w}_0 \tag{18.32}$$

Now, using the power equality $D_{\text{int}} = \dot{W}_{\text{ext}}$, we get an equation from which we may calculate a kinematically admissible load,

$$p_k = \frac{8M_p\dot{w}_0}{4a^2\dot{w}_0/3} = \frac{6M_p}{a^2} \tag{18.33}$$

providing an upper bound on the collapse load.

Instead of comparing the external and internal power, we can alternatively obtain the collapse load from the equilibrium condition of the segment BCD of the plate (Figure 18.4(c)). Considering the forces marked in the figure (the cross denotes the resultant of the distributed load p acting perpendicular to the plane xy), and taking the condition of moment equilibrium about the simply supported side BC, we get

$$-M_pa\sqrt{2}\frac{\sqrt{2}}{2} - M_pa\sqrt{2}\frac{\sqrt{2}}{2} + pa^2\frac{a}{3} = 0 \tag{18.34}$$

which again yields the same result for p_k.

Note that even though we have used an equilibrium equation, the calculated load is an upper bound because it corresponds to a kinematically admissible mechanism. The equilibrium equation is here used merely as an alternative to expressing the power equality, but it does not lead to a statically admissible solution. The lower bound (static) approach would consist in constructing a plastically admissible field of moments in equilibrium with a certain load p_s.

Example 18.2: Repeat the analysis from Example 18.1 for a plate with clamped edges.

Solution: In this case, we may consider the yield lines as shown in Figure 18.4(b), where the thick solid lines indicate negative yield hinges (tension in the top fibers). The rotation rate in these yield hinges along the edges of the plate is $\dot{\theta}_{AB} = -\dot{w}_0/a$. Thus, the expression for the dissipation rate (18.30) must be augmented by the term $4l_{AB}M_p|\dot{\theta}_{AB}| = 4 \times 2aM_p\dot{w}_0/a = 8M_p\dot{w}_0$. The value of $\dot{W}_{\text{ext}}$ remains unchanged. The resulting upper bound on the collapse load is thus twice as large as it was for the simply supported plate:

$$p_k = \frac{16M_p\dot{w}_0}{4a^2\dot{w}_0/3} = \frac{12M_p}{a^2} \tag{18.35}$$

□

The simple support is usually one-sided, i.e. the plate can lift up from the support. In fact, it is very difficult to fabricate a moment-free support in which the plate cannot lift up. Lifting of plate corners leads to collapse at a lower load.

Example 18.3: Try to obtain a better upper bound on the collapse load of a simply supported square plate using the pattern of 8 yield lines sketched in Figure 18.5(a).

Solution: Let ξ characterize the position of the yield line as shown in Figure 18.5(a). Rigid segment $1 \equiv$ OEF rotates about axis EF at rate $\dot{\theta}_{1x} = -\dot{w}_0/a$. Rigid segment $2 \equiv$ OEAG rotates about axis EG at rate $\dot{\theta}_2 = \dot{w}_0\sqrt{2}/(1+\xi)a$, which can be decomposed into components $\dot{\theta}_{2x} = -\dot{\theta}_2\sqrt{2}/2 = -\dot{w}_0/(1+\xi)a$ and $\dot{\theta}_{2y} = -\dot{\theta}_2\sqrt{2}/2 = -\dot{w}_0/(1+\xi)a$. According to (18.28), the relative rotation between the two segments is

$$\begin{aligned}\dot{\theta}_{12} &= \sqrt{\left(\dot{\theta}_{1x}-\dot{\theta}_{2x}\right)^2+\left(\dot{\theta}_{1y}-\dot{\theta}_{2y}\right)^2} \\ &= \sqrt{\left(-\frac{\dot{w}_0}{a}+\frac{\dot{w}_0}{(1+\xi)a}\right)^2+\left(0+\frac{\dot{w}_0}{(1+\xi)a}\right)^2} \\ &= \frac{\dot{w}_0}{(1+\xi)a}\sqrt{1+\xi^2} \end{aligned} \tag{18.36}$$

The total length of the yield hinges is $8l_{OE} = 8a\sqrt{1+\xi^2}$. Therefore, the dissipation rate is

$$D_{\text{int}} = 8l_{OE}M_p|\dot{\theta}_{12}| = 8a\sqrt{1+\xi^2}M_p\frac{\dot{w}_0}{(1+\xi)a}\sqrt{1+\xi^2} = \frac{1+\xi^2}{1+\xi}8M_p\dot{w}_0 \tag{18.37}$$

Decomposing the volume under the deflected middle plane into elementary tetrahedra, we can express the external power of uniform load p,

$$\begin{aligned}\dot{W}_{\text{ext}} &= \int_A p\dot{w}dA = p\times 4\times\frac{1}{3}\left\{\frac{1}{2}\left[\sqrt{2}a(1-\xi)\right]\left[(1+\xi)\frac{a}{\sqrt{2}}\right]\dot{w}_0\right. \\ &\quad \left.+\frac{1}{2}a(2\xi a)\dot{w}_0-\frac{1}{2}\left[\sqrt{2}a(1-\xi)\right]\left[(1-\xi)\frac{a}{\sqrt{2}}\right]\frac{1-\xi}{1+\xi}\dot{w}_0\right\} \\ &= \frac{2}{3}\left[(1-\xi^2)+2\xi-\frac{(1-\xi)^3}{1+\xi}\right]pa^2\dot{w}_0 = \frac{4}{3}\frac{3\xi-\xi^2}{1+\xi}pa^2\dot{w}_0 \end{aligned} \tag{18.38}$$

Setting $\dot{W}_{\text{ext}} = D_{\text{int}}$, we then get the load as a function of the parameter ξ:

$$p_k(\xi) = \frac{1+\xi^2}{3\xi-\xi^2}\frac{6M_p}{a^2} \tag{18.39}$$

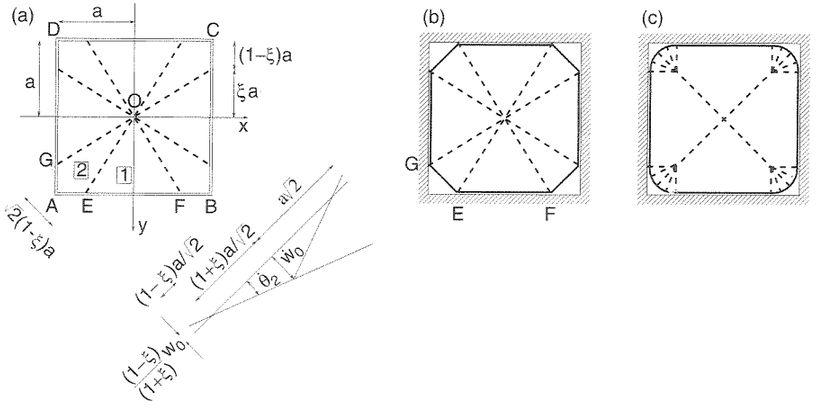


Figure 18.5 Plastic collapse of a square plate: a) yield line pattern for a simply supported plate with lifting corners, b–c) possible yield line patterns for a clamped plate

For any physically meaningful value of ξ, i.e. for any $\xi \in (0,1)$, the corresponding $p_k(\xi)$ is a kinematically admissible load. The best upper bound is obtained for the value of ξ which minimizes function $p_k(\xi)$ on the interval $(0,1)$. The minimum can occur either at points where $dp_k/d\xi = 0$, or at the boundary of the interval. Let us first check the former condition:

$$\frac{dp_k}{d\xi} = \frac{(2\xi)(3\xi-\xi^2)-(1+\xi^2)(3-2\xi)}{(3\xi-\xi^2)^2}\frac{6M_p}{a^2} = 0 \tag{18.40}$$

$$6\xi^2-2\xi^3-3-3\xi^2+2\xi+2\xi^3 = 0 \tag{18.41}$$

$$3\xi^2+2\xi-3 = 0 \tag{18.42}$$

$$\xi = \frac{-2+\sqrt{4+36}}{6} = 0.7208 \tag{18.43}$$

The other root of quadratic equation (18.42) is negative, and thus of no interest. The value of p_k at $\xi = 0$ tends to infinity. This is because if the yield lines in Figure 18.5(a) run parallel to the plate edges, the work of the uniform external load on such a mechanism is zero, and the mechanism ceases to be kinematically admissible. The value of p_k at $\xi = 1$ is $6M_p/a^2$, which corresponds to the basic solution from Example 18.1. For the value of ξ from (18.43) we get the best upper bound for this yield line pattern,

$$p_k = 5.550\frac{M_p}{a^2} \tag{18.44}$$

This is by 7.5% less than the previous result in (18.33). □

If the plate is prevented from lifting up at the simple support, we must consider further hinges along lines EG, etc. The best upper bound for such yield line pattern is obtained for $\xi = 1$, and so there is no improvement over the diagonal yield hinges in Figure 18.4(a). If the plate is clamped, yield hinges must be also considered along lines GE and EF, etc. (Figure 18.5(b)), and the calculations proceed similarly. We could also analyze an alternative failure pattern in Figure 18.5(c).

Example 18.4: Calculate the upper bound on the collapse load for the rectangular plate shown in Figure 18.6.

Solution: Segment $1 \equiv$ CDEF rotates about the x-axis at rate $\dot{\theta}_{1x} = \dot{w}_0/a$ while segment $2 \equiv$ ADE rotates about the y-axis at rate $\dot{\theta}_{2y} = -\dot{w}_0/\xi a$. So, the relative

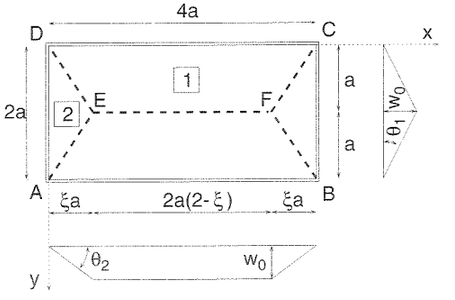


Figure 18.6 Yield line pattern for a rectangular plate

rotation rate between these segments (in yield line DE) is

$$\dot\theta_{DE}=\dot\theta_{12}=\sqrt{\left(\dot\theta_{1x}-0\right)^2+\left(0-\dot\theta_{2y}\right)^2}=\sqrt{\dot\theta_{1x}^2+\dot\theta_{2y}^2}$$
$$=\frac{\dot w_0}{a}\sqrt{1+\frac{1}{\xi^2}}=\frac{\dot w_0}{a}\frac{\sqrt{1+\xi^2}}{\xi} \tag{18.45}$$

We also need the length of the yield line DE, $l_{DE}=\sqrt{a^2+(\xi a)^2}=a\sqrt{1+\xi^2}$. The rotation rate in yield line EF is of course $\dot\theta_{EF}=2\dot\theta_{1x}=2\dot w_0/a$. The dissipation rate may now be calculated as

$$D_{\rm int}=4l_{DE}M_p|\dot\theta_{DE}|+l_{EF}M_p|\dot\theta_{EF}|$$
$$=4a\sqrt{1+\xi^2}M_p\frac{\dot w_0}{a}\frac{\sqrt{1+\xi^2}}{\xi}+2a(2-\xi)M_p\frac{2\dot w_0}{a}$$
$$=\left(\frac{4+4\xi^2}{\xi}+8-4\xi\right)M_p\dot w_0=4\left(2+\frac{1}{\xi}\right)M_p\dot w_0 \tag{18.46}$$

and the rate of the volume under the deflection surface is

$$\dot V=\int_A\dot w\,{\rm d}A=2\times\frac{1}{3}2a\xi a\dot w_0+\frac{1}{2}2a(2-\xi)2a\dot w_0=\left(4-\frac{2}{3}\xi\right)a^2\dot w_0 \tag{18.47}$$

from which

$$\dot W_{\rm ext}=p\dot V=\left(4-\frac{2}{3}\xi\right)pa^2\dot w_0 \tag{18.48}$$

Setting $\dot W_{\rm ext}=D_{\rm int}$, we get for the collapse load

$$p_k=\frac{6M_p}{a^2}\frac{2\xi+1}{6\xi-\xi^2} \tag{18.49}$$

Minimizing this with respect to parameter ξ, we obtain

$$\frac{dp_k}{d\xi}=\frac{6M_p}{a^2}\frac{2(6\xi-\xi^2)-(2\xi+1)(6-2\xi)}{(6\xi-\xi^2)^2}=0\quad\ldots\quad\xi^2+\xi-3=0 \tag{18.50}$$

One of the roots, $\xi=1.303$, is indeed in the interval of interest, $\xi\in(0,2)$. Upon substitution into (18.49) we finally have the upper bound

$$p_k=3.535M_p/a^2 \tag{18.51}$$

□

For a clamped rectangular plate, we must consider further hinges along the supported sides. In the case of a simple support which allows lifting of the corner, we must double the diagonal yield lines in a pattern similar to that in Figure 18.5(a) for the square plate, in order to get a better upper bound; see Problem 18.4.

Example 18.5: Calculate the upper bound on the collapse load for a simply supported regular polygonal plate with n sides inscribed in a circle of radius R.

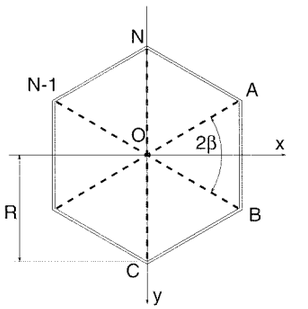


Figure 18.7 Regular polygonal plate

Solution: Assuming the diagonal yield hinge pattern shown in Figure 18.7, and denoting the center deflection as w_0, the rotation rate of segment OAB about the y-axis is $\dot\theta_1=\dot\theta_{1y}=\dot w_0/R\cos\beta$, in which $\beta=\pi/n$. The axis of rotation of the neighboring segment OBC is inclined by angle 2β with respect to the x-axis, and so the Cartesian components of its rotation rate are $\dot\theta_{2x}=-\dot\theta_1\sin2\beta$ and $\dot\theta_{2y}=\dot\theta_1\cos2\beta$. So, the relative rotation rate in the yield line is

$$\dot\theta=\sqrt{\left(\dot\theta_{1x}-\dot\theta_{2x}\right)^2+\left(\dot\theta_{1y}-\dot\theta_{2y}\right)^2}=\dot\theta_1\sqrt{\sin^2 2\beta+(1-\cos2\beta)^2}=\frac{2\dot w_0}{R}\tan\beta \tag{18.52}$$

Now we can express the dissipation rate,

$$D_{\rm int}=nRM_p\dot\theta=2nM_p\dot w_0\tan\beta \tag{18.53}$$

and the external power of applied uniform load p,

$$\dot W_{\rm ext}=pn\frac{1}{3}(R\sin\beta)(R\cos\beta)\dot w_0=\frac{pnR^2\dot w_0}{3}\sin\beta\cos\beta \tag{18.54}$$

Setting $\dot W_{\rm ext}=D_{\rm int}$, we obtain the upper bound

$$p_k=\frac{2nM_p\dot w_0\tan\beta}{(nR^2\dot w_0/3)\sin\beta\cos\beta}=\frac{6M_p}{R^2\cos^2\beta}=\frac{6M_p}{R^2\cos^2(\pi/n)} \tag{18.55}$$

Substituting $n=4$ and $R=a\sqrt{2}$, we obtain an upper bound on the limit load for a square plate, $p_k=6M_p/a^2$, which exactly agrees with (18.33).

Example 18.6: Calculate the upper bound on the collapse load for the circular plate of outer radius R and with a central hole of radius R_0.

Solution: In this case we must assume uniformly distributed yield hinges causing a uniform plastic curvature all over the plate. Denoting the deflection in the center (on the extension of the radial yield hinges) as w_0, the curvature of the plate may be figured out from the cross section sketched in Figure 18.8(a). We consider a circle of radius $r\in(R_0,R)$ and we find a sphere that is touching the deflection surface along the circle. From similarity of the triangles shown in Figure 18.8(b) we have

$$\frac{r}{\rho}=\frac{w_0\dfrac{r}{R}}{r\sqrt{1+\left(\dfrac{w_0}{R}\right)^2}} \tag{18.56}$$

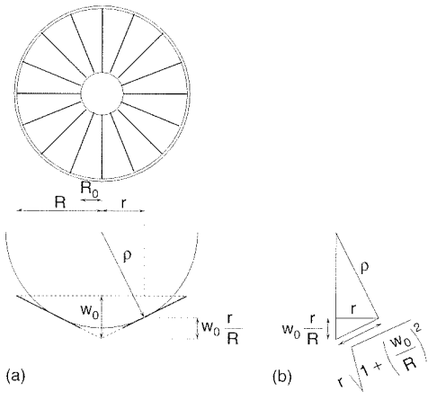


Figure 18.8 Radius of curvature of a conically deflected surface

where ρ is the radius of the tangent sphere. At the onset of collapse, the deflection w_0 is much smaller than the plate radius R. Neglecting terms of order $(w_0/R)^2$ and higher, we obtain the curvature of the sphere, and of the deflection surface as well:

$$\kappa = \frac{1}{\rho} = \frac{w_0}{rR} \tag{18.57}$$

The dissipation power is

$$D_{\text{int}} = \int_0^{2\pi}\int_{R_0}^{R} M_p \dot{\kappa} r \, d\phi \, dr = \int_0^{2\pi}\int_{R_0}^{R} M_p \frac{\dot{w}_0}{rR} r \, d\phi \, dr = 2\pi(R-R_0)M_p\frac{\dot{w}_0}{R} \tag{18.58}$$

and the external power of the applied uniform load p is

$$\dot{W}_{\text{ext}} = p\int_{R_0}^{R}\left(1-\frac{r}{R}\right)\dot{w}_0 2\pi r \, dr = \frac{p\pi\dot{w}_0}{3R}(R-R_0)^2(R+2R_0) \tag{18.59}$$

Setting $\dot{W}_{\text{ext}} = D_{\text{int}}$, we get the upper bound on the collapse load,

$$p_k = \frac{6M_p}{(R-R_0)(R+2R_0)} \tag{18.60}$$

□

For a plate without a hole, substituting $R_0 = 0$ into (18.60) gives

$$p_k = \frac{6M_p}{R^2} \tag{18.61}$$

which can also be obtained as a limit for $n \to \infty$ of the result in (18.55) for the regular polygonal plate.

For a circular plate that is clamped at its boundary, we must also consider a yield hinge along the boundary circle. The result for a plate without a hole is then $p = 12M_p/R^2$.

A conical, uniformly curved deflection surface in the corners of a rectangular plate may sometimes give better upper bounds than the basic yield line pattern. An example is shown in Figure 18.5(c).

18.3 STATIC APPROACH AND STRIP METHOD

Example 18.7: Calculate the lower bound on the collapse load for the simply supported square plate of side $2a$, shown in Figure 18.4(a).

Solution: We seek a moment field which is in equilibrium with the given uniform load p. A simple approach, called the *strip method*, is to imagine load p to be divided into two equal parts, one of which is carried by simply supported beams (or strips) in direction x and the other one by simply supported beams (or strips) in direction y. For these fictitious beams subjected to the load $p/2$ we have the bending moments $M_x = p(a^2-x^2)/4$, $M_y = p(a^2-y^2)/4$. The twisting moments are here assumed to be zero, $M_{xy} = 0$. This assumed moment field satisfies the equation of equilibrium (18.9). It must also satisfy the condition of plastic admissibility which, for the von Mises criterion (18.15), reads

$$M_x^2 + M_y^2 - M_x M_y + 3M_{xy}^2 \le M_0^2 \tag{18.62}$$

Substituting the assumed moment fields, we have

$$\frac{p^2}{16}\left[\left(a^2-x^2\right)^2 + \left(a^2-y^2\right)^2 - \left(a^2-x^2\right)\left(a^2-y^2\right)\right] \le M_0^2 \tag{18.63}$$

This inequality is satisfied for all $(x,y) \in [-a,a]\times[-a,a]$ if it is satisfied for those (x,y) that make the left-hand side maximum. Since the expression in the brackets can be written as $a^4-(a^2-x^2)x^2-(a^2-y^2)y^2-x^2y^2$, the maximum occurs at $x=y=0$. Substituting this in (18.63) and using the equality sign, we obtain the lower bound

$$p_s = \sqrt{16M_0^2/a^4} = 4M_0/a^2 = 3.464M_p/a^2 \tag{18.64}$$

Comparing this with the previous upper bound (18.33), $p_k = 6M_p/a^2$, we see that at least one of these bounds is not very accurate.

Following Prager (1959), we may obtain a better lower bound by assuming the twisting moment to be $M_{xy} = -kxy$, in which k is an undetermined constant. Substituting $\partial^2 M_{xy}/\partial x\partial y = -k$ into the differential equation of equilibrium (18.9) we have

$$-\left(\frac{\partial^2 M_x}{\partial x^2} + \frac{\partial^2 M_y}{\partial y^2}\right) = p - 2k \tag{18.65}$$

This differential equation is the same as in the previous case in which M_{xy} was neglected, except for the fact that p is replaced by $p-2k$. So the distribution of M_x and M_y could be taken as in the previous case, based on the idea of simply supported beams (or strips), except for replacing p by $p-2k$. Substituting the assumed fields $M_x = (p-2k)(a^2-x^2)/4$, $M_y = (p-2k)(a^2-y^2)/4$, and $M_{xy} = -kxy$ into (18.62) we get

$$\frac{(p-2k)^2}{16}\left[\left(a^2-x^2\right)^2 + \left(a^2-y^2\right)^2 - \left(a^2-x^2\right)\left(a^2-y^2\right)\right] + 3k^2x^2y^2 \le M_0^2 \tag{18.66}$$

The most critical situations can be expected at $x=y=0$ (plate center) and at $x=y=a$ (plate corner). So we must satisfy the inequalities

$$\text{for } x=y=0: \quad \frac{(p-2k)^2}{16}a^4 \le M_0^2 \quad \dots \quad p \le \frac{4M_0}{a^2} + 2k \tag{18.67}$$

$$\text{for } x=y=a: \quad 3k^2a^4 \le M_0^2 \quad \dots \quad k \le \frac{M_0}{\sqrt{3}a^2} \tag{18.68}$$

Therefore, the largest value of a statically admissible load that we can hope for with the present assumed moment fields is

$$p = \frac{4M_0}{a^2} + 2\frac{M_0}{\sqrt{3}a^2} = \frac{M_0}{a^2}\left(4 + \frac{2}{\sqrt{3}}\right) \tag{18.69}$$

It remains to be verified that, with $k = M_0/\sqrt{3}a^2$ and $p = (4+2/\sqrt{3})M_0/a^2$, condition (18.66) is satisfied for any x and y between $-a$ and a. This is left to the reader as an easy exercise. The load

$$p_s = \left(4 + \frac{2}{\sqrt{3}}\right)\frac{M_0}{a^2} = 5.155\frac{M_0}{a^2} = 4.464\frac{M_p}{a^2} \tag{18.70}$$

is indeed statically admissible and represents a lower bound on the actual collapse load. This is an improvement over the previous expression (18.64), but still not very close to our previously found upper bound (18.33). □

Example 18.8: Calculate the lower bound on the collapse load for a circular plate.

Solution: To make use of the axial symmetry of the problem, it is convenient to work in polar coordinates (r, ϕ). Bending moments M_r and M_ϕ are then functions of the radial coordinate r only, and the twisting moments $M_{r\phi}$ are zero.

Consider the internal forces acting on a small element of the plate (Figure 18.9(a)). Even though the effect of the shear forces on the deformation of the plate and on the yield condition is neglected, these forces must be taken into account in the equilibrium condition.[1] The shear force Q_r depends only on the radial coordinate r, and the shear force Q_ϕ is zero. Setting up the differential equations of equilibrium, we must multiply each internal force intensity by the width of the elementary area it acts on. Also, we must take into account that the circumferential moments M_ϕ on the opposite faces do not act in parallel planes, and so they contribute to the moment equilibrium condition in the radial plane. The resulting differential equations of equilibrium for the axisymmetric state read

$$\frac{\mathrm{d}}{\mathrm{d}r}(rM_r) = M_\phi - rQ_r \tag{18.71}$$

$$\frac{\mathrm{d}}{\mathrm{d}r}(rQ_r) = pr \tag{18.72}$$

Integration of the second equation under the condition $Q_r = 0$ at $r = 0$ provides $Q_r = pr/2$. Substituting this into the first equation we have

$$\frac{\mathrm{d}}{\mathrm{d}r}(rM_r) = M_\phi - \frac{pr^2}{2} \tag{18.73}$$

[1] The shear forces did not explicitly appear in the equilibrium equation (18.9) written in the Cartesian coordinates. This is because we used the principle of virtual work to derive directly the second-order equation containing only the moments. However, writing the equilibrium conditions of a differential plate element, we would have to take into account the shear forces and we would obtain a set of three first-order differential equations. Elimination of the shear forces would then provide the resulting equation (18.9).

1) First assume that the plate is *simply supported*. The boundary condition is then $M_r = 0$ at $r = R$. Within the plate, we expect $M_\phi > 0, M_r > 0$. Let us try as the solution $M_\phi = \text{const}$; integration of equation (18.73) gives $rM_r = rM_\phi - pr^3/6 + C$, in which $C =$ undetermined constant. So,

$$M_r = M_\phi - \frac{pr^2}{6} + \frac{C}{r} \tag{18.74}$$

Because the value of M_r at $r = 0$ must be finite, we must have $C = 0$. Furthermore, the boundary condition $M_r = 0$ at $r = R$ yields

$$M_\phi = \frac{pR^2}{6} \tag{18.75}$$

and so

$$M_r = \frac{p}{6}\left(R^2 - r^2\right) \tag{18.76}$$

To be statically admissible, the solution must be plastically admissible for every r between 0 and R. For the von Mises condition (18.15) we get

$$M_r^2 + M_\phi^2 - M_r M_\phi \le M_0^2 \tag{18.77}$$

$$\frac{p^2}{36}\left[\left(R^2 - r^2\right)^2 + R^4 - \left(R^2 - r^2\right)R^2\right] \le M_0^2 \tag{18.78}$$

$$\frac{p^2}{36}\left[R^4 - R^2r^2 + r^4\right] \le M_0^2 \tag{18.79}$$

The expression in the brackets is in the interval $(0, R)$ maximized at $r = 0$ (or $r = R$). Substituting this into condition (18.79) taken as an equality, we obtain a statically admissible load

$$p_s = \sqrt{\frac{36M_0^2}{R^4}} = \frac{6M_0}{R^2} = \frac{5.196M_p}{R^2} \tag{18.80}$$

This is a lower bound on the uniform collapse load p_0. It is lower by 13.5% than the previously derived upper bound (18.61). The corresponding distribution of the bending moments along the radius of the plate is shown in Figure 18.9(c). The circumferential moment M_ϕ is constant and equal to the 'beam' yield moment $M_0 = \sigma_0 h^2/4$ while the radial moment M_r varies quadratically between zero at the support and M_0 at the center. This solution at the same time satisfies the Tresca criterion, for which $M_p = M_0$; see (18.25). This means that if the material behaves according to the Tresca yield condition, the lower bound

$$p_s = \frac{6M_0}{R^2} = \frac{6M_p}{R^2} \tag{18.81}$$

is equal to the previously derived upper bound (18.61), and so we have in this case the exact solution.

2) Consider now a circular plate with a *clamped edge*, which is somewhat more complicated. To allow a simple solution, we must now restrict attention to the Tresca yield criterion. Moment M_r must obviously be negative at the boundary and positive at the center. So there must be a circle, of radius $r = a$, at which $M_r = 0$. The

previous solution can be used for $r \in (0, a)$ where both moments are positive. The point $r = a$ is characterized by the condition $M_r = 0$, from which

$$a = \sqrt{\frac{6M_0}{p}} \tag{18.82}$$

On the interval (a, R) the circumferential moment M_ϕ cannot be kept equal to M_0 because the radial moment M_r becomes negative and the stress state would not be plastically admissible. We can expect that the maximum load capacity will be achieved if the moments vary such that the yield condition be satisfied. The corresponding segment of the Tresca yield surface (Figure 18.9(b)) is described by

$$M_\phi - M_r = M_0 \tag{18.83}$$

At the same time, the moments must satisfy the equilibrium condition (18.73). Substituting $M_\phi = M_0 + M_r$, this equation becomes

$$M_r + r\frac{dM_r}{dr} = (M_0 + M_r) - \frac{pr^2}{2} \tag{18.84}$$

or

$$\frac{dM_r}{dr} = \frac{M_0}{r} - \frac{pr}{2} \tag{18.85}$$

Integrating, we have

$$M_r = M_0 \ln\frac{r}{R} - \frac{pr^2}{4} + C \tag{18.86}$$

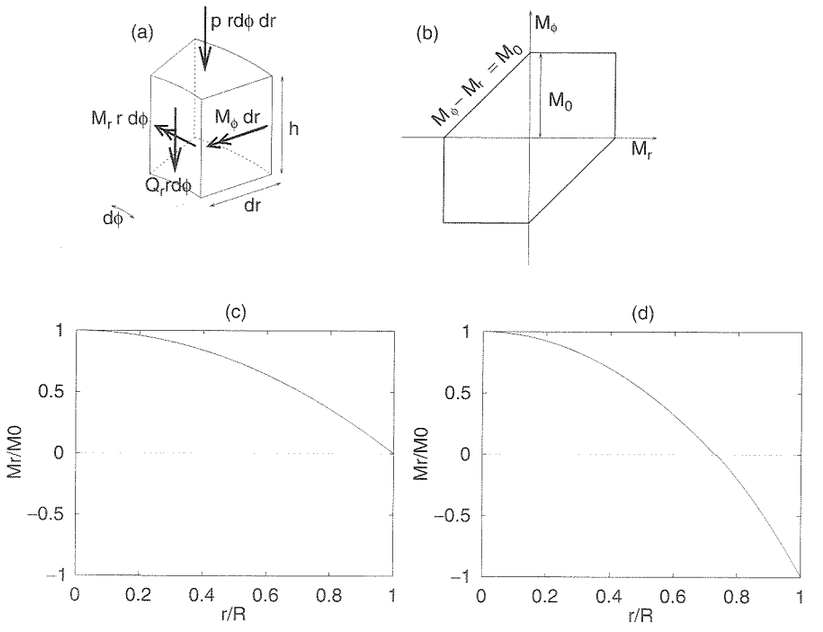


Figure 18.9 Circular plate: a) internal forces on a differential element, b) Tresca yield surface, c) moment distribution for a simply supported plate, d) moment distribution for a clamped plate

Here, the constant C may be expressed in terms of the (yet unknown) load p from the continuity condition, $M_r = 0$ at $r = a$. This yields

$$C = \frac{pa^2}{4} - M_0 \ln\frac{a}{R} = \frac{M_0}{2}\left(3 - \ln\frac{6M_0}{pR^2}\right) \tag{18.87}$$

and

$$M_r = \frac{M_0}{2}\left(2\ln\frac{r}{R} + 3 - \ln\frac{6M_0}{pR^2}\right) - \frac{pr^2}{4} = \frac{M_0}{2}\left(3 + \ln\frac{pr^2}{6M_0}\right) - \frac{pr^2}{4} \tag{18.88}$$

Now we have to look for the largest possible value p for which $M_r(r) \geq -M_0$ on the interval (a, R). As might be expected, the load is maximized if $M_r = -M_0$ at $r = R$, which leads to a nonlinear equation for the load p,

$$\frac{M_0}{2}\left(3 + \ln\frac{pR^2}{6M_0}\right) - \frac{pR^2}{4} + M_0 = 0 \tag{18.89}$$

or

$$\ln\frac{pR^2}{6M_0} - \frac{pR^2}{2M_0} + 5 = 0 \tag{18.90}$$

Numerical solution leads to a lower bound on the uniform collapse load for the clamped plate,

$$p_s = 11.3M_0/R^2 \tag{18.91}$$

Compared to the previously established upper bound $p_k = 12M_0/R^2$, this is relatively close. The corresponding distribution of bending moments is sketched in Figure 18.9(d).

18.4 REINFORCED CONCRETE PLATES

The cross sections of reinforced concrete plates usually fail by plastic extension of the reinforcement (which is known as the under-reinforced situation) rather than by crushing of concrete under compression. This allows us to consider ductile or plastic behavior in bending. Thus, the yield line theory is applicable for calculation of the collapse loads of reinforced concrete plates, even though concrete is not a plastic material. This has been amply confirmed by experiments (e.g. see the book by Park and Gamble, 1980). As a general concept, the yield line approach for reinforced concrete, particularly for circular plates, was proposed by Gvozdev (1938), in a Russian proceedings paper which long remained unknown in the West as well as in Russia.

The reinforcing net is usually orthogonal, characterized by steel ratios p_x and p_y in the directions x and y. Consider a triangular element of the plate, the hypotenuse of which is a yield line (Figure 18.10). To determine the plastic moment in the yield line, we must determine the resultant M of the yield forces of the bars normal to the cross section through the yield line. It is a relatively good approximation to neglect the shear forces in the bars (i.e. to neglect the dowel action and bar kinking), and to assume that the torque in the yield line has no effect. This means that we assume the reinforcing bars to transmit only axial forces. Bars in the x-direction develop a normal force per unit width $N_x = hp_x f_y$, in which h = thickness of the plate and f_y = yield stress of steel. The elementary force acting on the differential element in Figure 18.10 is $N_x dy$.

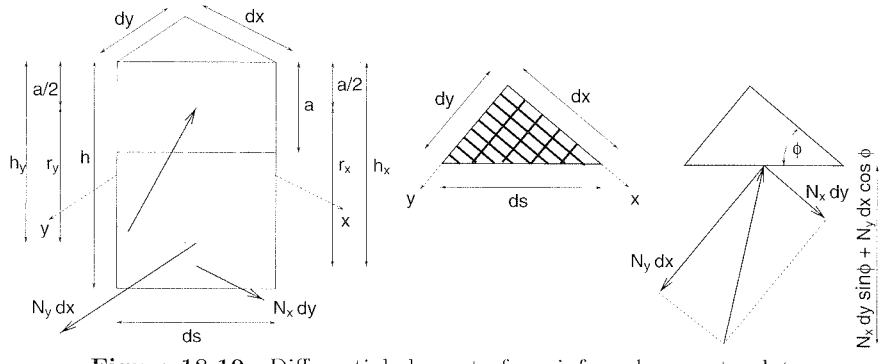


Figure 18.10 Differential element of a reinforced concrete plate

In the inclined yield line, this force contributes to the bending moment by its projection $N_x \mathrm{d}y \sin\phi$, where ϕ is the angle characterizing the direction of the yield line. Similarly, bars in the y-direction develop a normal force per unit width $N_y = hp_y f_y$.

As the total normal force is zero, the tensile forces in steel must be balanced by a compressive force in concrete. If the distance of the compressive resultant from the level of reinforcement in the x- and y-direction is r_x and r_y, respectively, the bending moment in the elementary segment of the yield line is

$$M_p \mathrm{d}s = N_x \mathrm{d}y\, r_x \sin\phi + N_y \mathrm{d}x\, r_y \cos\phi \tag{18.92}$$

Substituting $\mathrm{d}x = \mathrm{d}s \cos\phi$ and $\mathrm{d}y = \mathrm{d}s \sin\phi$, we finally get the yield moment

$$M_p = N_x r_x \sin^2\phi + N_y r_y \cos^2\phi \tag{18.93}$$

The internal lever arms r_x and r_y can be estimated as

$$r_x = h_x - \frac{a}{2},\ r_y = h_y - \frac{a}{2} \tag{18.94}$$

in which h_x and h_y are the distances of the x and y reinforcing bars from the top face of the plate,

$$a = \frac{N_x \sin^2\phi + N_y \cos^2\phi}{0.85 f_c'} = \frac{hf_y}{0.85 f_c'}(p_x \sin^2\phi + p_y \cos^2\phi) \tag{18.95}$$

is the depth of the equivalent rectangular compression stress block in the yield line, and $0.85\ f_c'$ is the effective compression stress according to design codes, f_c' being the standard cylinder strength (for the analysis of cross section, see Section 14.1).

Equation (18.93) can be written as

$$M_p = M_{px} \sin^2\phi + M_{py} \cos^2\phi \tag{18.96}$$

in which we introduce the notation

$$M_{px} = N_x r_x = hp_x f_y r_x, \qquad M_{py} = N_y r_y = hp_y f_y r_y \tag{18.97}$$

Equation (18.96) was proposed by Johansen (1943), with the minor difference that he interpreted M_{px} and M_{py} as the ultimate bending moments of the plate for bending

in the x and y directions, i.e. as quantities independent of the yield line orientation ϕ. In reality, the internal lever arms depend on ϕ through (18.94)–(18.95), and so M_{px} and M_{py} from (18.97) are not constants but functions of ϕ. Both interpretations give almost the same results; see Problem 18.8.

In any case, according to (18.96), the yield moment depends on the orientation of the yield hinge. This introduces a slight complication compared to metals. In the special case where $M_{px} = M_{py}$, which can be assumed when p_x and p_y are approximately the same, we have

$$M_p = M_{px}(\sin^2\phi + \cos^2\phi) = M_{px} \tag{18.98}$$

We see that in the case of equal reinforcement in two orthogonal directions, the yield moment in the yield hinge is independent of the hinge orientation.

Equation (18.96) for the yield moment in the yield hinge can alternatively be derived by writing the moment equilibrium condition about the hypotenuse of the triangular plate element sketched in Figure 18.10. The torsional moments on the sides of this element are neglected.

Example 18.9: Consider again the rectangular simply supported plate from Example 18.4. Find the optimal orthotropic reinforcement, for which the collapse load (determined approximately by the yield line theory) is maximized at given total volume of reinforcement.

Solution: The yield line pattern is sketched in Figure 18.6. The unknown position of point E is characterized by a variable ξ, to be determined by minimizing the value of the kinematically admissible load.

For given plastic moments in the principal directions of orthotropy, M_{px} and M_{py}, the yield moment in line DE is

$$\begin{aligned} M_{pDE} &= M_{px} \sin^2\phi_{DE} + M_{py} \cos^2\phi_{DE} \\ &= M_{px}\left(\frac{a}{l_{DE}}\right)^2 + M_{py}\left(\frac{\xi a}{l_{DE}}\right)^2 = \frac{M_{px} + \xi^2 M_{py}}{1+\xi^2} \end{aligned} \tag{18.99}$$

where $l_{DE} = \sqrt{a^2 + (\xi a)^2} = a\sqrt{1+\xi^2}$ is the length of segment DE and ϕ_{DE} is the angle between the line DE and the x-axis. For yield line EF we have $M_{pEF} = M_{py}$. Expressions for the rotation rates can be taken from Example 18.4, and the dissipation rate can be obtained by easy generalization of (18.46). Setting

$$\begin{aligned} D_{\mathrm{int}} &= 4 l_{DE} M_{pDE} |\dot\theta_{DE}| + l_{EF} M_{pEF} |\dot\theta_{EF}| \\ &= 4a\sqrt{1+\xi^2}\,\frac{M_{px} + \xi^2 M_{py}}{1+\xi^2}\,\frac{\sqrt{1+\xi^2}}{\xi}\,\frac{\dot w_0}{a} + 2a(2-\xi) M_{py}\frac{2\dot w_0}{a} \\ &= \frac{4M_{px} + 4\xi^2 M_{py}}{\xi} + 4(2-\xi) M_{py} = \frac{4}{\xi}(M_{px} + 2\xi M_{py})\dot w_0 \end{aligned} \tag{18.100}$$

equal to the external power (18.48),

$$\dot W_{\mathrm{ext}} = \left(4 - \frac{2}{3}\xi\right) p a^2 \dot w_0 \tag{18.101}$$

we obtain an expression for the kinematically admissible load

$$p_k = \frac{6}{a^2}\frac{M_{px} + 2\xi M_{py}}{\xi(6-\xi)} \tag{18.102}$$

For given M_{px} and M_{py}, the best estimate of the collapse load is obtained by minimizing p_k as a function of ξ. Then we should find the values of M_{px} and M_{py} such that the collapse load be maximized, subject to the condition that the total volume of reinforcement, $V_r = (p_x + p_y)8a^2h$, is constant.

As a first approximation, let us neglect the dependence of the internal lever arms r_x and r_y in (18.97) on the reinforcement ratio and on the direction of the yield line, i.e. let us assume that the plastic moment in each of the basic directions is proportional to the corresponding reinforcement ratio, with the same proportionality factor hf_yr for both basic directions. The condition of constant total volume of reinforcement, $p_x + p_y = V_r/8a^2h$, is then equivalent to the condition that the sum of plastic moments in the basic directions, $M_{px} + M_{py} = V_r f_y r/8a^2$, be constant. Denoting $M_p = V_r f_y r/16a^2$ the plastic moment in the isotropic case, we can characterize a general orthotropic reinforcement by the dimensionless ratio $\mu = M_{px}/M_p$ varying between 0 and 2. Substituting $M_{px} = \mu M_p$ and $M_{py} = (2-\mu)M_p$ into (18.102), the kinematically admissible load can be expressed as a function of two unknown variables,

$$p_k(\xi,\mu) = \frac{6M_p}{a^2}\frac{(1-2\xi)\mu + 4\xi}{\xi(6-\xi)} \tag{18.103}$$

The point at which p_k is minimized with respect to ξ and maximized with respect to μ is a saddle point at which both partial derivatives vanish. We tacitly assume that this point is located inside the admissible domain for ξ and μ. This assumption must later be verified. Conditions $\partial p_k/\partial\xi = 0$ and $\partial p_k/\partial\mu = 0$ lead to a set of equations

$$\mu(\xi^2 - \xi + 3) - 2\xi^2 = 0 \tag{18.104}$$

$$1 - 2\xi = 0 \tag{18.105}$$

These equations have a unique solution, namely $\xi = 0.5$ and $\mu = 2/11 = 0.182$, satisfying the conditions $0 \le \xi \le 2$ and $0 \le \mu \le 2$. This means that the plastic moment in the x-direction, $M_{px} = \mu M_p = 0.182M_p$, should be ten times smaller than the plastic moment in the y-direction, $M_{py} = (2-\mu)M_p = 1.818M_p$. The estimated collapse load for this orthotropic reinforcement is evaluated as

$$p_{k,\min} = p_k(0.5, 0.182) = 4.36\frac{M_p}{a^2} \tag{18.106}$$

The result shows that, for the present problem, the best orthotropic design increases the limit load by as much as 23 % compared to the collapse load for isotropic reinforcement given by (18.51). Of course, this is only an estimate because of our simplifying assumptions. We have based our analysis exclusively on maximization of the limit load. In a real design it is necessary to take into account additional code requirements specifying the minimum allowed reinforcement ratio, maximum allowed deflection, or maximum allowed crack opening. Such design issues are out of the scope of the present book and the reader is referred to specialized publications (e.g. Park and Gamble, 1980).

PROBLEMS

Problem 18.1: Find an upper bound on the collapse load of a square plate of side $2a$ loaded by a concentrated force at its center. Consider a) simply supported edges, b) clamped edges. Use the yield line theory and then check the result by equilibrium considerations.

Problem 18.2: Verify that if the simply supported square plate is prevented from lifting up at the support, the yield line pattern in Figure 18.5(a) with additional hinges along lines EG, etc., does not lead to any improvement over the diagonal yield hinges in Figure 18.4(a).

Problem 18.3: Calculate an upper bound on the collapse load of a clamped square plate using the yield line pattern from Figure 18.5(b)).

Problem 18.4: Find an upper bound on the collapse load of the rectangular plate from Example 18.4, considering that the corners can lift from free supports.

Problem 18.5: Find an upper bound on the collapse load of a rectangular plate of the same shape as in Example 18.4, but with clamped edges.

Problem 18.6: Find an upper bound on the collapse load of a uniformly loaded rectangular plate clamped at three sides and free (unsupported) at the fourth side; see Figure 18.11(a).

Problem 18.7: Construct a lower bound on the uniform collapse load of a simply supported rectangular plate of sides $2a$ and $2b$. Solve the problem first while assuming the twisting moments to vanish, and then try to improve the solution by adding the twisting moment field $M_{xy} = -kxy$.

Problem 18.8: Plot the dependence of the plastic moment M_p in an orthogonally reinforced concrete plate as a function of the orientation of the yield line, ϕ. Consider moments M_{px} and M_{py} from (18.96):

a) as constants representing the ultimate bending moments for bending in the respective directions,

b) as variables given by (18.97), in which the internal lever arms r_x and r_y depend on the angle ϕ according to (18.94)–(18.95).

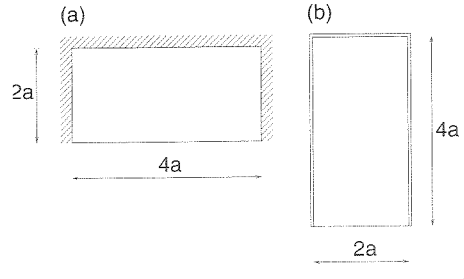


Figure 18.11 Plate to be solved in a) Problem 18.6, b) Problem 18.9

Study the difference between interpretations a) and b) as a function of the ratio p_y/p_x characterizing the deviation from isotropy, using the following values: $h = 200$ mm, $h_x = 170$ mm, $h_y = 150$ mm, $f_y = 280$ MPa, $f'_c = 30$ MPa, $p_x = 1\%$.

Problem 18.9: Find the collapse load for an isotropically reinforced square plate simply supported along three sides; see Figure 18.11(b). Then try to improve the design by using an orthotropic reinforcement, keeping the total volume of reinforcement constant. What is the optimal ratio of plastic moments M_{py}/M_{px}? Explain the physical meaning of the result.

19

Plane Problems

19.1 YIELDING UNDER PLANE STRAIN

Since finite element methods of plastic analysis is easily available today (Belytschko, Liu and Moran, 2000; Zienkiewicz and Taylor, 2000), we will limit our scope to explaining only the basic ideas and simple solutions. Powerful computational tools diminish the value of complicated analytical solutions. Only the simple solutions that provide intuitive insight and can be used as checks on numerical results retain their value.

Consider a plane strain problem described by a displacement field with components $u = u(x, y)$, $v = v(x, y)$ and $w = 0$, from which $\gamma_{xz} = \gamma_{yz} = \varepsilon_z = 0$. Note that the plastic part of the out-of-plane normal strain is not necessarily zero; only its sum with the elastic strain, $\varepsilon_z^e + \varepsilon_z^p = \varepsilon_z$, must vanish. At the onset of yielding at a certain material point, the direction of plastic flow is determined by the current stress state (through the flow rule) and, in general, it has a nonzero component $\dot{\varepsilon}_z^p$. In the plastic limit state, however, the elastic strain rate is zero, and so the plastic strain rate $\dot{\varepsilon}_z^p = \dot{\varepsilon}_z - \dot{\varepsilon}_z^e$ must also vanish. This means that, in the plastic limit state, the stress states in the plastic region are confined to that part of the yield surface on which the induced plastic flow satisfies the condition $\dot{\varepsilon}_z^p = 0$ (and of course also $\dot{\gamma}_{xz}^p = 0$ and $\dot{\gamma}_{yz}^p = 0$).

For the von Mises yield condition with an associated flow rule, the plastic strain rates are proportional to the components of the stress deviator. In the plastic limit state we must have $s_z = 0$, i.e. $\sigma_z = \sigma_V$. Since $\sigma_V = (\sigma_x + \sigma_y + \sigma_z)/3$ we obtain

$$\sigma_z = \frac{1}{2}(\sigma_x + \sigma_y) \tag{19.1}$$

$$s_x = \sigma_x - \sigma_V = \sigma_x - \sigma_z = \frac{1}{2}(\sigma_x - \sigma_y) \tag{19.2}$$

$$s_y = \sigma_y - \sigma_V = \sigma_y - \sigma_z = \frac{1}{2}(\sigma_y - \sigma_x) = -s_x \tag{19.3}$$

Thus the nonzero strain rates are

$$\dot{\varepsilon}_x = \dot{\varepsilon}_x^p = \dot{\lambda} s_x \tag{19.4}$$

$$\dot{\varepsilon}_y = \dot{\varepsilon}_y^p = \dot{\lambda} s_y = -\dot{\lambda} s_x = -\dot{\varepsilon}_x \tag{19.5}$$

$$\dot{\gamma}_{xy} = \dot{\gamma}_{xy}^p = 2\dot{\varepsilon}_{xy} = 2\dot{\lambda} s_{xy} = 2\dot{\lambda}\tau_{xy} \tag{19.6}$$

As expected, the flow is isochoric, i.e. $\dot{\varepsilon}_V = (\dot{\varepsilon}_x + \dot{\varepsilon}_y + \dot{\varepsilon}_z)/3 = 0$. The out-of-plane axis z is always one of the principal axes, say x_3, and so $\sigma_3 \equiv \sigma_z$. The in-plane axes

x and y are not necessarily the principal ones, but since the trace of the stress tensor is an invariant, it always holds $\sigma_1 + \sigma_2 = \sigma_x + \sigma_y$. In terms of principal stresses, relation (19.1) can be rewritten as $\sigma_3 = (\sigma_1 + \sigma_2)/2$. This condition characterizes the points on the yield surface at which the plastic flow direction has a zero component $\dot{\varepsilon}_3^p$. In the Haigh–Westergaard space, such points are located on two meridians of the von Mises cylinder. Meridians are in general intersections of the yield surface with the halfplanes starting from the hydrostatic axis. For the von Mises condition, the meridians are straight lines parallel to the hydrostatic axis. The intersection of the von Mises cylinder with a generic plane perpendicular to the σ_3-axis is shown in Figure 19.1(a). Since the volumetric part of the stress tensor remains unspecified, the axes are labeled by principal deviatoric stresses s_1 and s_2.

For the Tresca yield condition, $\dot{\varepsilon}_z^p$ vanishes if $\sigma_z \equiv \sigma_3$ is the intermediate principal stress, i.e. if it is between the in-plane principal stresses σ_1 and σ_2. The corresponding point in the principal stress space can be anywhere on the inclined segments of the Tresca hexagon; see Figure 19.1(b).

For both yield conditions, the plastic flow has the character of a shear flow on surfaces inclined by 45° with respect to the in-plane principal directions. These are the planes of maximum shear stress, $\tau_{\max} = |\sigma_1 - \sigma_2|/2$, which must be equal to the yield stress in shear, τ_0. For the Tresca condition the out-of-plane normal stress σ_z is undetermined. For simplicity, we may set $\sigma_z = \sigma_V = (\sigma_x + \sigma_y)/2$, as in the case of the von Mises yield condition. Thus, for the present case of plastic flow in the plastic limit state under plane strain, the von Mises condition gives the same result as the Tresca condition if we match the yield stresses in shear (instead of the uniaxial yield stresses); see Figure 19.1(c). The dissipation rate per unit volume can be expressed as

$$\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p, \quad \sigma_1\dot{\varepsilon}_1^p + \sigma_2\dot{\varepsilon}_2^p + \sigma_3\dot{\varepsilon}_3^p = (\sigma_1 - \sigma_2)\dot{\varepsilon}_1^p = 2\tau_0|\dot{\varepsilon}_1^p| \tag{19.7}$$

Alternatively, we can work in the coordinate system (ξ, η) rotated by 45° with respect to the principal one. Here, the normal stresses σ_ξ and σ_η are equal (this follows from the Mohr circle of stresses), the normal components of deviatoric stress s_ξ and s_η are zero, and the normal components of the plastic strain rate are also zero (this follows from the Mohr circle for the plastic strain rate tensor, or from equations (19.2)–(19.5)). The shear component $\dot{\gamma}_{\xi\eta}^p \quad 2\dot{\varepsilon}_{\xi\eta}^p$ is the largest among all possible orientations of the coordinate system and is equal to $\dot{\gamma}_{\max}^p = |\dot{\varepsilon}_1^p - \dot{\varepsilon}_2^p| = 2|\dot{\varepsilon}_1^p|$. Since the shear stress $\tau_{\xi\eta}$ is also the largest among all possible orientations and is equal in magnitude to the shear yield stress τ_0, formula (19.7) could alternatively be derived by writing

$$\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p = \sigma_\xi\dot{\varepsilon}_\xi^p + \sigma_\eta\dot{\varepsilon}_\eta^p + \tau_{\xi\eta}\dot{\gamma}_{\xi\eta}^p + \sigma_z\dot{\varepsilon}_z^p = \tau_{\xi\eta}\dot{\gamma}_{\xi\eta}^p = \tau_0\dot{\gamma}_{\max}^p = 2\tau_0|\dot{\varepsilon}_1^p| \tag{19.8}$$

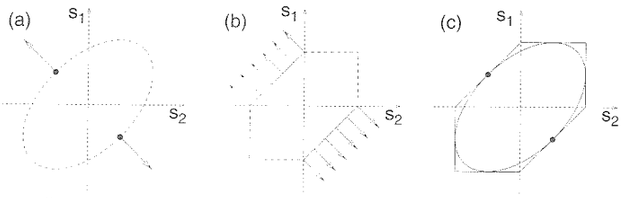


Figure 19.1 Intersection of the yield surface with a plane on which $\sigma_3 = \text{const.}$ for a) von Mises criterion, b) Tresca criterion. Condition $\dot{\varepsilon}_3^p = 0$ is satisfied by the two points marked by solid circles in a), and by all the states on the inclined segments in b). Part c) shows the von Mises and Tresca yield surfaces with equal yield stresses in shear

Let us emphasize that these formulae are valid only under the assumptions mentioned before – von Mises or Tresca yield criterion, associated flow, plane strain conditions, and yielding at constant stress (which occurs in the plastic limit state).

19.2 KINEMATIC APPROACH

19.2.1 Indentation Problem

Let us consider the plane strain indentation problem, in which a rigid block is pushed by force F into an elastic-perfectly plastic half space (Figure 19.2(a)). All calculations will be done per slice of unit thickness in the out-of-plane direction. First we discuss the upper bound estimates for the collapse load F. For this purpose we must select a suitable kinematically admissible velocity field and calculate the dissipation power in the half space for this field. Good upper bounds may often be obtained with velocity fields corresponding to a movement of rigid blocks with shear sliding between the blocks. For such assumed mechanisms, all deformation is concentrated into narrow bands of intense shear, inclined by 45° with respect to the in-plane principal directions. Let us introduce a local coordinate system with axes ξ and η, normal and tangential to the shear band, respectively; see Figure 19.2(b). The orientation of the local axes can always be chosen such that the shear stress $\tau_{\xi\eta}$ is positive inside the band. During plastic flow, $\tau_{\xi\eta}$ inside the band is equal to the shear yield stress, τ_0. The variation of the (plastic) shear strain rate across the band is described by a function $\dot{\gamma}_{\xi\eta}(\xi)$ equal to zero outside the band. In view of (19.8), the rate of energy dissipation per unit length of the band (and per unit slice in the out-of-plane direction) can be evaluated as

$$\bar{D}_{\text{int}} = \int_{-h/2}^{h/2} \tau_{\xi\eta}\dot{\gamma}_{\xi\eta}(\xi)\,\mathrm{d}\xi = \tau_0 \int_{-h/2}^{h/2} \dot{\gamma}_{\xi\eta}(\xi)\,\mathrm{d}\xi = \tau_0|\dot{u}_t| \tag{19.9}$$

where $\dot{u}_t$ is the tangential velocity difference across the shear band between the two adjacent rigid blocks. Note that the normal velocity difference (obtained by integrating $\dot{\varepsilon}_\xi$ across the band) vanishes, i.e. the vector of displacement jump across the shear band (relative velocity vector) must be parallel to the band. This is the compatibility condition that must be satisfied by any kinematically admissible mechanism.

Example 19.1: Estimate the limit indentation force using a very simple collapse mode – a cylindrical block rotating about the edge of the rigid punch; see Figure 19.3(a).

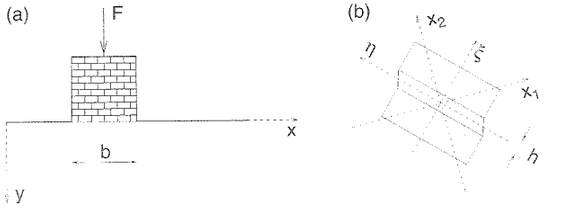


Figure 19.2 a) Indentation problem, b) shear band with local coordinate system (ξ, η) and principal coordinate system (x_1, x_2)

Solution: If the cylindrical block rotates at angular velocity $\dot{\theta}$, the tangential velocity difference across the shear band is $\dot{u}_t = \dot{\theta}b$. Setting the energy dissipation rate along the entire length of the shear band, i.e.

$$D_{\text{int}} = \pi b \bar{D}_{\text{int}} = \pi b \tau_0 |\dot{u}_t| = \pi b \tau_0 \dot{\theta} b = \pi b^2 \tau_0 \dot{\theta} \tag{19.10}$$

equal to the external power supplied by the load F,

$$\dot{W}_{\text{ext}} = F \dot{\theta} b/2 \tag{19.11}$$

we obtain an upper bound estimate of the collapse load,

$$F_k = 2\pi b \tau_0 = 6.28 b \tau_0 \tag{19.12}$$

Alternatively, this result could be obtained from the condition of moment equilibrium about the axis of rotation.

Example 19.2: Try to improve the estimate from the previous example by considering the mechanism depicted in Figure 19.3(b).

Solution: Let us denote the vertical velocity of the rigid punch as $\dot{v}_1$. We must now figure out the corresponding velocities of the triangular blocks shown in Figure 19.3(b); for the sake of simplicity we assume the triangles to be equilateral. The velocities of the blocks may be obtained by constructing the so-called hodograph of motion of this rigid body system; see Figure 19.3(c). In constructing the hodograph, we start by plotting the vector $\dot{\boldsymbol{u}}_1 = \{0, \dot{v}_1\}^T$. Then, starting from the end point of $\dot{\boldsymbol{u}}_1$, we plot the direction of the relative velocity vector between blocks 1 and 2; this vector, denoted as $\dot{\boldsymbol{u}}_{12}$, must have the direction of the boundary line between triangles 1 and 2. The velocity vector of block 2 must be horizontal, because the shear band separating this block from the non-yielding part of the half space is horizontal. Realizing that $\dot{\boldsymbol{u}}_2 = \dot{\boldsymbol{u}}_1 + \dot{\boldsymbol{u}}_{12}$, we see the end point of vector $\dot{\boldsymbol{u}}_{12}$ lies on a horizontal line passing through the origin of vector $\dot{\boldsymbol{u}}_1$. Consequently, the magnitude of velocity $\dot{\boldsymbol{u}}_2$ is $\|\dot{\boldsymbol{u}}_2\| = \dot{v}_1/\sqrt{3}$ and the magnitude of $\dot{\boldsymbol{u}}_{12}$ is $\|\dot{\boldsymbol{u}}_{12}\| = \dot{v}_1 2/\sqrt{3}$. Similarly, denoting the relative velocity between blocks 2 and 3 as $\dot{\boldsymbol{u}}_{23}$, the velocity of the block 3 is $\dot{\boldsymbol{u}}_3 = \dot{\boldsymbol{u}}_2 + \dot{\boldsymbol{u}}_{23}$. The relative velocity vector $\dot{\boldsymbol{u}}_{23}$ must have the direction of the boundary line between blocks 2 and 3. Plotting the lines of vector $\dot{\boldsymbol{u}}_{23}$ and block velocity $\dot{\boldsymbol{u}}_3$ in the hodograph in Figure 19.3(c), we again obtain the magnitude of these vectors $\|\dot{\boldsymbol{u}}_{23}\| = \dot{v}_1/\sqrt{3}$ and $\|\dot{\boldsymbol{u}}_3\| = \dot{v}_1/\sqrt{3}$. Therefore, the energy dissipation in the shear bands between the triangular blocks and between the blocks and the rest

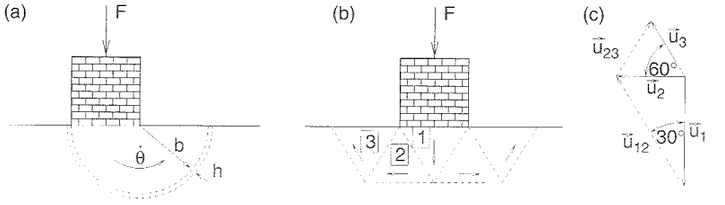


Figure 19.3 Indentation problem: a, b) admissible mechanisms, c) hodograph of motion

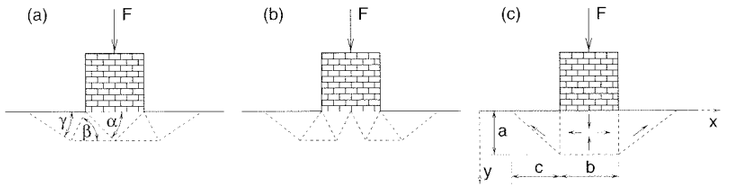


Figure 19.4 Admissible mechanisms for the indentation problem

of the half space is

$$\begin{aligned} D_{\text{int}} &= \tau_0 \left(2b\|\dot{\boldsymbol{u}}_{12}\| + 2b\|\dot{\boldsymbol{u}}_2\| + 2b\|\dot{\boldsymbol{u}}_{23}\| + 2b\|\dot{\boldsymbol{u}}_3\|\right) \\ &= 2\tau_0 b \dot{v}_1 \left(\frac{2}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}}\right) = \frac{10}{\sqrt{3}} \tau_0 b \dot{v}_1 \end{aligned} \tag{19.13}$$

The external power of load F is $\dot{W}_{\text{ext}} = F\dot{v}_1$. Setting $\dot{W}_{\text{ext}} = D_{\text{int}}$ we obtain

$$F_k = \frac{10}{\sqrt{3}} b \tau_0 = 5.76 b \tau_0 \tag{19.14}$$

□

More generally, one may consider triangular blocks of arbitrary angles α, β, λ as shown in Figure 19.4(a); see Problem 19.2. The energy dissipation equality then yields the collapse load as a function of the angles α, β, λ, and the best upper bound may be found by minimizing this function with regard to α, β, λ. An alternative collapse mode also consisting of triangular blocks is shown in Figure 19.4(b).

As a collapse mode of a different type, consider two rigid triangular blocks and a rectangular region of homogeneous plastic deformation rate; see Figure 19.4(c). The energy dissipation rate within the narrow shear bands at the boundaries of the rectangular region and of the triangular regions is calculated from (19.9), taking into account that the relative slip rate on the boundaries of the rectangular region is not constant along the length of the shear band. Sliding also takes place on the surface of the indenter. Assuming a frictionless sliding, i.e. neglecting the energy dissipated at the interface between the indenter and the half space, we remain on the safe side. The contribution of plastic flow in the rectangular region must be added to the rate of dissipation in the shear bands. The rate of vertical normal strain in the rectangular region is $\dot{\varepsilon}_y = -\dot{v}_1/a$, in which $\dot{v}_1$ is the downward velocity of the rigid indentor and a is the depth of the rectangular block. Due to symmetry, $\dot{\varepsilon}_y$ is one of the principal strain rates (the other one is $\dot{\varepsilon}_x = -\varepsilon_y$) and, according to (19.7), the dissipation rate per unit volume is $\mathcal{D} = 2\tau_0 |\dot{\varepsilon}_y| = 2\tau_0 \dot{v}_1/a$. Multiplying this by ab we obtain the total dissipation rate in the rectangular region (per unit slice),

$$D_{\text{int}} = \mathcal{D} ab = 2\tau_0 b \dot{v}_1 \tag{19.15}$$

Adding to (19.15) the rate of energy dissipation in the shear bands, the kinematically admissible load is obtained as $\tau_0 b$ times a function of the ratios a/b and c/a. Minimizing with regard to these ratios, one obtains the best upper bound for this mechanism, which turns out to be worse than that obtained in Example 19.2; see Problem 19.3.

19.2.2 Applications to Soil Mechanics

The problem of failure of a foundation is similar to the indentation problem we just discussed. There are, however, some important differences which complicate the solution. One is that in the failure of a foundation one must usually take the friction of the soil into account. Also, inelastic shear is accompanied by an inelastic normal strain, either dilatancy or densification. The coefficient of dilatancy is usually different from the coefficient of friction, which causes violation of the normality rule and, consequently, the upper and lower bound theorems cease to be valid. It is possible to derive certain bounds on the limit load even for the nonassociated case, but the best lower bound is not equal to the best upper bound and they are usually quite far apart; so the practical value of such estimates is rather limited. Finally, another complication is the fact that, in the analysis of failure of foundations of large structures, one must usually take the own weight of the material into account. For a detailed exposition of plasticity problems of soil mechanics, see the book by Salençon (1977).

19.2.3 Applications to Reinforced Concrete Structures

The theory of plasticity has also been extensively applied to the prediction of limit loads of *reinforced concrete structures* in two and three dimensions. Compared to the yield line theory for plates, such applications of plasticity rest, however, on a much shakier foundation. While the collapse of plates usually consists strictly in the plastic extension of the reinforcing bars, the collapse of two and three-dimensional bodies often involves inelastic deformation of concrete, which is certainly not plastic. Nevertheless, after calibrating the material coefficients according to the measurements of structures rather than of small laboratory specimens, it is possible to obtain some reasonable formulae for the collapse loads in certain problems; e.g. see Braestrup (1974), Nielsen and Braestrup (1978), Marti (1979) and Chen (1982).

The greatest shortcomings of these formulae is that they cannot capture the deterministic size effect, inherent to quasibrittle collapse with a large fracture process zone. Therefore, the plasticity formulae for failures of reinforced concrete structures not caused exclusively by the yielding of steel should not be used for any larger structures unless at least some empirical correction for size effect is introduced; see Bažant and Planas (1998).

Recently, it has become fashionable to predict all kinds of failure of reinforced concrete structures by the truss analogy, also called the *strut-and-tie model* (Collins, 1978; Vecchio and Collins, 1986; Schlaich, Schafer and Jannewein, 1987; Collins and Mitchell, 1991; CEB-FIP, 1991), first proposed by Ritter (1899) and Mörsch (1920, 1922) for shear failure of reinforced concrete beams, and developed by Wagner (1929). When these model predict failure by the yielding steel, their use is justified. Often, however, they are used to predict failures in which the so-called 'compression struts' of concrete are crushed. The fact that compression crushing of concrete is brittle (no yield plateau) and introduces a size effect is ignored by the CEB Code and usually in practice (ACI does not sanction this approach). Revisions are in order, and only applications to small structures, not afflicted by size effect, are justified.

The limitations of the strut-and-tie model, especially the lack of size effect and the inherent hypothesis of simultaneous (non-propagating) failure in all parts of the structure, can be overcome in an approximate sense by the recently proposed fracturing truss model (Bažant, 1997; Bažant and Becq-Giraudon, 2000). In that

model, the failure of the compression struts is considered as progressive, due to propagation of a localized softening band of shear-compression crushing across the strut. The equilibrium conditions in the statically determinate truss are used to calculate the stored complementary energy or potential energy (rather than directly the collapse load), and differentiation with respect to the length of the shear-compression band yields the energy release rate, which must be equal to the energy supply rate required to produce the cracking in the band. This condition yields the maximum load of the structure, and the size effect is a natural outcome of this approach.

19.3 APPLICATION OF LINEAR PROGRAMMING

The upper bound estimate constructed by the kinematic approach can be improved by dividing the region where plastic flow potentially takes place into a large number of small triangular blocks. The increased size of the problem then calls for an automatic procedure. Suppose that the discretization into N triangles has been selected. Any assumed failure mechanism can be described by the displacement rates $\dot{\boldsymbol{u}}_i = (\dot{u}_i, \dot{v}_i)$, $i = 1, 2, \ldots N$, characterizing the motion of individual blocks. In general, the blocks might also experience rotation, but this can be excluded for the type of problems that we have in mind here. In this section it is assumed that all plastic deformation is concentrated into shear bands and that no gaps or overlaps can form between blocks. The sliding blocks are usually surrounded by a region that does not move during plastic collapse. Taking into account that the relative rotation between two blocks separated by a potential shear band (or between a block and the immobile region) is zero, it is often possible to prove that none of the blocks can rotate.

A mechanism is kinematically admissible if the neighboring blocks slide with respect to each other with zero normal relative velocity (here we neglect the dilatancy associated with friction, which is often inadmissible in soil mechanics). The normal and tangential components of the relative velocity between blocks i and j separated by segment k are uniquely determined by the absolute velocities $\dot{u}_i$, $\dot{v}_i$, $\dot{u}_j$ and $\dot{v}_j$; they can be expressed as

$$\dot{e}_{nk} = (\dot{u}_j - \dot{u}_i)\cos\alpha + (\dot{v}_j - \dot{v}_i)\sin\alpha \tag{19.16}$$

$$\dot{e}_{tk} = -(\dot{u}_j - \dot{u}_i)\sin\alpha + (\dot{v}_j - \dot{v}_i)\cos\alpha \tag{19.17}$$

where α is the angle between the normal to the shear band and the x-axis (see Figure 19.1), and subscripts k label the segments whose total number is denoted as M. Collecting the kinematic equations for the entire model we obtain

$$\dot{\boldsymbol{e}}_n = \boldsymbol{B}_n \dot{\boldsymbol{d}} \tag{19.18}$$

$$\dot{\boldsymbol{e}}_t = \boldsymbol{B}_t \dot{\boldsymbol{d}} \tag{19.19}$$

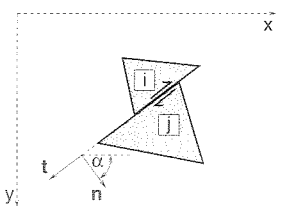


Figure 19.5 Two blocks separated by a shear band

where column matrices $\boldsymbol{e}_n$ and $\boldsymbol{e}_t$ contain the normal and tangential relative displacements, respectively, column matrix $\boldsymbol{d}$ contains the displacement components, and $\boldsymbol{B}_n$ and $\boldsymbol{B}_t$ are the kinematic matrices. The external power is given by

$$\dot{W}_{\text{ext}} = \mu \bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} \tag{19.20}$$

where $\bar{\boldsymbol{f}}$ is a column matrix containing the reference values of external forces work-conjugate to displacements $\boldsymbol{d}$, and μ is the load multiplier. The dissipation rate is

$$D_{\text{int}} = \boldsymbol{s}_0^T |\dot{\boldsymbol{e}}_t| \tag{19.21}$$

where each entry of column matrix $\boldsymbol{s}_0$ is a product of the shear yield stress, τ_0, and the length of the kth shear band segment, ℓ_k, $k = 1, 2, \ldots M$.

The resulting problem has the same format as in the frame analysis with standard plastic hinges.[1] It can be cast as a problem of linear programming

$$\text{Minimize } \mu_k(\dot{\boldsymbol{d}}, \dot{\boldsymbol{e}}_t^+, \dot{\boldsymbol{e}}_t^-) \equiv \boldsymbol{s}_0^T \dot{\boldsymbol{e}}_t^+ + \boldsymbol{s}_0^T \dot{\boldsymbol{e}}_t^- \tag{19.22}$$

subject to

$$\boldsymbol{B}_n \dot{\boldsymbol{d}} = \boldsymbol{0} \tag{19.23}$$

$$\dot{\boldsymbol{e}}_t^+ - \dot{\boldsymbol{e}}_t^- - \boldsymbol{B}_t \dot{\boldsymbol{d}} = \boldsymbol{0} \tag{19.24}$$

$$\bar{\boldsymbol{f}}^T \dot{\boldsymbol{d}} = 1 \tag{19.25}$$

$$\dot{\boldsymbol{e}}_t^+ \geq \boldsymbol{0}, \quad \dot{\boldsymbol{e}}_t^- \geq \boldsymbol{0} \tag{19.26}$$

Conversion to the standard form and solution by the simplex method are completely analogous to the procedures presented in Chapter 7.

Example 19.3 Set up and solve the linear programming problem resulting from the triangularization in Figure 19.6(a).

Solution: The assumed mechanisms consists of four triangular blocks, separated from each other and from the rest of the foundation by seven potential shear line segments, numbered according to Figure 19.6(b). Using (19.16) we construct the conditions of zero normal relative velocity presented in Table 19.1(a), and (19.17) then leads to the expressions for the tangential relative velocities shown in Table 19.1(b).

After the elimination of eight displacement rates from 14 kinematic equations in Table 19.1 we obtain $14 - 8 = 6$ compatibility conditions for the relative tangential velocities $\dot{e}_{ti}$, $i = 1, 2, \ldots 7$. These constraints must be supplemented by the condition of unit external power, $\dot{v}_1 = 1$, also expressed in terms of the relative tangential velocities. Each of the tangential velocities must be decomposed into the positive and negative part. The coefficients in the objective function reflect the lengths of individual shear band segments. The resulting initial

[1] In Chapter 7 we have set up problems (7.7)–(7.11) for trusses. For frames with standard yield hinges (experiencing only plastic rotation but no plastic extension) the kinematic equations have a structure similar to (19.23)–(19.24). The first set of equations expresses inextensibility of individual members while the second relates the rotation rates in plastic hinges to the displacement and rotation rates at joints.

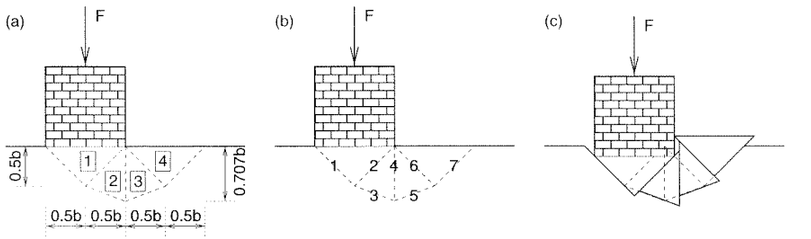


Figure 19.6 Assumed mechanism consisting of four triangular blocks: a) block numbers, b) shear line numbers, c) displaced blocks

Table 19.1 a) Conditions of zero normal relative velocity, b) expressions for tangential relative velocities

(a)

$\dot{u}_1$	$\dot{v}_1$	$\dot{u}_2$	$\dot{v}_2$	$\dot{u}_3$	$\dot{v}_3$	$\dot{u}_4$	$\dot{v}_4$	
0.707	−0.707	0.000	0.000	0.000	0.000	0.000	0.000	0
−0.707	−0.707	0.707	0.707	0.000	0.000	0.000	0.000	0
0.000	0.000	0.371	−0.928	0.000	0.000	0.000	0.000	0
0.000	0.000	−1.000	0.000	1.000	0.000	0.000	0.000	0
0.000	0.000	0.000	0.000	−0.371	−0.928	0.000	0.000	0
0.000	0.000	0.000	0.000	−0.707	0.707	0.707	−0.707	0
0.000	0.000	0.000	0.000	0.000	0.000	−0.707	−0.707	0

(b)

$\dot{u}_1$	$\dot{v}_1$	$\dot{u}_2$	$\dot{v}_2$	$\dot{u}_3$	$\dot{v}_3$	$\dot{u}_4$	$\dot{v}_4$	
0.707	0.707	0.000	0.000	0.000	0.000	0.000	0.000	$\dot{e}_{t1}$
0.707	−0.707	−0.707	0.707	0.000	0.000	0.000	0.000	$\dot{e}_{t2}$
0.000	0.000	0.928	0.371	0.000	0.000	0.000	0.000	$\dot{e}_{t3}$
0.000	0.000	0.000	−1.000	0.000	1.000	0.000	0.000	$\dot{e}_{t4}$
0.000	0.000	0.000	0.000	0.928	−0.371	0.000	0.000	$\dot{e}_{t5}$
0.000	0.000	0.000	0.000	−0.707	−0.707	0.707	0.707	$\dot{e}_{t6}$
0.000	0.000	0.000	0.000	0.000	0.000	0.707	−0.707	$\dot{e}_{t7}$

simplex tableau is shown in Table 19.2. The simplex algorithm leads to the optimal solution in Table 19.3. As might have been expected, all the potential shear bands are activated; see Figure 19.6(c). The kinematically admissible load

$$F_k = 5.314 b \tau_0 \tag{19.27}$$

is the best upper bound on the collapse load obtained so far.

19.4 STATIC APPROACH

Let us now demonstrate the determination of lower bounds on the limit indentation load acting on an elastic-perfectly plastic half space. Here we need to find a plastically admissible and equilibrated stress field satisfying the traction boundary conditions. It is convenient to assume that this field consists of regions under uniform stress.

Table 19.2 Initial simplex tableau

$\dot{e}_{t1}^{+}$	$\dot{e}_{t2}^{+}$	$\dot{e}_{t3}^{+}$	$\dot{e}_{t4}^{+}$	$\dot{e}_{t5}^{+}$	$\dot{e}_{t6}^{+}$	$\dot{e}_{t7}^{+}$	$\dot{e}_{t1}^{-}$	$\dot{e}_{t2}^{-}$	$\dot{e}_{t3}^{-}$	$\dot{e}_{t4}^{-}$	$\dot{e}_{t5}^{-}$	$\dot{e}_{t6}^{-}$	$\dot{e}_{t7}^{-}$	RHS
0.000	1.000	0.394	0.000	0.000	0.000	0.000	0.000	−1.000	−0.394	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.690	0.928	0.000	0.000	0.000	0.000	0.000	−0.690	−0.928	0.000	0.000	0.000	0.000
0.000	0.000	0.394	−0.707	0.000	0.000	−1.000	0.000	0.000	−0.394	0.707	0.000	0.000	1.000	0.000
0.000	0.000	−0.724	0.371	1.000	0.000	0.000	0.000	0.000	0.724	−0.371	−1.000	0.000	0.000	0.000
0.000	0.000	0.919	0.707	0.000	1.000	0.000	0.000	0.000	−0.919	−0.707	0.000	−1.000	0.000	0.000
1.000	0.000	−0.919	0.000	0.000	0.000	0.000	−1.000	0.000	0.919	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.650	0.000	0.000	0.000	0.000	0.000	0.000	−0.650	0.000	0.000	0.000	0.000	1.000
0.707	0.707	0.539	0.700	0.539	0.707	0.707	0.707	0.707	0.539	0.700	0.539	0.707	0.707	0.000

Table 19.3 Simplex tableau with the optimal solution

$\dot{e}_{t1}^{+}$	$\dot{e}_{t2}^{+}$	$\dot{e}_{t3}^{+}$	$\dot{e}_{t4}^{+}$	$\dot{e}_{t5}^{+}$	$\dot{e}_{t6}^{+}$	$\dot{e}_{t7}^{+}$	$\dot{e}_{t1}^{-}$	$\dot{e}_{t2}^{-}$	$\dot{e}_{t3}^{-}$	$\dot{e}_{t4}^{-}$	$\dot{e}_{t5}^{-}$	$\dot{e}_{t6}^{-}$	$\dot{e}_{t7}^{-}$	RHS
0.000	−1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.606
0.000	0.000	0.000	−1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	1.143
0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	−1.000	1.414
0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	−1.000	0.000	0.000	0.000	0.000	1.539
0.000	0.000	0.000	0.000	0.000	−1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.606
1.000	0.000	0.000	0.000	0.000	0.000	0.000	−1.000	0.000	0.000	0.000	0.000	0.000	0.000	1.414
0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	−1.000	0.000	0.000	1.539
0.000	1.414	0.000	1.400	0.000	1.414	0.000	1.414	0.000	1.077	0.000	1.077	0.000	1.414	−5.314

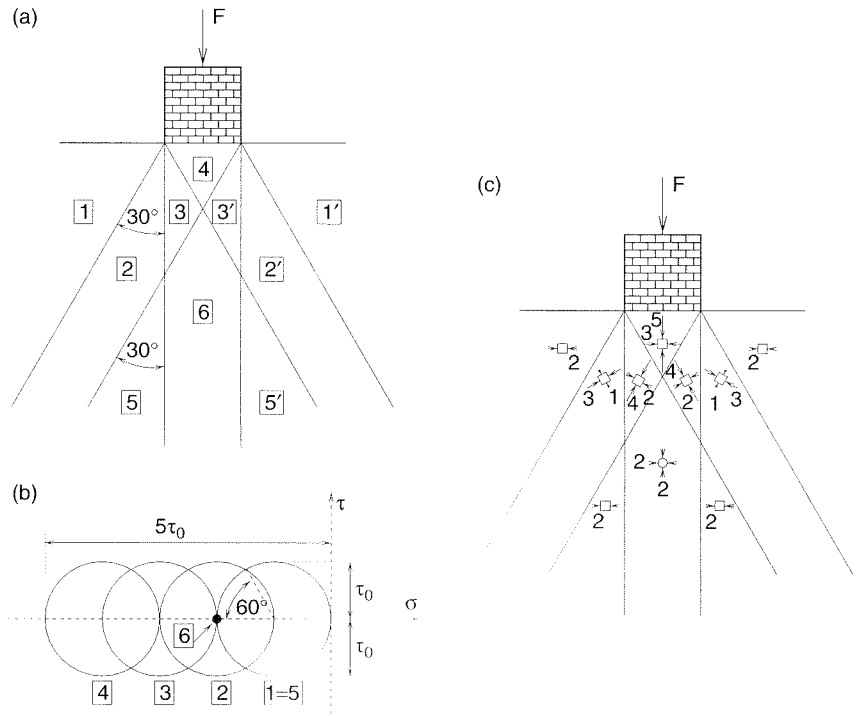


Figure 19.7 Statically admissible stress distribution: a) regions of constant stress, b) Mohr circles, c) principal directions and principal stresses (in multiples of τ_0)

Example 19.4: Construct a lower bound on the limit indentation force using the regions of uniform stress sketched in Figure 19.7(a).

Solution: The equilibrium conditions at the boundaries between the regions 1–6 may be easily figured out with the help of the Mohr circles for the states of stress in the individual regions; see Figure 19.7(b). According to the Tresca criterion, the Mohr circles cannot have a radius exceeding the shear yield stress τ_0. As explained before, the von Mises criterion for plane strain gives the same limit loads as the Tresca criterion with the same shear yield stress. In region 1, bounded by the free surface of the half space, the vertical stress σ_y and the shear stress τ_{xy} must vanish. Consequently, $\sigma_y = \sigma_1$ is one of the principal stresses, and the Mohr circle corresponding to this region must pass through the origin. The other principal stress, $\sigma_x = \sigma_2$, can be determined from the yield condition $\tau_{\max} = |\sigma_1 - \sigma_2|/2 = \tau_0$. Since σ_x is expected to be compressive, we obtain $\sigma_x = -2\tau_0$. Geometrically this means that the Mohr circle must touch the line $\tau = \tau_0$, and its left intersection with the σ-axis is at the distance $2\tau_0$ from the origin. If the stress state in region 2 is also at the plastic limit, it will be represented again by a circle of radius τ_0 but shifted horizontally with regard to circle 1. The stress state at the boundary between regions 1 and 2 corresponds to the intersection of these two circles. Suppose that the inclined boundaries in Figure 19.7(a) form the angle 30°. The intersection of circles 1 and 2 must lie on the ray inclined by 60° with respect to the direction of principal stress $\sigma_2 = \sigma_x$ in region 1; see Figure 19.7(b). This fixes the position of circle 2 as shown, and the corresponding principal stresses are $\sigma_1 = -\tau_0$ and $\sigma_2 = -3\tau_0$. The principal

directions are rotated by 30° with respect to the global axes; see Figure 19.7(c). Assuming the stress state in region 3 to again be at the plastic limit, and noting that the stress state corresponding to the boundary regions 2 and 3 must lie at the radial ray inclined by 60°, we again fix the position of circle 3 as shown in Figure 19.7(c), and we obtain the principal stresses in region 3, $\sigma_1 = -2\tau_0$ and $\sigma_2 = -4\tau_0$. Finally, in the same stereotype, we fix the position of circle 4 and determine the principal stresses in region 4, $\sigma_1 = -3\tau_0$ and $\sigma_2 = -5\tau_0$. The Mohr circle for region 5 is the same as for region 1, and the stress state in region 6 is represented by a circle that is shrunken to a point and lies at point 6 in Figure 19.7(b) (at the contact of circles 3 and 5 for the adjacent stress regions).

We have constructed a stress distribution that satisfies the equilibrium conditions (with zero body forces) inside each region and also traction equilibrium conditions on the boundaries between regions.[2] The shear stress does not exceed τ_0 anywhere. Moreover, static boundary conditions (of zero tractions) are satisfied on the free part of the boundary. It remains to apply the boundary conditions on the contact with the indentor. The vertical stress σ_y under the rigid block equals $5\tau_0$. Since its resultant $b\sigma_y$ must be in equilibrium with the applied force F, we obtain the lower bound estimate of the collapse load,

$$F_s = 5b\tau_0 \tag{19.28}$$

This is only 6% below the upper bound estimate (19.27). □

19.5 SLIP LINE THEORY

Good lower bounds and sometimes exact solutions can be obtained by constructing the so-called slip line fields. The slip line theory is only applicable to planar problems and to the Tresca or von Mises yield condition. From the mathematical point of view, the slip lines are the characteristic lines of a hyperbolic differential equation governing the problem. Physically, the slip lines (sometimes called the shear lines) are defined as the trajectories of the directions of maximum shear stresses. Because the maximum shear stresses act on orthogonal planes, the slip lines represent an orthogonal net, which defines curvilinear orthogonal coordinates denoted as (α, β); see Figure 19.8(a).

While constructing a statically admissible stress field, we look for three independent stress components. In the plastic region, the yield condition $f(\boldsymbol{\sigma}) = 0$ provides an additional equation, which can be used for the elimination of one of the unknowns. This is especially easy with a special choice of three independent stress variables – the mean stress $\sigma = (\sigma_1 + \sigma_2)/2$, the maximum shear stress $\tau_{\max} = |\sigma_1 - \sigma_2|/2$, and the angle ϕ characterizing the direction of maximum shear according to Figure 19.8(b). In a local coordinate system rotated by ϕ with respect to the global one, the normal stresses are both equal to σ and the shear stress is equal to $\tau_{\max}$. The yield condition simply reads $\tau_{\max} = \tau_0$ and, consequently, any stress distribution in the plastic region is fully described by the fields $\sigma(\alpha, \beta)$ and $\phi(\alpha, \beta)$. Now we can use the equilibrium conditions to set up two differential equations for the unknown functions σ and ϕ. These equations turn out to be particularly simple if we write them at each point with

[2] Note that, on the internal boundaries, only the stress component normal to the boundary and the shear component must be continuous, but the stress component parallel to the boundary can have a jump.

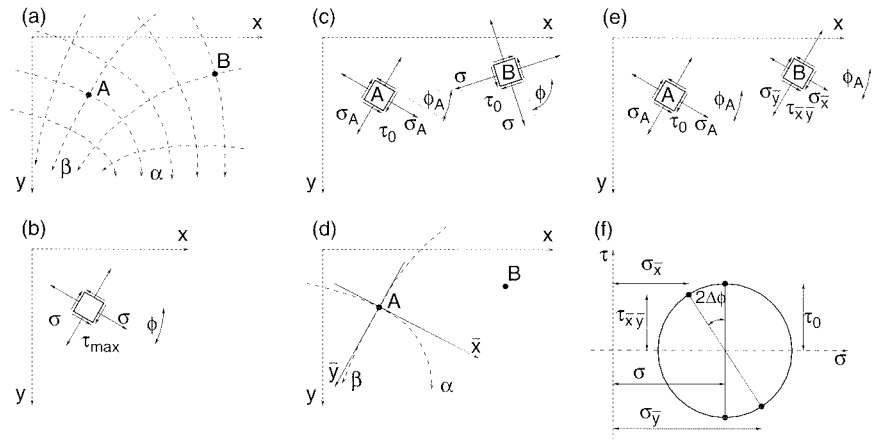


Figure 19.8 a) Trajectories of maximum shear, b) stress components in a local coordinate system aligned with directions of maximum shear, c) stresses at points A and B, d) local coordinate system $(\bar{x}, \bar{y})$ aligned with directions of maximum shear at point A, e) stress components at point B with respect to the local coordinate system $(\bar{x}, \bar{y})$, f) Mohr circle characterizing the stress state at point B

respect to a local coordinate system $(\bar{x}, \bar{y})$ aligned with the directions of maximum shear.

Consider a given material point A inside the plastic region, and denote by σ_A and ϕ_A the values of σ and ϕ at this point (Figure 19.8(c)). In the local Cartesian coordinate system $(\bar{x}, \bar{y})$ aligned with the directions of maximum shear at point A (Figure 19.8(d)), the equations of equilibrium read

$$\frac{\partial \sigma_{\bar{x}}}{\partial \bar{x}} + \frac{\partial \tau_{\bar{x}\bar{y}}}{\partial \bar{y}} = 0 \tag{19.29}$$

$$\frac{\partial \tau_{\bar{x}\bar{y}}}{\partial \bar{x}} + \frac{\partial \sigma_{\bar{y}}}{\partial \bar{y}} = 0 \tag{19.30}$$

To compute the partial derivatives, we need to express the components $\sigma_{\bar{x}}$, $\sigma_{\bar{y}}$ and $\tau_{\bar{x}\bar{y}}$ of the stress tensor at an arbitrary point B with respect to the local coordinate system at point A. This involves a rotation of the local coordinate system at B by the angle $\Delta\phi = \phi - \phi_A$ where ϕ defines the orientation of slip lines at B (Figures 19.8(c),(e)). From the Mohr circle in Figure 19.8(f), we easily get the transformation rules

$$\sigma_x = \sigma - \tau_0 \sin 2(\phi - \phi_A) \tag{19.31}$$

$$\sigma_y = \sigma + \tau_0 \sin 2(\phi - \phi_A) \tag{19.32}$$

$$\tau_{xy} = \tau_0 \cos 2(\phi - \phi_A) \tag{19.33}$$

Differentiating these expressions with respect to $\bar{x}$ and $\bar{y}$ we obtain

$$\frac{\partial \sigma_{\bar{x}}}{\partial \bar{x}} = \frac{\partial \sigma}{\partial \bar{x}} - 2\tau_0 \frac{\partial \phi}{\partial \bar{x}} \cos 2(\phi - \phi_A) \tag{19.34}$$

$$\frac{\partial \sigma_{\bar{y}}}{\partial \bar{y}} = \frac{\partial \sigma}{\partial \bar{y}} + 2\tau_0 \frac{\partial \phi}{\partial \bar{y}} \cos 2(\phi - \phi_A) \tag{19.35}$$

$$\frac{\partial \tau_{\bar{x}\bar{y}}}{\partial \bar{x}} = -2\tau_0 \frac{\partial \phi}{\partial \bar{x}} \sin 2(\phi - \phi_A) \tag{19.36}$$

$$\frac{\partial \tau_{\bar{x}\bar{y}}}{\partial \bar{y}} = -2\tau_0 \frac{\partial \phi}{\partial \bar{y}} \sin 2(\phi - \phi_A) \tag{19.37}$$

The equilibrium equations are being set up at point A and, therefore, the derivatives must be evaluated at $(\bar{x}, \bar{y}) = (0, 0)$. Consequently, we set $\phi = \phi_A$, $\cos 2(\phi - \phi_A) = 1$, and $\sin 2(\phi - \phi_A) = 0$, which greatly simplifies the above expressions. After substitution into (19.29) (19.30) one gets

$$\frac{\partial \sigma}{\partial \bar{x}} - 2\tau_0 \frac{\partial \phi}{\partial \bar{x}} = 0 \tag{19.38}$$

$$\frac{\partial \sigma}{\partial \bar{y}} + 2\tau_0 \frac{\partial \phi}{\partial \bar{y}} = 0 \tag{19.39}$$

Since the local coordinate axes are tangent to the curves $\alpha = \text{const.}$ and $\beta = \text{const.}$, we may replace in (19.38)–(19.39) the derivatives with respect to the local Cartesian coordinates $\bar{x}$ and $\bar{y}$ by the derivatives with respect to the curvilinear coordinates α and β. This is clear from the chain rule $\partial f/\partial \bar{x} = (\partial f/\partial \alpha)(\partial \alpha/\partial \bar{x}) + (\partial f/\partial \beta)(\partial \beta/\partial \bar{x})$ and from the fact that $\partial \beta/\partial \bar{x} = 0$ and $\partial \alpha/\partial \bar{y} = 0$. The final form of the differential equations for σ and ϕ is

$$\frac{\partial \sigma}{\partial \alpha} = 2\tau_0 \frac{\partial \phi}{\partial \alpha} \tag{19.40}$$

$$\frac{\partial \sigma}{\partial \beta} = -2\tau_0 \frac{\partial \phi}{\partial \beta} \tag{19.41}$$

Due to the decoupled character of (19.40)–(19.41), it is possible to integrate these equations along the slip lines and write

$$\sigma(\alpha_2, \beta_1) = \sigma(\alpha_1, \beta_1) + 2\tau_0 \left[\phi(\alpha_2, \beta_1) - \phi(\alpha_1, \beta_1)\right] \tag{19.42}$$

$$\sigma(\alpha_1, \beta_2) = \sigma(\alpha_1, \beta_1) - 2\tau_0 \left[\phi(\alpha_1, \beta_2) - \phi(\alpha_1, \beta_1)\right] \tag{19.43}$$

This means that if we know the orthogonal system of slip lines and the value of σ at one single point, we can directly evaluate σ at any other point just by marching along the slip lines and using (19.42) (19.43). Since the result must be independent of the path (first along an α-line and then along a β-line or vice versa), we get the identities

$$\phi(\alpha_2, \beta_2) - \phi(\alpha_1, \beta_2) = \phi(\alpha_2, \beta_1) - \phi(\alpha_1, \beta_1) \tag{19.44}$$

$$\phi(\alpha_2, \beta_2) - \phi(\alpha_2, \beta_1) = \phi(\alpha_1, \beta_2) - \phi(\alpha_1, \beta_1) \tag{19.45}$$

known as the Hencky theorem. Thus, the difference $\phi(\alpha_2, \beta) - \phi(\alpha_1, \beta)$ depends only on α_1 and α_2 but not on β, and the difference $\phi(\alpha, \beta_2) - \phi(\alpha, \beta_1)$ depends only on β_1 and β_2 but not on α. In other words, the difference in slope between any two slip lines of the same system is constant. Consequently, if a certain α-line is straight for $\alpha \in (\alpha_1, \alpha_2)$, then any other α-line must also be straight in that range. If the straight α-lines are parallel the transverse β-lines are also straight and parallel (Figure 19.9(a)), and if the straight α-lines have a common intersection at point C the β-lines are circular arches centered at C (Figure 19.9(b)). Along a straight slip line, the stress state remains constant. Along a curved slip line, the principal directions rotate and the mean stress varies at a rate proportional to the curvature of the line.

Using these simple rules, it is often possible to geometrically construct the system of slip lines by 'marching out' from the boundary of known stress. From a load-free surface, the slip lines start at inclination 45°, since these are the maximum shear stress directions (the principal directions run parallel and perpendicular to the

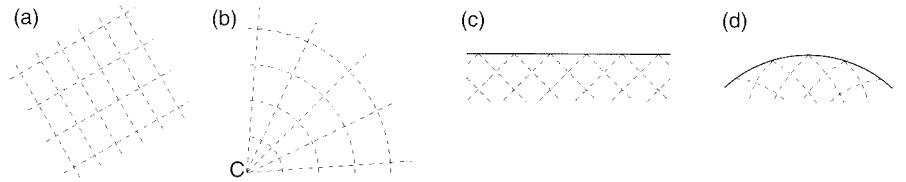


Figure 19.9 Orthogonal systems of α- and β-lines

surface). If the free surface is straight, the slip lines must have a zero curvature near the surface (Figure 19.9(c)). If the free surface is curved, the slip lines must also be curved (Figure 19.9(d)).

We illustrate the applications of the slip lines again with the indentation problem.

Example 19.5: Construct the slip line field for the indentation problem of a half space and estimate the limit indentation force.

Solution: The state of stress at the free surface on the side of the indenter is represented by Mohr circle 1 in Figure 19.10(a); the radius of the circle is τ_0 and its center is at $\sigma = -\tau_0$. As we move away from the surface along the α-line, the Mohr circle keeps the same radius but shifts along the horizontal axis as the hydrostatic pressure component, indicated by the center of the circle, varies along the slip line. We assume the slip lines to be of a constant slope of 45° in region 1, then continue as concentric circles in region 2, and further continue up to the indenter again in straight lines in region 3; see Figure 19.10(b). Since the slope of the slip lines in region 1 is constant, the hydrostatic stress component must be constant, and so this entire region is in a state of uniform stress and represented by the same Mohr circle 1, with $\sigma = -\tau_0$. Within region 2, the hydrostatic stress σ varies proportionally to the slope of the line, and the change of the hydrostatic stress along the α-lines from one boundary of region 2 to the other must therefore be

$$\Delta\sigma = 2\tau_0\Delta\phi = 2\tau_0 \times (-\pi/2) = -\pi\tau_0 \tag{19.46}$$

This change of σ is reflected by the shift of the Mohr circle by a distance $\pi\tau_0$ to the left. Within region 3, the slip lines have a constant 45° inclination (because we assume a frictionless indenter, i.e. zero shear stress under the indenter). Therefore, there is no change in σ within region 3, i.e. this region is in a state of uniform stress represented by circle 3 in Figure 19.10(a). The vertical normal stress under the indenter is represented by the left extreme point of Mohr circle 3. From the figure, the horizontal coordinate

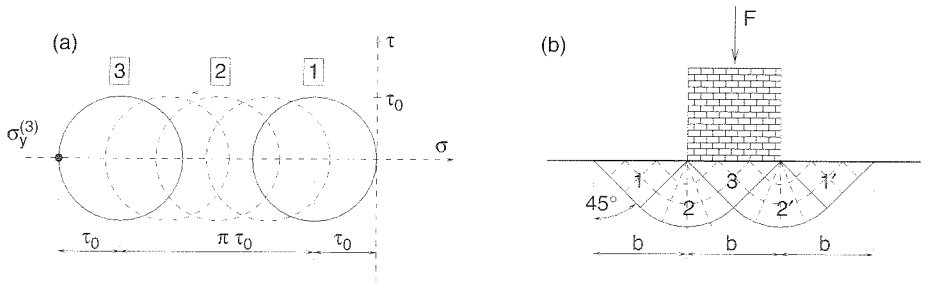


Figure 19.10 Indentation problem: a) Mohr circles, b) slip lines

of this point is $\sigma_y^{(3)} = -2\tau_0 - \pi\tau_0$, and this must be equal to F/b. Thus we obtain the lower bound

$$F_s = b\sigma_y^{(3)} = (2+\pi)b\tau_0 = 5.14b\tau_0 \tag{19.47}$$

□

The slip lines provide an exact solution if the strain velocities associated with the slip lines are kinematically admissible. The principal directions of stress are at the same time the principal directions of the plastic strain rate. Since the plastic flow takes place at constant volume ($\dot\varepsilon_1 = -\dot\varepsilon_2$) and the slip lines are inclined by 45° with respect to the principal directions, the rate of normal plastic strain in the direction tangential to a slip line must be zero. Consequently, the condition of a kinematically admissible field is that the α- and β-slip lines do not change length. A second, obvious condition is that the sign of the shear stress τ must be the same as the sign of the shear strain velocity $\dot\gamma$. The shear strain rate associated with the slip line directions, $\dot\gamma_{\max} = |\dot\varepsilon_1 - \dot\varepsilon_2| = 2|\dot\varepsilon_1|$ is the largest one among all the directions.

Example 19.6: Construct a kinematically admissible mechanism corresponding to the slip lines from the previous example.

Solution: One kinematically admissible mechanism is shown in Figure 19.11(a). Regions 1, 1', and 3 move as rigid triangular blocks, with concentrated shear slip bands along their inclined boundaries. The circular segments 2 and 2' cannot then behave as rigid blocks because the velocity of each of the concentrated circles must be the same (and therefore the angular velocity is different). Thus, in addition to a concentrated shear band along the circular boundary of regions 2 and 2', we also have continuous shearing deformation within those regions.

The deformation mode is not easy to visualize because of a singularity at points S and S' (the points separating the free part of the surface from the part loaded by the indentor). It is therefore challenging to analyze this mechanism and show that the corresponding kinematically admissible load can still be evaluated.

Triangular segment 3 moves as a rigid body downwards at the $\dot v_3$. The right inclined boundary of segment 2 is pushed by segment 3 in the inward direction at rate $\dot u = \dot v_3\sqrt{2}/2$, constant along the boundary (Figure 19.12(a)). As already explained, for each slip line the tangential velocity component must be constant along the entire slip line. In our case, the β-lines in region 2 are straight rays emanating from the singular point S and, at their intersections with the curved shear band at the bottom, the displacement rate normal to the band (i.e. tangential to the β-line) must vanish. Consequently, no point of segment 2 can displace in the radial direction. On each α-line, the displacement rate imposed by block 3 on the right inclined boundary is the same, and so the circumferential velocity component is constant within segment

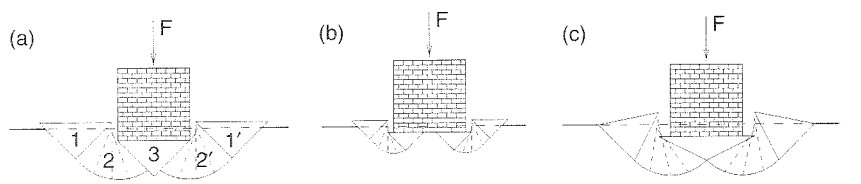


Figure 19.11 Three possible failure mechanisms corresponding to the same set of slip lines from Figure 19.10(b)

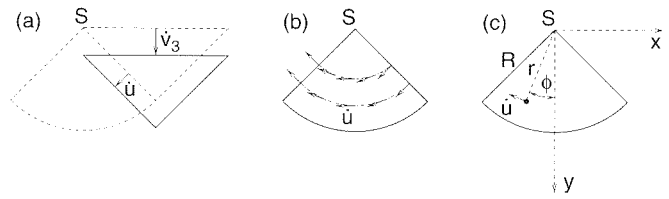


Figure 19.12 Analysis of the velocity field

2 (Figure 19.12(b)). For convenience, let us put the origin of the global coordinate system (x, y) into the singular point S, and let us also introduce polar coordinates (ϕ, r) where ϕ is at the same time the angle characterizing the orientation of the slip lines; see Figure 19.12(c). The radius of the circular segment is $R = b\sqrt{2}/2$. The velocity field can be described by its Cartesian components

$$\dot{u}_x = -\dot{u}\cos\phi = -\dot{u}\frac{y}{\sqrt{x^2+y^2}} \tag{19.48}$$

$$\dot{u}_y = -\dot{u}\sin\phi = \dot{u}\frac{x}{\sqrt{x^2+y^2}} \tag{19.49}$$

Note that there is no unique limit of $\dot{u}_x$ and $\dot{u}_y$ as (x, y) approaches the origin, but both velocity components remain bounded. The strain rates could be evaluated by differentiating (19.48)–(19.49). However, our goal is to express the dissipation rate, and for that purpose it is sufficient to find the maximum shear strain rate. Since we know the directions of maximum shear (the slip lines), the corresponding shear strain rate is easily obtained from Figure 19.13. The motion of a curved differential element can be decomposed into a rigid body translation and shearing at rate

$$\dot{\gamma} \qquad \frac{\dot{u}\,\mathrm{d}\phi}{r\,\mathrm{d}\phi} = \frac{\dot{u}}{r} \tag{19.50}$$

where $r = \sqrt{x^2+y^2}$ is the radial coordinate of the point of interest. Equation (19.50) indicates that the strain rate is unbounded and grows proportionally to $1/r$ as $r \to 0$, i.e. as the singular point is approached. Nevertheless, the total dissipation in segment 2 is finite, and it can be evaluated as

$$D^{(2)}_{\mathrm{int}} = \int_0^R \int_{-\pi/4}^{\pi/4} \tau_0 |\dot{\gamma}| r\,\mathrm{d}\phi\,\mathrm{d}r = \int_0^R \int_{-\pi/4}^{\pi/4} \tau_0 \frac{\dot{u}}{r} r\,\mathrm{d}\phi\,\mathrm{d}r = \tau_0 \dot{u} R \frac{\pi}{2} = \frac{\pi}{4} b\tau_0 \dot{v}_3 \tag{19.51}$$

To this we have to add the power dissipated in the shear bands along the boundaries of all segments. Note that, for this particular mechanism, there is no shear band between segments 1 and 2, and the tangential slip rate in all the other shear bands is $\dot{u}_t = \dot{v}_3\sqrt{2}/2$. Thus the total dissipation rate is

$$\begin{aligned} D_{\mathrm{int}} &= 2D^{(2)}_{\mathrm{int}} + 2(\ell_1 + \ell_2 + \ell_{23})\tau_0 \dot{u}_t \\ &= 2\frac{\pi}{4} b\tau_0 \dot{v}_3 + 2\left(R + \frac{\pi}{2}R + R\right)\tau_0 \frac{\dot{v}_3\sqrt{2}}{2} = (2+\pi) b\tau_0 \dot{v}_3 \end{aligned} \tag{19.52}$$

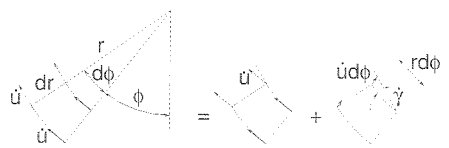


Figure 19.13 Deformation of a differential element

Setting this equal to $\dot{W}_{\mathrm{ext}} = F\dot{v}_3$ we obtain a kinematically admissible load $F_k = (2+\pi)b\tau_0$, which proves that the previously obtained lower bound (19.47) is exactly equal to the collapse load. Interestingly, however, the kinematically admissible displacement field which corresponds to the slip line pattern in Figure 19.10(b) (Prandtl, 1923) is not unique. The mode of deformation shown in Figure 19.11(b) is equally possible (Hill, 1949). There are even other possible modes, e.g. that in Figure 19.11(c), which involves continuous shearing in all regions but no concentrated shear bands at all. □

The foregoing account of the slip line theory is rather incomplete. Additional examples of slip line fields and plastic zones for selected problems are given in Figure 19.14. A detailed exposition of the theory may be found, for example, in the classical work by Hill (1950). A succint but insightful presentation is given in the small book by Calladine (1969).

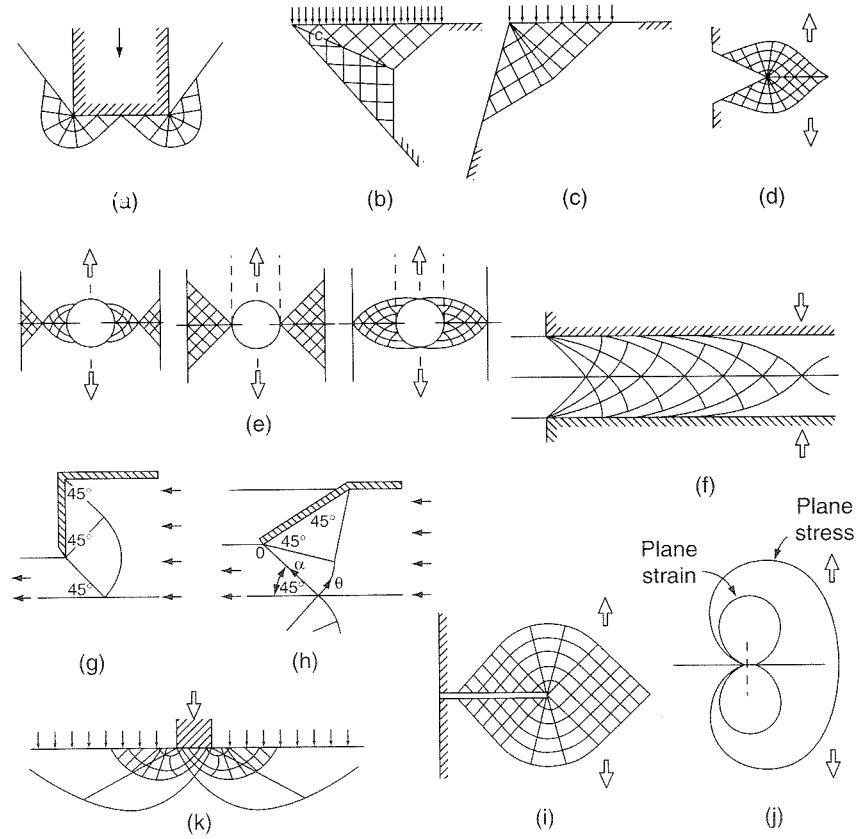


Figure 19.14 Examples of slip line fields and plastic zones: (a) punch in a trench, (b) corner of acute angle, (c) failure of a slope, (d) notch under tension, (e) various solutions for a circular hole, (f) transverse compression of a plastic plate between two rigid plates, (g, h) extrusion of a plane through a slit, (i) slip line field at a crack (Prandtl's solution), (j) realistic boundaries of plastic zones at a crack for plane strain and plane stress obtained for plastic hardening material, (k) Salençon's example of a difference between Prandtl's slip line field for a punch and bounding slip surface for Coulomb frictional material (soil foundation failure)

PROBLEMS

Problem 19.1*: In plane-strain elasticity, the out-of-plane normal stress is given by

$$\sigma_z = \nu(\sigma_x + \sigma_y) \tag{19.53}$$

This relation directly follows from the condition $\varepsilon_z = 0$, which remains valid even after the onset of plastic flow. In the plastic limit state, there can be no plastic flow in the out-of-plane direction, i.e. $\dot{\varepsilon}_z^p = 0$. Since $\dot{\varepsilon}_z = 0$ and $\dot{\varepsilon}_z^e = \dot{\varepsilon}_z - \dot{\varepsilon}_z^p$, the elastic strain rate $\dot{\varepsilon}_z^e$ also vanishes. From this it may seem that (19.53) still holds. However, from the flow rule we have derived a different expression (19.1). Explain this contradiction.

Problem 19.2: Find the best upper bound on the limit indentation force using the general mechanism shown in Figure 19.4(a).

Problem 19.3: Find the best upper bound on the limit indentation force using the mechanism with a rectangular region of homogeneous strain shown in Figure 19.4(c). Hint: Take into account that, on the boundary of the region of homogeneous strain, the relative slip rate is variable along the shear bands.

Problem 19.4: Homogeneous plastic flow in a rectangular region of size $a \times b$ (aligned with the directions of principal stress) can be approximated by a mechanism consisting of four rigid triangular blocks separated by shear bands (Figure 19.15(c)). Check whether such a mechanism properly represents the condition of a zero volumetric strain rate and whether it gives the correct energy dissipation inside the rectangular region.

Problem 19.5: Find the best upper bound on the limit indentation force using the mechanisms shown in Figures 19.15(a),(b).

Problem 19.6: Construct the dual problem to (19.22)–(19.26) and find the physical meaning of the dual variables and constraints.

Problem 19.7*: The problem dual to (19.22)–(19.26) corresponds to the static approach, and since its optimal solution gives the same load factor as the optimal solution of the primal problem, this load factor should be equal to the exact collapse load. Is this reasoning correct?

Problem 19.8: Evaluate the kinematically admissible loads corresponding to the mechanisms in Figures 19.11(b),(c) and show that they are both equal to the exact collapse load.

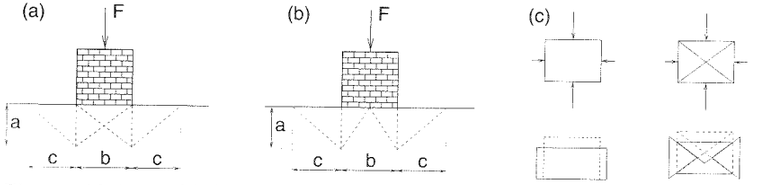


Figure 19.15 a-b) Admissible mechanisms for the indentation problem, c) approximation of a rectangular region with homogeneous plastic flow by four rigid triangular blocks separated by shear bands

Part IV

Advanced Topics in Plasticity

20

General Elastoplastic Constitutive Models

In the first three parts of this book we focused on the simplest type of plasticity model – an isotropic, perfectly elastoplastic material with an associated flow rule. Such models are sufficiently realistic only under certain conditions and should be used with caution. Now it is time to introduce more advanced concepts.

20.1 HARDENING

Let us recall that a perfectly (or ideally) elastoplastic material subjected to uniaxial loading yields at a constant stress. During plastic flow under general multiaxial loading, the stress state can move along the yield surface, but the surface itself remains unchanged. However, in reality the microstructure of the material changes as plastic flow continues, and this results in a change of the properties observable on the macroscale. Under uniaxial loading, the stress transmitted by a yielding material can increase or decrease. An increase of the yield stress is referred to as *hardening*, and its decrease is called *softening*. Typically, many materials initially harden and later soften. For convenience, however, we will sometimes use the term 'hardening' in a broader sense, meaning yield stress changes of any sign, negative hardening being equivalent to softening.

During hardening (in the broad sense), the elastic domain undergoes a certain evolution. The elastic domain of a virgin material is bounded by the initial yield surface, also called the elastic limit envelope. Due to microstructural changes in the material induced by plastic flow, the elastic domain changes its size or position, or both. Its boundary at an intermediate state is usually called a *loading surface*. Some models work with a maximum possible elastic domain, which is bounded by the *strength envelope* (failure surface), representing the largest possible resistance of the material.

20.1.1 Isotropic Hardening

In order to describe the evolution of the loading surface, we need one or several new parameters that characterize the effect of hardening. The simplest approach, introduced by Odqvist (1933a), deals with a one-parameter family of loading surfaces that are all similar and affine with regard to the origin (Figure 20.1). This is called *isotropic hardening*. The loading surfaces can be derived from the same basic form of

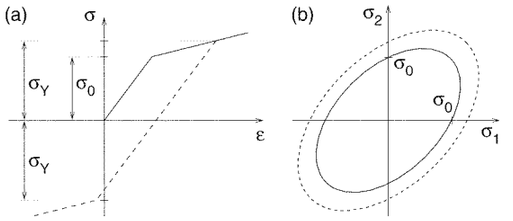


Figure 20.1 Isotropic hardening: a) uniaxial stress-strain diagram, b) evolution of the yield surface in the biaxial stress plane

the yield function, in which different yield stress values are used. Any yield criterion originally stated in the form

$$F(\boldsymbol{\sigma}) - \sigma_0 = 0 \tag{20.1}$$

where σ_0 is a constant yield stress, can be reformulated as

$$F(\boldsymbol{\sigma}) - \sigma_Y = 0 \tag{20.2}$$

where σ_Y is a new variable, the current yield stress, initially equal to the material parameter σ_0.

The evolution of the yield stress during plastic flow must be described by an additional equation, the *hardening law*. Simple examples of hardening in one dimension have already been presented in Chapter 1. Hardening laws under uniaxial monotonic loading can be postulated as an explicit dependence of the yield stress on the plastic strain,

$$\sigma_Y = h(\varepsilon_{11}^p) \tag{20.3}$$

which is called the *strain hardening*. Here, ε_{11}^p is the plastic part of the normal strain ε_{11} in the direction of applied stress σ_{11}. The function h is easily extracted from the monotonic uniaxial stress-strain curve. Its derivative,

$$h'(\varepsilon_{11}^p) \equiv H(\varepsilon_{11}^p) \tag{20.4}$$

is called the *plastic modulus*. In the simplest case, the hardening law is linear, $\sigma_Y = \sigma_0 + H\varepsilon_{11}^p$, and the plastic modulus is constant, $H(\varepsilon_{11}^p) = H$. For a positive plastic modulus we get true hardening in the sense of an increase of the yield stress; $H = 0$ corresponds to perfect plasticity; and for $H < 0$ the material exhibits softening. A negative plastic modulus is often called the softening modulus.

In a general multiaxial setting, the plastic strain $\boldsymbol{\varepsilon}_p$ is a second-order tensor, and the one-dimensional hardening law (20.3) can be extended to this case if we introduce a scalar measure that reflects the amount of changes in the material microstructure. The simplest choice might seem to be the norm of the plastic strain tensor, $\|\boldsymbol{\varepsilon}_p\| = \sqrt{\boldsymbol{\varepsilon}_p : \boldsymbol{\varepsilon}_p}$, but this variable does not always increase during plastic flow. The reason is that the plastic strain rate can in general have any direction, which causes that the norm of $\boldsymbol{\varepsilon}_p$ can decrease even when the plastic flow continues. Therefore it makes more sense to characterize strain hardening by the *cumulative plastic strain* (also called the effective plastic strain or the equivalent plastic strain), $\bar{\varepsilon}_p$, defined by the rate equation

$$\dot{\bar{\varepsilon}}_p = \sqrt{\frac{2}{3}}\,\|\dot{\boldsymbol{\varepsilon}}_p\| = \sqrt{\frac{2}{3}\dot{\boldsymbol{\varepsilon}}_p : \dot{\boldsymbol{\varepsilon}}_p} \tag{20.5}$$

This was first proposed, without the scaling factor $\sqrt{2/3}$, by Odqvist (1993a). The scaling factor is chosen such that, under monotonic uniaxial loading, $\dot{\bar{\varepsilon}}_p$ would coincide with the component $\dot{\varepsilon}_{11}^p$ of $\dot{\boldsymbol{\varepsilon}}_p$, provided that the plastic flow is isochoric (purely deviatoric), i.e. no plastic change of volume takes place. Definition (20.5) can be integrated to yield

$$\bar{\varepsilon}_p(t) = \sqrt{\frac{2}{3}} \int_0^t \|\dot{\boldsymbol{\varepsilon}}_p(\tau)\|\, \mathrm{d}\tau \tag{20.6}$$

where the time-like variable t can be any monotonically increasing parameter controling the loading process.

An alternative to the strain hardening hypothesis (i.e. to the assumption that the yield stress depends on the cumulative plastic strain) is the *work hardening* hypothesis (Taylor and Quinney, 1931), stating that the yield stress depends on the plastic work,

$$W_p(t) = \int_0^t \boldsymbol{\sigma}(\tau) : \dot{\boldsymbol{\varepsilon}}_p(\tau)\, \mathrm{d}\tau \tag{20.7}$$

The choice between strain and work hardening depends on multiaxial experiments. If only uniaxial test results are available, there is no way to decide which hypothesis is better. Both are widely used, and even in many multiaxial stress situations they lead to similar results. For associated J_2-plasticity they are completely equivalent (after an appropriate transformation of the hardening function h).

To cover both cases by a unique description, we will deal with the notion of a hardening variable, denoted as κ, which can be either the cumulative plastic strain (for strain hardening) or the plastic work (for work hardening). It is convenient, albeit not necessary, to scale down the plastic work such that it has the dimension of strain and that, under uniaxial tension, it corresponds to ε_{11}^p. This can be achieved by defining the work hardening variable by the rate equation

$$\dot{\kappa} = \frac{1}{h(\kappa)}\, \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p \tag{20.8}$$

Equation (20.3) is then written as

$$\sigma_Y = h(\kappa) \tag{20.9}$$

Note that, during plastic flow under uniaxial tension, $\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p = \sigma_{11}\dot{\varepsilon}_{11}^p = \sigma_Y \dot{\varepsilon}_{11}^p$, and so (20.8) combined with (20.9) gives $\dot{\kappa} = \dot{\varepsilon}_{11}^p$. Consequently, the strain hardening variable and the work hardening variable are identical under uniaxial stress conditions, and the same hardening function h can then be used for either type of hardening.

Let us now explore the effect of hardening on the elastoplastic stiffness tensor and on the uniqueness of the model response. Here we deal only with uniqueness in the local sense, i.e. we check whether the response of a material point to any prescribed strain evolution is unique. Compared to the perfectly plastic model analyzed in Chapter 15, we now have two additional variables – the current yield stress, σ_Y, and the hardening variable, κ. The corresponding additional equations are the two components of the hardening law – the dependence of the current yield stress on the hardening variable, and the definition of the hardening variable. Since (20.9) explicitly relates the current

yield stress to the hardening variable, it is sufficient to introduce only the hardening variable as a primary unknown. The yield function is reformulated as

$$f(\boldsymbol{\sigma}, \kappa) = F(\boldsymbol{\sigma}) - h(\kappa) \tag{20.10}$$

and its time derivative is

$$\dot{f} = \frac{\partial F}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} - h' \dot{\kappa} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} - H\dot{\kappa} \tag{20.11}$$

Note that the gradients of f and F with respect to $\boldsymbol{\sigma}$ are identical, $\partial F/\partial \boldsymbol{\sigma} = \partial f/\partial \boldsymbol{\sigma}$, and so we can use for $\partial F/\partial \boldsymbol{\sigma}$ the symbol $\boldsymbol{f}_\sigma$, introduced in Chapter 15 for the gradient $\partial f/\partial \boldsymbol{\sigma}$. Of course, $\boldsymbol{f}_\sigma$ depends on the current value of $\boldsymbol{\sigma}$, and H may depend on the current value of κ (if the hardening law is nonlinear) but, for the sake of simplicity, we do not mark these dependencies explicitly.

Some important criteria, e.g. those due to von Mises, or Drucker and Prager, are written in their most natural form in terms of an alternative yield stress measure, e.g. the yield stress in shear, τ_0. For hardening materials, such criteria can be written in the general form

$$f(\boldsymbol{\sigma}, \kappa) = F(\boldsymbol{\sigma}) - \bar{h}(\kappa) \tag{20.12}$$

where

$$\bar{h}(\kappa) = \frac{\sigma_0^{ch}}{\sigma_0} h(\kappa) \tag{20.13}$$

is the hardening function rescaled by the ratio of the characteristic yield stress, σ_0^{ch}, to the yield stress in uniaxial tension, σ_0. For example, a criterion originally written as $F(\boldsymbol{\sigma}) - \tau_0 = 0$ will be reformulated as $F(\boldsymbol{\sigma}) - \bar{h}(\kappa) = 0$, where $\bar{h}(\kappa) = (\tau_0/\sigma_0)h(\kappa)$. The time derivative of the yield function is then

$$\dot{f} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} - \bar{H}\dot{\kappa} \tag{20.14}$$

where

$$\bar{H} = \bar{h}' = \frac{\sigma_0^{ch}}{\sigma_0} h' = \frac{\sigma_0^{ch}}{\sigma_0} H \tag{20.15}$$

is the rescaled plastic modulus.

For strain hardening as well as work hardening, the rate of the hardening variable, $\dot{\kappa}$, is proportional to the rate of the plastic multiplier, $\dot{\lambda}$, with a proportionality factor dependent on the current stress. Indeed, using the associated flow rule (15.28), for strain hardening we obtain

$$\dot{\kappa} = \sqrt{\frac{2}{3}}\, \|\dot{\boldsymbol{\varepsilon}}_p\| = \sqrt{\frac{2}{3}}\, \|\boldsymbol{f}_\sigma\| \dot{\lambda} \tag{20.16}$$

and for work hardening

$$\dot{\kappa} = \frac{1}{h(\kappa)}\, \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p = \frac{\boldsymbol{\sigma} : \boldsymbol{f}_\sigma}{h(\kappa)}\, \dot{\lambda} \tag{20.17}$$

Both cases can be covered by a general relation

$$\dot{\kappa} = k\dot{\lambda} \tag{20.18}$$

where k is a scalar factor;

$$k = \sqrt{\frac{2}{3}}\, \|\boldsymbol{f}_\sigma\| \tag{20.19}$$

for strain hardening and

$$k = \frac{\boldsymbol{\sigma} : \boldsymbol{f}_\sigma}{h(\kappa)} \tag{20.20}$$

for work hardening. Recall that if the elastic domain is convex and contains the origin (the initial stress-free state), then $\boldsymbol{\sigma} : \boldsymbol{f}_\sigma > 0$. The factor k is, therefore, always positive.

Example 20.1: Evaluate the factor k from (20.18) for associated J_2-plasticity.

Solution: Hardening J_2-plasticity is described by the yield function

$$f(\boldsymbol{\sigma}, \kappa) = \sqrt{J_2(\boldsymbol{\sigma})} - \bar{h}(\kappa) \tag{20.21}$$

where $\bar{h}(\kappa) = (\tau_0/\sigma_0)h(\kappa) = h(\kappa)/\sqrt{3}$. The gradient of the yield function with respect to the stress tensor is $\boldsymbol{f}_\sigma = \boldsymbol{s}/(2\sqrt{J_2})$. For strain hardening we obtain

$$k = \sqrt{\frac{2}{3}} \|\boldsymbol{f}_\sigma\| = \frac{1}{\sqrt{6J_2}} \|\boldsymbol{s}\| = \frac{1}{\sqrt{6J_2}} \sqrt{2J_2} = \frac{1}{\sqrt{3}} \tag{20.22}$$

while for work hardening we have

$$k = \frac{\boldsymbol{\sigma} : \boldsymbol{f}_\sigma}{h(\kappa)} = \frac{\boldsymbol{s} : \boldsymbol{s}}{2\sqrt{J_2} h(\kappa)} = \frac{\sqrt{J_2}}{h(\kappa)} = \frac{\bar{h}(\kappa)}{h(\kappa)} = \frac{\tau_0}{\sigma_0} = \frac{1}{\sqrt{3}} \tag{20.23}$$

because, during plastic flow, $\sqrt{J_2} = \bar{h}(\kappa)$. We can see that, in this specific case, the factor k is constant and has the same value for strain hardening as for work hardening. □

In summary, isotropically hardening elastoplasticity is described by the elastic stress-strain law,

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) \tag{20.24}$$

the associated flow rule,

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda} \boldsymbol{f}_\sigma \tag{20.25}$$

the evolution law for the hardening variable,

$$\dot{\kappa} = \dot{\lambda} k \tag{20.26}$$

and by the loading-unloading conditions,

$$f \le 0, \quad \dot{\lambda} \ge 0, \quad f\dot{\lambda} = 0 \tag{20.27}$$

in which f is a function of $\boldsymbol{\sigma}$ and κ having the form (20.10), and k is a function of $\boldsymbol{\sigma}$ and κ given by (20.19) or (20.20). Our aim is to find the rates $\dot{\boldsymbol{\sigma}}$, $\dot{\boldsymbol{\varepsilon}}_p$, $\dot{\lambda}$ and $\dot{\kappa}$ for a given strain rate $\dot{\boldsymbol{\varepsilon}}$, provided that the current values of all the variables are known.

Following the same line of reasoning as in Chapter 15, we analyze the cases of plastic loading and of unloading from a plastic state separately. The case of an elastic process inside the elastic domain is trivial.

1. During plastic flow, the stress state must remain on the yield surface, and the consistency condition

$$\dot{f} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} - \bar{H}\dot{\kappa} = 0 \tag{20.28}$$

must be satisfied. Unlike perfect plasticity, the yield surface does not remain stationary but evolves as the plastic flow continues. This is reflected by the term $-\bar{H}\dot{\kappa}$ in (20.28). Now we substitute (20.24)–(20.26) into the consistency condition (20.28) so as to eliminate all the unknowns except $\dot{\lambda}$. The first term, $\boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}}$, can be handled in exactly the same manner as in Chapter 15; cf. equation (15.40). In the second term, we simply substitute (20.26) for $\dot{\kappa}$. Thus the consistency condition is transformed into the equation

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\lambda}\boldsymbol{f}_\sigma) - \bar{H}k\dot{\lambda} = 0 \tag{20.29}$$

which contains a single unknown, $\dot{\lambda}$. The solution

$$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}k} \tag{20.30}$$

is admissible if the resulting value of $\dot{\lambda}$ is nonnegative. Back-substitution into the flow rule (20.25) and the elastic law (20.24) then leads to the rate form (or tangential form) of the stress-strain equations,

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_{ep} : \dot{\boldsymbol{\varepsilon}} \tag{20.31}$$

where

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}k} \tag{20.32}$$

is the elastoplastic material stiffness tensor. The only difference compared to (15.47) is the presence of the term $\bar{H}k$ in the denominator. For $\bar{H} = 0$ we recover formula (15.47) valid for a perfectly elastoplastic material.

2. During elastic unloading from a plastic state we have $\dot{\lambda} = 0$, and so $\dot{\boldsymbol{\varepsilon}}_p = \boldsymbol{0}$, $\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$, and $\dot{\kappa} = 0$. However, the solution is plastically admissible only if $\dot{f} \leq 0$. Substituting into (20.14), we obtain

$$\dot{f} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} - \bar{H}\dot{\kappa} = \boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \tag{20.33}$$

Having derived the solutions valid for plastic loading and for unloading from a plastic state, we can address the issue of uniqueness. The former solution is admissible if the rate of the plastic multiplier is nonnegative, i.e. if

$$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}k} \geq 0 \tag{20.34}$$

The latter solution is admissible if the time derivative of the yield function is nonpositive, i.e. if

$$\dot{f} = \boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \leq 0 \tag{20.35}$$

The denominator in (20.34) depends only on the current state and can be evaluated independently of the prescribed strain rate. If this denominator is positive, then condition (20.34) is equivalent with

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} \geq 0 \tag{20.36}$$

and thus is complementary to (20.35). This means that, for any prescribed strain rate, exactly one of the conditions is satisfied and the solution is unique. Both conditions are simultaneously satisfied only in the case of neutral loading when $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} = 0$, but then both solutions are identical and uniqueness is not violated.

On the other hand, if the denominator in (20.34) is negative, conditions (20.34) and (20.35) are equivalent and the problem either has two different solutions or no solution at all, depending on whether the prescribed strain rate satisfies or violates (20.35). This is clearly an undesirable situation. So we can conclude that the response to any prescribed strain evolution is unique if

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}k > 0 \tag{20.37}$$

for any state of the material. Note that the first term, $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma$, is always positive due to the positive definiteness of $\boldsymbol{D}_e$. The factor k is also always positive. If the plastic modulus is positive or zero, the condition of uniqueness (20.37) is satisfied for any possible state of the material. This means that the response of hardening (in the narrow sense, $\bar{H} > 0$) or perfectly plastic ($\bar{H} = 0$) materials with an associated flow is always locally unique, i.e. any prescribed strain history generates a unique response in terms of stress, plastic strain, etc. However, the condition of uniqueness might be satisfied by softening materials as well, as is demonstrated by the following example.

Example 20.2: Consider an isotropically hardening or softening von Mises material with an associated flow rule. Express the condition of uniqueness in terms of the plastic modulus.

Solution: The yield function for von Mises material is given by

$$f(\boldsymbol{\sigma}, \kappa) = \sqrt{J_2(\boldsymbol{\sigma})} - \bar{h}(\kappa) \tag{20.38}$$

According to (20.37), uniqueness is guaranteed if $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}k > 0$. Substituting the gradient $\boldsymbol{f}_\sigma = \partial f/\partial\boldsymbol{\sigma} = \boldsymbol{s}/(2\sqrt{J_2})$ into the first term, we obtain

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma = \frac{1}{4J_2}\boldsymbol{s} : \boldsymbol{D}_e : \boldsymbol{s} = \frac{1}{4J_2}\boldsymbol{s} : (2G\boldsymbol{s}) = \frac{G}{2J_2}\boldsymbol{s} : \boldsymbol{s} = G \tag{20.39}$$

In Example 20.1 we have shown that, for J_2-plasticity, $k = 1/\sqrt{3}$, independently of the type of hardening hypothesis. The condition of uniqueness thus reads

$$G + \frac{1}{\sqrt{3}}\bar{H} > 0 \tag{20.40}$$

i.e., $\bar{H} > -\sqrt{3}G$. So local uniqueness is preserved even for a softening material provided that the magnitude of the softening modulus does not exceed a certain critical value related to the elastic shear modulus. In terms of the standard softening modulus, $H = (\sigma_0/\tau_0)\bar{H} = \sqrt{3}\bar{H}$, the condition of uniqueness would read $H > -3G$. □

20.1.2 Linear Kinematic Hardening

So far we have considered only isotropic hardening, characterized by a single parameter. It appears, however, that more complicated hardening rules are necessary.

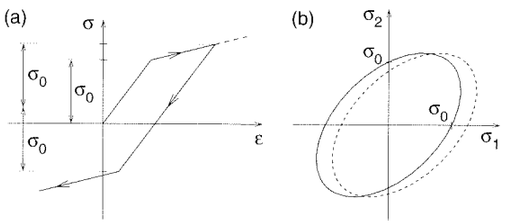


Figure 20.2 Kinematic hardening: a) uniaxial stress-strain diagram, b) evolution of the yield surface in the biaxial stress plane

especially for the case of unloading and cyclic loading. As an alternative hardening rule, the current loading surface is assumed not to expand but to move as a rigid body within the stress space (Figure 20.2(b)); this is known as *kinematic hardening*. The use of kinematic hardening is, for example, necessary to model the so-called *Bauschinger effect* (Bauschinger, 1881). This effect is often observed in metals subjected to cyclic loading. Even if the magnitudes of the yield stress in tension and in compression are initially the same, this is no longer the case when the material is preloaded into the plastic range and then unloaded. For example, after previous yielding in tension, yielding in compression may start at a stress level lower than the initial yield stress (Figure 20.2(a)).

The hardening behavior rule of most materials appears to be a combination of the isotropic and kinematic type of hardening, sometimes accompanied by a change of shape of the yield surface (see the discussion of the vertex effect in Section 25.4).

Kinematic hardening leads to a translation of the loading surface, i.e. to a shift of the origin of the initial yield surface. If the initial yield surface is described by a yield function of the form

$$f(\boldsymbol{\sigma}) = F(\boldsymbol{\sigma}) - \sigma_0 \tag{20.41}$$

the shifted surface is obviously described by

$$f(\boldsymbol{\sigma}, \boldsymbol{\sigma}_b) = F(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b) - \sigma_0 \tag{20.42}$$

where $\boldsymbol{\sigma}_b$ is the so-called *back stress* that represents the center of the shifted elastic domain and plays the role of a tensorial hardening variable. Now we need a kinematic hardening law that governs the evolution of the back stress. Melan (1938b) proposed a law of the form

$$\dot{\boldsymbol{\sigma}}_b = \bar{H}_K \, \dot{\boldsymbol{\varepsilon}}_p \tag{20.43}$$

according to which the rate of the back stress is proportional to the plastic strain rate. The proportionality factor $\bar{H}_K$ is directly related to the plastic modulus; see Problem 20.5. The linear hardening law (20.43) is often credited to Prager (1955, 1956); we will call it the *Melan–Prager hardening rule*. Generalizations are due to Backhaus (1968), who considered $\bar{H}_K$ as a variable dependent on the cumulative plastic strain, and Lehmann (1972), who replaced it by a tensor $\bar{\boldsymbol{H}}_K$.

Following the usual procedure, we construct the consistency condition

$$\frac{\partial f}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial f}{\partial \boldsymbol{\sigma}_b} : \dot{\boldsymbol{\sigma}}_b = 0 \tag{20.44}$$

and then use the elastic stress-strain law, flow rule and hardening law so as to eliminate all the unknowns except the rate of the plastic multiplier. Denoting again $\partial f/\partial \boldsymbol{\sigma} = \partial F/\partial \boldsymbol{\sigma} = \boldsymbol{f}_\sigma$, and realizing that $\partial f/\partial \boldsymbol{\sigma}_b = -\partial F/\partial \boldsymbol{\sigma} = -\boldsymbol{f}_\sigma$, we can write the resulting equation as

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\lambda} \boldsymbol{f}_\sigma) - \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma \bar{H}_K \dot{\lambda} = 0 \tag{20.45}$$

from which

$$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}_K \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma} \tag{20.46}$$

This result is very similar to that obtained for isotropic hardening, with the only difference that the term $\bar{H}k$ is replaced by the term $\bar{H}_K \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma$; cf. (20.30). It is now easy to obtain the elastoplastic stiffness tensor

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}_K \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma} \tag{20.47}$$

and the condition of local uniqueness

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}_K \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma > 0 \tag{20.48}$$

For nonnegative values of the plastic modulus, this condition is always satisfied.

Negative values of $\bar{H}_K$ would lead to softening. A model with purely kinematic softening does not make physical sense but a model with mixed isotropic and kinematic softening might be useful, e.g. for materials in which tensile loading induces a degradation of both tensile and compressive strength, but the degradation in tension is faster than in compression.

Example 20.3: Derive the specific form of the elastoplastic stiffness matrix for associated J_2-plasticity with kinematic hardening according to the Melan–Prager rule.

Solution: Partial differentiation of the yield function $f(\boldsymbol{\sigma}, \boldsymbol{\sigma}_b) = \sqrt{J_2(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)} - \tau_0$ with respect to the stress tensor $\boldsymbol{\sigma}$ at points where $f(\boldsymbol{\sigma}, \boldsymbol{\sigma}_b) = 0$ gives $\boldsymbol{f}_\sigma = (\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)_{\text{dev}}/2\tau_0$, where $(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)_{\text{dev}}$ denotes the deviatoric part of $\boldsymbol{\sigma} - \boldsymbol{\sigma}_b$. Realizing that the rate of the back stress is proportional to the rate of plastic strain, which is purely deviatoric, we conclude that the back stress $\boldsymbol{\sigma}_b$ is also purely deviatoric, and so $(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)_{\text{dev}} = \boldsymbol{s} - \boldsymbol{\sigma}_b$. Substituting this into (20.47) and using the relations $\boldsymbol{D}_e : (\boldsymbol{s} - \boldsymbol{\sigma}_b) = 2G(\boldsymbol{s} - \boldsymbol{\sigma}_b)$ and $(\boldsymbol{s} - \boldsymbol{\sigma}_b) : (\boldsymbol{s} - \boldsymbol{\sigma}_b) = 2J_2(\boldsymbol{s} - \boldsymbol{\sigma}_b) = 2\tau_0^2$, we obtain

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{2G^2}{\tau_0^2(2G + \bar{H}_K)} (\boldsymbol{s} - \boldsymbol{\sigma}_b) \otimes (\boldsymbol{s} - \boldsymbol{\sigma}_b) \tag{20.49}$$

□

Ziegler (1959) proposed a modification of the Melan–Prager kinematic hardening rule. He observed that if the original equation (20.43) is reduced to a subspace of the stress space (e.g. to the subspace corresponding to the plane stress), the yield surface does not always move in the direction of its normal at the current stress point. In some cases, e.g. for the Tresca condition in the σ_x–τ_{xy} space, the yield surface even deforms during kinematic hardening according to the Melan–Prager rule. Of course, the complete yield surface in the nine-dimensional stress space does not deform but,

as it moves, the shape of its intersection with a given stress subspace can change. This might be annoying and even confusing.

Ziegler also observed that the von Mises yield surface following the Melan–Prager rule always moves in the direction of the vector connecting its center with the current stress point, and that this property holds in any subspace of the stress space. Therefore, he suggested a modified kinematic hardening law

$$\dot{\boldsymbol{\sigma}}_b = \dot{\mu}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b) \tag{20.50}$$

where $\dot{\mu}$ is a new multiplier, to be determined from a suitable additional condition. Equation (20.50) only defines the direction in which the yield surface moves, but the rate at which this happens remains unspecified. The consistency condition

$$\boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} - \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}}_b = 0 \tag{20.51}$$

provides (after the usual substitution from the elastic stress-strain law, flow rule, and hardening law) only one equation for two unknowns – the plastic multiplier, $\dot{\lambda}$, and the multiplier from the hardening law, $\dot{\mu}$:

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\lambda}\boldsymbol{f}_\sigma) - \dot{\mu}\boldsymbol{f}_\sigma : (\boldsymbol{\sigma} - \boldsymbol{\sigma}_b) = 0 \tag{20.52}$$

Note that no parameter playing the role of a plastic modulus has been introduced yet, and the relationship between the stress rate and the plastic strain rate remains unspecified. It therefore makes sense to postulate a condition that the projection of the stress rate onto the direction of the normal to the yield surface is proportional to the projection of the plastic strain rate. This supplementary condition,

$$\dot{\boldsymbol{\sigma}} : \boldsymbol{f}_\sigma = \bar{H}_K\, \dot{\boldsymbol{\varepsilon}}_p : \boldsymbol{f}_\sigma \tag{20.53}$$

is in fact satisfied by the model with Melan–Prager hardening as well, which is verified by substituting the hardening law (20.43) into the consistency condition (20.51). Thus, the parameter $\bar{H}_K$ has the same meaning as in the Melan–Prager hardening rule.

Again, using the elastic stress-strain law and the flow rule, we obtain an equation with a single unknown,

$$\boldsymbol{f}_\sigma : \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\lambda}\boldsymbol{f}_\sigma) = \bar{H}_K \dot{\lambda} \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma \tag{20.54}$$

which is easily solved to give

$$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}_K \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma} \tag{20.55}$$

Upon substitution into (20.52), we finally obtain

$$\dot{\mu} = \bar{H}_K \frac{\boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma}{\boldsymbol{f}_\sigma : (\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)} \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}_K \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma} \tag{20.56}$$

Thus, for a prescribed strain rate, the rate of the back stress can be evaluated from (20.50) with $\dot{\mu}$ given by (20.56).

For the von Mises condition, Ziegler's rule leads to exactly the same results as the Melan–Prager rule. However, for other yield conditions the response of the model in general depends on the specific form of the hardening rule. According to Ziegler's rule, the yield surface moves in the direction connecting its current center with the current stress point, and this remains true in any reduced stress space.

20.1.3 Mixed Hardening

Mixed hardening models combine isotropic and kinematic hardening, which results into a modification of the loading surface by simultaneous translation and expansion (or contraction). It is easy to design a yield function

$$f(\boldsymbol{\sigma}, \boldsymbol{\sigma}_b, \kappa) = F(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b) - \bar{h}(\kappa) \tag{20.57}$$

that covers functions (20.12) and (20.42) as special cases. Derivation of the elastoplastic stiffness tensor

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H}_K \boldsymbol{f}_\sigma : \boldsymbol{f}_\sigma + \bar{H}k} \tag{20.58}$$

is left to the reader as an easy excercise. Generalization of the conditions of local uniqueness (20.37) and (20.48) is also straightforward.

Example 20.4: Develop an associated J_2-plasticity model combining linear isotropic and linear kinematic hardening.

Solution: The yield condition is given by

$$\sqrt{J_2(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)} - \bar{h}(\kappa) = 0 \tag{20.59}$$

For convenience, we rewrite it in the equivalent form

$$f(\boldsymbol{\sigma}, \boldsymbol{\sigma}_b, \kappa) \equiv \frac{1}{2}(\boldsymbol{s} - \boldsymbol{\sigma}_b) : (\boldsymbol{s} - \boldsymbol{\sigma}_b) - \bar{h}^2(\kappa) = 0 \tag{20.60}$$

The gradient of the yield function (20.60) with respect to the stress tensor is $\boldsymbol{f}_\sigma = \boldsymbol{s} - \boldsymbol{\sigma}_b$, and so the associated flow rule reads

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda}(\boldsymbol{s} - \boldsymbol{\sigma}_b) \tag{20.61}$$

Linear isotropic hardening can be described by

$$\bar{h}(\kappa) = \frac{\tau_0}{\sigma_0} h(\kappa) = \frac{\sqrt{3}}{3}(\sigma_0 + H_I \kappa) \tag{20.62}$$

where H_I is the (constant) isotropic plastic modulus, and $\kappa \equiv \bar{\varepsilon}_p$ is the cumulative plastic strain defined by (20.6). Linear kinematic hardening according to the Melan–Prager rule is characterized by a linear dependence of the back stress on the plastic strain,

$$\boldsymbol{\sigma}_b = \bar{H}_K \boldsymbol{\varepsilon}_p \tag{20.63}$$

□

20.1.4 Nonlinear Kinematic Hardening

Realistic modeling of engineering materials often requires nonlinear hardening laws. For isotropic hardening, nonlinearity can be easily incorporated through the function $h(\kappa)$, but for kinematic hardening this is not so simple. At first, one might think that it is sufficient to replace the constant $\bar{H}_K$ in (20.43) by a function of the cumulative

plastic strain. However, this would seldom lead to realistic shapes of the stress-strain diagrams under cyclic loading.

A suitable form of the nonlinear kinematic hardening law was proposed by Armstrong and Frederick (1966). They enriched the linear Melan–Prager rule (20.43) by a term proportional to the current back stress multiplied by the norm of the plastic strain rate. According to the *Armstrong–Frederick rule*, the evolution of the back stress is governed by the differential equation

$$\dot{\boldsymbol{\sigma}}_b = \bar{H}_K \dot{\boldsymbol{\varepsilon}}_p - \gamma \sqrt{\frac{2}{3}} \|\dot{\boldsymbol{\varepsilon}}_p\| \boldsymbol{\sigma}_b \tag{20.64}$$

where $\bar{H}_K$ and γ are constant material parameters. At the onset of yielding, the back stress is still zero and (20.64) gives the same response as the linear hardening law (20.43). As the back stress develops, the additional term becomes activated and slows down the rate at which the back stress grows (i.e. reduces the tangent plastic modulus), as long as the loading remains monotonic. Upon a load reversal, the back stress and its rate have opposite directions (in the sense that their scalar product is negative) and the additional term increases the plastic modulus. The following example shows that this interplay between the value of the back stress and its rate leads to reasonable shapes of the cyclic stress-strain diagrams.

Example 20.5: Plot the uniaxial cyclic stress-strain diagram for an elastoplastic model with a pressure-independent yield condition, associated flow rule and nonlinear kinematic hardening of the Armstrong Frederick type. The material is characterized by Young's modulus $E = 200$ GPa, initial uniaxial yield stress $\sigma_0 = 100$ MPa, initial plastic modulus $H_K = 150$ GPa, and parameter $\gamma = 3000$.

Solution: A flow rule associated with a pressure-insensitive yield condition results into purely deviatoric plastic flow. Therefore, the back stress will be a purely deviatoric tensor. Under uniaxial stress σ_x, the largest principal deviatoric stress, acting in the x-direction, is $s_x = \frac{2}{3}\sigma_x$. The corresponding principal value of the back stress will be conveniently denoted as $\sigma_{bx} = \frac{2}{3}\sigma_b$, because this notation leads to the uniaxial yield condition in the form $|\sigma_x - \sigma_b| = \sigma_0$.

If ε_{px} denotes the plastic strain in the x-direction, it follows from axial symmetry and from the condition of plastic incompressibility that the other principal plastic strains are $\varepsilon_{py} = \varepsilon_{pz} = -\frac{1}{2}\varepsilon_{px}$. The 'physical' plastic modulus H_K gives the slope of the curve relating the uniaxial stress σ_x to the plastic strain ε_{px}. To obtain the parameter $\bar{H}_K$, it must be multiplied by a factor 2/3; see Problem 20.5. The hardening law (20.64) written for the normal components in the x-direction and scaled by the factor 3/2 gives

$$\dot{\sigma}_b = H_K \dot{\varepsilon}_{px} - \gamma |\dot{\varepsilon}_{px}| \sigma_b \tag{20.65}$$

During each interval in which the plastic strain rate keeps the same sign, the particular solution of this differential equation satisfying the initial condition $\sigma_b(\varepsilon_{p0}) = \sigma_{b0}$ is given by

$$\sigma_b(\varepsilon_{px}) = \frac{H_K}{\gamma} \operatorname{sgn} \dot{\varepsilon}_{px} + \left(\sigma_{b0} - \frac{H_K}{\gamma} \operatorname{sgn} \dot{\varepsilon}_{px} \right) \exp\left[-\gamma (\varepsilon_{px} - \varepsilon_{p0}) \operatorname{sgn} \dot{\varepsilon}_{px} \right] \tag{20.66}$$

During plastic flow we have $\sigma_x - \sigma_b = \sigma_0 \operatorname{sgn} \dot{\varepsilon}_{px}$. Substituting $\sigma_b = \sigma_b(\varepsilon_{px})$ according to (20.66), we obtain an explicit formula for the stress σ_x as a function of the plastic

strain ε_{px}. The total strain is of course obtained by adding the elastic strain $\sigma_x(\varepsilon_{px})/E$ to the plastic strain ε_{px}.

For example, during monotonic tensile loading ($\operatorname{sgn} \dot{\varepsilon}_{px} = 1$) that starts from the initial virgin state ($\varepsilon_{p0} = 0$, $\sigma_{b0} = 0$), we have

$$\sigma_x(\varepsilon_{px}) = \sigma_0 + \sigma_b(\varepsilon_{px}) = \sigma_0 + \frac{H_K}{\gamma} \left[1 - \exp(-\gamma \varepsilon_{px}) \right] \tag{20.67}$$

$$\varepsilon_x(\varepsilon_{px}) = \frac{1}{E} \sigma_x(\varepsilon_{px}) + \varepsilon_{px} \tag{20.68}$$

This equations provide a parametric description of the stress-strain diagram plotted by the solid curve in Figure 20.3(a). For $\varepsilon_{px} \to \infty$, the stress asymptotically approaches a limit value $\sigma_0 + H_K/\gamma$. This means that the hardening process becomes saturated and the back stress cannot exceed $\sigma_{b\infty} = H_K/\gamma = 50$ MPa. This observation helps to clarify the physical meaning of the parameter γ.

If the loading direction is reversed at plastic strain ε_{p1} and back stress $\sigma_{b1} = \sigma_{b\infty}[1 - \exp(-\gamma \varepsilon_{p1})]$, the yielding in compression starts at $\sigma_x = \sigma_{b1} - \sigma_0$ and the subsequent evolution of stress is given by

$$\sigma_x(\varepsilon_{px}) = -\sigma_0 - \sigma_{b\infty} + (\sigma_{b1} + \sigma_{b\infty}) \exp\left[\gamma (\varepsilon_{px} - \varepsilon_{p1}) \right] \tag{20.69}$$

For $\varepsilon_{px} \to -\infty$, the stress approaches the asymptotic limit $-\sigma_0 - \sigma_{b\infty}$, which has the same magnitude as the asymptotic limit that corresponds to monotonic loading

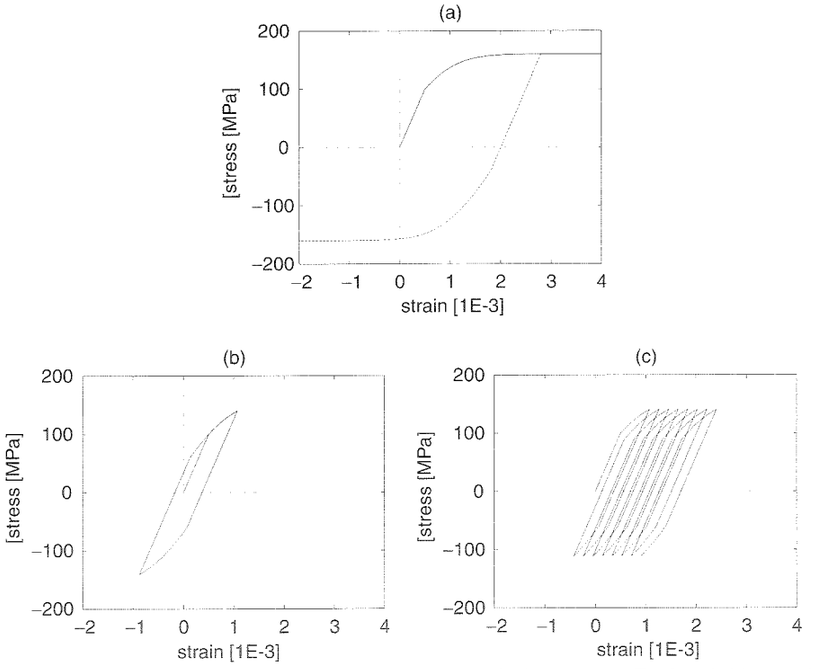


Figure 20.3 Stress-strain diagrams for a model with Armstrong–Frederick kinematic hardening: (a) monotonic loading and load reversal, (b) cyclic loading with a zero mean stress, (c) cyclic loading with a positive mean stress

from the virgin state; see the dashed curve in Figure 20.3(a). So, even though the hardening is kinematic and the onset of yielding upon load reversal occurs at a stress smaller in magnitude than the largest stress previously reached, the memory effect, after a sufficiently long monotonic yielding, fades away and the stress approaches a value independent of the previous loading history. This characteristic feature of the Armstrong–Frederick model can be used to check the suitability of that model for a particular material. In the general triaxial case, the state of saturated hardening (which is never reached exactly, only approached asymptotically) is characterized by a vanishing rate of the back stress. Substituting $\dot{\boldsymbol{\sigma}}_b = \mathbf{0}$ into (20.64), we obtain

$$\boldsymbol{\sigma}_b = \sqrt{\frac{2}{3}} \frac{H_K}{\gamma} \frac{\dot{\boldsymbol{\varepsilon}}_p}{\|\dot{\boldsymbol{\varepsilon}}_p\|} \tag{20.70}$$

from which

$$\|\boldsymbol{\sigma}_b\| = \sqrt{\frac{2}{3}} \frac{H_K}{\gamma} = \sqrt{\frac{2}{3}} \sigma_{b\infty} \tag{20.71}$$

So the parameter $\sigma_{b\infty}$ scaled by $\sqrt{2/3}$ characterizes the maximum distance by which the center of the yield surface can move in the stress space.

The stress-strain diagrams corresponding to cyclic loading, with stress varying in the range from −140 MPa to 140 MPa and from −110 MPa to 140 MPa are shown in Figures 20.3(b),(c). For symmetric stress cycles, the response immediately stabilizes and the strain history becomes periodic (Figure 20.3(b)). However, for stress cycles oscillating around a nonzero mean value, the net increment of plastic strain over one cycle is not zero (Figure 20.3(c)). This phenomenon, indeed observed in experiments, is called *ratchetting*. Note that this type of behavior could not be captured by the linear Melan–Prager hardening rule because a stabilized response ensues from this rule even if the mean value of stress is nonzero. In reality, of course, the net plastic strain increment over one cycle does not remain constant, and further refinements of the basic Armstrong–Frederick rule are needed to take that into account. □

20.1.5 General Hardening

All the hardening models discussed so far can be covered by the general yield condition

$$f(\boldsymbol{\sigma}, \boldsymbol{\kappa}) = 0 \tag{20.72}$$

and the general hardening law

$$\dot{\boldsymbol{\kappa}} = \dot{\lambda} \boldsymbol{k}(\boldsymbol{\sigma}, \boldsymbol{\kappa}) \tag{20.73}$$

where $\boldsymbol{\kappa}$ collects the hardening variables and $\boldsymbol{k}$ is a suitable function of the hardening variables and the stress. This function is in (20.73) multiplied by the plastic multiplier rate $\dot{\lambda}$, to make sure that the response of the model is not affected by any rescaling of the 'time' parameter, i.e. that the model remains rate-independent.

The mixed hardening model from Section 20.1.3 requires a scalar variable controling the isotropic hardening and a second-order tensor controling the kinematic hardening. The most straightforward choice would be $\boldsymbol{\kappa} = (\boldsymbol{\sigma}_b, \kappa)$, where $\boldsymbol{\sigma}_b$ is the back stress and κ is the cumulative plastic strain. However, the thermodynamic formulation of the model (Chapter 23) is more transparent if all the hardening variables have the

dimension of strain. Therefore, we introduce a tensorial variable $\boldsymbol{\varepsilon}_b = \boldsymbol{\sigma}_b / \bar{H}_K$, and we set $\boldsymbol{\kappa} = (\boldsymbol{\varepsilon}_b, \kappa)$. The yield condition can be written in the form

$$f(\boldsymbol{\sigma}, \boldsymbol{\varepsilon}_b, \kappa) \equiv F(\boldsymbol{\sigma} - \bar{H}_K \boldsymbol{\varepsilon}_b) - \bar{h}(\kappa) = 0 \tag{20.74}$$

The evolution of the hardening variables is defined by the rate equations

$$\dot{\boldsymbol{\varepsilon}}_b = \dot{\lambda} \boldsymbol{f}_\sigma(\boldsymbol{\sigma}, \boldsymbol{\varepsilon}_b) \tag{20.75}$$

$$\dot{\kappa} = \dot{\lambda} k(\boldsymbol{\sigma}, \boldsymbol{\varepsilon}_b, \kappa) \tag{20.76}$$

with the initial conditions $\boldsymbol{\varepsilon}_b = \mathbf{0}$ and $\kappa = 0$. With $k(\boldsymbol{\sigma}, \boldsymbol{\varepsilon}_b, \kappa)$ defined by (20.19) or (20.20), equation (20.76) covers the strain hardening and the work hardening hypotheses. Equation (20.75) has exactly the same form as the associated flow rule (20.25), and so the rate $\dot{\boldsymbol{\varepsilon}}_b$ is equal to the plastic strain rate $\dot{\boldsymbol{\varepsilon}}_p$.

Even though the values of $\boldsymbol{\varepsilon}_p$ and $\boldsymbol{\varepsilon}_b$ always remain equal, it is useful to consider them as two different physical quantities. This facilitates the generalization to nonlinear kinematic hardening. For example, the Armstrong–Frederick hardening rule (20.64) is obtained if (20.75) is replaced by

$$\dot{\boldsymbol{\varepsilon}}_b = \dot{\lambda} \left(\boldsymbol{f}_\sigma - \gamma \sqrt{\frac{2}{3}} \|\boldsymbol{f}_\sigma\| \boldsymbol{\varepsilon}_b \right) \tag{20.77}$$

To set up the consistency condition, we need to express the rate of the yield function in terms of the rates $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\kappa}}$. The notation is greatly simplified by the abstract operator $\bullet$, denoting the scalar product of $\dot{\boldsymbol{\kappa}}$ with the partial gradient $\boldsymbol{f}_\kappa = \partial f / \partial \boldsymbol{\kappa}$ (or any other similarly structured object). For instance, if $\boldsymbol{\kappa} = (\boldsymbol{\varepsilon}_b, \kappa)$, this operation must be interpreted as

$$\frac{\partial f}{\partial \boldsymbol{\kappa}} \bullet \dot{\boldsymbol{\kappa}} = \frac{\partial f}{\partial \boldsymbol{\varepsilon}_b} : \dot{\boldsymbol{\varepsilon}}_b + \frac{\partial f}{\partial \kappa} \dot{\kappa} \tag{20.78}$$

Using this compact notation, we can write the consistency condition in the convenient form

$$\dot{f} = \frac{\partial f}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial f}{\partial \boldsymbol{\kappa}} \bullet \dot{\boldsymbol{\kappa}} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} + \boldsymbol{f}_\kappa \bullet \dot{\boldsymbol{\kappa}} = 0 \tag{20.79}$$

After the usual substitutions from (20.24), (20.25) and (20.73), we get the formulae for the plastic multiplier rate,

$$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma - \boldsymbol{f}_\kappa \bullet \boldsymbol{k}} \tag{20.80}$$

and for the elastoplastic stiffness tensor,

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma - \boldsymbol{f}_\kappa \bullet \boldsymbol{k}} \tag{20.81}$$

Local uniqueness is guaranteed if $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma - \boldsymbol{f}_\kappa \bullet \boldsymbol{k} > 0$.

20.2 DRUCKER'S POSTULATE AND UNIQUENESS

In previous chapters, we have repeatedly encountered the postulate of maximum plastic dissipation, which is equivalent to the conditions of convexity and normality

and plays a key role in the derivation of powerful bound theorems of plastic analysis. Recall that this postulate is not a universally valid physical principle; it characterizes a certain class of material models with convenient mathematical properties. These so-called *standard materials* include many important models used in the engineering practice.

Another special class of elastoplastic models was defined by Drucker (1952, 1959) and called by him *stable materials*. Stability in Drucker's sense is stronger than convexity and normality. It permits us to extend the global uniqueness theorem from Section 16.3 to materials for which the yield surface evolves during the plastic flow, and showing that the response of a structure is unique not only in terms of stress but also in terms of strain and displacement.

Drucker's definition of stable materials can be motivated by the analysis of the work done in a quasi-cyclic process. Consider an infinitesimal volume of an elastoplastic material in a state characterized by stress $\boldsymbol{\sigma}^*$ and strain $\boldsymbol{\varepsilon}^*$. Suppose that, during the time interval $[0, T]$, an 'external agency' slowly applies some additional stress and then removes it. If the material remains elastic, it returns to its initial state. However, when plastic yielding takes place, the final strain $\boldsymbol{\varepsilon}(T)$ in general differs from the initial strain $\boldsymbol{\varepsilon}(0) = \boldsymbol{\varepsilon}^*$, even if the additional stress is completely removed, i.e. if $\boldsymbol{\sigma}(T) = \boldsymbol{\sigma}(0) = \boldsymbol{\sigma}^*$. This is why such a process is called a *stress quasi-cycle*.

Drucker's stability postulate requires the *excessive work* (or net work)

$$\Delta W = \int_0^T [\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^*] : \dot{\boldsymbol{\varepsilon}}(t)\, \mathrm{d}t \tag{20.82}$$

done by the additional stress, to be non-negative in any stress quasi-cycle. The physical idea behind this requirement is that if ΔW is negative, the 'external agency' can gain energy from the material and use it for further amplification of the disturbances. The postulate also tacitly assumes that the stress $\boldsymbol{\sigma}^*$ remains plastically admissible, at least if the stress $\boldsymbol{\sigma}(t)$ stays in some neighborhood of $\boldsymbol{\sigma}^*$. This assumption guarantees that the quasi-cycle can be closed in the stress space.

The strain increments can be decomposed into the elastic and plastic parts, and this decomposition is transfered to the excessive work,

$$\begin{aligned} \Delta W &= \int_0^T [\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^*] : \dot{\boldsymbol{\varepsilon}}(t)\, \mathrm{d}t \qquad (20.83)\\ &= \int_0^T [\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^*] : \dot{\boldsymbol{\varepsilon}}_e(t)\, \mathrm{d}t + \int_0^T [\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^*] : \dot{\boldsymbol{\varepsilon}}_p(t)\, \mathrm{d}t = \Delta W_e + \Delta W_p \end{aligned}$$

Due to the reversibility of elastic processes, the elastic strain $\boldsymbol{\varepsilon}_e(t) = \boldsymbol{C}_e : \boldsymbol{\sigma}(t)$ returns at the end of the cycle to its initial value, and the elastic part of the excessive work ΔW_e vanishes. This can be formally proven as follows:

$$\begin{aligned} \Delta W_e &= \int_0^T [\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^*] : \dot{\boldsymbol{\varepsilon}}_e(t)\, \mathrm{d}t = \int_0^T [\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^*] : \boldsymbol{C}_e : \dot{\boldsymbol{\sigma}}(t)\, \mathrm{d}t \qquad (20.84)\\ &= \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{1}{2} \boldsymbol{\sigma}(t) : \boldsymbol{C}_e : \boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^* : \boldsymbol{C}_e : \boldsymbol{\sigma}(t) \right] \mathrm{d}t\\ &= \tfrac{1}{2} \boldsymbol{\sigma}(T) : \boldsymbol{C}_e : \boldsymbol{\sigma}(T) - \tfrac{1}{2} \boldsymbol{\sigma}(0) : \boldsymbol{C}_e : \boldsymbol{\sigma}(0) - \boldsymbol{\sigma}^* : \boldsymbol{C}_e : [\boldsymbol{\sigma}(T) - \boldsymbol{\sigma}(0)] = 0 \end{aligned}$$

Consequently, the excessive work can be computed as the work done by the additional stress on the plastic strain increments;

$$\Delta W = \Delta W_p = \int_0^T [\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^*] : \dot{\boldsymbol{\varepsilon}}_p(t)\, \mathrm{d}t \tag{20.85}$$

In the absence of plastic flow, ΔW_p always vanishes, and the stress quasi-cycle is a true cycle, because both stress and strain return to their initial values. If ΔW_p is strictly positive in any stress quasi-cycle that is not purely elastic, the material is called *stable in Drucker's sense*. If ΔW_p is always non-negative, but vanishes for some stress quasi-cycle that involves plastic flow, the material is called *neutrally stable in Drucker's sense*. Both stable and neutrally stable materials are classified as *materials satisfying Drucker's postulate*.

The previous definitions are physically motivated but, for a given material model, it would be hard to check the sign of ΔW_p in all possible stress quasi-cycles. Fortunately, there are simpler criteria of stability in Drucker's sense. Stable materials are those that satisfy the inequalities

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}^*) : \dot{\boldsymbol{\varepsilon}}_p \geq 0 \tag{20.86}$$

$$\dot{\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}}_p > 0 \tag{20.87}$$

for any stress $\boldsymbol{\sigma}$ and stress rate $\dot{\boldsymbol{\sigma}}$ generating a nonzero plastic strain rate $\dot{\boldsymbol{\varepsilon}}_p$, and for any plastically admissible stress $\boldsymbol{\sigma}^*$.

Condition (20.86) is equivalent to the postulate of maximum plastic dissipation, which is in turn equivalent to the conditions of convexity and normality. Condition (20.87) makes sure that the elastic domain during plastic flow expands, at least locally (i.e. in a neighborhood of the current stress state), so that it is possible to return to the stress state from which the stress quasi-cycle started. Consequently, (20.87) can be considered as the definition of hardening (in the narrow sense, meaning the opposite of softening). In summary, material stability in Drucker's sense is equivalent to convexity, normality and hardening. Elastic-perfectly plastic materials satisfying convexity and normality are only neutrally stable.

Example 20.6: Show that, for elastic-perfectly plastic materials satisfying convexity and normality, (20.87) turns into an equality.

Solution: For elastic-perfectly plastic materials, the yield function depends only on the stress, and the rate of its change is

$$\dot{f} = \frac{\partial f}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} \tag{20.88}$$

Substituting the associated flow rule (15.34) into the left-hand side of (20.87) and using (20.88), we obtain

$$\dot{\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}}_p = \dot{\boldsymbol{\sigma}} : \boldsymbol{f}_\sigma \dot{\lambda} = \dot{f} \dot{\lambda} \tag{20.89}$$

According to the consistency condition (15.36), the product $\dot{f}\dot{\lambda}$ vanishes, and so $\dot{\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}}_p = 0$. □

Example 20.7: Check whether an associated elastoplasticity model with isotropic hardening satisfies condition (20.87).

Solution: Using the rate form of the elastic stress-strain law (20.24) and the associated flow rule (20.25), we can write

$$\dot{\boldsymbol{\varepsilon}}_p : \dot{\boldsymbol{\sigma}} = \dot{\lambda} \boldsymbol{f}_\sigma : \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\lambda} \boldsymbol{f}_\sigma) = \dot{\lambda} \left(\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} - \dot{\lambda} \boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma \right) \tag{20.90}$$

During plastic yielding we have $\dot{\lambda} > 0$, and so the sign of $\dot{\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}}_p$ depends on the expression in the parentheses on the right-hand side of (20.90). Since $\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma$ is always positive, condition (20.87) is equivalent to

$$\dot{\lambda} < \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma} \tag{20.91}$$

We recognize that the expression on the right-hand side of (20.91) corresponds to the formula for $\dot{\lambda}$ valid in perfect elastoplasticity, which confirms the result of the previous example. For associated elastoplasticity with isotropic hardening, the rate of the plastic multiplier is given by (20.30), and (20.91) can be rewritten as

$$\frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma + \bar{H} k} < \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{f}_\sigma} \tag{20.92}$$

In Section 20.1 we have established condition (20.37), which is necessary for uniqueness of the model response on the local level of one material point. If it is violated, the stress rate is not uniquely determined by the strain rate and the model cannot be used. Therefore, we can restrict attention to the case of a positive denominator on the left-hand side of (20.92). The numerator is also positive, or else the assumption of plastic yielding would not be valid. Consequently, (20.92) is equivalent to

$$\bar{H} k > 0 \tag{20.93}$$

Since k is a positive scaling constant, the model satisfies condition (20.87) if and only if the plastic modulus $\bar{H}$ is positive. So, Drucker's definition of hardening is consistent with the 'natural' definition in terms of the plastic modulus. □

The result of the previous example can be easily extended to the general class of hardening models described in Section 20.1.5. The plastic multiplier rate $\dot{\lambda}$ is in the general case given by (20.80), and so (20.93) is replaced by

$$-\boldsymbol{f}_\kappa \bullet \boldsymbol{k} > 0 \tag{20.94}$$

Drucker's postulate has important implications for global uniqueness. For stable materials, the structural response remains unique not only in terms of stresses but also in terms of strains and displacements. To show that, let us return to the proof of the global uniqueness theorem for perfectly plastic materials given in Section 16.3. The key step was the proof of the lemma asserting that if $\Delta\dot{\boldsymbol{\sigma}} = \dot{\boldsymbol{\sigma}}^A - \dot{\boldsymbol{\sigma}}^B$ is the difference between two stress rates and $\Delta\dot{\boldsymbol{\varepsilon}}_p = \dot{\boldsymbol{\varepsilon}}_p^A - \dot{\boldsymbol{\varepsilon}}_p^B$ is the difference between the corresponding plastic strain rates, then $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p \geq 0$. Recall that both solutions labeled by A and B start from the same initial stress state. The derivation of the lemma was based on the identity

$$\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p = \Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{f}_\sigma \, \Delta\dot{\lambda} = \Delta\dot{f} \, \Delta\dot{\lambda} \tag{20.95}$$

For hardening materials, $\Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{f}_\sigma$ is not equal to $\Delta\dot{f}$, because the rate of the yield function involves an additional term due to hardening:

$$\dot{f} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} + \boldsymbol{f}_\kappa \bullet \dot{\boldsymbol{\kappa}} = \boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} + \boldsymbol{f}_\kappa \bullet \boldsymbol{k} \dot{\lambda} \tag{20.96}$$

Let us denote $\tilde{H} = -\boldsymbol{f}_\kappa \bullet \boldsymbol{k}$ and rewrite (20.96) as

$$\boldsymbol{f}_\sigma : \dot{\boldsymbol{\sigma}} = \dot{f} + \tilde{H} \dot{\lambda} \tag{20.97}$$

Since $\boldsymbol{f}_\sigma$ and $\tilde{H}$ are the same for both solutions A and B, the difference between (20.97) written for $\boldsymbol{\sigma}^A$, $\dot{f}^A$, $\dot{\lambda}^A$ and is $\boldsymbol{\sigma}^B$, $\dot{f}^B$, $\dot{\lambda}^B$

$$\boldsymbol{f}_\sigma : \Delta\dot{\boldsymbol{\sigma}} = \Delta\dot{f} + \tilde{H} \Delta\dot{\lambda} \tag{20.98}$$

Consequently, we can write

$$\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p = \Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{f}_\sigma \, \Delta\dot{\lambda} = \Delta\dot{f} \, \Delta\dot{\lambda} + \tilde{H} (\Delta\dot{\lambda})^2 \tag{20.99}$$

The product $\Delta\dot{f} \, \Delta\dot{\lambda}$ was in Section 16.3 shown to be non-negative. If the material is stable in Drucker's sense, it follows from (20.94) that $\tilde{H} = -\boldsymbol{f}_\kappa \bullet \boldsymbol{k}$ is indeed a positive variable; it can be considered as a generalized hardening modulus. Therefore, the expression $\tilde{H} (\Delta\dot{\lambda})^2$ is positive, unless $\Delta\dot{\lambda} = 0$, in which case $\Delta\dot{\boldsymbol{\varepsilon}}_p = \boldsymbol{f}_\sigma \, \Delta\dot{\lambda} = \boldsymbol{0}$. Consequently, the product $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p$ is always non-negative, and vanishes only if $\Delta\dot{\boldsymbol{\varepsilon}}_p = \boldsymbol{0}$. As already mentioned in Section 16.3, the expression $\Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{C}_e : \Delta\dot{\boldsymbol{\sigma}}$ is always nonnegative and vanishes only if $\Delta\dot{\boldsymbol{\sigma}} = \boldsymbol{0}$. Combining these results, we conclude that

$$\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}} = \Delta\dot{\boldsymbol{\sigma}} : \boldsymbol{C}_e : \Delta\dot{\boldsymbol{\sigma}} + \Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}}_p \tag{20.100}$$

is always non-negative and vanishes only if $\Delta\dot{\boldsymbol{\sigma}} = \boldsymbol{0}$ and $\Delta\dot{\boldsymbol{\varepsilon}}_p = \boldsymbol{0}$. Since $\Delta\dot{\boldsymbol{\sigma}}$ is self-equilibrated and $\Delta\dot{\boldsymbol{\varepsilon}}$ is compatible, we have

$$\int_V \Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}} \, \mathrm{d}V = 0 \tag{20.101}$$

and, due to the non-negativity of the integrand, this is possible only if $\Delta\dot{\boldsymbol{\sigma}} : \Delta\dot{\boldsymbol{\varepsilon}} = 0$. For a perfectly plastic material we could conclude only that $\dot{\boldsymbol{\sigma}}^A = \dot{\boldsymbol{\sigma}}^B$, but for a hardening material we have also $\dot{\boldsymbol{\varepsilon}}_p^A = \dot{\boldsymbol{\varepsilon}}_p^B$, and from $\dot{\boldsymbol{\varepsilon}} = \boldsymbol{C}_e : \dot{\boldsymbol{\sigma}} + \dot{\boldsymbol{\varepsilon}}_p$ it follows that $\dot{\boldsymbol{\varepsilon}}^A = \dot{\boldsymbol{\varepsilon}}^B$. The response remains unique in terms of the stress and strain histories. Uniqueness of the displacement history follows from the uniqueness of the strain history, provided that the kinematic boundary conditions are sufficient to suppress rigid-body motions.

20.3 NONASSOCIATED FLOW

From the purely mathematical point of view it would be convenient to use only associated flow rules. Unfortunately, such rules do not always describe the real physical processes in a sufficiently realistic manner, especially if the material is pressure-sensitive. For example, a flow rule associated with the Mohr–Coulomb yield condition for soils usually overestimates the volumetric part of plastic strain. The source of the problem can be illustrated by a simple mechanical model of dry Coulomb friction. Slip on a contact surface between two bodies in Figure 20.4(a) is initiated when

$$|T| = -\mu N \tag{20.102}$$

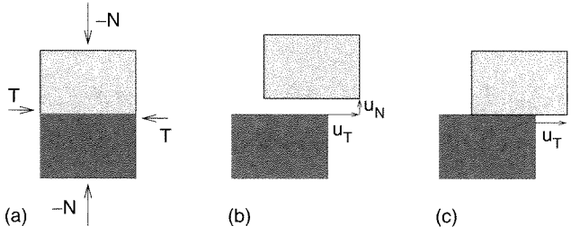


Figure 20.4 Coulomb friction model

where T is the tangential component of the contact force, N is the normal contact force ($N < 0$), and μ is the static coefficient of friction. The corresponding yield function can be written as

$$f(T, N) = |T| + \mu N \tag{20.103}$$

Kinematic variables work-conjugate to forces T and N are the opening (or gap) u_N and the relative slip u_T (the normal and tangential component of the relative displacement between the bodies), shown in Figure 20.4(b). The associated flow rule would read

$$\dot{u}_T = \dot{\lambda}\frac{\partial f}{\partial T} = \dot{\lambda}\ \text{sign}\ T \tag{20.104}$$

$$\dot{u}_N = \dot{\lambda}\frac{\partial f}{\partial N} = \dot{\lambda}\mu \tag{20.105}$$

This means that an associated flow should take place with a fixed ratio $\dot{u}_N/|\dot{u}_T| = \mu$, i.e. the normal relative displacement would be proportional to the tangential relative displacement. However, this is at odds with the 'natural' assumption that the bodies remain in contact and only tangential slip takes place as shown in Figure 20.4(c). A model exhibiting such behavior requires a nonassociated flow rule derived from a plastic potential

$$g(T, N) = |T| \tag{20.106}$$

that differs from the yield function (20.103).

The friction model with a zero gap is the most rudimentary one, based on our macroscopic 'everyday experience'. It assumes that the surfaces are 'rough' but the asperities are so small compared to the scale of observation that we still consider the surfaces as ideal planes that allow sliding in the tangential direction without any normal displacement. The roughness is reflected only by the frictional resistance that develops under normal pressure. A closer look at the interface motivates a different type of model that reflects the asperities directly as deviations from the ideal shape of the surface. One strongly idealized model of this type is sketched in Figure 20.5(a). The interface is represented by a zig-zag line consisting of straight segments with a constant deviation from the 'macroscopic' contact line. Locally, on the level of asperities, we could assume that the contact is frictionless and only forces normal to the contact segments can be transmitted. Due to the inclination of the segments, the interface can even transmit (without any slip) a force that is not perpendicular to the mean contact line, provided that its tangential component is sufficiently small; see Figure 20.5(b). Slip is initiated when the deviation of the global contact force from the normal to the mean contact line equals the deviation of the contact segments from

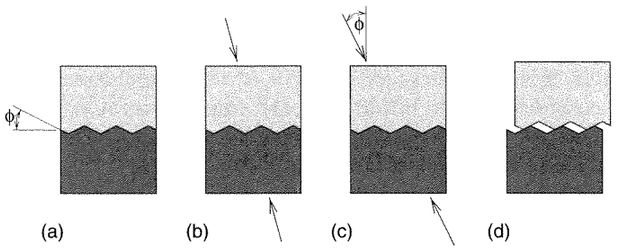


Figure 20.5 Friction model with asperities

the mean contact line (Figure 20.5(c)), i.e. if

$$\frac{|T|}{-N} = \tan\phi \tag{20.107}$$

This is the same condition as (20.102) if we identify $\tan\phi$ with the friction coefficient μ. The angle ϕ can be called the *friction angle*. A remarkable feature of the present model is that the slip not only has a component parallel to the mean contact line but also a component normal to it, and their ratio is obviously $\tan\phi = \mu$; see Figure 20.5(d). The model therefore obeys the associated flow rule (20.104)–(20.105).

The two models just described are certainly very schematic but they offer some insight into the behavior of frictional materials. Note that the model with geometric asperities would only give an associated flow in a certain range limited by the size of the asperities. After a sufficiently large relative displacement the behavior would approach that of the model with only tangential slip. The situation is also complicated by the fact that real asperities have variable sizes and shapes, and that they experience some plastic deformation and damage during sliding.

A number of additional factors should be taken into account when extending the models of friction on an interface between two blocks to materials that have no pre-existing internal interfaces. Here, a part of a potential slip plane may already be slipping and resisting the slip by friction, and the rest of that plane still retains cohesion. The cohesion is reflected by an additional constant term in the yield function. If the plastic deformation were exclusively due to slip, the plastic strain would be purely deviatoric. However, the presence of a relative displacement component normal to the sliding plane results into plastic changes of volume. The simple mechanical model of a surface with rigid asperities predicts an increase of volume during plastic flow. This effect is called *dilatancy* (occasionally also dilatation or dilation) and is observed for example in concrete under low confinement, in overconsolidated clays, or in dense sands. On the other hand, some materials such as concrete under very high confinement, normally consolidated clays, or loose sands exhibit a plastic decrease of volume, which is called *contractancy* (or negative dilatancy). It is often possible to assume that the flow is associated in the projection onto the deviatoric plane and only its volumetric part is nonassociated. A suitable plastic potential can be derived from the yield function by modifying only the term that reflects pressure sensitivity of the material. For example, from the Drucker–Prager criterion (15.20) we could derive a plastic potential

$$g(I_1, J_2) = \alpha_\psi I_1 + \sqrt{J_2} \tag{20.108}$$

where α_ψ is the *dilatancy coefficient*, in general different from the friction coefficient α that appears in the yield function. For $\alpha_\psi = \alpha$ we have an associated flow, for $\alpha_\psi = 0$ a purely deviatoric flow, for $\alpha_\psi \in (0, \alpha)$ a nonassociated dilatant flow, and for $\alpha_\psi < 0$ a contractive flow. In (20.108), the constant term $-\tau_0$ from the original yield function has been omitted because it has no effect on the gradient $\partial g/\partial \boldsymbol{\sigma}$ that determines the direction of plastic flow.

It is not difficult to generalize the formulae for the rate of plastic multiplier and the elastoplastic stiffness to the case of a nonassociated flow described by the flow rule

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda} \frac{\partial g}{\partial \boldsymbol{\sigma}} \tag{20.109}$$

To simplify the notation, let us denote the gradient of the plastic potential as $\boldsymbol{g}_\sigma$. Substituting (20.24), (20.73) and (20.109) into the consistency condition we get

$$\dot{f} = \frac{\partial f}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial f}{\partial \kappa} \bullet \dot{\kappa} = \boldsymbol{f}_\sigma : \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\lambda} \boldsymbol{g}_\sigma) + \boldsymbol{f}_\kappa \bullet k \dot{\lambda} = 0 \tag{20.110}$$

from which

$$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{g}_\sigma - \boldsymbol{f}_\kappa \bullet k} \tag{20.111}$$

The elastoplastic stiffness tensor is given by

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{g}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{g}_\sigma - \boldsymbol{f}_\kappa \bullet k} \tag{20.112}$$

Example 20.8: Derive the specific expression for the elastoplastic stiffness of a Drucker–Prager material with isotropic strain hardening and a nonassociated flow rule derived from the plastic potential (20.108). Discuss local uniqueness.

Solution: Using the chain rule from Appendix D.3, we obtain the gradient of the Drucker–Prager yield function,

$$\boldsymbol{f}_\sigma = \frac{\partial f}{\partial \boldsymbol{\sigma}} = \frac{\partial f}{\partial I_1} \frac{\partial I_1}{\partial \boldsymbol{\sigma}} + \frac{\partial f}{\partial J_2} \frac{\partial J_2}{\partial \boldsymbol{\sigma}} = \alpha \boldsymbol{\delta} + \frac{1}{2\sqrt{J_2}} \boldsymbol{s} \tag{20.113}$$

The gradient of the plastic potential,

$$\boldsymbol{g}_\sigma = \alpha_\psi \boldsymbol{\delta} + \frac{1}{2\sqrt{J_2}} \boldsymbol{s} \tag{20.114}$$

is obtained simply by replacing the friction coefficient α by the dilatancy coefficient α_ψ. For isotropic strain hardening we have $\boldsymbol{f}_\kappa \bullet k = -\bar{H} k$, where

$$k = \sqrt{\frac{2}{3}} \|\boldsymbol{g}_\sigma\| = \sqrt{\frac{2}{3} \left(\alpha_\psi \boldsymbol{\delta} + \frac{1}{2\sqrt{J_2}} \boldsymbol{s} \right) : \left(\alpha_\psi \boldsymbol{\delta} + \frac{1}{2\sqrt{J_2}} \boldsymbol{s} \right)} = \sqrt{2\alpha_\psi^2 + \frac{1}{3}} \tag{20.115}$$

because $\boldsymbol{\delta} : \boldsymbol{\delta} = 3$, $\boldsymbol{\delta} : \boldsymbol{s} = 0$, and $\boldsymbol{s} : \boldsymbol{s} = 2J_2$. Substituting into the general formula (20.112) and using the relations $\boldsymbol{D}_e : \boldsymbol{\delta} = 3K\boldsymbol{\delta}$ and $\boldsymbol{D}_e : \boldsymbol{s} = 2G\boldsymbol{s}$, we obtain the desired expression

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{9K^2 \alpha \alpha_\psi \boldsymbol{\delta} \otimes \boldsymbol{\delta} + \frac{3KG}{\sqrt{J_2}} (\alpha \boldsymbol{s} \otimes \boldsymbol{\delta} + \alpha_\psi \boldsymbol{\delta} \otimes \boldsymbol{s}) + \frac{G}{J_2} \boldsymbol{s} \otimes \boldsymbol{s}}{9K\alpha\alpha_\psi + G + \bar{H} \sqrt{2\alpha_\psi^2 + \frac{1}{3}}} \tag{20.116}$$

An inspection of the denominator reveals that the condition of local uniqueness $9K\alpha\alpha_\psi + G + \bar{H}k > 0$ can be violated even for a non-softening material ($\bar{H} \geq 0$) if the flow is contractive ($\alpha_\psi < 0$). The critical value of α_ψ for a perfectly plastic material is

$$\alpha_{\psi,crit} = -\frac{G}{9K\alpha} = -\frac{1}{6\alpha} \frac{1 - 2\nu}{1 + \nu} \tag{20.117}$$

□

Although for plasticity models with an associated flow rule the tangential stiffness tensor is symmetric, nonsymmetry should characterize a more realistic model. For two-dimensional behavior, this can be illustrated by a model of frictionally sliding blocks connected by springs, presented in Bažant and Cedolin (1991, Section 10.7, Figure 10.22). The nonsymmetry reflecting friction, however, must be of a form that causes no material instability. Generalizing from Mandel's (1964a) ingeneous example of a spring-loaded frictionally-sliding block, Bažant (1980) formulated the sufficient (albeit not necessary) conditions under which the tangential stiffness tensor can be nonsymmetric while causing no material instability.

Comparison with the aforementioned spring-block model further reveals that the modeling of internal friction as an influence of I_1 on $\sqrt{J_2}$, as done by the Drucker–Prager criterion, is a gross simplification. It cannot reflect the fact that frictional slip occurs on one or several preferred planes. It is an advantage of the microplane model (to be presented in Section 25.2) that it can capture this fact.

20.4 NON-SMOOTH AND MULTI-SURFACE PLASTICITY

Another issue that deserves attention is the treatment of singularities of a yield surface, e.g. of the corners (edges) of the Tresca hexagon (hexahedral prism). At such points, the yield surface does not have a uniquely defined normal because the gradient $\boldsymbol{f}_\sigma$ vanishes or even does not exist. If the flow rule is associated, the direction of plastic flow is not properly defined. These problems can be handled if the elastic domain bounded by a non-smooth yield surface is represented as an intersection of several domains bounded by smooth auxiliary surfaces.

20.4.1 Perfect Plasticity

We start by looking at the simple case of a perfectly elastoplastic material. For illustration, consider the criterion due to Tresca, with the yield function given by the product-like expression (15.14). The inequality $f(\boldsymbol{\sigma}) \leq 0$ is not an unambiguous description of the Tresca hexahedral prism because it also holds at points for which, for example, $\sigma_1 - \sigma_2 > 2\tau_0$, $\sigma_2 - \sigma_3 > 2\tau_0$, and $|\sigma_3 - \sigma_1| < 2\tau_0$. Therefore it is preferable not to describe the elastic domain in Figure 20.6 by a single inequality $f(\boldsymbol{\sigma}) < 0$ but by a set of three inequalities

$$f_1(\boldsymbol{\sigma}) \equiv (\sigma_2 - \sigma_3)^2 - 4\tau_0^2 < 0 \tag{20.118}$$

$$f_2(\boldsymbol{\sigma}) \equiv (\sigma_3 - \sigma_1)^2 - 4\tau_0^2 < 0 \tag{20.119}$$

$$f_3(\boldsymbol{\sigma}) \equiv (\sigma_1 - \sigma_2)^2 - 4\tau_0^2 < 0 \tag{20.120}$$

which checks the signs of functions f_1, f_2 and f_3 separately. Plastic flow occurs if at least one of these functions vanishes. The auxiliary yield surfaces for which

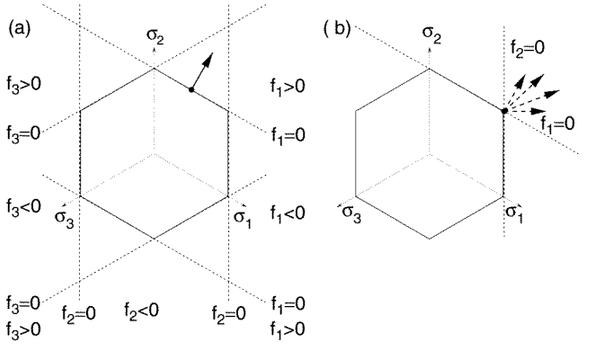


Figure 20.6 Deviatoric section of the Tresca yield surface: a) plastic flow at a regular point, b) plastic flow at a vertex

$f_I = 0$ are then called *active*. If, for example, only surface 1 is active ($f_1 = 0$ while $f_2 < 0$ and $f_3 < 0$, see Figure 20.6(a)), the stress state corresponds to a regular point of the yield surface, which has a uniquely defined normal determined by the gradient $\boldsymbol{f}_\sigma \equiv \partial f/\partial\boldsymbol{\sigma} = f_3 f_2 \partial f_1/\partial\boldsymbol{\sigma} + f_1\partial(f_2 f_3)/\partial\boldsymbol{\sigma} = f_3 f_2 \boldsymbol{f}_{1,\sigma}$, colinear with $\boldsymbol{f}_{1,\sigma} \equiv \partial f_1/\partial\boldsymbol{\sigma}$. However, it can happen that two surfaces are active simultaneously, e.g. $f_1 = 0$ and $f_2 = 0$ while $f_3 < 0$ (Figure 20.6(b)). Then the stress point is located at an edge of the hexahedral prism, where the original yield function f has a zero gradient, but the auxiliary functions f_1 and f_2 have nonzero gradients $\boldsymbol{f}_{1,\sigma}$ and $\boldsymbol{f}_{2,\sigma}$. It has already been explained in Section 15.2.1 that the postulate of maximum plastic dissipation remains valid if the direction of plastic flow is anywhere within the fan of directions between $\boldsymbol{f}_{1,\sigma}$ and $\boldsymbol{f}_{2,\sigma}$. This fan is formed by vectors that are linear combinations of $\boldsymbol{f}_{1,\sigma}$ and $\boldsymbol{f}_{2,\sigma}$ with non-negative coefficients. The corresponding flow rule reads

$$\dot{\boldsymbol{\varepsilon}}_p = \dot\lambda_1 \boldsymbol{f}_{1,\sigma} + \dot\lambda_2 \boldsymbol{f}_{2,\sigma} \tag{20.121}$$

where $\dot\lambda_1 \geq 0$ and $\dot\lambda_2 \geq 0$ are the rates of plastic multipliers associated with the auxiliary surfaces $f_1 = 0$ and $f_2 = 0$. To cover the general situation of yielding at an arbitrary point of the yield surface, we may write the flow rule as

$$\dot{\boldsymbol{\varepsilon}}_p = \sum_{I=1}^{N} \dot\lambda_I \boldsymbol{f}_{I,\sigma} \tag{20.122}$$

where N is the number of auxiliary yield surfaces. This is the famous *Koiter's rule* (Koiter, 1953b). A multiplier $\dot\lambda_I$ can be nonzero only if the stress point lies on the corresponding surface $f_I = 0$. This leads us to the generalized loading-unloading conditions

$$f_I \leq 0, \qquad \dot\lambda_I \geq 0, \qquad f_I \dot\lambda_I = 0, \qquad I = 1, 2, \ldots N \tag{20.123}$$

which again correspond to the Karush–Kuhn–Tucker conditions known from optimization theory (see Section 15.2.4). In $f_I\dot\lambda_I$ and similar expressions, the summation convention does not apply because the subscript I does not refer to tensorial components, but simply labels individual yield surfaces.

If, during a certain time interval, the Ith plastic multiplier is increasing ($\dot\lambda_I > 0$), the stress point must stay on the Ith auxiliary yield surface, and consequently $\dot f_I = 0$.

We thus obtain N consistency conditions

$$\dot f_I \dot\lambda_I = 0, \qquad I = 1, 2, \ldots N \tag{20.124}$$

which can be exploited while computing the rates of plastic multipliers $\dot\lambda_I$, $I = 1, 2, \ldots N$. Following the standard procedure, we can obtain the rates

$$\dot f_I = \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \left(\dot{\boldsymbol{\varepsilon}} - \sum_{J=1}^{N} \dot\lambda_J \boldsymbol{f}_{J,\sigma} \right), \qquad I = 1, 2, \ldots N \tag{20.125}$$

Note that the rate $\dot f_I$ depends not only on $\dot\lambda_I$ but, in general, on all $\dot\lambda_J$, $J = 1, 2, \ldots N$. If several yield surfaces are active simultaneously, the consistency conditions become coupled.

Suppose that the yield surfaces are renumbered such that surfaces number $1, 2, \ldots M$ are currently active while $M+1, M+2, \ldots N$ are not. The currently active surfaces are characterized by $f_I = 0$. They can either remain active, in which case $\dot f_I = 0$ and $\dot\lambda_I > 0$, or start unloading, in which case $\dot f_I < 0$ and $\dot\lambda_I = 0$. Both situations are covered by the consistency conditions

$$\dot f_I \leq 0, \qquad \dot\lambda_I \geq 0, \qquad \dot f_I \dot\lambda_I = 0, \qquad I = 1, 2, \ldots M \tag{20.126}$$

which are formally similar to (20.123). The difference is that (1) the values of the yield functions are replaced by their rates, and (2) conditions (20.123) hold for all the auxiliary yield surfaces while (20.126) only for the active ones.

It is convenient to introduce column matrices

$$\dot{\boldsymbol{f}} = \{\dot f_1, \dot f_2, \ldots \dot f_M\}^T, \qquad \dot{\boldsymbol{\lambda}} = \{\dot\lambda_1, \dot\lambda_2, \ldots \dot\lambda_M\}^T \tag{20.127}$$

and rewrite conditions (20.126) as

$$\dot{\boldsymbol{f}} \leq \boldsymbol{0}, \qquad \dot{\boldsymbol{\lambda}} \geq \boldsymbol{0}, \qquad \dot{\boldsymbol{f}}^T \dot{\boldsymbol{\lambda}} = 0 \tag{20.128}$$

Note that $\dot{\boldsymbol{f}}^T\dot{\boldsymbol{\lambda}} = \sum \dot f_I \dot\lambda_I$, and so the isolated condition $\dot{\boldsymbol{f}}^T\dot{\boldsymbol{\lambda}} = 0$ is not equivalent to $\dot f_I \dot\lambda_I$ vanishing for all $I = 1, 2, \ldots M$ (no sum over M). However, the ensemble of conditions (20.128) is equivalent to (20.126) because, if all $\dot f_I$ are nonpositive and all $\dot\lambda_I$ are nonnegative, then each term $\dot f_I \dot\lambda_I$ is nonpositive and the sum of such terms can vanish only if each individual term vanishes.

According to (20.125), the rate $\dot{\boldsymbol{f}}$ can be expressed as

$$\dot{\boldsymbol{f}} = \boldsymbol{b} - \boldsymbol{A}\dot{\boldsymbol{\lambda}} \tag{20.129}$$

where $\boldsymbol{b}$ is a column matrix with entries $b_I = \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}$, $I = 1, 2, \ldots M$, and $\boldsymbol{A}$ is an $M \times M$ matrix with entries

$$a_{IJ} = \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{J,\sigma}, \qquad I, J = 1, 2 \ldots M \tag{20.130}$$

Note that $a_{ij} = a_{ji}$, i.e. matrix $\boldsymbol{A}$ is symmetric. Substituting (20.129) into (20.128) we obtain the *linear complementarity problem* (LCP)

$$\boldsymbol{A}\dot{\boldsymbol{\lambda}} - \boldsymbol{b} \geq \boldsymbol{0}, \qquad \dot{\boldsymbol{\lambda}} \geq \boldsymbol{0}, \qquad \dot{\boldsymbol{\lambda}}^T(\boldsymbol{A}\dot{\boldsymbol{\lambda}} - \boldsymbol{b}) = \boldsymbol{0} \tag{20.131}$$

Linear complementarity problems have been extensively studied in optimization theory, because they are closely related to *quadratic programming* (minimization of quadratic functions subject to linear constraints). Sophisticated algorithms are available for the solution of such problems. The following theorem provides a necessary and sufficient condition of existence and uniqueness. Its proof can be found, for instance, in Cottle, Pang and Stone (1992).

Theorem 20.1
The LCP (20.131) has exactly one solution $\dot{\boldsymbol{\lambda}}$ *for any vector* $\boldsymbol{b}$ *if and only if all the principal minors of matrix* $\boldsymbol{A}$ *are positive.*

Recall that the principal minors of an $M \times M$ matrix are the determinants of all submatrices constructed by selecting a set $\mathcal{I} \subset \{1, 2, \dots M\}$ and deleting all the rows and columns whose number is contained in $\mathcal{I}$. For symmetric matrices, positivity of all the principal minors is equivalent to the *positive definiteness* of the matrix, i.e. to the condition that

$$\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > 0 \qquad \forall \boldsymbol{x} \neq \boldsymbol{0} \tag{20.132}$$

For general matrices, positivity of all the principal minors is equivalent to the so called *P-positivity*, i.e. to the condition

$$\forall \boldsymbol{x} \neq \boldsymbol{0} \;\; \exists\, i \in \{1, 2, \dots M\} : \qquad x_i (\boldsymbol{A}\boldsymbol{x})_i > 0 \qquad \text{(no sum over } i\text{)} \tag{20.133}$$

Here, x_i is the ith component of vector $\boldsymbol{x}$, and $(\boldsymbol{A}\boldsymbol{x})_i$ is the ith component of vector $\boldsymbol{A}\boldsymbol{x}$. Condition (20.133) means that there is no nonzero vector $\boldsymbol{x}$ for which multiplication by $\boldsymbol{A}$ would reverse the signs of all its nonzero components.

Theorem 20.1 facilitates the checks of uniqueness for plasticity models with multiple yield surfaces. Let us now apply the theorem to our plasticity model with a composite yield surface. Matrix $\boldsymbol{A}$ given by (20.130) is symmetric, and it is positive definite if

$$\begin{aligned} \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} &= \sum_{I,J=1}^{M} x_I a_{IJ} x_J = \sum_{I,J=1}^{M} x_I \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{J,\sigma} x_J \\ &= \left(\sum_{I=1}^{M} x_I \boldsymbol{f}_{I,\sigma} \right) : \boldsymbol{D}_e : \left(\sum_{J=1}^{M} x_J \boldsymbol{f}_{J,\sigma} \right) > 0 \qquad \forall \boldsymbol{x} \neq \boldsymbol{0} \end{aligned} \tag{20.134}$$

The final expression in (20.134) has the form $\boldsymbol{f} : \boldsymbol{D}_e : \boldsymbol{f}$ where

$$\boldsymbol{f} = \sum_{I=1}^{M} x_I \boldsymbol{f}_{I,\sigma} \tag{20.135}$$

Since $\boldsymbol{D}_e$ is a positive definite tensor, we have $\boldsymbol{f} : \boldsymbol{D}_e : \boldsymbol{f} > 0$ whenever $\boldsymbol{f} \neq \boldsymbol{0}$. So the positive definiteness of $\boldsymbol{A}$ can be proven if the sum in (20.135) is nonzero for any set of coefficients x_I, $I = 1, 2, \dots M$, that are not simultaneously zero. This condition means that the gradients $\boldsymbol{f}_{I,\sigma}$, $I = 1, 2, \dots M$, must be linearly independent. We have derived the following theorem:

Theorem 20.2
Consider associated perfect plasticity with a composite yield surface described by a set of auxiliary surfaces. If, at any singular point, the normals to all those auxiliary yield functions that intersect at that point are linearly independent, then the model gives a unique response (in terms of stress and plastic strain evolution) to any prescribed strain evolution.

The condition of linear independence limits the number of yield surfaces that may intersect at the same point. In three spatial dimensions, the normals belong to the six-dimensional space of symmetric second-order tensors, and so the maximum admissible number of simultaneously active surfaces is six. For plane stress, this number reduces to three.

If at most one yield surface is active, the problem reduces to the standard one, and the solution is given by (15.40). Let us now explore the case of two initially active yield surfaces.

Example 20.9: Set up and solve the linear complementarity problem describing associated plastic flow at an edge of the Tresca hexahedral prism, assuming that the material is perfectly elastoplastic.

Solution: The elastic domain corresponding to the Tresca criterion is described by (20.118)–(20.120). At most, two surfaces can be active simultaneously. Without any loss of generality, let us assume that the stress point is at the edge characterized by

$$\sigma_2 - \sigma_3 = 2\tau_0 \tag{20.136}$$

$$\sigma_1 - \sigma_3 = 2\tau_0 \tag{20.137}$$

and so surfaces 1 and 2 are active; see Figure 20.6(b). Differentiating the yield function $f_1(\boldsymbol{\sigma})$ defined in (20.118), we obtain

$$\frac{\partial f_1}{\partial \boldsymbol{\sigma}} = \sum_{I=1}^{3} \frac{\partial f_1}{\partial \sigma_I} \frac{\partial \sigma_I}{\partial \boldsymbol{\sigma}} = 2(\sigma_2 - \sigma_3) \left(\frac{\partial \sigma_2}{\partial \boldsymbol{\sigma}} - \frac{\partial \sigma_3}{\partial \boldsymbol{\sigma}} \right) \tag{20.138}$$

According to formula (D.61) derived in Appendix D, we have

$$\frac{\partial \sigma_2}{\partial \boldsymbol{\sigma}} = \boldsymbol{n}_2 \otimes \boldsymbol{n}_2 \qquad \text{and} \qquad \frac{\partial \sigma_3}{\partial \boldsymbol{\sigma}} = \boldsymbol{n}_3 \otimes \boldsymbol{n}_3 \tag{20.139}$$

where $\boldsymbol{n}_2$ and $\boldsymbol{n}_3$ are unit vectors in the principal directions associated with principal stresses σ_2 and σ_3, respectively. The gradient of f_1 is thus given by

$$\boldsymbol{f}_{1,\sigma} = 2(\sigma_2 - \sigma_3)(\boldsymbol{n}_2 \otimes \boldsymbol{n}_2 - \boldsymbol{n}_3 \otimes \boldsymbol{n}_3) \tag{20.140}$$

and a similar expression

$$\boldsymbol{f}_{2,\sigma} = 2(\sigma_1 - \sigma_3)(\boldsymbol{n}_1 \otimes \boldsymbol{n}_1 - \boldsymbol{n}_3 \otimes \boldsymbol{n}_3) \tag{20.141}$$

can be derived for the gradient of f_2.

Evaluating the entries of matrix $\boldsymbol{A}$, we have to deal with terms of the type $(\boldsymbol{n}_I \otimes \boldsymbol{n}_I) : \boldsymbol{D}_e : (\boldsymbol{n}_J \otimes \boldsymbol{n}_J)$. Such expressions are scalars, i.e. they are invariant

with respect to the choice of the coordinate system, and they can be conveniently evaluated in the principal stress coordinates. For the isotropic elastic stiffness tensor $\boldsymbol{D}_e$ we obtain

$$(\boldsymbol{n}_2 \otimes \boldsymbol{n}_2) : \boldsymbol{D}_e : (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) = D^e_{2211} = K - \frac{2}{3}G \tag{20.142}$$

$$(\boldsymbol{n}_3 \otimes \boldsymbol{n}_3) : \boldsymbol{D}_e : (\boldsymbol{n}_3 \otimes \boldsymbol{n}_3) = D^e_{3333} = K + \frac{4}{3}G \tag{20.143}$$

Now it is easy to evaluate

$$\begin{aligned} a_{11} &= \boldsymbol{f}_{1,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{1,\sigma} = 4(\sigma_2 - \sigma_3)^2 (D^e_{2222} + D^e_{3333} - D^e_{2233} - D^e_{3322}) \\ &= 16G(\sigma_2 - \sigma_3)^2 \end{aligned} \tag{20.144}$$

$$\begin{aligned} a_{12} &= \boldsymbol{f}_{1,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{2,\sigma} = 4(\sigma_2 - \sigma_3)(\sigma_1 - \sigma_3)(D^e_{2211} + D^e_{3333} - D^e_{2233} - D^e_{3311}) \\ &= 8G(\sigma_2 - \sigma_3)(\sigma_1 - \sigma_3) \end{aligned} \tag{20.145}$$

$$\begin{aligned} a_{22} &= \boldsymbol{f}_{2,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{2,\sigma} = 4(\sigma_1 - \sigma_3)^2 (D^e_{1111} + D^e_{3333} - D^e_{1133} - D^e_{3311}) \\ &= 16G(\sigma_1 - \sigma_3)^2 \end{aligned} \tag{20.146}$$

Making use of (20.136) and (20.137), we finally obtain

$$\boldsymbol{A} = 32G\tau_0^2 \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{20.147}$$

This matrix is positive definite, and so the problem has a unique solution for any prescribed strain rate.

The first component of the column matrix $\boldsymbol{b}$ is given by

$$\begin{aligned} b_1 &= \boldsymbol{f}_{1,\sigma} : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} = 2(\sigma_2 - \sigma_3)(\boldsymbol{n}_2 \otimes \boldsymbol{n}_2 - \boldsymbol{n}_3 \otimes \boldsymbol{n}_3) : \boldsymbol{D}_e : (\dot{\varepsilon}_V \boldsymbol{\delta} + \dot{\boldsymbol{e}}) \\ &= 4\tau_0 (\boldsymbol{n}_2 \otimes \boldsymbol{n}_2 - \boldsymbol{n}_3 \otimes \boldsymbol{n}_3) : (3K\dot{\varepsilon}_V \boldsymbol{\delta} + 2G\dot{\boldsymbol{e}}) \end{aligned} \tag{20.148}$$

Realizing that $(\boldsymbol{n}_2 \otimes \boldsymbol{n}_2) : \boldsymbol{\delta} = \boldsymbol{n}_2 \cdot \boldsymbol{\delta} \cdot \boldsymbol{n}_2 = \boldsymbol{n}_2 \cdot \boldsymbol{n}_2 = 1$, $(\boldsymbol{n}_2 \otimes \boldsymbol{n}_2) : \dot{\boldsymbol{e}} = \boldsymbol{n}_2 \cdot \dot{\boldsymbol{e}} \cdot \boldsymbol{n}_2 = \dot{e}_{22}$, etc., we can simplify (20.148) to $b_1 = 8G\tau_0(\dot{e}_{22} - \dot{e}_{33})$. Here, $\dot{e}_{22}$ and $\dot{e}_{33}$ are the components of the deviatoric strain rate expressed in the principal stress coordinates. Evaluating b_2 in an analogous way, we finally obtain

$$\boldsymbol{b} = 8G\tau_0 \begin{Bmatrix} \dot{e}_{22} - \dot{e}_{33} \\ \dot{e}_{11} - \dot{e}_{33} \end{Bmatrix} \tag{20.149}$$

The linear complementarity problem (20.131) is now fully specified, and we can analyze its solutions. Three basic situations can be distinguished:

1. *Both surfaces remain active*, i.e. the stress point remains at the edge of the yield surface. In this case, $\dot{\lambda}_1$ and $\dot{\lambda}_2$ can be solved from the set of linear equations $\boldsymbol{A}\dot{\boldsymbol{\lambda}} = \boldsymbol{b}$ with $\boldsymbol{A}$ and $\boldsymbol{b}$ given by (20.147) and (20.149), respectively. The solution

$$\dot{\boldsymbol{\lambda}} = \frac{1}{12\tau_0} \begin{Bmatrix} 2\dot{e}_{22} - \dot{e}_{11} - \dot{e}_{33} \\ 2\dot{e}_{11} - \dot{e}_{22} - \dot{e}_{33} \end{Bmatrix} \tag{20.150}$$

is admissible if both $\dot{\lambda}_1$ and $\dot{\lambda}_2$ are non-negative, i.e. if

$$2\dot{e}_{22} - \dot{e}_{11} - \dot{e}_{33} \geq 0 \tag{20.151}$$

$$2\dot{e}_{11} - \dot{e}_{22} - \dot{e}_{33} \geq 0 \tag{20.152}$$

2. (a) Only *one surface remains active*, say surface 1. In this case, the stress point leaves the edge and moves to the regular part of the yield surface on which $f_1 = 0$. The rate $\dot{\lambda}_2$ vanishes, and $\dot{\lambda}_1$ is computed from $a_{11}\dot{\lambda}_1 = b_1$. The solution

$$\dot{\boldsymbol{\lambda}} = \frac{1}{8\tau_0} \begin{Bmatrix} \dot{e}_{22} - \dot{e}_{33} \\ 0 \end{Bmatrix} \tag{20.153}$$

is admissible if $\dot{\lambda}_1 \geq 0$ and $a_{21}\dot{\lambda}_1 - b_2 \geq 0$, which is equivalent to

$$\dot{e}_{22} - \dot{e}_{33} \geq 0 \tag{20.154}$$

$$2\dot{e}_{11} - \dot{e}_{22} - \dot{e}_{33} \leq 0 \tag{20.155}$$

(b) A complementary situation arises when only surface 2 remains active. The admissibility conditions for this solution are

$$2\dot{e}_{22} - \dot{e}_{11} - \dot{e}_{33} \leq 0 \tag{20.156}$$

$$\dot{e}_{11} - \dot{e}_{33} \geq 0 \tag{20.157}$$

3. *No surface remains active*, i.e. the stress point moves to the elastic domain and no plastic flow takes place. Both rates $\dot{\lambda}_1$ and $\dot{\lambda}_2$ are zero. The solution $\dot{\boldsymbol{\lambda}} = \boldsymbol{0}$ is admissible if $\boldsymbol{A}\dot{\boldsymbol{\lambda}} - \boldsymbol{b} = -\boldsymbol{b} \geq \boldsymbol{0}$, i.e. if

$$\dot{e}_{22} - \dot{e}_{33} \leq 0 \tag{20.158}$$

$$\dot{e}_{11} - \dot{e}_{33} \leq 0 \tag{20.159}$$

Comparing the admissibility conditions for the previous solutions, we realize that they decompose the strain rate space into four sectors with disjoint interiors. This confirms that the response to any prescribed strain rate is unique, as was expected owing to the *positive definiteness* of matrix $\boldsymbol{A}$.

The results have a clear geometrical interpretation. The four sectors are plotted in the deviatoric section of the principal stress space in Figure 20.7. Note that, in the present case, the solution depends only on the deviatoric part of the strain rate, $\dot{\boldsymbol{e}}$. This is natural since the yield criterion is pressure-insensitive and, if an associated flow rule is used, the volumetric response is purely elastic and completely decomposed from the deviatoric response. Moreover, only the normal components of $\dot{\boldsymbol{e}}$ with respect

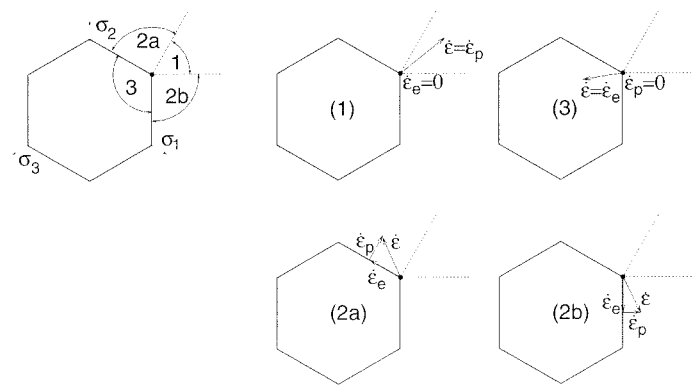


Figure 20.7 Plastic flow at a corner of Tresca yield surface

to the principal stress coordinates have an effect on the solution. These components are linked by the condition $\dot{e}_{11}+\dot{e}_{22}+\dot{e}_{33}=0$, and so only two of them are independent and the presentation in the two-dimensional deviatoric plane covers all the possible situations. The deviatoric part of the elastic stress rate is colinear with the deviatoric part of the strain rate because the elastic law has the simple form $\dot{\boldsymbol{s}} = 2G\dot{\boldsymbol{e}}$. We can, therefore, interpret the vectors (starting from the corner of the yield surface in Figure 20.7) alternatively as representing the strain rates or the stress rates.

Now we can discuss the four types of solutions. In case (1), the strain rate (and also the elastic stress rate) points to the sector between the normals to the straight segments of the yield surface. It is therefore possible to express the strain rate as a linear combination of the normal vectors with nonnegative coefficients. The total strain rate is equal to the plastic strain rate and the elastic strain rate vanishes. Consequently, the stress rate must also vanish and the stress point remains at the corner. In case (3), the elastic stress rate points to the interior of the elastic domain and is identical to the actual stress rate. No plastic flow takes place and the process is fully elastic. Finally, in case (2), the strain rate points to one of the remaining two sectors. Its orthogonal projection on the yield surface is located either on the segment where $f_1 = 0$ or on that where $f_2 = 0$. The projection on the yield surface represents the elastic part of the strain rate and the component normal to the yield surface is the plastic part. The stress point moves from the corner along one or the other straight segment of the yield surface.

20.4.2 Hardening Plasticity

The next challenge is an extension of the previous results to hardening plasticity. A crucial point to note is that, during hardening, two cases can occur: (1) either the auxiliary surfaces expand or move simultaneously, driven by the same hardening variables; or (2) each of the surfaces is associated with its own hardening variables that only grow if the surface is active. The former case corresponds to plasticity theory with a single non-smooth surface, in which the auxiliary surfaces only describe individual smooth portions of the basic surface. An example is Tresca or Mohr–Coulomb plasticity with isotropic hardening or softening. The latter case is referred to as *multi-surface plasticity*, or plasticity theory with multiple yield conditions. Here, each surface represents a different mechanism of plastic flow, independent of the others and controled by its own internal variables (though couplings can be introduced, too). This class of models includes plasticity-based microplane models (Carol and Bažant, 1997). Another example is a model based on a combination of the Drucker–Prager yield surface (representing failure under predominantly compressive stresses) with the Rankine surface (representing tensile cracking), recently proposed for concrete by Feenstra and de Borst (1996).

Multiple yield surfaces are also used to model nonlinear kinematic hardening by a set of nested surfaces, each of which only exhibits linear hardening (Iwan, 1967; Mróz, 1967, 1969). However, this is a somewhat different problem because the individual surfaces do not intersect each other, and all the active surfaces have a common tangent at the current stress point.

Let us start with the *non-smooth single-surface plasticity*, only for convenience described by a set of auxiliary functions

$$f_I(\boldsymbol{\sigma},\kappa) = F_I(\boldsymbol{\sigma}) - \bar{h}(\kappa), \qquad I = 1,2,\dots N \tag{20.160}$$

The time derivatives of these functions are

$$\dot{f}_I(\boldsymbol{\sigma},\kappa) = \frac{\partial F_I}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} - \frac{\mathrm{d}\bar{h}}{\mathrm{d}\kappa}\dot{\kappa} = \boldsymbol{f}_{I,\sigma} : \dot{\boldsymbol{\sigma}} - \bar{H}\dot{\kappa}, \qquad I = 1,2,\dots N \tag{20.161}$$

We will need an expression for the rate of the hardening parameter, $\dot{\kappa}$, in terms of the plastic multiplier rates, $\dot{\lambda}_I$ $(I = 1,2,\dots N)$, that appear in the associated flow rule (20.122). If we decided to use the strain hardening model, the rate of the hardening parameter

$$\dot{\kappa} = \sqrt{\frac{2}{3}\dot{\boldsymbol{\varepsilon}}_p : \dot{\boldsymbol{\varepsilon}}_p} = \sqrt{\frac{2}{3}\sum_{I,J=1}^{N} \dot{\lambda}_I\dot{\lambda}_J \boldsymbol{f}_{I,\sigma} : \boldsymbol{f}_{J,\sigma}} \tag{20.162}$$

would be given by an expression that is nonlinear in terms of the plastic multiplier rates. This would certainly complicate the solution procedure. A more convenient choice is the work hardening hypothesis analogous to (20.8), which leads to a linear expression

$$\dot{\kappa} = \frac{\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p}{h(\kappa)} = \sum_{I=1}^{N} \frac{\boldsymbol{\sigma} : \boldsymbol{f}_{I,\sigma}}{h(\kappa)}\dot{\lambda}_I = \sum_{I=1}^{N} k_I\dot{\lambda}_I \tag{20.163}$$

where $k_I = \boldsymbol{\sigma} : \boldsymbol{f}_{I,\sigma}/h(\kappa)$.

Substituting (20.122), (20.163) and the elastic stress-strain law into (20.161), we obtain

$$\begin{aligned}\dot{f}_I &= \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \left(\dot{\boldsymbol{\varepsilon}} - \sum_{J=1}^{M} \boldsymbol{f}_{J,\sigma}\dot{\lambda}_J\right) - \bar{H}\sum_{J=1}^{M} k_J\dot{\lambda}_J \\ &= \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} - \sum_{J=1}^{M}(\boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{J,\sigma} + \bar{H}k_J)\dot{\lambda}_J, \qquad I = 1,2,\dots M\end{aligned} \tag{20.164}$$

We took into account only the active surfaces $1,2,\dots M$ because $\dot{\lambda}_I = 0$ for $I = M+1, M+2,\dots N$.

The consistency conditions (20.128) again lead to a linear complementarity problem of the form (20.131) but this time with matrix $\boldsymbol{A}$ having components $a_{IJ} = \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{J,\sigma} + \bar{H}k_J$ which could be nonsymmetric. However, for Tresca, and Mohr–Coloumb yield criteria, the product $\sigma : f_{J,\sigma}$ is the same for all the active surfaces J, and so all coefficients $K_J (J = 1,2,\dots M)$ have the same value and symmetry preserved. The postulate of maximum plastic dissipation extended to hardening materials naturally leads to a modified form of the hardening law, $\dot{\kappa} = \sum \dot{\lambda}_I$, for which the matrix $\boldsymbol{A}$ is always symmetric; see Simo and Hughes (1998) and Problem 23.9.

Let us shift attention to *multi-surface plasticity*. Each yield surface now represents a specific failure mechanism, and the evolution of the surface is controled by a hardening variable associated only with this mechanism. For the sake of simplicity, we only consider isotropic hardening of each individual loading surface. However, note that as each surface evolves at its own pace, the resulting change of the composite surface is not isotropic. The individual loading surfaces are described by

$$f_I(\boldsymbol{\sigma},\kappa_I) \equiv F_I(\boldsymbol{\sigma}) - \bar{h}_I(\kappa_I) = 0, \qquad I = 1,2,\dots N \tag{20.165}$$

where the hardening functions $\bar{h}_I$ $(I = 1,2,\dots N)$ are in general different. We can start from the assumption that the inactive surfaces remain stationary, i.e. they exhibit

neither hardening nor softening. Consequently, the hardening variable associated with an inactive surface should remain constant. It is therefore natural to relate the rate of hardening variable associated with a certain surface only to the part of the plastic strain rate (or of the plastic work rate) due to yielding on that surface. This leads us to the evolution equation for the hardening variables,

$$\dot\kappa_I = \|\dot\lambda_I \boldsymbol{f}_{I,\sigma}\| = \dot\lambda_I \|\boldsymbol{f}_{I,\sigma}\| \equiv \dot\lambda_I k_I(\boldsymbol{\sigma}), \qquad I = 1, 2, \dots N \tag{20.166}$$

for strain hardening, and

$$\dot\kappa_I = \frac{1}{\bar h(\kappa_I)} \boldsymbol{\sigma} : \dot\lambda_I \boldsymbol{f}_{I,\sigma} = \dot\lambda_I \frac{\boldsymbol{\sigma} : \boldsymbol{f}_{I,\sigma}}{\bar h(\kappa_I)} \equiv \dot\lambda_I k_I(\boldsymbol{\sigma}, \kappa_I), \qquad I = 1, 2, \dots N \tag{20.167}$$

for work hardening. Evaluating the time derivatives of the yield functions,

$$\dot f_I = \boldsymbol{f}_{I,\sigma} : \dot{\boldsymbol{\sigma}} - \frac{\mathrm{d}\bar h_I}{\mathrm{d}\kappa_I}\dot\kappa_I = \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} - \sum_{J=1}^{M} (\boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{J,\sigma}) \dot\lambda_J - \bar H_I k_I \dot\lambda_I \tag{20.168}$$

we find that the coefficients of matrix $\boldsymbol{A}$ from (20.129) are now given by

$$a_{IJ} = \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{J,\sigma} + \bar H_I k_I \delta_{IJ}, \qquad I, J = 1, 2, \dots M \tag{20.169}$$

where δ_{IJ} is Kronecker delta but no summation over repeated subscripts is implied. Again, we only took into account the active surfaces $1, 2, \dots M$ because $\dot\lambda_I = 0$ for $I = M+1, M+2, \dots N$. Now the matrix $\boldsymbol{A}$ is again symmetric and, compared to the case of perfect plasticity with a_{IJ} given by (20.130), only the diagonal terms are modified. For an arbitrary vector $\boldsymbol{x}$ we have

$$\begin{aligned} \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} &= \sum_{I,J=1}^{M} x_I \boldsymbol{f}_{I,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{J,\sigma} x_J + \sum_{I=1}^{M} \bar H_I k_I x_I^2 \\ &= \left(\sum_{I=1}^{M} x_I \boldsymbol{f}_{I,\sigma} \right) : \boldsymbol{D}_e : \left(\sum_{I=1}^{M} x_I \boldsymbol{f}_{I,\sigma} \right) + \sum_{I=1}^{M} \bar H_I k_I x_I^2 \end{aligned} \tag{20.170}$$

If all the plastic moduli H_I are positive, matrix $\boldsymbol{A}$ with components given by (20.169) is positive definite even for linearly dependent gradients $\boldsymbol{f}_{I,\sigma}$, and the corresponding linear complementarity problem (20.131) has a unique solution. With zero or negative plastic moduli, uniqueness might be lost.

Example 20.10: Find the condition of local uniqueness for a model combining the Drucker–Prager criterion with the Rankine criterion. Assume that the evolution of each yield surface is described by a work hardening law with no interaction between the two hardening mechanisms.

Solution: The given model uses two yield functions[1]

$$f_1(\boldsymbol{\sigma}, \kappa_1) = \alpha I_1(\boldsymbol{\sigma}) + \sqrt{J_2(\boldsymbol{\sigma})} - \bar h_1(\kappa_1) \tag{20.171}$$

[1] Strictly speaking, the Rankine surface is also a composite surface, because it has singularities at points where two or three principal stresses are equal. Our aim here is to illustrate the methodology, and so we focus on the intersection of the Drucker–Prager surface with the Rankine surface (at a regular point of the Rankine surface), and we leave a detailed analysis of other possible singular points to the reader as an exercise; see Problem 20.8.

$$f_2(\boldsymbol{\sigma}, \kappa_2) = \sigma_1(\boldsymbol{\sigma}) - \bar h_2(\kappa_2) \tag{20.172}$$

Their gradients with respect to the stress tensor are easily computed as

$$\boldsymbol{f}_{1,\sigma} = \alpha \frac{\partial I_1}{\partial \boldsymbol{\sigma}} + \frac{1}{2\sqrt{J_2}} \frac{\partial J_2}{\partial \boldsymbol{\sigma}} = \alpha \boldsymbol{\delta} + \frac{1}{2\sqrt{J_2}} \boldsymbol{s} \tag{20.173}$$

$$\boldsymbol{f}_{2,\sigma} = \frac{\partial \sigma_1}{\partial \boldsymbol{\sigma}} = \boldsymbol{n}_1 \otimes \boldsymbol{n}_1 \tag{20.174}$$

where $\boldsymbol{n}_1$ is a unit vector in the direction of maximum principal stress. Now we evaluate

$$\boldsymbol{D}_e : \boldsymbol{f}_{1,\sigma} = 3K\alpha \boldsymbol{\delta} + \frac{G}{\sqrt{J_2}} \boldsymbol{s} \tag{20.175}$$

$$\boldsymbol{D}_e : \boldsymbol{f}_{2,\sigma} = \boldsymbol{D}_e : (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) \tag{20.176}$$

$$\boldsymbol{f}_{1,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{1,\sigma} = 9K\alpha^2 + G \tag{20.177}$$

$$\boldsymbol{f}_{2,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{1,\sigma} = 3K\alpha (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) : \boldsymbol{\delta} + \frac{G}{\sqrt{J_2}} (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) : \boldsymbol{s} = 3K\alpha + \frac{G s_1}{\sqrt{J_2}} \tag{20.178}$$

$$\boldsymbol{f}_{2,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{2,\sigma} = (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) : \boldsymbol{D}_e : (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) = D^e_{1111} = K + \frac{4}{3} G \tag{20.179}$$

$$\boldsymbol{\sigma} : \boldsymbol{f}_{1,\sigma} = \alpha I_1 + \frac{1}{2\sqrt{J_2}} 2J_2 = \alpha I_1 + \sqrt{J_2} \tag{20.180}$$

$$\boldsymbol{\sigma} : \boldsymbol{f}_{2,\sigma} = \boldsymbol{\sigma} : (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) = \sigma_1 \tag{20.181}$$

where s_1 is the maximum principal deviatoric stress. If surface 1 is active, we have $\alpha I_1 + \sqrt{J_2} - \bar h_1 = 0$, and so

$$k_1 = \frac{\boldsymbol{\sigma} : \boldsymbol{f}_{1,\sigma}}{\bar h_1} = \frac{\alpha I_1 + \sqrt{J_2}}{\bar h_1} = 1 \tag{20.182}$$

A similar result, $k_2 = 1$, is obtained also for surface 2. We can see that, in the present case, the hardening variables κ_I are identical with the plastic multipliers λ_I.

The existence and uniqueness of the solution is guaranteed if all the principal minors of the matrix

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} \boldsymbol{f}_{1,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{1,\sigma} + \bar H_1 k_1 & \boldsymbol{f}_{1,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{2,\sigma} \\ \boldsymbol{f}_{2,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{1,\sigma} & \boldsymbol{f}_{2,\sigma} : \boldsymbol{D}_e : \boldsymbol{f}_{2,\sigma} + \bar H_2 k_2 \end{bmatrix} \\ &= \begin{bmatrix} 9K\alpha^2 + G + \bar H_1 & 3K\alpha + G s_1/\sqrt{J_2} \\ 3K\alpha + G s_1/\sqrt{J_2} & K + 4G/3 + \bar H_2 \end{bmatrix} \end{aligned} \tag{20.183}$$

are positive. The principal minors of a 2×2 matrix are the diagonal coefficients and the determinant. This leads to three inequalities

$$9K\alpha^2 + G + \bar H_1 > 0 \tag{20.184}$$

$$K + \tfrac{4}{3} G + \bar H_2 > 0 \tag{20.185}$$

$$(9K\alpha^2 + G + \bar H_1)(K + \tfrac{4}{3} G + \bar H_2) > \left(3K\alpha + \frac{G s_1}{\sqrt{J_2}} \right)^2 \tag{20.186}$$

The first two conditions set a lower bound on $\bar H_1$ and $\bar H_2$, respectively. They correspond to the yield modes in which only one of the surfaces remains active. For

given material parameters, the bounds can easily be evaluated. The last condition is quadratic and contains both $\bar{H}_1$ and $\bar{H}_2$. It corresponds to the mode in which both surfaces evolve simultaneously. In addition to material parameters, the right-hand side contains also the first principal deviatoric stress, s_1, and the second deviatoric invariant, J_2. To guarantee uniqueness for all possible situations, we have to consider the most unfavorable case, in which the ratio $s_1/\sqrt{J_2}$ is maximized (because then the right-hand side of (20.186) has the largest possible value and the resulting restriction on $\bar{H}_1$ and $\bar{H}_2$ is the most severe). Fortunately, the task of maximizing $s_1/\sqrt{J_2}$ is easy if we use the Haigh–Westergaard coordinates. According to (D.71) in Appendix D, we have $s_1 = \sqrt{2/3}\,\rho\cos\theta$, where $\rho = \sqrt{2J_2}$, and so $s_1/\sqrt{J_2} = \sqrt{4/3}\cos\theta$ attains the largest possible value for $\theta = 0$, in which case $s_1/\sqrt{J_2} = \sqrt{4/3}$. Condition (20.186) can now be written exclusively in terms of material parameters;

$$\left(9K\alpha^2 + G + \bar{H}_1\right)\left(K + \frac{4}{3}G + \bar{H}_2\right) > \left(3K\alpha + \frac{2G}{\sqrt{3}}\right)^2 \tag{20.187}$$

In the $(\bar{H}_1, \bar{H}_2)$-plane, the boundary of the region in which (20.187) holds is a hyperbola centered at $\bar{H}_1 = -(9K\alpha^2 + G)$ and $\bar{H}_2 = -(K + 4G/3)$.

For illustration we plot conditions (20.184), (20.185) and (20.187) in the plane of the plastic moduli normalized by Young's modulus, with a specific choice of Poisson's ratio, $\nu = 0.2$, and the friction coefficient, $\alpha = 0.21$ (these values are representative of concrete). Figure 20.8 indicates that the third condition, which takes into account simultaneous yielding, reduces the region of admissible plastic moduli as compared to the case in which each yield surface is considered separately. For example, if we consider only the Rankine surface, the plastic modulus $\bar{H}_2$ must be larger than $-1.111E$. If we combine the Rankine criterion with softening and the Drucker–Prager criterion with perfect plasticity ($\bar{H}_1 = 0$), the critical value of the softening modulus $\bar{H}_2$ becomes $-0.027E$. Additional restrictions could be obtained by analyzing the edges and the vertex of the Rankine surface (i.e. the cases for which two or three principal stresses are equal); see Problem 20.8.

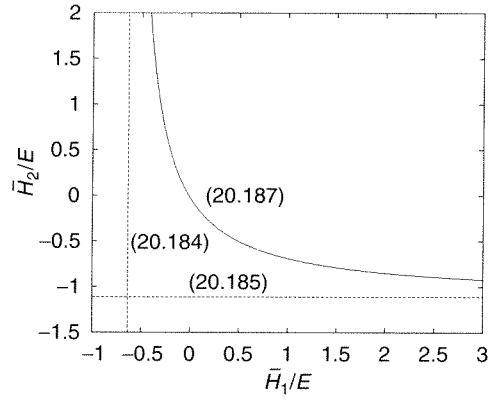


Figure 20.8 Restrictions on plastic moduli for the model combining the Rankine and Drucker–Prager surfaces

20.5 ANISOTROPIC YIELD CRITERIA

Some materials (e.g. oriented fiber-matrix composites or wood) have a strongly anisotropic structure, which must be reflected not only by the elastic moduli, but also by the yield condition. The description of the material is no longer invariant with respect to arbitrary rotations of the coordinate system, and so the yield criterion should be expressed in terms of the stress components with respect to a certain (fixed) coordinate system rather than in terms of the invariants of the stress tensor.

20.5.1 Hill Criterion

We will restrict our attention to *orthotropic materials*, i.e. to materials with three mutually orthogonal planes of symmetry. It is natural to relate the material description to the axes of orthotropy x_1, x_2 and x_3, defined as the intersections of the planes of symmetry. The stress components with respect to the axes of orthotropy are denoted as σ_{11}, σ_{12}, etc. The orthotropic yield condition

$$\left(\frac{\sigma_{22}-\sigma_{33}}{k_{11}}\right)^2 + \left(\frac{\sigma_{33}-\sigma_{11}}{k_{22}}\right)^2 + \left(\frac{\sigma_{11}-\sigma_{22}}{k_{33}}\right)^2 + \left(\frac{\sigma_{23}}{k_{23}}\right)^2 + \left(\frac{\sigma_{31}}{k_{31}}\right)^2 + \left(\frac{\sigma_{12}}{k_{12}}\right)^2 - 1 = 0 \tag{20.188}$$

came first, proposed by Hill (1947). It is a straightforward generalization of the von Mises condition; cf. expression (D.40) for the invariant J_2. As for the von Mises condition, Hill's condition is insensitive to volumetric stress and possesses tension-compression symmetry.

The six material parameters k_{11}, k_{22}, k_{33}, k_{12}, k_{23} and k_{31} can be related to the yield stresses in tension and compression along the three axes of orthotropy, σ^0_{11}, σ^0_{22} and σ^0_{33}, and to the yield stresses in shear in the three planes of symmetry, τ^0_{12}, τ^0_{23} and τ^0_{31}. For uniaxial tension along axis x_1,

$$\left(\frac{\sigma^0_{11}}{k_{22}}\right)^2 + \left(\frac{\sigma^0_{11}}{k_{33}}\right)^2 = 1 \tag{20.189}$$

and two similar conditions can be obtained by cyclic permutation of the subscripts. Solving for the material parameters, we identify

$$k_{11} = \sqrt{2}\left(\frac{1}{(\sigma^0_{22})^2} + \frac{1}{(\sigma^0_{33})^2} - \frac{1}{(\sigma^0_{11})^2}\right)^{-1/2} \tag{20.190}$$

and analogous expressions can be written for k_{22} and k_{33}.

For pure shear in the plane x_1–x_2 we get

$$\left(\frac{\tau^0_{12}}{k_{12}}\right)^2 = 1 \tag{20.191}$$

from which

$$k_{12} = \tau^0_{12} \tag{20.192}$$

and similar relations $k_{23} = \tau^0_{23}$ and $k_{31} = \tau^0_{31}$ hold for the other two planes of symmetry. If the material is isotropic, we have $\sigma^0_{11} = \sigma^0_{22} = \sigma^0_{33} = \sqrt{3}\tau_0$ and $\tau^0_{12} = \tau^0_{23} = \tau^0_{31} = \tau_0$. The reader can verify that (20.188) is then equivalent to the von Mises yield condition.

20.5.2 Hoffman Criterion

The Hill criterion does not allow modeling of materials with different values of the yield stress in tension and in compression. Hoffman (1967) added a linear combination of the normal stresses σ_{11}, σ_{22} and σ_{33}, to make the yield surface nonsymmetric with respect to the origin. To facilitate parameter identification, we write the Hoffman criterion in the form

$$\left(\frac{\sigma_{22}-\sigma_{33}-c_{22}+c_{33}}{k_{11}}\right)^2+\left(\frac{\sigma_{33}-\sigma_{11}-c_{33}+c_{11}}{k_{22}}\right)^2 + \left(\frac{\sigma_{11}-\sigma_{22}-c_{11}+c_{22}}{k_{33}}\right)^2+\left(\frac{\sigma_{23}}{k_{23}}\right)^2 + \left(\frac{\sigma_{31}}{k_{31}}\right)^2+\left(\frac{\sigma_{12}}{k_{12}}\right)^2-1=0 \tag{20.193}$$

so that the new parameters c_{11}, c_{22}, and c_{33} correspond to the shifted center of the yield surface. They can be easily identified as

$$c_{11}=\frac{\sigma_{11}^{0+}-\sigma_{11}^{0-}}{2},\qquad c_{22}=\frac{\sigma_{22}^{0+}-\sigma_{22}^{0-}}{2},\qquad c_{33}=\frac{\sigma_{33}^{0+}-\sigma_{33}^{0-}}{2} \tag{20.194}$$

where σ_{11}^{0+} is the tensile yield stress in direction 1, $-\sigma_{11}^{0-}<0$ is the compressive yield stress in direction 1, etc. Equations of the type (20.190) and (20.192) remain valid if we replace σ_{11}^0 by $(\sigma_{11}^{0+}+\sigma_{11}^{0-})/2$, etc.

20.5.3 Tsai–Wu Criterion

The most general quadratic yield condition was proposed by Tsai and Wu (1971). They defined the yield function by

$$f(\boldsymbol{\sigma})=\boldsymbol{f}:\boldsymbol{\sigma}+\boldsymbol{\sigma}:\boldsymbol{F}:\boldsymbol{\sigma}-1 \tag{20.195}$$

where $\boldsymbol{f}$ is a symmetric second-order tensor and $\boldsymbol{F}$ is a fourth-order tensor with both major and minor symmetry. In engineering notation, $\boldsymbol{f}$ can be represented by a column matrix with six coefficients and $\boldsymbol{F}$ by a symmetric 6×6 matrix with 21 independent coefficients (for a general anisotropic material). If the material is orthotropic, the number of independent material parameters can be reduced because the yield function must be invariant with respect to reflections about the axes of symmetry. For example, after reflection about the plane (x_1,x_2), σ_{23} and σ_{31} change sign but σ_{12} does not, and so the coefficients multiplying the products $\sigma_{23}\sigma_{12}$ and $\sigma_{31}\sigma_{12}$ must vanish. By similar arguments we arrive at the Tsai–Wu criterion for orthotropic materials,

$$\begin{Bmatrix} f_1\\ f_2\\ f_3\end{Bmatrix}^T\begin{Bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\end{Bmatrix}+\begin{Bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{Bmatrix}^T\begin{bmatrix} F_{11}&F_{12}&F_{13}&0&0&0\\ F_{12}&F_{22}&F_{23}&0&0&0\\ F_{13}&F_{23}&F_{33}&0&0&0\\ 0&0&0&F_{44}&0&0\\ 0&0&0&0&F_{55}&0\\ 0&0&0&0&0&F_{66}\end{bmatrix}\begin{Bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{Bmatrix}=1 \tag{20.196}$$

This criterion contains 12 material parameters.

The Tsai–Wu criterion can be reduced to the Hoffman and Hill criteria. In other words, conditions (20.193) and (20.188) can be written in the form (20.196), with some internal dependencies among the material parameters. Verification of this fact is left to the reader as an exercise.

20.5.4 Comparison

In summary, we have three orthotropic yield criteria with an increasing level of complexity. They are due to Hill, Hoffman, and Tsai and Wu, and they use (respectively) 6, 9 and 12 parameters. Hill's criterion is insensitive to the volumetric stress and gives the same yield stresses in uniaxial tension and compression. Hoffman's criterion allows for different behaviors in tension and in compression, but is still insensitive to the volumetric part of the stress tensor. The Tsai–Wu criterion is the most general quadratic yield condition, and appears to be the most rational one. Unfortunately, it requires the determination of 12 material parameters, which is often impossible due to the scarcity of experimental data. In its most general form (20.195), this criterion is applicable to materials with general anisotropy, but the number of parameters then increases to 27.

Example 20.11: Reduce the orthotropic yield conditions to a plane-stress situation (in one of the planes of material symmetry) and present them graphically. Discuss the differences.

Solution: Let us start from the simplest case. *Hill's criterion* (20.188) reduces under plane-stress conditions to

$$\left(\frac{\sigma_{22}}{k_{11}}\right)^2+\left(\frac{\sigma_{11}}{k_{22}}\right)^2+\left(\frac{\sigma_{11}-\sigma_{22}}{k_{33}}\right)^2+\left(\frac{\sigma_{12}}{k_{12}}\right)^2-1=0 \tag{20.197}$$

This equation describes an ellipsoid in the $(\sigma_{11},\sigma_{22},\sigma_{12})$-space. The ellipsoid has its center at the origin and one of its principal axes coincides with the axis σ_{12}. The intersection with the plane $(\sigma_{11},\sigma_{22})$ is an ellipse with principal axes in a general position. The criterion (20.197) has four parameters. Parameter k_{12} has the meaning of the yield stress in shear. Formula (20.190) for parameter k_{11} and analogous formulae for k_{22} and k_{33} contain the yield stress for tension and compression in the out-of-plane direction x_3. This does not appear to be logical from the physical point of view. For identifying the parameters to be used in the criterion for plane stress, it is more reasonable to ignore the test results in the out-of-plane direction and supplement the tensile yield stresses in directions x_1 and x_2 by a yield stress corresponding to another loading path, e.g. by the yield stress σ_{et}^0 corresponding to equibiaxial tension with $\sigma_{11}=\sigma_{22}$. Substituting this into (20.197) we obtain the relation

$$\left(\frac{\sigma_{et}^0}{k_{11}}\right)^2+\left(\frac{\sigma_{et}^0}{k_{22}}\right)^2=1 \tag{20.198}$$

and so the parameters k_{11}, k_{22} and k_{33} can be evaluated from formulae analogous to (20.190), with σ_{33}^0 replaced by σ_{et}^0.

Hoffman's criterion (20.193) reduces under plane-stress conditions to

$$\left(\frac{\sigma_{22}-c_{22}}{k_{11}}\right)^2+\left(\frac{\sigma_{11}-c_{11}}{k_{22}}\right)^2+\left(\frac{\sigma_{11}-\sigma_{22}-c_{11}+c_{22}}{k_{33}}\right)^2+\left(\frac{\sigma_{12}}{k_{12}}\right)^2-1=0 \tag{20.199}$$

This is still an ellipsoid in the space $(\sigma_{11}, \sigma_{22}, \sigma_{12})$ with one principal axis parallel to σ_{12} but its center is now shifted to the point $(c_{11}, c_{22}, 0)$. The criterion contains six material parameters. Instead of using the formulae that depend on the yield stresses in the out-of-plane direction, it would again be more appropriate to identify the parameters from six experiments performed under plane stress. Parameter k_{12} is equal to the yield stress in shear, and the remaining parameters can be related to the four yield stresses in tension and in compression along axes x_1 and x_2, supplemented by one additional yield stress, e.g. that obtained under equibiaxial tension.

The *Tsai–Wu criterion* (20.196) reduced to a plane-stress situation reads

$$F_{11}\sigma_{11}^2 + 2F_{12}\sigma_{11}\sigma_{22} + F_{22}\sigma_{22}^2 + F_{66}\sigma_{12}^2 + f_1\sigma_{11} + f_2\sigma_{22} - 1 = 0 \qquad (20.200)$$

Note also that it also involves six materials parameters, the same as the reduced Hoffman criterion. If $F_{11}F_{22} > F_{12}^2$ and $F_{66} > 0$, (20.200) describes an ellipsoid and the elastic domain is convex. In this case, criteria (20.200) and (20.199) are fully equivalent. The only difference is that they are expressed in terms of different material parameters.

PROBLEMS

Problem 20.1: Show that, for a purely deviatoric flow under monotonic uniaxial loading, the rate of cumulative plastic strain $\dot{\bar{\varepsilon}}_p$ defined by (20.5) is equal to the usual plastic strain rate $\dot{\varepsilon}_{11}^p$.

Problem 20.2: Plot the uniaxial stress-strain curve that would correspond to linear work hardening, defined by a hardening law of the form $\sigma_Y = \sigma_0 + cW_p$, where c is a constant and W_p is the plastic work (20.7). Is the stress-strain curve linear?

Problem 20.3: Consider a strain-hardening or strain-softening Rankine material with an associated flow rule. Derive the condition of uniqueness in terms of the plastic modulus.

Problem 20.4*: Consider an elastoplastic material with linear strain softening subjected to uniaxial tension. Combining the basic equations $\varepsilon = \varepsilon_e + \varepsilon_p$, $\varepsilon_e = \sigma/E$, and $\dot{\varepsilon}_p = \dot{\sigma}/H$, show that the rate form of the stress-strain law reads

$$\dot{\sigma} = \frac{EH}{E+H}\dot{\varepsilon} \qquad (20.201)$$

The uniqueness of response of the model is preserved only if $E+H > 0$. This condition is more restrictive than condition (20.40) derived in Example 20.2. Explain why.

Problem 20.5: Consider a material with isochoric plastic flow and with purely kinematic hardening. Find the relation between the Melan–Prager hardening parameter, $\bar{H}_k$, and the plastic modulus, H, defined as the proportionality factor between the stress rate and the plastic strain rate under uniaxial tension.

Problem 20.6: Figure 20.3(b) shows the evolution of the von Mises yield surface during kinematic hardening under uniaxial stress. The center of the ellipse moves

along the axis σ_1, but the flow direction has a nonzero component $\dot{\varepsilon}_2$. Is this in contradiction with the Melan–Prager kinematic hardening rule?

Problem 20.7*: Consider an elastic domain with a non-smooth boundary, described by $f_I(\boldsymbol{\sigma}) < 0$ $(I = 1, 2)$, and a nonassociated flow rule $\dot{\boldsymbol{\varepsilon}} = \dot{\lambda}\boldsymbol{g}_\sigma$ with a *smooth* plastic potential. (a) Formulate the consistency condition(s). (b) Find the expression for the rate of plastic multiplier in a situation in which $f_1 = 0$ and $f_2 = 0$. (c) Discuss the local uniqueness of the model.

Problem 20.8: For the composite yield surface combining the Rankine and Drucker–Prager conditions, explore singular points that were not considered in Example 20.10. Set up the conditions of local uniqueness and plot them in the plane of plastic moduli.

Problem 20.9: Explore the singular points of the composite yield surface combining the Rankine and Drucker–Prager conditions for plane stress problems. Set up the conditions of local uniqueness and plot them in the plane of plastic moduli. Make a comparison to the three-dimensional case.

Problem 20.10: In the proof of the global uniqueness theorem for perfectly plastic materials given in Section 16.3 it was tacitly assumed that the yield surface is smooth. Show that the theorem remains valid for yield surfaces with corners.

Problem 20.11: Write the Hill criterion (20.188) in the form of the Tsai–Wu criterion (20.196). What are the internal dependencies among the 12 Tsai–Wu parameters?

Problem 20.12: Write the Hoffman criterion (20.193) in the form of the Tsai–Wu criterion (20.196). What are the internal dependencies among the 12 Tsai–Wu parameters?

Problem 20.13: An orthotropic material is called transversely isotropic if its properties are invariant with respect to arbitrary rotations about one axis of orthotropy, say x_1. Write the Tsai–Wu criterion with a reduced number of independent parameters, suitable for transversely isotropic materials.

21

Plastic Material Models for Concrete and Soils

There are several classes of engineering materials, for example, metals, concretes, or soils, that exhibit rather different basic features of their constitutive behavior. Plasticity models based on elementary yield conditions, such as those due to von Mises or Drucker and Prager, can often be regarded only as very crude approximations for limit load analysis. Realistic analysis requires models that reflect the particularities of each specific class of materials.

In this chapter we will give a brief introduction to plasticity-based modeling of two material classes important in civil engineering – concretes and soils. However, before starting it is important to point out that, in the case of concrete and soils, the applicability of plastic models is limited. Except for extreme hydrostatic pressures, concretes, and in some instances also soils, exhibit no plastic yield plateau. Instead, the material *softens* after the peak stress, which causes *localization of inelastic strain* and causes the failure to be non-simultaneous, propagating. This gives rise to size effects, important for large structures. Therefore, the yield criteria that follow in this chapter must be used in tandem with some realistic model for localization and propagation of softening inelastic zones, and must be tied to some characteristic length that is a material property (e.g. Bažant and Planas, 1998). Objective modeling of localization will be discussed in more detail in Chapter 26.

21.1 CONCRETE

21.1.1 Material Behavior

Concrete is a pressure-sensitive material with a dramatically lower strength in tension than in compression. Tensile loading typically leads to the formation and propagation of cracks normal to the axis of maximum principal stress. Such failure is reasonably well described by the Rankine criterion. Tensile cracking usually leads to a degradation of elastic stiffness, which is not reflected by standard plasticity models that unload with the initial slope. Stiffness degradation can be handled within the framework of *fracture mechanics* or *damage mechanics* (with a proper localization limiter). The so-called smeared crack models, popular in engineering applications, can be interpreted as a special type of damage model. However, cracking of concrete under monotonic loading can also be approximated by models based exclusively on the theory of plasticity, provided that the energy dissipation due to localized fracture and the existence of a characteristic length are properly taken into account.

If tensile loading in one direction is combined with compression in one or two perpendicular directions, the development of cracks is promoted, and the maximum tensile stress that can be transmitted by the material decreases with respect to the uniaxial tensile strength. Such a combination of principal stresses corresponds, for instance, to pure shear.

In concrete under *uniaxial compression*, cracks or microcracks, called splitting cracks, propagate mainly in planes parallel to the direction of loading. After the compressive stress exceeds the elasticity limit, the material hardens with a gradually decreasing slope of the stress-strain curve until the compressive strength is exhausted. In larger structures, the splitting cracks are organized in a relatively narrow band that propagates laterally, and gives the impression of shear failure, although there is hardly any shear slip at maximum stress. The slip develops only later, in the softening regime, when the compressive stress decreases in magnitude with increasing magnitude of the applied strain. Lateral positive strains are in the elastic range proportional to the applied negative strain, with Poisson's ratio as the proportionality factor. When the splitting cracks start growing, the increase of lateral strains sharply accelerates, and soon after the peak stress the total volumetric strain becomes positive and grows rapidly. This means that the material exhibits dilatancy.

Under *confined compression*, i.e. if the axial splitting is inhibited or slowed down by lateral compressive stresses, the strength and ductility of concrete increases as compared to the unconfined case. Such a situation arises, e.g. in concrete columns with a very large amount of transverse reinforcement (stirups). Highly confined concrete exhibits remarkable ductility. For illustration, a recently devised 'tube-squash' test (Bažant, Kim and Brocca, 1999; Caner and Bažant, 2000a) has revealed that a core drilled out from a concrete cylinder cast inside a thick-wall steel tube and squashed between rigid platens to about 60% of its original length still keeps about 1/3 of its initial uniaxial compressive strength, despite the fact that the shear angle in the core exceeded 70°. Under purely hydrostatic pressure, as well as uniaxial compressive strain (with zero lateral strain), concrete never softens but keeps hardening. When major pores close, the tangential stiffness (current bulk modulus) increases even above the initial elastic value, which is naturally expected because matter cannot be compacted into a zero volume.

The typical experimental results for concrete indicate that the *strength envelope* (failure surface) in the principal stress space is a deformed cone with three planes of symmetry that all intersect at the hydrostatic axis. The deviatoric sections (octahedral plane projections) have the form of a rounded triangle whose shape changes from almost triangular for tensile and low compressive hydrostatic pressures to almost circular for high compressive hydrostatic pressures (Launay and Gachon, 1971). Recall that a triangular shape of the deviatoric sections corresponds to the Rankine criterion, while a circular shape of these sections corresponds to the Drucker–Prager criterion. The change of shape of the deviatoric section can also be described in terms of the so-called meridians, i.e. the intersections of the failure surface with halfplanes that start from the hydrostatic axis.

As the deviatoric section of the strength envelope is not circular, the meridians in halfplanes characterized by different values of the Lode angle are in general different (see Section D.4 for the definition of the Lode angle and Haigh–Westergaard coordinates). Two extreme cases are represented by the meridians with Lode angles $\theta = 0$ and $\theta = \pi/3$. The former corresponds to stress states with two equal principal stresses smaller than the third one, and is called the *tensile meridian*. The latter corresponds to stress states with two equal principal stresses larger than the third one, and is called the *compressive meridian*. Together, both meridians are sometimes called the *Rendulić sections*. Note that, for the Drucker–Prager criterion, all the meridians are identical and, if the friction coefficient α is constant, are straight lines. The experimentally determined meridians for concrete are curved and, in the region of small hydrostatic pressures, the tensile meridian is below the compressive one. This is related to the noncircular shape of the deviatoric section. All the meridians have of course a common intersection with the hydrostatic axis, and for large hydrostatic pressure they approach each other.

21.1.2 Failure Criteria

The failure criteria designed specifically for concrete should reflect the properties of the experimentally determined failure surface (strength envelope) described in the preceding section. The Drucker–Prager criterion can be regarded as the initial approximation that needs further refinement. The meridians should be curved, and the compressive meridian should be different from the tensile one.

The basic idea of the criterion proposed by Willam and Warnke (1974) is that the extreme meridians can be represented by two different parabolas and the intermediate meridians can be interpolated using a smooth function of the Lode angle. A somewhat different approach was suggested by Ottosen (1977), who started from a parabolic approximation of only one meridian and then incorporated the dependence on the Lode angle.

We will now demonstrate the derivation of a rather general failure criterion for concrete based on logical arguments. Let us start from the Drucker–Prager criterion (15.21). Expressing the invariants I_1 and J_2 in terms of the Haigh–Westergaard coordinates ξ, ρ and θ (see (D.68)–(D.70) for their definition), we obtain

$$\frac{\alpha\sqrt{3}}{\tau_0}\xi + \frac{1}{\tau_0\sqrt{2}}\rho - 1 = 0 \tag{21.1}$$

This criterion is linear in terms of ξ and ρ and independent of the polar coordinate, the Lode angle θ. It has the general form

$$c_1\xi + c_2\rho - 1 = 0 \tag{21.2}$$

where c_1 and c_2 are constant parameters. For any Lode angle, the meridian is the same straight line in the ξ–ρ plane. By adding a term quadratic in ρ we can obtain parabolic meridians, but still independent of the Lode angle. The failure criterion would then have the form

$$c_1\xi + c_2\rho + c_3\rho^2 - 1 = 0 \tag{21.3}$$

Now we consider the Rankine criterion (15.27). Substituting expression (D.71) that relates the maximum principal stress to the Haigh–Westergaard coordinates, we obtain the condition

$$\frac{1}{\sigma_0\sqrt{3}}\xi + \sqrt{\frac{2}{3}}\frac{1}{\sigma_0}\rho\cos\theta - 1 = 0 \tag{21.4}$$

which has the general form

$$c_1\xi + c_2\rho r(\theta) - 1 = 0 \tag{21.5}$$

where $r(\theta)$ is a given function of the Lode angle, in our case $r(\theta) = \cos\theta$. Now each meridian is a straight line, which is the same as for (21.2), but its slope depends on the value of θ. The function $r(\theta)$ sets the shape of the deviatoric section. For a fixed value of the hydrostatic coordinate ξ, the deviatoric section is described in the polar coordinates ρ and θ by

$$\rho(\theta) = \frac{1 - c_1\xi}{c_2 r(\theta)} = \frac{\text{const}}{r(\theta)} \tag{21.6}$$

When $r(\theta) = 1$, we obtain a circular deviatoric section and (21.5) reduces to (21.2). When $r(\theta) = \cos\theta$, we obtain a triangular deviatoric section corresponding to the Rankine criterion. However, we may also design a function $r(\theta)$ that gives the deviatoric section in the form of a rounded triangle. Possible definitions of $r(\theta)$ will be given later.

The desired form of the failure criterion can be constructed by combining the modifications of (21.2) that lead (1) to a parabolic shape of the meridians, and (2) to a noncircular shape of the deviatoric section. Looking for a common generalization of (21.3) and (21.5), we bear in mind that, for high compressive hydrostatic pressures, the effect of the Lode angle should be negligible. This is true for the criterion

$$c_1\xi + c_2\rho r(\theta) + c_3\rho^2 - 1 = 0 \tag{21.7}$$

that covers both (21.3) and (21.5) as special cases. For large negative values of ξ, the deviatoric radius ρ solved from (21.7) is also large. Consequently, the term $c_3\rho^2$ dominates over $c_2\rho r(\theta)$, and the deviatoric section is almost circular. This would not be true if we decided to generalize (21.3) by replacing ρ by $\rho r(\theta)$ in both the linear and the quadratic term, which would lead to the modified Leon–Pramono model, studied by Weihe (1989).

The deviatoric (octahedral) section of the failure surface has three axes of symmetry inclined by 60° with respect to each other, and so the function $r(\theta)$ is fully determined by its values in the interval $[0, \pi/3]$. It follows from (21.6) that the deviatoric section is smooth only if the derivatives of $r(\theta)$ at both $\theta = 0$ and $\theta = \pi/3$ are zero. This condition is satisfied by the constant function $r(\theta) = 1$ that corresponds to a circular section, but not by the function that corresponds to the triangular section, $r(\theta) = \cos\theta$, which only has a zero derivative at $\theta = 0$ but not at $\theta = \pi/3$. This is natural to expect since the triangle has sharp corners. A failure criterion of the form (21.7) with a noncircular smooth deviatoric section was proposed by Ottosen (1977) who used

$$r(\theta) = \begin{cases} \cos\left(\dfrac{1}{3}\arccos(K\cos 3\theta)\right) & \text{if } \cos 3\theta \geq 0 \\ \cos\left(\dfrac{\pi}{3} - \dfrac{1}{3}\arccos(-K\cos 3\theta)\right) & \text{if } \cos 3\theta \leq 0 \end{cases} \tag{21.8}$$

where $K \leq 1$ is a shape factor affecting the 'out-of-roundedness' of the deviatoric section; see Figure 21.1(a).

Another function $r(\theta)$ with the desired properties was constructed by Willam and Warnke (1974),but they exploited it in a somewhat different manner. Their function,

$$r(\theta, e) = \frac{4(1-e^2)\cos^2\theta + (2e-1)^2}{2(1-e^2)\cos\theta + (2e-1)\sqrt{4(1-e^2)\cos^2\theta + 5e^2 - 4e}} \tag{21.9}$$

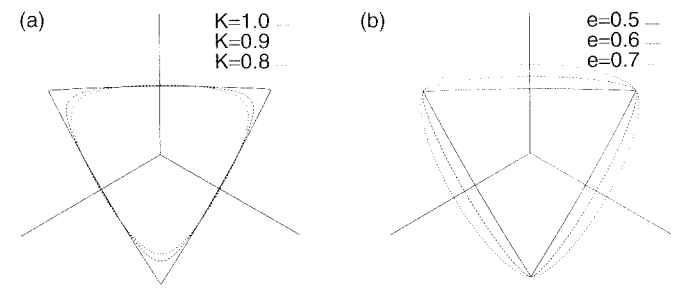


Figure 21.1 Influence of the shape parameter on the deviatoric section corresponding to zero mean pressure according to a) Ottosen, b) Menétrey and Willam

is derived from an ellipse in the deviatoric plane. Parameter e is the eccentricity, again related to the shape of the deviatoric section, as illustrated in Figure 21.1(b). Note that $r(0, e) = 1/e$ and $r(\pi/3, e) = 1$. Menétrey and Willam (1995) presented a failure criterion of the form (21.7), using function $r(\theta)$ given by (21.9) with the eccentricity e considered as a constant.

The surface described by (21.7) is smooth everywhere, except for the intercept with the hydrostatic axis. Both definitions (21.8) and (21.9) lead to convex deviatoric sections for all physically meaningful values of parameters K or e, covering the entire range of shapes between an equilateral triangle and a circle. In contrast to that, the simpler definition

$$r(\theta) = \frac{2K}{1 + K + (1 - K)\cos 3\theta} \tag{21.10}$$

proposed by Gudehus (1973) and Argyris *et al.* (1974) preserves convexity only for $K \geq 0.777$, i.e. only for deviatoric sections close to a circle; see Lin and Bažant (1986). Such a function can be useful for materials with a weak dependence on the Lode angle (such as soils), but it cannot be recommended for concrete.

Criterion (21.7) has three adjustable parameters, c_1, c_2 and c_3, and one additional parameter is usually hidden in the definition of the function $r(\theta)$, e.g. parameter K in (21.8) or parameter e in (21.9). The failure surface is therefore determined by four parameters that need to be identified from experimental results. Motivated by the Hoek–Brown criterion for rock (Hoek and Brown, 1980), Menétrey and Willam (1995) proposed a three-parameter model that is a special case of (21.7) with an internal dependence between the parameters, $c_1 = \sqrt{2}c_2$. The parameters of this model can be evaluated, for example, from the uniaxial tensile strength, uniaxial compressive strength, and equibiaxial compressive strength. If only the uniaxial compressive strength is available, it is possible to use empirical relations suggested by design codes. For example, the CEB-FIP Model Code 1990 (CEB-FIP, 1991) recommends estimating the uniaxial tensile strength from the empirical relation

$$\bar{f}_t = \left(\frac{\bar{f}_c - 8\text{ MPa}}{10\text{ MPa}}\right)^{2/3} \times 1.40\text{ MPa} \tag{21.11}$$

where $\bar{f}_t$ and $\bar{f}_c$ are the statistically mean values of tensile and compressive strength, respectively (the original notation used by the code, f_{ctm} and f_{cm}, has been adapted to ensure consistency with the present notation). The code uses the Ottosen criterion with parameters $c_1 = 0.468/(\bar{f}_c k^{1.1})$, $c_2 = 1.01/(\bar{f}_c k^{0.9})$, $c_3 = 0.0556/(\bar{f}_c^2 k^{1.4})$ and $K = 1 - 6.8(k - 0.07)^2$, where $k = \bar{f}_t/\bar{f}_c$ is the ratio of the tensile and compressive

strength. ACI Standard 318 (ACI, 1999) recommends

$$\bar{f}_t = \sqrt{\frac{\bar{f}_c}{1\ \text{psi}}} \times 6\ \text{psi}, \qquad (1\ \text{psi} = 6895\ \text{Pa}) \tag{21.12}$$

Example 21.1: For concrete with a given compressive strength, $\bar{f}_c$, determine the biaxial (plane stress) failure envelope that corresponds to the (a) von Mises criterion, (b) Drucker–Prager criterion, (c) Mohr–Coulomb criterion, (d) parabolic Leon criterion, $a(\sigma_1-\sigma_3)^2+b(\sigma_1+\sigma_3)-1=0$ (see Problem 15.9 for motivation of this criterion), (e) Hoek–Brown criterion, $a(\sigma_1-\sigma_3)^2+b\sigma_1-1=0$, (f) Menétrey–Willam criterion, (g) Ottosen criterion. If more than one material parameter is needed for identification, assume the tensile strength, $\bar{f}_t = 0.1\bar{f}_c$, and the equibiaxial compressive strength, $\bar{f}_b = 1.16\bar{f}_c$. Compare the biaxial envelope to the experimental results of Kupfer and Gerstle (1973), and discuss the suitability of individual criteria for concrete.

Solution: First, note that, under plane stress, the invariants can be expressed as $I_1 = \sigma_1 + \sigma_2$ and $J_2 = (\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma_2)/3$. The values of principal stresses, invariants I_1 and J_2, and Haigh–Westergaard coordinates ξ, ρ and θ are listed in Table 21.1.

(a) *Von Mises criterion*, $\sqrt{J_2} - \tau_0 = 0$.
Substituting the value of J_2 at failure under uniaxial compression, we identify the parameter $\tau_0 = \bar{f}_c/\sqrt{3}$. Under plane stress conditions the criterion can be presented in the form

$$\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma_2 - \bar{f}_c^2 = 0 \tag{21.13}$$

This equation describes an ellipse symmetric with respect to the origin of the principal stress plane. As might have been expected, such a criterion would completely misrepresent the tensile part of the actual failure envelope for concrete; see Figure 21.2(a). However, in the compressive quadrant it could be accepted as the first approximation. A reasonable criterion (limited to plane-stress problems) can be obtained by combining the von Mises criterion for compression with the Rankine criterion for tension, as suggested by Feenstra and de Borst (1995a).

(b) *Drucker–Prager criterion*, $\alpha I_1 + \sqrt{J_2} - \tau_0 = 0$.
This criterion has two parameters, α and τ_0, and so we can match it to two characteristic strength values. For example, substituting the values of I_1 and J_2 at peak stress under uniaxial compression and under uniaxial tension, and solving for

Table 21.1 Principal stresses, stress invariants, and Haigh–Westergaard coordinates at failure

stress state	σ_1	σ_2	I_1	J_2	ξ	ρ	θ
uniaxial tension	$\bar{f}_t$	0	$\bar{f}_t$	$\bar{f}_t^2/3$	$\bar{f}_t/\sqrt{3}$	$\bar{f}_t\sqrt{2/3}$	0
uniaxial compression	0	$-\bar{f}_c$	$-\bar{f}_c$	$\bar{f}_c^2/3$	$-\bar{f}_c/\sqrt{3}$	$\bar{f}_c\sqrt{2/3}$	$\pi/3$
equibiaxial compression	$-\bar{f}_b$	$-\bar{f}_b$	$-2\bar{f}_b$	$\bar{f}_b^2/3$	$-2\bar{f}_b/\sqrt{3}$	$\bar{f}_b\sqrt{2/3}$	0

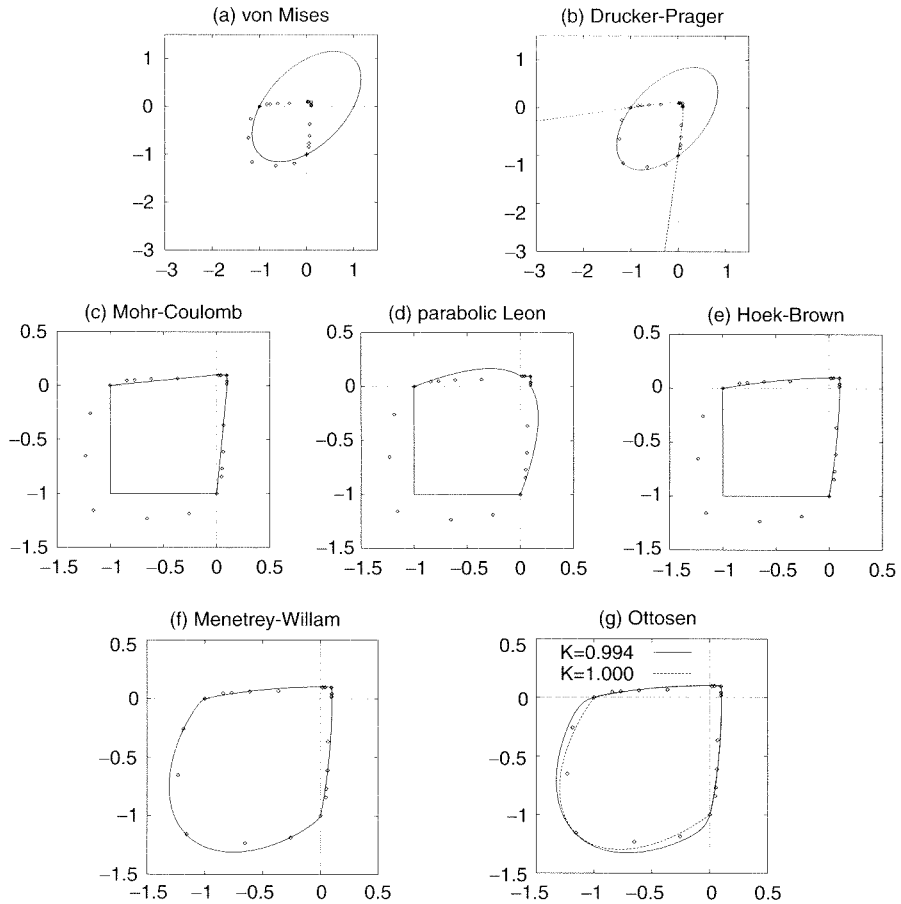


Figure 21.2 Biaxial failure envelopes resulting from various yield criteria, plotted in the normalized principal stress plane $(\sigma_1/\bar{f}_c, \sigma_2/\bar{f}_c)$ and compared to experimental data due to Kupfer and Gerstle (marked by isolated points)

the model parameters, we obtain

$$\alpha = \frac{1}{\sqrt{3}}\frac{\bar{f}_c - \bar{f}_t}{\bar{f}_c + \bar{f}_t} \tag{21.14}$$

$$\tau_0 = \frac{2}{\sqrt{3}}\frac{\bar{f}_c\bar{f}_t}{\bar{f}_c + \bar{f}_t} \tag{21.15}$$

Under plane stress conditions, the criterion can be presented in the form

$$4(\sigma_1^2 + \sigma_2^2) + \left(2 - 3\frac{\bar{f}_c^2 + \bar{f}_t^2}{\bar{f}_c\bar{f}_t}\right)\sigma_1\sigma_2 + 4(\bar{f}_c - \bar{f}_t)(\sigma_1 + \sigma_2) - 4\bar{f}_c\bar{f}_t = 0 \tag{21.16}$$

Depending on the ratio $\bar{f}_t/\bar{f}_c$, this can be an ellipse, parabola or hyperbola. The critical value for which (21.16) describes a parabola is $\bar{f}_t/\bar{f}_c = 1/3$. For smaller ratios (typical of concrete), the intersection of the Drucker–Prager cone with the biaxial stress plane is a hyperbola and the equibiaxial strength becomes infinite. For $\bar{f}_t/\bar{f}_c = 0.1$ this hyperbola is shown as the dashed curve in Figure 21.2(b). So, by matching the uniaxial strengths in tension and in compression, we would obtain

quite a reasonable approximation of the actual failure envelope in the quadrants that correspond to biaxial tension and to tension-compression. On the other hand, a good description in the biaxial compression range can be obtained by matching the uniaxial and equibiaxial compressive strengths. This leads to

$$\alpha = \frac{1}{\sqrt{3}} \frac{\bar{f}_b - \bar{f}_c}{2\bar{f}_b - \bar{f}_c} \tag{21.17}$$

$$\tau_0 = \frac{1}{\sqrt{3}} \frac{\bar{f}_b \bar{f}_c}{2\bar{f}_b - \bar{f}_c} \tag{21.18}$$

The corresponding biaxial failure envelope is an ellipse with a shifted center; see the solid curve in Figure 21.2(b). In the compressive region, the approximation of the actual failure envelope is better than that for the von Mises ellipse, but large deviations in the tensile region are still observed. Feenstra and de Borst developed a model combining this criterion with the Rankine cutoff in tension (Feenstra and de Borst, 1996).

(c) *Mohr–Coulomb criterion*, $(1 - \sin\phi)\sigma_1 \quad (1 + \sin\phi)\bar{\sigma}_3 - 2c\cos\phi = 0$.
Here we must be careful because the criterion in the above form assumes the principal stresses to be ordered, $\bar{\sigma}_1 \geq \bar{\sigma}_2 \geq \sigma_3$. We have labeled the ordered principal stresses by overbars, to avoid confusion with the principal stresses describing a plane stress state, $\sigma_1 \geq \sigma_2$ and $\sigma_3 = 0$. Depending on the sign of σ_1 and σ_2 we must map them appropriately on $\bar{\sigma}_1$, $\bar{\sigma}_2$ and σ_3. For example, if $\sigma_1 \geq 0 \geq \sigma_2$ we must set $\bar{\sigma}_1 = \sigma_1$, $\bar{\sigma}_2 = 0$, and $\bar{\sigma}_3 = \sigma_2$. Using the stress values at uniaxial tensile and compressive failure, we obtain the relations

$$f_t = \frac{2c\cos\phi}{1 - \sin\phi} \tag{21.19}$$

$$f_c = \frac{2c\cos\phi}{1 + \sin\phi} \tag{21.20}$$

which help to convert the criterion to the form

$$\frac{\bar{\sigma}_1}{\bar{f}_t} - \frac{\bar{\sigma}_3}{\bar{f}_c} - 1 = 0 \tag{21.21}$$

Under plane stress, the failure envelope is described by three segments:

$$\begin{array}{ll} \sigma_1 - \bar{f}_t = 0 & \text{if } 0 \leq \sigma_2 \leq \sigma_1 \\ \sigma_1/\bar{f}_t - \sigma_2/\bar{f}_c - 1 = 0 & \text{if } \sigma_2 \leq 0 \leq \sigma_1 \\ \sigma_2 + \bar{f}_c = 0 & \text{if } \sigma_2 \leq \sigma_1 \leq 0 \end{array} \tag{21.22}$$

Note that, in each segment, the criterion is linear in terms of the principal stresses. The resulting failure envelope is a hexagon plotted in Figure 21.2(c). It underestimates the actual strength in the quadrant corresponding to biaxial compression. This deficiency is a consequence of the fact that the envelope of Mohr circles is independent of the medium principal stress, which is the main weakness of the Mohr–Coulomb criterion.

(d) *Parabolic Leon criterion*, $a(\bar{\sigma}_1 - \bar{\sigma}_3)^2 + b(\bar{\sigma}_1 + \bar{\sigma}_3) - 1 = 0$.
Again, we map σ_1 and σ_2 on $\bar{\sigma}_1$ and $\bar{\sigma}_3$ depending on the current quadrant in the

principal stress plane. The model parameters can be identified as

$$a = \frac{1}{\bar{f}_c \bar{f}_t} \tag{21.23}$$

$$b = \frac{\bar{f}_c - \bar{f}_t}{\bar{f}_c \bar{f}_t} \tag{21.24}$$

and the criterion reduced to plane stress reads

$$\begin{array}{ll} \sigma_1 - \bar{f}_t = 0 & \text{if } 0 \leq \sigma_2 \leq \sigma_1 \\ (\sigma_1 - \sigma_2)^2 + (\bar{f}_c - \bar{f}_t)(\sigma_1 + \sigma_2) - \bar{f}_c \bar{f}_t = 0 & \text{if } \sigma_2 \leq 0 \leq \sigma_1 \\ \sigma_2 + \bar{f}_c = 0 & \text{if } \sigma_2 \leq \sigma_1 \leq 0 \end{array} \tag{21.25}$$

In the quadrants of biaxial tension and biaxial compression, the criterion is identical to the Mohr–Coulomb criterion. In the tensile-compressive quadrant, the failure envelope is a part of a parabola; see Figure 21.2(d). For small ratios $\bar{f}_t/\bar{f}_c < 1/3$, typical of concrete, the elastic domain turns out to be nonconvex.

(e) *Hoek–Brown criterion*, $a(\bar{\sigma}_1 - \bar{\sigma}_3)^2 + b\bar{\sigma}_1 - 1 = 0$.
This is a simple modification of the previous criterion. The only difference from (21.25) arises in the tensile-compressive quadrant where the failure criterion now reads

$$\bar{f}_t(\sigma_1 - \sigma_2)^2 + (\bar{f}_c^2 - \bar{f}_t^2)\sigma_1 - \bar{f}_c^2 \bar{f}_t = 0, \qquad \text{if } \sigma_2 \leq 0 \leq \sigma_1 \tag{21.26}$$

This is also a parabola but the elastic domain is now convex for any ratio $\bar{f}_t/\bar{f}_c$; see Figure 21.2(e).

(f) *Menétrey–Willam criterion*, $c_2[\sqrt{2}\xi + \rho r(\theta)] + c_3\rho^2 - 1 = 0$, with $r(\theta)$ given by (21.9).
This criterion has three parameters, c_2, c_3 and e (eccentricity from (21.9)). They can be adjusted such that the model exactly reproduces the strength in uniaxial tension, uniaxial compression and biaxial compression. Substituting the values from Table 21.1 corresponding to failure under uniaxial compression, we identify $c_3 = 3/(2\bar{f}_c^2)$. For uniaxial tension and equibiaxial compression, we obtain additional equations that can be solved for the model parameters;

$$c_2 = \frac{1}{\sqrt{6}} \left(\frac{1}{\bar{f}_t} - \frac{1}{\bar{f}_b} + \frac{\bar{f}_b - \bar{f}_t}{\bar{f}_c^2} \right) \tag{21.27}$$

$$e = \frac{1 + \epsilon}{2 - \epsilon}, \qquad \text{where } \epsilon = \frac{\bar{f}_t}{\bar{f}_b} \frac{\bar{f}_b^2 - \bar{f}_c^2}{\bar{f}_c^2 - \bar{f}_t^2} \tag{21.28}$$

It is interesting to observe that, if $\bar{f}_b > \bar{f}_c > \bar{f}_t$, then $\epsilon > 0$ and $e > 0.5$, because the last inequality is a necessary condition for convexity of the failure surface. For the present material characteristics, we obtain $\epsilon = 0.03$, $e = 0.523$ and $c_2 = 4.163/\bar{f}_c$. The biaxial failure envelope plotted in Figure 21.2(f) nicely matches the experimental results.

(g) *Ottosen criterion*, $c_1\xi + c_2\rho r(\theta) + c_3\rho^2 - 1 = 0$, with $r(\theta)$ given by (21.8).
This criterion uses four independent parameters, c_1, c_2, c_3 and K (shape factor from (21.8)). Identification of the fourth parameter requires some additional information.

In his original paper, Ottosen fitted exactly three material characteristics, $\bar{f}_c$, $\bar{f}_t$, and $\bar{f}_b$, and then he applied the method of least squares to obtain the best fit of the compressive meridian to the test results of Balmer (1949) and Richart, Brandtzaeg and Brown (1928). He found the following optimal set of parameters: $c_1 = 5.54/\bar{f}_c$, $c_2 = 8.3/\bar{f}_c$, $c_3 = 0.638/\bar{f}_c^2$ and $K = 0.98$. Parameter K is usually very close to 1 but, for $K = 1$, the failure surface would not be smooth. It seems reasonable to estimate K from the CEB code formula,

$$K = 1 - 6.8\left(\frac{\bar{f}_t}{\bar{f}_c} - 0.07\right)^2 \tag{21.29}$$

and then determine parameters c_1, c_2 and c_3 from $\bar{f}_c$, $\bar{f}_t$ and $\bar{f}_b$. This step involves the solution of a set of three linear equations. For the present material characteristics, we obtain $K = 0.994$, $c_1 = 5.8/\bar{f}_c$, $c_2 = 8.04/\bar{f}_c$ and $c_3 = 1.29/\bar{f}_c^2$. The biaxial failure envelope plotted in Figure 21.2(g) is only slightly different from that which corresponds to the Menétrey-Willam criterion, and it matches the experimental results quite well; see the solid curve in Figure 21.2(g). An even better fit would be obtained with $K = 1$, $c_1 = 6.1/\bar{f}_c$, $c_2 = 7.76/f_c$ and $c_3 = 2.03/\bar{f}_c$; see the dashed curve with sharp corners. Note that the parameters c_i are quite sensitive to the choice of K. □

Another general criterion, somewhat different from (21.7), can be obtained by interpolation between two extreme meridians. Suppose that the tensile and compressive meridians are, respectively, described by the explicit formulae

$$\rho = \bar{\rho}_t(\xi), \qquad \rho = \bar{\rho}_c(\xi) \tag{21.30}$$

where $\bar{\rho}_t(\xi)$ and $\bar{\rho}_c(\xi)$ are given functions with common roots (to make sure that both meridians intersect the hydrostatic axis at the same points). Defining the pressure-sensitive eccentricity,

$$e(\xi) = \frac{\bar{\rho}_t(\xi)}{\bar{\rho}_c(\xi)} \tag{21.31}$$

we can construct an interpolated failure surface described by the equation

$$\rho r(\theta, e(\xi)) - \bar{\rho}_c(\xi) = 0 \tag{21.32}$$

where $r(\theta, e)$ is, for example, the Willam–Warnke function (21.9), or any other suitable function with similar properties. Since $r(0, e) = 1/e$ and $r(\pi/3, e) = 1$, it is easy to verify that, for $\theta = 0$ and $\theta = \pi/3$, condition (21.32) reduces to the first and second part of (21.30), respectively.

Failure criteria of this type have been used by several authors. Willam and Warnke (1974) defined the tensile and compressive meridians as parabolas while Lin *et al.* (1987) used slanted ellipses. Kang and Willam (2000) defined the compressive meridian by a power function and, instead of defining the tensile meridian, they prescribed directly the variation of eccentricity.

Net-Reinforced Concrete

When a concrete panel is reinforced by a dense reinforcing net and the change of internal forces from one bar to the next is very small, the net may be considered as smeared (continuously distributed). Then it makes sense to construct the yield

surface for the steel-concrete composite. Although local slip between the concrete and the bars must occur near the intersections of cracks with the bars, the overall average slip over many bar spacings can usually be assumed to be zero. The yield criterion for the bars may be combined with various yield criteria for concrete. The Mohr–Coulomb criterion was used to deduce the yield surface (Marti and Thürlimann, 1977; Marti, 1979) and the limit design envelope (Bažant and Tsubaki, 1979; Bažant, Tsubaki and Belytschko, 1980). Considering only plane-strain failure behavior, and the Mohr–Coulomb failure criterion for concrete with zero cohesion and a constant internal friction angle ϕ, one obtains the following anisotropic yield criterion for the composite of concrete and an orthogonal reinforcing net (Bažant and Tsubaki, 1979):

$$[(N_x^0 - N_x) - \beta_1(N_y^0 - N_y)][(N_y^0 - N_y) - \beta_1(N_x^0 - N_x)] - (2\beta_2 N_{xy})^2 = 0 \tag{21.33}$$

where

$$\beta_1 = \frac{1 - \sin\phi}{1 + \sin\phi}, \qquad \beta_2 = \frac{1}{1 + \sin\phi} \tag{21.34}$$

N_x^0 and N_y^0 are the steel yield forces (per unit area of composite) in bars of x and y directions; and N_x, N_y, N_{xy} are the normal and shear forces in the composite.

21.1.3 Hardening and Softening

Loading Surfaces for Plastic Constitutive Law

The preceding exposition has dealt with the failure surface describing the strength limits of a small volume of the material (or representative volume), roughly corresponding to standard small laboratory specimens. The usefulness of such a loading surface for larger structures is limited because the material softens after the peak stress, which causes that one must analyze the effects of localization of damage and take into account the size effect on the strength limits.

To handle such problems, the yield surfaces just described are often used in an incremental elasto-plastic constitutive law which describes the entire response curves of the material and characterizes the strength envelope only indirectly. For that purpose, it is necessary to define a family of surfaces parameterized by one or more variables that evolve during the hardening-softening process. For concrete, it is not realistic to construct such loading surfaces by simple radial scaling of the strength envelope, identified from the peak stress states under various triaxial loadings. The failure envelope is always open along the hydrostatic axis because, under compressive loadings close to hydrostatic pressure, there is no limit on the strength of concrete (the hardening process never turns into softening). Nevertheless, the response curves generated by such loadings are far from linear, and therefore the initial 'yield' surface, marking the onset of nonlinearity and enclosing the elastic domain, must be closed. It is thus appropriate to determine the loading surfaces of the elasto-plastic constitutive law from the measured response curves, including data points before and after the peak stress states. When this is done, the resulting optimal loading surface may be quite different from that identified solely from the the peak load states.

Several effective models of this kind are available. Some of them combine the plasticity concepts with those of damage mechanics, and thus are beyond the scope of this book. We will outline three representative formulations based exclusively on plasticity. To unify notation with the previously discussed models, we use variables and

parameters different from those used in the original papers, but the basic equations presented here are equivalent to the original ones.

Model of Lin et al.

An effective model for concrete was developed by Lin *et al.* (1987) and subsequently used in large finite element structural analyses at Argonne National Laboratory. The *loading surfaces* have the form of (21.32) and are parameterized by a variable σ_c, which reflects the current level of hardening and has the physical meaning of the current yield stress in uniaxial compression. The loading function is defined as

$$f(\boldsymbol{\sigma}, \sigma_c) = f(\xi, \rho, \theta; \sigma_c) = \rho r(\theta, e(\xi, \sigma_c)) - \bar{\rho}_c(\xi, \sigma_c) \tag{21.35}$$

where $r(\theta, e)$ is the Willam–Warnke function given by (21.9), and

$$e(\xi, \sigma_c) = \frac{\bar{\rho}_t(\xi, \sigma_c)}{\bar{\rho}_c(\xi, \sigma_c)} \tag{21.36}$$

is the variable eccentricity.

The Rendulić section is obtained by an appropriate modification of the Cam-Clay model for soils, to be explained in Section 21.2.3. The meridians have the shape of a 'slanted ellipse', constructed by taking the square root of a cubic function (Figure 21.3). The compressive meridian is described by the equation

$$\rho = \bar{\rho}_c(\xi, \sigma_c) \equiv \sigma_c \sqrt{c_{c0} + c_{c1}\frac{\xi}{\sigma_c} + c_{c2}\left(\frac{\xi}{\sigma_c}\right)^2 + c_{c3}\left(\frac{\xi}{\sigma_c}\right)^3} \tag{21.37}$$

where c_{ci} $(i = 0, 1, 2, 3)$ are constant dimensionless parameters. The tensile meridian is described by a function $\rho_t(\xi, \sigma_c)$, given by the same formula as $\bar{\rho}_c(\xi, \sigma_c)$ in (21.37), however, with parameters c_{ci} replaced by c_{ti} $(i = 0, 1, 2, 3)$. Since the intersections of the tensile and compressive meridians with the hydrostatic axis must coincide, only six of the eight parameters c_{ti} and c_{ci} are independent.

The loading surface is convex and smooth (even at its intersections with the hydrostatic axis), which facilitates the development of the corresponding stress return algorithms; see Section 22.3.

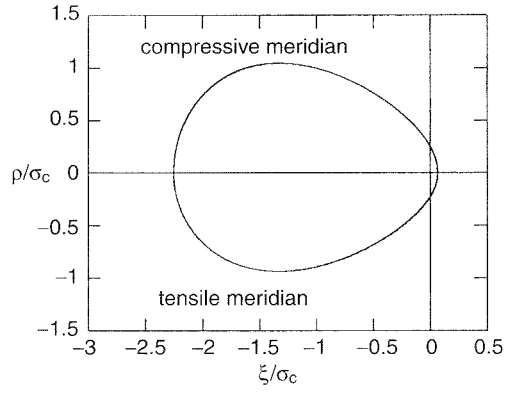


Figure 21.3 Rendulić section used by Lin *et al.*

The development of a realistic *flow rule* for concrete is not an easy task. The shortcomings of the associated flow rule were illustrated experimentally by Smith (1987). Loading surfaces obtained by radial scaling of a failure envelope that is open along the hydrostatic axis typically require a nonassociated flow rule, which is often defined in terms of a plastic potential constructed by modifying the volumetric terms in the yield function. The simplest example is the Drucker–Prager plastic potential (20.108) with the dilatancy coefficient α_ψ smaller than the friction coefficient α that appears in the yield criterion (15.21); see Section 20.3. Nevertheless, Lin *et al.* (1987) used the flow rule

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda}\frac{\partial f(\boldsymbol{\sigma}, \sigma_c)}{\partial \boldsymbol{\sigma}} \tag{21.38}$$

associated with their loading function (21.35), and they obtained reasonable fits of the volumetric expansion under uniaxial compression (Figure 21.4(a)).

The *hardening law* proposed by Lin *et al.* is formulated in such a manner that, at the low-pressure side of the slanted ellipse, up to maximum ρ, the hardening is followed by postpeak softening, while at the high-pressure side, after the maximum ρ, only hardening takes place. This ensures that standard triaxial tests end up with softening at low pressures but not at high pressures. Thus, despite the fact that the surface is not open in the direction of the hydrostatic axis, the response in the hydrostatic compression test (as well as in the oedometric test, i.e. with $\varepsilon_x < 0$ and $\varepsilon_y = \varepsilon_z = 0$) is always hardening.

According to Lin *et al.* (1987), the hardening process has two stages, governed by different hardening laws. During the first stage, the variable σ_c (current yield stress in uniaxial compression) increases from its initial value $\sigma_{c0} = 0.6\bar{f}_c$ to the value $\sigma_c = \bar{f}_c$, which corresponds to the peak of the uniaxial compressive curve. The hardening law is based on the work hardening hypothesis, i.e. σ_c is driven by the plastic work. As explained in Section 20.1.1, it is convenient to introduce an effective plastic strain κ defined by the rate equation

$$\dot{\kappa} = \frac{\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p}{\sigma_c} \tag{21.39}$$

Due to the presence of σ_c in the denominator, κ is in the uniaxial compressive test equal to the plastic strain ε_{px}, and the relation between κ and σ_c can be identified from the uniaxial compressive stress-strain curve. Lin *et al.* (1987) proposed to approximate the hardening curve by an ellipse (Figure 21.4(a)), described by the relation

$$\sigma_c = h_1(\kappa) \equiv \sigma_{c0} + \frac{\bar{f}_c - \sigma_{c0}}{\kappa_p}\sqrt{\kappa(2\kappa_p - \kappa)} \tag{21.40}$$

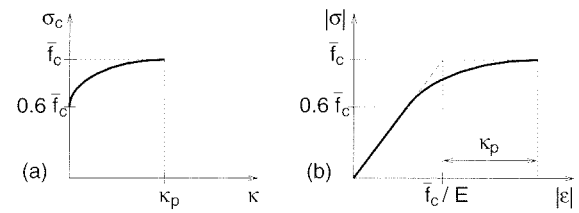


Figure 21.4 Hardening law used by Lin *et al.*: a) hardening curve $\sigma_c = h(\kappa)$, b) pre-peak part of the uniaxial compressive stress-strain curve

where κ_p is a parameter defining the value of κ for which σ_c reaches its peak value, $\bar{f}_c$. With κ_p identified from the uniaxial test and considered as a constant parameter, the peak point in multiaxial tests would be achieved too early. This can be corrected by taking κ_p as a function of the Lode angle,

$$\kappa_p(\theta) = \kappa_{p2} + (\kappa_{p1} - \kappa_{p2}) \sin^2 \tfrac{3}{2}\theta \tag{21.41}$$

where $\kappa_{p1} = \kappa_p(\pi/3)$ is the plastic strain at peak stress under uniaxial compression, and $\kappa_{p2} = \kappa_p(0)$ is the effective plastic strain at peak stress under biaxial compression. Equation (21.40) remains valid in the total form only if the Lode angle θ does not vary during the loading. For a general case, it is rewritten in the rate form

$$\dot{\sigma}_c = H(\kappa, \theta)\, \dot{\kappa} \tag{21.42}$$

where

$$H(\kappa, \theta) = \frac{\bar{f}_c - \sigma_{c0}}{\kappa_p(\theta)} \frac{\kappa_p(\theta) - \kappa}{\sqrt{\kappa[2\kappa_p(\theta) - \kappa]}} \tag{21.43}$$

is the tangent plastic modulus.

The hardening law (21.42) is used only until σ_c reaches the uniaxial compressive strength, $\bar{f}_c$. Subsequently, the material is either hardening or softening, depending on the sign of the volumetric plastic strain rate $\dot{\varepsilon}_{Vp} = (\dot{\boldsymbol{\varepsilon}}_p : \boldsymbol{\delta})/3$ (which in turn depends on the position of the stress point with respect to the 'critical state line', connecting the peaks of the egg-shaped meridians; see Section 21.2.3 for a discussion in the context of critical state soil mechanics). The hardening-softening law valid for this second stage is given by

$$\dot{\sigma}_c = \lambda_1 \sigma_c \dot{\varepsilon}_{Vp} \qquad \text{if} \quad \dot{\varepsilon}_{Vp} \le 0 \quad \text{(hardening)} \tag{21.44}$$

$$\dot{\sigma}_c = -2\lambda \sigma_c \varepsilon_{Vp} \dot{\varepsilon}_{Vp} \qquad \text{if} \quad \dot{\varepsilon}_{Vp} > 0 \quad \text{(softening)} \tag{21.45}$$

where λ_1 is a constant and

$$\lambda = \lambda_t + \tfrac{1}{8}(\lambda_c - \lambda_t)\,[1 + \tanh(10\,\sigma_V/\sigma_c)]^3 \tag{21.46}$$

is a variable parameter. The variation of this parameter with the volumetric stress $\sigma_V = \xi/\sqrt{3}$ achieves continuous transition between softening under tension ($\sigma_V > 0$, $\dot{\varepsilon}_{Vp} > 0$) and softening under compression ($\sigma_V < 0$, $\dot{\varepsilon}_{Vp} > 0$). Except for states very close to the transition, (21.46) gives $\lambda \approx \lambda_t$ or $\lambda \approx \lambda_c$, which are material constants. Typically, $\lambda_t \approx 10\lambda_c$. For a constant value of λ, the differential equation (21.45) has a closed-form solution,

$$\sigma_c = h_2(\Delta\varepsilon_{Vp}) \equiv \bar{f}_c \exp\left(-\lambda\, \Delta\varepsilon_{Vp}^2\right) \tag{21.47}$$

where $\Delta\varepsilon_{Vp}$ is the increment of the volumetric plastic strain measured from the start of softening.

An important advantage of this model is that the material parameters can be identified from a set of well defined characteristic response states and that the identification can be carried out sequentially, rather than by simultaneous optimization of the fits of all the data sets considered collectively. The data fitting procedure involves two systems of four linear equations, and the required test data include: the uniaxial compression and tensile strengths, biaxial compression strength,

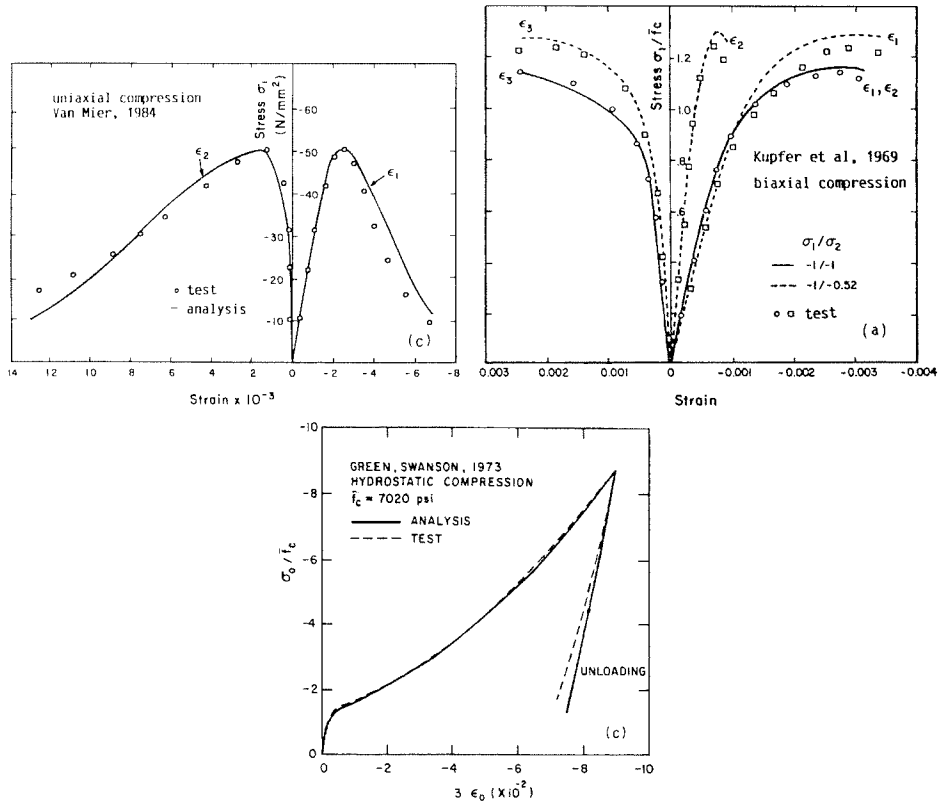


Figure 21.5 Fits of basic test data for concrete achieved with Lin *et al.*'s (1987) plasticity model

hydrostatic elastic limit, dilatancy-free states for uniaxial and biaxial compression, and the strains at uniaxial and biaxial peak stresses. For a detailed procedure, see Lin *et al.* (1987). The calibrated model was shown to exhibit good agreement with a broad range of monotonic uni-, bi- and triaxial tests. Figure 21.5 reproduces from Lin *et al.* (1987) the fits of selected test data for concrete (van Mier, 1984; Kupfer, Hilsdorf and Rüsch, 1969; Green and Swanson, 1973), which are seen to be quite satisfactory.

Model of Feenstra and de Borst

The 'composite plasticity model for concrete' proposed by Feenstra and Borst (1996) is an example of a plasticity-based model with multiple yield surfaces. Its simplified version was enriched by interaction with steel reinforcement and applied in simulations of reinforced concrete structures under plane stress (Feenstra and de Borst, 1995b).

The basic framework of multi-surface plasticity has been outlined in Section 20.4. The model of Feenstra and de Borst is in fact a nonassociated version of the two-surface plasticity model from Example 20.10. Nonlinear behavior under compression is controled by radially scaled Drucker–Prager loading surfaces,

$$f_1(\boldsymbol{\sigma}, \sigma_c) \equiv \alpha I_1(\boldsymbol{\sigma}) + \sqrt{J_2(\boldsymbol{\sigma})} - (\alpha + 1/\sqrt{3})\sigma_c = 0 \tag{21.48}$$

and tensile cracking is interpreted as another 'yielding' mechanism controled by radially scaled Rankine surfaces,

$$f_2(\boldsymbol{\sigma}, \sigma_t) \equiv \sigma_1(\boldsymbol{\sigma}) - \sigma_t = 0 \tag{21.49}$$

Here, σ_1 denotes the maximum principal stress, variables σ_t and σ_c have the meaning of (current) uniaxial tensile and compressive yield stresses, resp., α is the coefficient of internal friction, and the purpose of the factor $\alpha + 1/\sqrt{3}$ in (21.48) is to give the desired physical meaning to the variable σ_c.

The *flow rule* for the compressive mechanism is nonassociated and is derived from the plastic potential

$$g_1(\boldsymbol{\sigma}) = \alpha_\psi I_1(\boldsymbol{\sigma}) + \sqrt{J_2(\boldsymbol{\sigma})} \tag{21.50}$$

This means that the flow is associated in the deviatoric plane but its volumetric part uses a dilatancy coefficient α_ψ different from the coefficient of internal friction α. The flow rule for the tensile mechanism is associated, i.e. the yield function f_2 is taken as the plastic potential. Summing up, the contribution of both mechanisms leads to the plastic strain rate

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\boldsymbol{\varepsilon}}_{p1} + \dot{\boldsymbol{\varepsilon}}_{p2} = \dot{\lambda}_1 \frac{\partial g_1(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} + \dot{\lambda}_2 \frac{\partial f_2(\boldsymbol{\sigma}, \sigma_t)}{\partial \boldsymbol{\sigma}} \tag{21.51}$$

The Drucker–Prager surface evolves by isotropic hardening followed by softening, while the Rankine surface evolves by isotropic softening. The model works with two *hardening variables*, κ_1 and κ_2, defined by Feenstra and de Borst (1996) by the rate equations

$$\dot{\kappa}_1 = \dot{\lambda}_1 + \zeta_{12}\dot{\lambda}_2 \tag{21.52}$$

$$\dot{\kappa}_2 = \zeta_{21}\dot{\lambda}_2 + \dot{\lambda}_2 \tag{21.53}$$

where ζ_{12} and ζ_{21} are constant parameters describing the coupling between the two inelastic mechanisms. Due to the lack of experimental data, the authors set $\zeta_{12} = \zeta_{21} = 0$ for actual applications. This decouples the softening mechanism in tension from the hardening-softening mechanism in compression. In other words, if only one mechanism is active, the other yield surface does not evolve. This is of course not completely realistic, the same as the assumption of an isotropic evolution of the yield surfaces.

If parameters ζ_{12} and ζ_{21} are set to zero, the hardening variables κ_1 and κ_2 can be interpreted as effective plastic strains in the sense of the work hardening hypothesis, adapted for nonassociated flow. Indeed, since

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{p1} = \dot{\lambda}_1 \boldsymbol{\sigma} : \frac{\partial g_1(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} = \dot{\lambda}_1 \boldsymbol{\sigma} : \left(\alpha_\psi \boldsymbol{\delta} + \frac{\boldsymbol{s}}{2\sqrt{J_2(\boldsymbol{\sigma})}} \right) = \dot{\lambda}_1 g_1(\boldsymbol{\sigma}) \tag{21.54}$$

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{p2} = \dot{\lambda}_2 \boldsymbol{\sigma} : \frac{\partial f_1(\boldsymbol{\sigma}, \sigma_c)}{\partial \boldsymbol{\sigma}} = \dot{\lambda}_2 \boldsymbol{\sigma} : (\boldsymbol{n}_1 \otimes \boldsymbol{n}_1) = \dot{\lambda}_2 \sigma_1(\boldsymbol{\sigma}) = \dot{\lambda}_2 \sigma_t \tag{21.55}$$

the hardening variables can be defined by the rate equations

$$\dot{\kappa}_1 = \frac{\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{p1}}{g_1(\boldsymbol{\sigma})} \tag{21.56}$$

$$\dot{\kappa}_2 = \frac{\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{p2}}{\sigma_t} \tag{21.57}$$

which are equivalent to (21.52) (21.53) with $\zeta_{12} = \zeta_{21} = 0$. Under uniaxial tension, one has $\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{p2} = \sigma_t \dot{\varepsilon}_{p2x}$, and so (21.57) gives $\dot{\kappa}_2 = \dot{\varepsilon}_{p2x}$, which means that κ_2 is in this uniaxial case equal to the plastic strain in the direction of applied stress. Under uniaxial compression, one has $\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{p1} = \sigma_c \dot{\varepsilon}_{p1x}$ and $g_1(\boldsymbol{\sigma}) = \alpha_\psi \sigma_c + \sigma_c/\sqrt{3}$, and so (21.56) gives $\dot{\kappa}_1 = \dot{\varepsilon}_{p1x}/(\alpha_\psi + 1/\sqrt{3})$, which means that κ_1 is in this uniaxial case proportional to the plastic strain in the direction of applied stress. This interpretation of the hardening variables κ_1 and κ_2 makes it possible to identify the hardening laws from uniaxial stress-strain curves only.

Feenstra and de Borst (1996) proposed describing the *hardening and softening curves* by explicit relations

$$\sigma_c = \bar{h}_1(\kappa_1) \equiv \begin{cases} \dfrac{\bar{f}_c}{3}\left(1 + \dfrac{4\kappa_1}{\kappa_p} - \dfrac{2\kappa_1^2}{\kappa_p^2}\right) & \text{if } \kappa_1 \le \kappa_p \\ \bar{f}_c\left[1 - \left(\dfrac{\kappa_1 - \kappa_p}{\kappa_{1u} - \kappa_p}\right)^2\right] & \text{if } \kappa_p \le \kappa_1 \le \kappa_{1u} \\ 0 & \text{if } \kappa_{1u} \le \kappa_1 \end{cases} \tag{21.58}$$

$$\sigma_t = h_2(\kappa_2) \equiv \bar{f}_t \exp\left(-\frac{\kappa_2}{\kappa_{2u}}\right) \tag{21.59}$$

where $\bar{f}_c$, $\bar{f}_t$, κ_p, κ_{1u} and κ_{2u} are material parameters that can be deduced from the uniaxial stress-strain diagrams. Initially, $\kappa_1 = \kappa_2 = 0$, $\sigma_c = \bar{h}_1(0) = \bar{f}_c/3$ and $\sigma_t = h_2(0) = \bar{f}_t$. Function h_2 is monotonically decreasing, and so $\bar{f}_t$ is the uniaxial tensile strength of the material. The pre-peak behavior is linear elastic in tension, but not in compression; $\bar{h}_1$ is first increasing, and then at $\kappa_1 = \kappa_p$, it attains its maximum value $\bar{f}_c$. This means that yielding under uniaxial compression begins at one-third of the compressive strength, and then the material is hardening until the compressive strength is reached, with parameter κ_p governing the plastic strain at peak stress. Subsequently, the material softens; the stress decreases until it vanishes at $\kappa_1 = \kappa_{1u}$. Figure 21.6 shows the graphs of $\bar{h}_1(\kappa_1)$ and $h_2(\kappa_2)$ for $\bar{f}_c = 38$ MPa, $\bar{f}_t = 2.9$ MPa, $\kappa_p = 2.2 \times 10^{-3}$, $\kappa_{1u} = 6 \times 10^{-3}$ and $\kappa_{2u} = 0.431 \times 10^{-3}$.

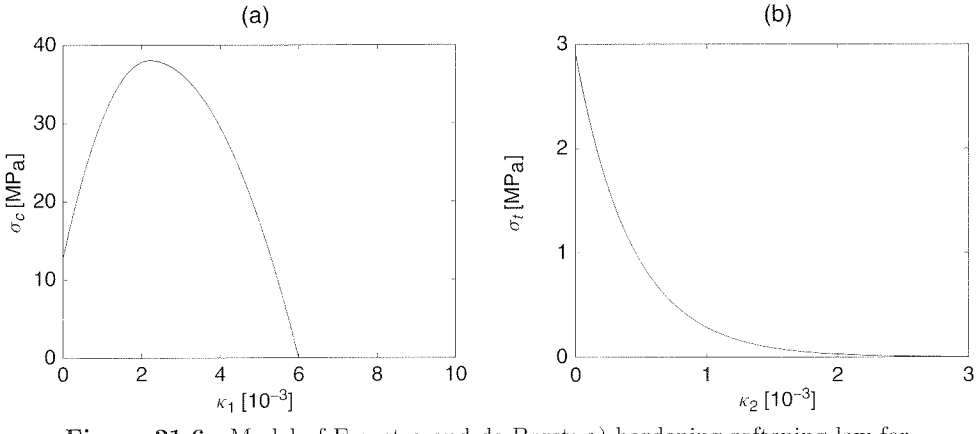


Figure 21.6 Model of Feenstra and de Borst: a) hardening-softening law for compression, b) softening law for tension

To take into account *strain localization* occuring after the onset of softening and to avoid pathological sensitivity of the numerical results to the spatial density of discretization (e.g. to the size of finite elements), the parameters that control softening cannot be considered as pure material parameters but must be related to the width of the numerically resolved localized softening band, L_s. When the finite element method is used, the width of the band depends on the type, size and shape of finite elements. As a rough approximation, it can be taken as the square root of the element area (in two dimensions) or the cubic root of the element volume (in three dimensions). The softening parameters are then evaluated as

$$\kappa_{1u} = \kappa_p + 1.5\frac{G_c}{L_s \bar{f}_c} \tag{21.60}$$

$$\kappa_{2u} = \frac{G_f}{L_s \bar{f}_t} \tag{21.61}$$

where G_f is the *fracture energy* and G_c is the (nonstandard) fracture energy in compression, which is about two orders of magnitude larger than G_f.

Feenstra and de Borst (1996) developed their model primarily for applications to plane-stress problems. The intersection of the composite yield surface with the principal stress plane can approximate the actual biaxial strength envelope for concrete reasonably well. Figure 21.7 shows the initial loading surface ($\sigma_t = \bar{f}_t$, $\sigma_c = \bar{f}_c/3$) and the strength envelope ($\sigma_t = \bar{f}_t$, $\sigma_c = \bar{f}_c$). Under highly confined triaxial stress states (exemplified by the standard triaxial test data of Balmer (1949) and Green and Swanson (1973), and including oedometric and hydrostatic compression), the response would be linear elastic for arbitrarily large stresses, which is not realistic.

Model of Kang and Willam

A number of plasticity models for concrete have been developed by Willam and coworkers (Willam and Warnke, 1974; Pramono and Willam, 1989; Etse and Willam, 1994). The most recent version is described in the dissertation of Kang (1997), and

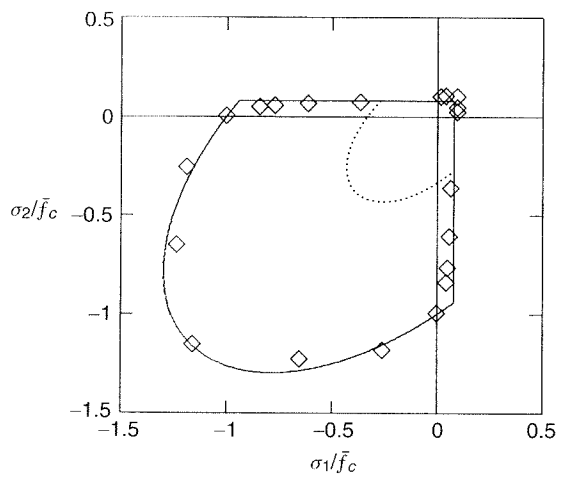


Figure 21.7 Initial loading surface (dotted) and biaxial strength envelope (solid) for model of Feenstra and de Borst (1996), compared to the test data by Kupfer and Gerstle (1973)

in Kang and Willam (1999). It uses *loading functions* defined by the sum

$$f(\xi, \rho, \theta; k, c) = f_f(\xi, \rho, \theta) + f_h(\xi; k) + f_s(\xi; c) \tag{21.62}$$

where the subscripts f, h and s stand for 'failure', 'hardening' and 'softening'. As usual, ξ, ρ and θ are the Haigh–Westergaard coordinates. Variables k and c parameterize the loading surfaces and evolve during the hardening and softening processes. The role of the stress invariants and of the variables k and c is different, which is emphasized by semicolons separating them in the lists of arguments.

The failure surface is described by the equation

$$f_f(\xi, \rho, \theta) \equiv \rho r(\theta, e(\xi)) - \rho_{uc}\left(\frac{\xi - \xi_0}{\xi_{uc} - \xi_0}\right)^{\alpha} = 0 \tag{21.63}$$

where $r(\theta, e)$ is again the Willam–Warnke function (21.9), and the pressure-dependent eccentricity parameter is given by

$$e(\xi) = 1 - \frac{5.5\bar{f}_c - \xi_0}{11\bar{f}_c - 2\xi} \tag{21.64}$$

Constant parameters $\xi_{uc} = -\bar{f}_c/\sqrt{3}$ and $\rho_{uc} = \bar{f}_c\sqrt{2/3}$ are the Haigh–Westergaard coordinates of the point that corresponds to the peak stress under uniaxial compression, $\xi_0 = \bar{f}_t\sqrt{3}$ is the equi-triaxial tensile strength based on the Rankine criterion, and $\alpha = 0.77$ is a constant exponent. As usual, $\bar{f}_c$ and $\bar{f}_t$ denote the uniaxial compressive and tensile strengths, respectively. Note that the failure criterion (21.63) has the general form of (21.32). The deviatoric section of the failure surface corresponding to (21.63) changes its shape from a rounded triangle to a circle as the hydrostatic pressure is increased (Figure 21.8(a)).

The effects of *hardening and softening* are incorporated through functions

$$f_h(\xi; k) = -\rho_1(k)\left[\left(\frac{\xi - \xi_1(k)}{\xi_0 - \xi_1(k)}\right)^{\beta(k)} - 1\right] \tag{21.65}$$

$$f_s(\xi; c) = (1 - c)\,\rho_1(k)\left(\frac{\xi_0}{\xi_0 - \xi_1(k)}\right)^{\alpha}\frac{\langle \xi - \xi_c\rangle^2}{\xi_c^2} \tag{21.66}$$

where

$$\xi_1(k) = -\sqrt{3}S\bar{f}_c k \tag{21.67}$$

$$\rho_1(k) = \rho_{uc}\left(\frac{\xi_1(k) - \xi_0}{\xi_{uc} - \xi_0}\right)^{\alpha} \tag{21.68}$$

$$\beta(k) = \frac{1 - k^2}{4(1 - k_0^2)} \tag{21.69}$$

Parameters S, ξ_c and k_0 are constant, and all the other symbols have been defined before.

In the initial state, $k = k_0 > 0$ and $c = 1$, which means that f_s vanishes and the initial yield surface is described by $f_f + f_h = 0$. The hardening function f_h introduces a compressive cap. The meridians of the initial yield surface, shown as the dotted curves in Figure 21.8(b), are similar to the slanted ellipse proposed by Lin *et al.*, but

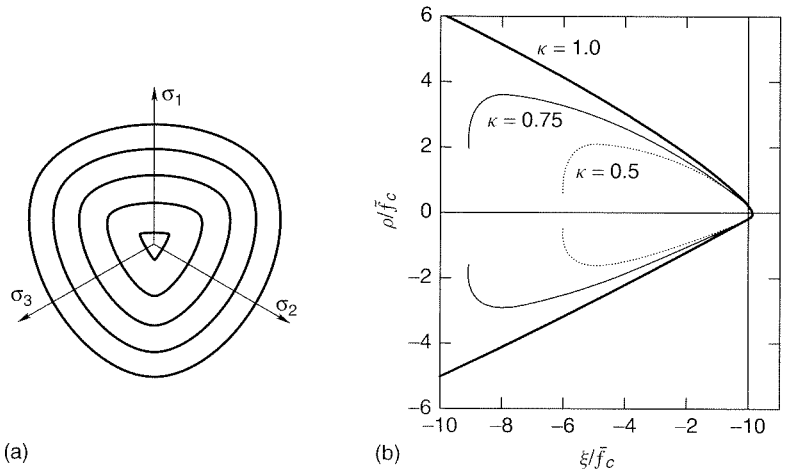


Figure 21.8 Model of Kang and Willam: a) deviatoric sections for different values of hydrostatic pressure, b) evolution of Rendulić section during hardening

the points with maximum ρ are shifted closer to the intersection with the hydrostatic axis in the high-pressure domain. During the hardening stage of evolution, k grows and c remains constant, which causes the yield surface to expand. When $k = 1$, the function f_h also vanishes and the current yield surface, described by $f_f = 0$, coincides with the failure envelope whose meridians are shown in Figure 21.8(b) as the thick solid curves. This envelope is open along the hydrostatic axis, which means that for highly confined loading paths the failure envelope is never reached and hardening can continue without any limits. If the current stress state reaches the failure surface below the so-called transition point of brittle ductile failure, i.e. if $\xi > \xi_c$, hardening turns into softening and the parameter c decreases from 1 to 0. This activates the softening part f_s of the loading function (21.62), and the low-pressure part of the loading surface shrinks.

The *flow rule* is nonassociated and is derived from the plastic potential $g(\xi, \rho, \theta; k, c)$, which has exactly the same form as the loading function (21.62), except that the exponent $\alpha = 0.77$ in (21.63) is replaced by $\bar{\alpha} = 0.23$. This modification reduces the excessive dilatancy that would be observed for an associated flow rule.

Hardening is driven by the equivalent plastic strain, κ_1, obtained by integrating the rate equation

$$\dot{\kappa}_1 = \|\dot{\boldsymbol{\varepsilon}}_p\| \tag{21.70}$$

Note that κ_1 is in fact the cumulative plastic strain from (20.5), but without the rescaling factor $\sqrt{2/3}$, which means that, under uniaxial compression, κ_1 is the magnitude of the plastic strain multiplied by $\sqrt{3/2}$. The hardening law is given in the total form

$$k = \bar{h}_1(\kappa_1, \xi) \equiv k_0 + \frac{1 - k_0}{\kappa_p(\xi)} \left(\sqrt{4\kappa_p(\xi)\kappa_1} - \kappa_1 \right) \tag{21.71}$$

For constant κ_p, the hardening curve would correspond to the so-called Hognestadt parabola with a vertical tangent at $\kappa_1 = 0$ and a horizontal tangent at $\kappa_1 = \kappa_p$. The purpose of the ductility function $\kappa_p(\xi)$, defined as a quadratic function of the pressure coordinate ξ, is to take into account the effect of confinement on ductility.

This adjustment is similar to that in the model of Lin *et al.* (1987), except that Lin *et al.* define κ_p as a function of θ; see (21.41).

Softening is driven by the equivalent fracture strain, κ_2, obtained by integrating the norm of the positive part of the plastic strain rate. In other words, the rate of the equivalent fracture strain is given by

$$\dot{\kappa}_2 = \|\langle \dot{\boldsymbol{\varepsilon}}_p \rangle\| \tag{21.72}$$

and the initial value at the start of softening is $\kappa_2 = 0$. To take into account *localization*, the relative cohesion c is not related to the fracture strain, κ_2, but to its product with the width of the softening band, $L_s\kappa_2$, which can be interpreted as the equivalent opening of a fictitious crack. Furthermore, the softening law

$$c = \bar{h}_2(\kappa_2, \xi) \equiv \exp\left(-\delta \left[\frac{L_s \kappa_2}{\kappa_s(\xi)} \right]^2 \right) \tag{21.73}$$

takes into account the volumetric stress, which enters through the ductility function

$$\kappa_s(\xi) = 1 + A_s(\xi - B_s \bar{f}_t)^4 \tag{21.74}$$

In (21.73)–(21.74), A_s, B_s and δ are constant parameters. Note the formal similarity of (21.73) to the softening law (21.47) formulated by Lin *et al.* (1987).

Other Models

Another group of models in which the loading surface is closed in the direction of hydrostatic compression is the so-called 'cap' model (Levine, 1982); see the later discussion of soils in Section 21.2.4.

Further plasticity models for concrete, which are intended essentially for biaxial plane stress and share various characteristics of the models already discussed, were proposed by Chen and Chen (1975), Dragon and Mróz (1979) and Lade and Kim (1995). Another useful model, geared toward cyclic loading, was that of Yang, Dafalias and Herrmann (1985). A concrete model based on plasticity formulated in the strain space was developed by Pekau and Zhang (1994).

The Hencky-type deformation theory (Hencky, 1924) can also be effectively generalized for concrete. The total strain model of Bažant and Tsubaki (1980) can well describe the triaxial response of concrete for monotonic proportional and nearly proportional loading, and at the same time it can correctly represent the response for the initial incremental deviation from such loading. Moreover, unlike all the models described so far, the Hencky deformation theory (total strain theory) simulates essentially correctly the vertex effect (for metals, this was shown by Budianski (1959)). In other words, the Hencky-type deformation theory correctly captures the great reduction of the tangential stiffness that is experimentally observed for the initial loading increment parallel to the loading surface (loading to the side); see the experiments reported by Caner, Bažant and Červenka (2001) and discussed in Section 25.4.3.

A number of models combine the plasticity concepts with those of damage mechanics, which is for concrete quite important, because standard plasticity alone cannot capture the progressive loss of stiffness due to cracking (it can only take into account the loss of strength). One early example is the plastic-fracturing theory

of Bažant and Kim (1979b), and there are numerous more recent ones (Resende, 1987; Lubliner *et al.*, 1989; Yazdani and Schreyer, 1990; Oller *et al.*, 1990, 1996; Yazdani and Karnawat, 1996; Faria, Oliver and Cervera, 1998; Červenka, Červenka and Eligehausen, 1998). Damage mechanics, however, is outside the scope of this book.

21.1.4 Orthotropic Models and Form Invariance

During the 1970s, it was popular to formulate the nonlinear triaxial constitutive models for geomaterials in the hypoelastic form

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D} : \dot{\boldsymbol{\varepsilon}} \quad \text{or} \quad \dot{\boldsymbol{\varepsilon}} = \boldsymbol{C} : \dot{\boldsymbol{\sigma}} \tag{21.75}$$

where the tangent stiffness tensor $\boldsymbol{D}$ and its inverse $\boldsymbol{C}$, the tangent compliance tensor, are functions of $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ that are figured out intuitively and then fit to test data. For isotropic materials with tangent compliance defined as a function of the stress tensor, one must then check the so-called *condition of form invariance* (Truesdell, 1955a), which requires that

$$C_{pqrs}(\boldsymbol{\sigma}', \boldsymbol{\varepsilon}') = c_{pi}c_{qj}c_{rk}c_{sm}C_{ijkm}(\boldsymbol{\sigma}, \boldsymbol{\varepsilon}) \tag{21.76}$$

for any coordinate rotation $x_i' = c_{ij}x_j$ where $c_{ij} = \cos(x_i', x_j)$; here $\sigma_{ij}' = c_{ik}c_{jm}\sigma_{km}$ and $\varepsilon_{ij}' = c_{ik}c_{jm}\varepsilon_{km}$. This condition is in plasticity satisfied automatically, by virtue of using scalar potentials (the yield function or loading function). Many early hypoelastic models for concretes and soils violated this condition (Bažant, 1983).

Coon and Evans (1972) assumed the most general linear dependence of $\boldsymbol{C}$ on $\boldsymbol{\sigma}$ and did satisfy this condition. Many subsequent investigators assumed the tangent compliance to exhibit orthotropic symmetry, i.e. to have the component form

$$\boldsymbol{C} \quad \begin{bmatrix} 1/E_x & \nu_y/E_x & -\nu_z/E_x & 0 & 0 & 0 \\ & 1/E_y & -\nu_z/E_y & 0 & 0 & 0 \\ & & 1/E_z & 0 & 0 & 0 \\ & \text{sym.} & & 1/G_{yz} & 0 & 0 \\ & & & & 1/G_{zx} & 0 \\ & & & & & 1/G_{xy} \end{bmatrix} \tag{21.77}$$

This form, however, is generally admissible only when the principal directions of stress and strain coincide and do not rotate during loading (this occurs in the so-called 'true' triaxial tests on cubes, but not in the axial-torsional tests). In that case, one can place the material orthotropy axes into the principal stress directions, which means that matrix (21.77) can be reduced to a 3×3 matrix, the rows and columns with $1/G$ being dropped.

According to Mohr circle, infinitesimal stress increments cause the principal directions of stress and strain to rotate, respectively, around axis x_3 by the angles $\mathrm{d}\theta_\sigma = \mathrm{d}\sigma_{12}/(\sigma_{11} - \sigma_{22})$ and $\mathrm{d}\theta_\varepsilon = \mathrm{d}\varepsilon_{12}/(\varepsilon_{11} - \varepsilon_{22})$. Equating both angles, one has the restriction (Bažant, 1983);

$$2G_{12} = \frac{\sigma_{11} - \sigma_{22}}{\varepsilon_{11} - \varepsilon_{22}} \quad \text{for } \varepsilon_{11} \neq \varepsilon_{22} \tag{21.78}$$

Two similar restrictions hold for shear moduli G_{23} and G_{31}. These restrictions are now used in the so-called *rotating crack models*, which evolved from the idea of Cope

(1980), and were consistently formulated by Gupta and Akbar (1984) and Rots (1988). Such models are widely used to simulate the evolution of distributed tensile cracking of concrete. The cracks, of course, do not actually rotate, rather, dominant cracks of a formerly dominant orientation close and new cracks of another orientation become dominant. As shown by Bažant (1983), many older orthotropic models for concrete and soils in the literature violate restriction (21.78).

21.2 SOILS

21.2.1 Material Behavior

Soils are multi-phase materials consisting of a solid skeleton and voids filled by two or more fluids, usually water and air. The skeleton is composed of small particles whose sizes can range from nanometers (in clays) to centimeters (for coarse gravel). The properties of soils are strongly affected by the saturation degree, i.e. by the relative volume of voids occupied by water. The soil can be saturated (if water completely fills all voids), partially saturated, or dry (no free water present).

An important aspect is the ability of the material to develop cohesion among the particles. In clay-like soils, layers of water adsorbed on the surface of their fine grains penetrate into the contact between the grains and, together with capillary tension, make the material capable of resisting tensile stresses and cause the shear response to be almost independent of normal stresses (provided that the loading is fast enough so as to prevent the water from leaving the pore space). As a result, such materials, classified as *cohesive*, keep their integrity even for large slip between the grains. The effect of capillary tension is, for instance, responsible for the fact that a clay specimen can freely stand up, without the need for a support or container.

On the other hand, gravels, sands, and to some extent also silts, do not support tensile stresses. The relative motion of the grains must overcome friction, and so the shear resistance strongly depends on the normal stresses. Such materials are classified as *frictional*.

There are also natural clay-like or sand-like materials which are cemented. In this case, true bonds exist between the material particles.

Based on experimental observations, Terzaghi (1936) postulated that all measurable effects of a change of stress, such as compression, distortion and a change of shearing resistance, are exclusively due to changes in the *effective stress*, defined as

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} - u\,\boldsymbol{\delta} \tag{21.79}$$

where $\boldsymbol{\sigma}$ is the total stress, u is the pore water pressure, and $\boldsymbol{\delta}$ is the unit hydrostatic tensor (Kronecker delta). The effective stress characterizes the resultant of the contact forces transmitted between the grains across a plane that is microscopically tortuous, passing not through the grains but around them and through the contacts which are regarded as point-like. The effective stress concept must be distinguished from Biot's two-phase material concept (refining an idea of Fillunger), in which the stress in the solid characterizes the resultant of forces in the solid on a perfectly plane section cutting through the grains; see de Boer (1996). Biot's concept is appropriate for constitutive laws well before failure.

In soil mechanics, compressive stresses and strains are usually denoted by a positive sign. However, for the present purpose we prefer to remain consistent with the standard sign convention of general continuum mechanics, i.e. we consider compressive

stresses and strains as negative. Furthermore, as our attention is focused on the constitutive description of the solid skeleton, we denote the effective stress by $\boldsymbol{\sigma}$ rather than $\boldsymbol{\sigma}'$, to keep the notation simple. While the modeling of soils as multi-phase materials is treated in specialized publications (Lewis and Schrefler, 1987, 1998; Craig, 1992; Gens, Jouanna and Schrefler, 1995), our aim is to describe an elastoplastic model that can serve as the solid-phase component of more sophisticated models which take into account the complicated interaction between the solid and fluid phases of soils.

Modeling of the plastic behavior of soils is usually based on the concept of *critical state*, developed by Roscoe and his coworkers at the University of Cambridge. The first complete constitutive model of this type, presented in 1963 by Roscoe and Schofield, was given the name 'Cam-Clay' two years later. Subsequent improvements led to the modified Cam-Clay model (Burland, 1965; Roscoe and Burland, 1968), which is now the most popular elastoplastic model for clays. From now on, we will leave out the adjective 'modified' and refer simply to the 'Cam-Clay model', always meaning its modified version. It is relatively simple yet captures a number of phenomena that occur in cohesive soils, e.g. an increasing stiffness under hydrostatic compression, hardening and softening, volume changes produced by deviatoric loading (contractancy and dilatancy), incremental nonlinearity (in the sense of different stiffnesses for loading and for unloading), and the tendency to reach at large deformation a state in which the flow is purely deviatoric and perfectly plastic.

21.2.2 Hydrostatic Compression

Let us begin by describing the behavior of a soil sample during a hydrostatic compression test. It is a common practice to present the results in terms of the hydrostatic pressure $p = -\sigma_{ii}/3 = -I_1/3$ (positive if the stresses are compressive) and the *specific volume* v defined as the total volume of the sample (including voids) divided by the volume of the solid material (soil particles). The p-v plot in Figure 21.9(a) shows a typical experimental curve including its unloading and reloading branches (dashed curves). In the present context, unloading and reloading are referred to as *swelling* and *recompression*, respectively. Replotting the diagram with the logarithm of pressure on the horizontal axis (Figure 21.9(b)) we realize that, at least within a certain range, the unloading-reloading branches can be idealized as a set of parallel straight lines and the loading branch as a straight line with a different slope. The loading branch is described by

$$v = v_{\mathrm{ref}} - \lambda \ln \frac{p}{p_{\mathrm{ref}}} \tag{21.80}$$

where λ is a certain material parameter corresponding to the slope of the p-v diagram in semi-logarithmic scale, and $(p_{\mathrm{ref}}, v_{\mathrm{ref}})$ is an arbitrary point of that diagram chosen as a state of reference.

Now we can convert the p–v diagram to a representation in terms of the mean stress, $\sigma_m = \sigma_{ii}/3$, and the mean strain, $\varepsilon_m = \varepsilon_{ii}/3$. The mean stress differs from the pressure only by the sign, i.e. $\sigma_m = -p$. The strains are measured from some initial state, which in general is not stress free. For soils, the initial state is usually taken as the in situ state of the material, or as the state of a sample at the beginning of a test. Let us denote the specific volume in the initial state as v_0 and the corresponding pressure as p_0. In the range of small strains, the mean strain is equal to one-third of

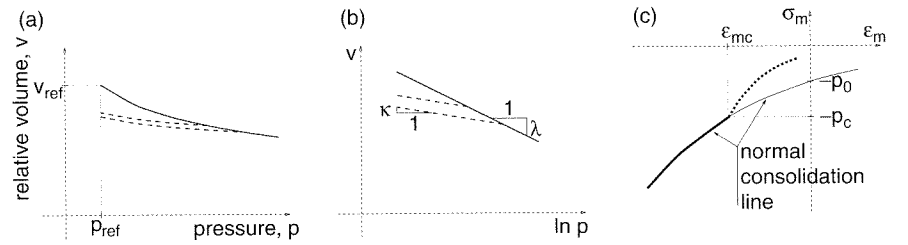


Figure 21.9 a) Volume-pressure plot in natural scale, b) volume-pressure plot in semilogarithmic scale, c) hydrostatic stress-strain diagram

the relative change of volume, i.e.[1]

$$\varepsilon_m = \frac{1}{3}\frac{v - v_0}{v_0} \tag{21.81}$$

Substituting v from (21.76), we obtain

$$\varepsilon_m = \frac{v_{\mathrm{ref}} - v_0}{3v_0} - \frac{\lambda}{3v_0} \ln \frac{p}{p_{\mathrm{ref}}} = \varepsilon_{m,\mathrm{ref}} - \frac{\lambda}{3v_0} \ln \frac{p}{p_{\mathrm{ref}}} \tag{21.82}$$

where $\varepsilon_{m,\mathrm{ref}}$ is the mean strain that would correspond to the reference specific volume used in (21.80). Since the choice of the reference point is completely arbitrary, we may simplify things by setting $v_{\mathrm{ref}} = v_0$, which leads to $\varepsilon_{m,\mathrm{ref}} = 0$ and $p_{\mathrm{ref}} = p_0$. Replacing the pressure by the mean stress with a negative sign, we finally get the volumetric stress-strain law,

$$\varepsilon_m = -\frac{\lambda^*}{3} \ln \frac{-\sigma_m}{p_0} \tag{21.83}$$

where $\lambda^* = \lambda/v_0$ is a material parameter called the *compression index*. Differentiating (21.83), we obtain the rate form of the law,

$$\dot{\varepsilon}_m = -\frac{\lambda^*}{3\sigma_m}\dot{\sigma}_m \tag{21.84}$$

Although we have started from the notation frequently used in critical state soil mechanics, from now on we consider λ^* as a basic material parameter and we reserve the symbol λ for the plastic multiplier.

Equations (21.83) and (21.84) describe only the virgin branch of the stress-strain diagram, which in soil mechanics is referred to as the *normal compression line* or *normal consolidation line*. The branches that correspond to unloading and reloading (swelling and recompression) are described by similar equations with the compression index λ^* replaced by the *swelling index* κ^*, which is related to the unloading slope of the v-p plot. Furthermore, it is natural to select as the reference point the intersection of the unloading branch with the normal compression line. Equations analogous to (21.83) and (21.84) then read

$$\varepsilon_m = \varepsilon_{mc} - \frac{\kappa^*}{3} \ln \frac{-\sigma_m}{p_c} \tag{21.85}$$

[1] It is implicitly assumed that the change of volume is mainly due to the closure of pores (from which water is expelled) while the contribution of the solid particle deformation is negligible. Under this assumption, the total volume is directly proportional to the specific volume.

and

$$\dot{\varepsilon}_m = -\frac{\kappa^*}{3\sigma_m}\dot{\sigma}_m \tag{21.86}$$

where p_c is the *pre-consolidation pressure*, i.e. the largest previously reached pressure, and ε_{mc} is the mean strain corresponding to that pressure.

Now suppose that we test a specimen that has previously experienced pressures up to p_c. The stress-strain diagram in the representation usual in plasticity is plotted in Figure 21.9(c). As long as the current pressure remains below p_c, the evolution of the mean strain is governed by (21.86), independently of the sign of the increment. As such strain changes are reversible, the response may be classified as (nonlinear) elastic. When the applied pressure reaches the largest value that has ever occured in the material, the response begins to follow the normal consolidation line and the strain evolves according to (21.84). However, this is true only if the pressure keeps increasing. For decreasing pressures, the stress-strain curve follows an unloading branch described by (21.86). This means that the strain changes are no longer reversible. We may define the plastic strain rate, $\dot{\varepsilon}_{mp}$, as the difference between the total strain rate, $\dot{\varepsilon}_m$, and the (fictitious) strain rate $\dot{\varepsilon}_{me}$ that would occur if the response were governed by the elastic law (21.86):

$$\dot{\varepsilon}_{mp} = \dot{\varepsilon}_m - \dot{\varepsilon}_{me} = -\frac{\lambda^*}{3\sigma_m}\dot{\sigma}_m - \left(-\frac{\kappa^*}{3\sigma_m}\dot{\sigma}_m\right) = -\frac{\lambda^* - \kappa^*}{3\sigma_m}\dot{\sigma}_m \tag{21.87}$$

It is worth noting that, despite the nonlinear character of the elastic response, the sum of plastic strain increments is equal to the permanent strain that would remain in the material after unloading to the initial pressure. During plastic loading, the pre-consolidation pressure, p_c, is equal to the growing applied pressure, $-\sigma_m$, and so we can rewrite (21.87) as

$$\frac{\dot{p}_c}{p_c} = -\frac{3\dot{\varepsilon}_{mp}}{\lambda^* - \kappa^*} \tag{21.88}$$

and interpret this relation as a special type of hardening law.

Clearly, the range of applicability of the simple exponential stress-strain laws (21.83) and (21.85) is not unlimited. For unloading to very small pressures, equation (21.85) would predict an excessively large volumetric expansion. In fact, for complete unloading to zero pressure the volumetric strain would tend to infinity, which is physically unrealistic. To remedy this problem, it is possible to replace (21.85) by a linear stress-strain law in the range of hydrostatic pressures below a certain threshold, say 4–5 kPa. Usually, it is assumed that the material cannot transmit any hydrostatic tension, and so a perfectly plastic flow would start at $\sigma_m = 0$.

Additional modification would be needed for extremely high hydrostatic pressures, because the basic model (21.83) could lead to physically unacceptable values of volumetric strain (the material cannot be compacted to zero volume). However, for typical applications in geotechnical engineering, the model just presented performs reasonably well.

21.2.3 Three-Dimensional Model (Cam-Clay)

The behavior of soils under hydrostatic compression clearly shows that a realistic model for soils must abandon one of the standard assumptions of classical plasticity, namely that the elastic (reversible) part of the response is linear. For purely

hydrostatic compression, the problem is essentially one-dimensional and can be handled quite easily. Generalization of the nonlinear elastic law (21.85) to general three-dimensional stress states is more difficult. The Cam-Clay model postulates the law governing the elastic response in the rate form

$$\dot{\boldsymbol{\varepsilon}}_e = \boldsymbol{C}_{et}(\boldsymbol{\sigma}) : \dot{\boldsymbol{\sigma}} \tag{21.89}$$

where the tangent elastic compliance tensor $\boldsymbol{C}_{et}$ is not constant but depends on the current stress. Equation (21.89) represents a *hypoelastic formulation*, which is not always convertible to a *hyperelastic formulation* based on an elastic potential. As a consequence, a hypoelastic formulation can result in spurious energy dissipation or energy generation in closed loading cycles, which is not consistent with the thermodynamic interpretation of elastic processes as reversible. In spite of this shortcoming, the hypoelastic approach is widely used in practical engineering models because of its simplicity. In the context of the Cam-Clay model, the compliance tensor $\boldsymbol{C}_{et}$ is assumed to remain isotropic, and so it has the same structure as the standard isotropic elastic compliance tensor, $\boldsymbol{C}_e$, except that the elastic moduli are now variable. Recall that the isotropic elastic properties are fully specified by two elastic moduli. The dependence of the elastic tangent bulk modulus on the mean stress,

$$K_{et} = -\frac{\sigma_m}{\kappa^*} \tag{21.90}$$

follows directly from equation (21.86) describing the volumetric part of the elastic stress-strain law. This relation is usually supplemented by the condition that the Poisson ratio, ν, remains constant. Such an assumption leads to a variable shear modulus

$$G_{et} = \frac{3(1-2\nu)}{2(1+\nu)}K_{et} = -\frac{3(1-2\nu)}{2(1+\nu)}\frac{\sigma_m}{\kappa^*} \tag{21.91}$$

and the elastic tangent stiffness can be expressed as

$$\boldsymbol{D}_{et} = \left(K_{et} - \frac{2}{3}G_{et}\right)\boldsymbol{\delta}\otimes\boldsymbol{\delta} + 2G_{et}\boldsymbol{I}_s = \frac{\sigma_m}{\sigma_{m,\mathrm{ref}}}\boldsymbol{D}_{e,\mathrm{ref}} \tag{21.92}$$

where $\boldsymbol{D}_{e,\mathrm{ref}}$ is the stiffness at a reference value of the mean stress, $\sigma_{m,\mathrm{ref}}$.

Alternatively, one may assume that the shear modulus G remains constant, which leads to a variable Poisson ratio. This formulation, by contrast, can be derived from an elastic potential, and is thus thermodynamically consistent; see Problems 21.5–21.7. The tangential stiffness,

$$\boldsymbol{D}_{et} = \left(K_{et} - \frac{2}{3}G\right)\boldsymbol{\delta}\otimes\boldsymbol{\delta} + 2G\boldsymbol{I}_s = -\frac{\sigma_m}{\kappa^*}\boldsymbol{\delta}\otimes\boldsymbol{\delta} + 2G\left(\boldsymbol{I}_s - \frac{1}{3}\boldsymbol{\delta}\otimes\boldsymbol{\delta}\right) \tag{21.93}$$

is no longer directly proportional to σ_m, but has a constant part and a part growing proportionally to σ_m.

Now we can turn our attention to the plastic response under a general multiaxial stress state. To be consistent with the model for purely volumetric compression, the yield surface should intersect the hydrostatic axis at the origin and at the point corresponding to the pre-consolidation pressure p_c. The modified Cam-Clay model neglects the effect of the Lode angle and uses a yield condition dependent only on invariants I_1 and J_2. In soil mechanics literature, the yield condition is usually

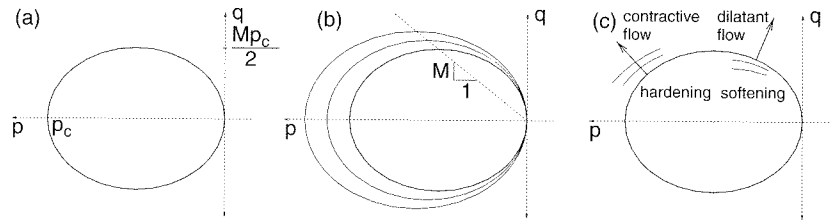


Figure 21.10 a) Elliptical yield surface of the modified Cam-Clay model, b) family of subsequent yield surfaces generated by hardening, c) regions of yield surface corresponding to contractive and dilatant flow

formulated in terms of the hydrostatic pressure, $p = -\sigma_m = -I_1/3$, and a deviatoric stress invariant, $q = \sqrt{3J_2} = \sqrt{3/2}\,\|\boldsymbol{s}\|$. In the (p, q) plane, the yield surface is an ellipse symmetric about the p-axis and passing through the origin and through the point $(p_c, 0)$; see Figure 21.10(a). The shape of the ellipse is described by the ratio of the minor to major axis, M (we tacitly assume that the ellipse is elongated in the p-direction, i.e. its major axis is the one in the p-direction). The shape of the ellipse is assumed not to change during hardening, i.e. M is a constant material parameter; see Figure 21.10(b). The pre-consolidation pressure plays the role of the yield stress that increases during hardening. The yield function is therefore written as

$$f(p, q; p_c) = q^2 - M^2 p(p_c - p) \tag{21.94}$$

Further, it is assumed that during plastic flow the evolution of the pre-consolidation pressure is related to the volumetric part of the plastic strain rate through equation (21.88), originally derived for purely volumetric loading. Note that the yield function (21.94) does not have the form $f(\boldsymbol{\sigma}, p_c) = F(\boldsymbol{\sigma}) - p_c$ typical of isotropic hardening. This is due to the fact that the yield surface is not symmetric with respect to the origin and its center moves during hardening.

To complete the theory, it remains to postulate that the flow rule is associated. The gradient of the yield function with respect to the stress tensor is easily evaluated as

$$\begin{aligned}\boldsymbol{f}_\sigma = \frac{\partial f}{\partial \boldsymbol{\sigma}} &= \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}I_1}\frac{\partial I_1}{\partial \boldsymbol{\sigma}} + \frac{\partial f}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}J_2}\frac{\partial J_2}{\partial \boldsymbol{\sigma}} \\ &= M^2(2p - p_c)\left(-\frac{1}{3}\right)\boldsymbol{\delta} + 2q\frac{\sqrt{3}}{2\sqrt{J_2}}\boldsymbol{s} \\ &= \frac{M^2}{3}(2\sigma_m + p_c)\boldsymbol{\delta} + 3\boldsymbol{s}\end{aligned} \tag{21.95}$$

and so the flow rule reads

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda}\left[3\boldsymbol{s} + \frac{M^2}{3}(2\sigma_m + p_c)\boldsymbol{\delta}\right] \tag{21.96}$$

Now we have all necessary ingredients of the model – the hypoelastic stress-strain law (21.89), yield function (21.94), flow rule (21.96) and hardening law (21.88). The hardening law indicates that the signs of $\dot{p}_c$ and of $\dot{\varepsilon}_{mp}$ are opposite, because $\lambda^* - \kappa^*$ is a positive material parameter and p_c is also always positive. This means that hardening (softening), characterized by $\dot{p}_c > 0$ ($\dot{p}_c < 0$), takes place if the plastic mean strain decreases (increases), i.e. if the flow is contractive (dilatant). Since the

flow rule (21.96) is associated, the flow is contractive in the left half of the elliptical yield surface, and dilatant in its right half (Figure 21.10(c)). At the top apex of the ellipse, the flow is purely deviatoric and, consequently, the material exhibits neither hardening nor softening. For any value of pre-consolidation pressure p_c, the apex of the yield surface lies on the straight line characterized by $q = Mp$; see Figure 21.10(c). This line is the projection on the (p, q)-plane of the so-called *critical state line*, which plays a very important role, as we will see in Example 21.3.

Example 21.2: Derive a formula for the elastoplastic stiffness tensor of the modified Cam-Clay model.

Solution: The model resembles associated elastoplasticity with isotropic hardening but has two nonstandard features: (1) The elastic stress-strain law is nonlinear, and (2) the yield function does not have the form $f(\boldsymbol{\sigma}, \kappa) = F(\boldsymbol{\sigma}) - \bar{h}(\kappa)$. The consistency condition reads

$$\dot{f} = \frac{\partial f}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial f}{\partial p_c}\dot{p}_c = 0 \tag{21.97}$$

where $\partial f/\partial \boldsymbol{\sigma} \equiv \boldsymbol{f}_\sigma$ is given by (21.95) and

$$\frac{\partial f}{\partial p_c} = -M^2 p = M^2 \sigma_m \tag{21.98}$$

From the flow rule (21.96), we can obtain the plastic rate of the mean strain,

$$\dot{\varepsilon}_{mp} = \frac{1}{3}\dot{\boldsymbol{\varepsilon}}_p : \boldsymbol{\delta} = \frac{M^2}{3}(2\sigma_m + p_c)\dot{\lambda} \tag{21.99}$$

and substitute it into the hardening law (21.88) to get a relation between the rates of the pre-consolidation pressure and the plastic multiplier,

$$\dot{p}_c = k\dot{\lambda} \tag{21.100}$$

where

$$k = -\frac{M^2}{\lambda^* - \kappa^*} p_c(2\sigma_m + p_c) \tag{21.101}$$

The usual substitution from the elastic stress-strain law, flow rule and hardening law leads to the expression for the plastic multiplier,

$$\dot{\lambda} = \frac{\boldsymbol{f}_\sigma : \boldsymbol{D}_{et} : \dot{\boldsymbol{\varepsilon}}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_{et} : \boldsymbol{f}_\sigma - \dfrac{\partial f}{\partial p_c}k} \tag{21.102}$$

This formula is analogous to (20.30), with the elastic stiffness $\boldsymbol{D}_e$ replaced by the tangent elastic stiffness $\boldsymbol{D}_{et}$ and the generalized hardening modulus $\bar{H}$ replaced by $-\partial f/\partial p_c$. The same modifications are needed in formula (20.32) for the elastoplastic stiffness:

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_{et} - \frac{\boldsymbol{D}_{et} : \boldsymbol{f}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_{et}}{\boldsymbol{f}_\sigma : \boldsymbol{D}_{et} : \boldsymbol{f}_\sigma - \dfrac{\partial f}{\partial p_c}k} \tag{21.103}$$

Substitution from (21.92), (21.95), (21.98), and (21.101) leads to

$$\boldsymbol{D}_{ep} = \left(K_{et} - \frac{2}{3}G_{et}\right)\boldsymbol{\delta}\otimes\boldsymbol{\delta} + 2G_{et}\boldsymbol{I}_s - \tag{21.104}$$

$$- \frac{M^4K_{et}^2(2\sigma_m + p_c)^2\boldsymbol{\delta}\otimes\boldsymbol{\delta} + 6M^2K_{et}G_{et}(2\sigma_m + p_c)(\boldsymbol{\delta}\otimes\boldsymbol{s} + \boldsymbol{s}\otimes\boldsymbol{\delta}) + 36G_{et}^2\boldsymbol{s}\otimes\boldsymbol{s}}{M^4K_{et}(2\sigma_m + p_c)^2 + 36G_{et}J_2 + M^4(\lambda^* - \kappa^*)^{-1}p_c\sigma_m(2\sigma_m + p_c)}$$

where K_{et} and G_{et} are functions of σ_m given by (21.90) and (21.91), respectively.

Example 21.3: Consider a conventional drained triaxial compression test on a cylindrical soil sample. The specimen is initially subjected to hydrostatic compression. Afterwards, the magnitude of normal stress in the axial direction is progressively increased while the lateral pressure is kept constant. Analyze the response of the modified Cam-Clay model.

Solution: During a drained test, the pore water can flow out freely and no excess pore pressure builds up (to make sure that the specimen is saturated, it is a standard practice to apply a certain pore pressure called the back pressure, which is kept constant during the experiment). Consequently, the changes of the effective stress in the solid skeleton are equal to the changes of the applied (total) stress. The evolution of principal stresses during the experiment can be described as $\sigma_1 = -p_0 - q$ and $\sigma_2 = \sigma_3 = -p_0$ where p_0 is the (given) initial hydrostatic pressure and q is a progressively increasing loading parameter that has the physical meaning of the added pressure in the axial direction. It is easy to verify that $\sqrt{3J_2} = q$, and so the choice of the symbol q is consistent with the definition of the deviatoric invariant q that appears in the yield function (21.91). The evolution of mean pressure during the experiment is obviously given by

$$p = p_0 + \frac{1}{3}q \tag{21.105}$$

If the initial pressure p_0 is smaller than the pre-consolidation pressure p_c, the response will be (nonlinear) elastic until the stress point reaches the yield surface. The hypoelastic law (21.89) can be written separately for the volumetric and deviatoric components:

$$\dot{\varepsilon}_m = \frac{\dot{\sigma}_m}{3K_{et}(\sigma_m)} \tag{21.106}$$

$$\dot{\boldsymbol{e}} = \frac{\dot{\boldsymbol{s}}}{2G_{et}(\sigma_m)} \tag{21.107}$$

Substituting the expressions from (21.90), (21.91) and (21.105) and integrating, we obtain explicit formulae for the strain components as functions of the loading parameter q,

$$\varepsilon_m = -\frac{\kappa^*}{3}\ln\left(1 + \frac{q}{3p_0}\right) \tag{21.108}$$

$$\boldsymbol{e} = -\frac{1+\nu}{1-2\nu}\frac{\kappa^*}{3}\ln\left(1 + \frac{q}{3p_0}\right)\begin{bmatrix}2 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix} \tag{21.109}$$

$$\varepsilon_1 = \varepsilon_m + e_1 = -\frac{\kappa^*}{1-2\nu}\ln\left(1 + \frac{q}{3p_0}\right) \tag{21.110}$$

$$\varepsilon_2 = \varepsilon_m + e_2 = \frac{\kappa^*\nu}{1-2\nu}\ln\left(1 + \frac{q}{3p_0}\right) \tag{21.111}$$

The reference state from which the strains are measured has been chosen as the state just after the application of the hydrostatic pressure p_0. Note that $\varepsilon_2/\varepsilon_1 = -\nu$, as dictated by the assumption of constant Poisson's ratio in the elastic range. Relations (21.108)–(21.111) remain valid until the stress point reaches the yield surface. The value of the loading parameter q at the onset of yielding can be solved from the yield condition into which we substitute p as a function of q according to (21.105). This procedure leads to the quadratic equation

$$\left(\frac{1}{3} + \frac{3}{M^2}\right)q^2 + (2p_0 - p_c)q + 3p_0(p_0 - p_c) = 0 \tag{21.112}$$

and the root of interest is the positive one.

After the onset of yielding, the response is governed by the yield condition (or its rate form – the consistency condition), flow rule, and hardening law. A closed-form solution is difficult to obtain and would be represented by complicated formulae that do not increase insight into the problem. Instead, let us present a qualitative analysis and its geometrical interpretation.

First, suppose that the intersection of the stress path with the initial yield surface is located below the critical state line, i.e. in the left half of the elliptical yield locus; see Figure 21.11(a), in which the initial yield surface is plotted by a thicker line than the subsequent loading surfaces. The associated flow rule (21.96) indicates a contractive behavior and, according to the hardening law (21.88), the yield surface will expand. The stress point then travels upwards along the prescribed stress path, and this process continues until the path intersects the critical state line. After that, the flow becomes purely deviatoric and the yield surface stops evolving, which means

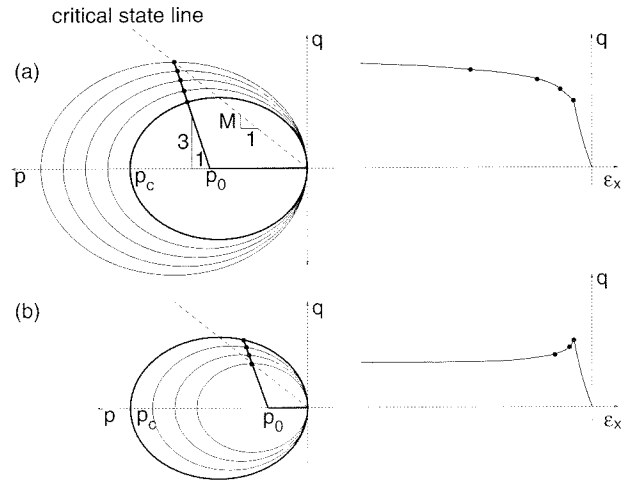


Figure 21.11 Conventional drained triaxial test: a) lightly overconsolidated case, b) heavily overconsolidated case

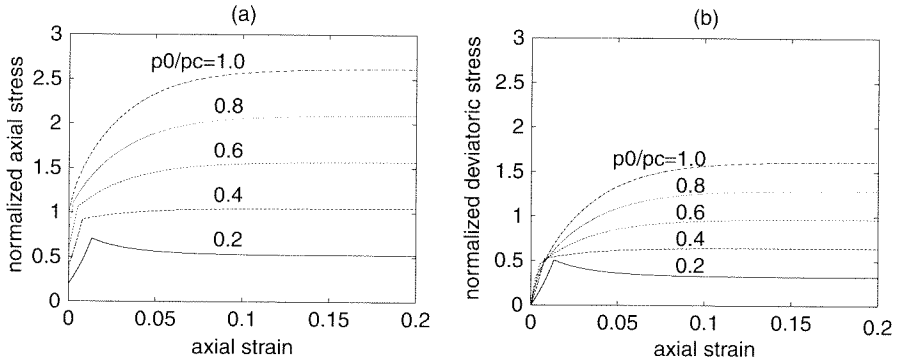


Figure 21.12 Stress-strain curves for drained triaxial test for various ratios of initial hydrostatic pressure p_0 to consolidation pressure p_c: a) normalized axial stress $-\sigma_x/p_c$ as a function of the magnitude of axial strain $-\varepsilon_x$, b) normalized deviatoric stress $q/p_c \equiv \sigma_y - \sigma_x/p_c$ as a function of $-\varepsilon_x$

that the loading process has reached a steady state in which the strength and volume of the material remain constant.

Secondly, suppose that the intersection of the stress path with the initial yield surface is located above the critical state line, i.e. in the right half of the elliptical yield locus; see Figure 21.11(b). The associated flow rule indicates a dilatant behavior and, according to the hardening law, the yield surface will shrink. The stress point then travels downwards along the prescribed stress path, and this process continues until the stress point returns to the critical state line. After that, the flow becomes purely deviatoric and the yield surface stops evolving.

Figure 21.12 shows typical stress-strain curves for ratios p_0/p_c ranging from 0.2 to 1. For convenience, all quantities are plotted with a positive sign and the stresses are normalized by the initial consolidation pressure p_c. The numerical results have been obtained with the following set of parameters, taken from Borja (1991) and representing the Boston Blue Clay: $\lambda^* = 0.032$, $\kappa^* = 0.013$, $M = 1.05$, $\nu = 0.2$. In Figure 21.12(a), the magnitude of the stress in the axial direction, $|\sigma_x| = p_0 + q$, is plotted as a function of the magnitude of axial strain, $|\varepsilon_x|$. The strain is measured from the state after the application of hydrostatic pressure p_0. Figure 21.12(b) shows only the deviatoric part of strain, $|\sigma_x - \sigma_y| = q$, again as a function of $|\varepsilon_x|$. The figures illustrate the qualitative difference between the response to loading paths with a large lateral confining stress, close to the consolidation pressure, and to those with a relatively small confining stress. □

Analyzing the results of the above example, we realize that the type of response is determined by the position of the intersection of the stress path with the initial yield surface. If this intersection is below (or above) the critical state line, the material will exhibit hardening (softening). The limit situation, in which the intersection is exactly on the critical state line, is characterized by $p = p_0 + q/3 = p_c/2$ and $q = Mp_c/2$, from which

$$\frac{p_c}{p_0} = \frac{6}{3 - M} \tag{21.113}$$

If the ratio p_c/p_0 is smaller (or larger) than this critical value, the soil is said to be *lightly overconsolidated* (*heavily overconsolidated*), and the conventional triaxial test under drained conditions leads to hardening (softening) asymptotically approaching

a critical state. The typical values of parameter M are around 1, and so the critical ratio p_c/p_0 is around 3.

As already mentioned in Section 21.1.3, the Cam-Clay yield surface was adapted for concrete by Lin *et al.* (1987). They transformed the elliptical shape of the meridians into slanted ellipses and introduced the dependence on the third invariant. The resulting yield surface remains smooth, convex and bounded.

21.2.4 Other Models

Simple failure surfaces such as the Drucker–Prager surface have been extended to deal with the compression response of soils by attaching a so-called 'cap' – an additional loading surface that provides closure in the direction of hydrostatic compression (DiMaggio and Sandler, 1971). The cap is not a failure surface, but a loading surface that governs strain hardening at high hydrostatic compression, a situation especially important in simulating explosions and impact. A number of refined cap models have been developed and used extensively in dynamic finite element programs for soils, e.g. by Baron, Nelson and Sandler (1973), Sandler, DiMaggio and Baladi (1976) and Desai and Siriwardane (1984). An extension to concrete has also been proposed (Levine, 1982).

Iwan (1967) introduced a special kinematic hardening rule which is effective for hysteresis at cyclic loading. It uses the concept of nested circular yield surface in the deviatoric section. This approach was developed for soils by Prevost (1977, 1978b, 1978a). However, despite the use of many surfaces, the vertex effect (to be discussed in Section 25.4) is not captured because each surface is tangent to all others at the current state point in the stress space. Capturing the vertex effect would require the nested surfaces to intersect at an angle.

Other interesting models for soils have been presented (Lade and Duncan, 1975; Lade, 1977; Cuellar *et al.*, 1977; Nova and Wood, 1979; Mróz, Norris and Zienkiewicz, 1979; Mróz, 1980; Lade and Kim, 1995). Two major groups of incrementally nonlinear hypoplastic models evolved from the ideas of Kolymbas (1987) and Chambon (1989); they are compared in a review article by Tamagnini, Viggiani and Chambon (2000).

Detailed exposition of soil mechanics can be found in the classical textbooks (Terzaghi, 1946; Terzaghi and Peck, 1948; Terzaghi, Peck and Mesri, 1996), and in more recent textbooks and monographs (Schofield and Wroth 1968; Atkinson and Bransby, 1978; Bolton, 1979; Desai and Siriwardane, 1984; Chen and Baladi, 1985; Derski *et al.*, 1989; Wood, Chen and Mizuno, 1990; Chen and Liu, 1990; Das, 1997; Powrie, 1997). Soil dynamics is treated by Das (1983) and rock mechanics by Charlez (1991). The mechanics of soils as multi-phase porous media, including computational aspects, is covered by Lewis and Schrefler (1998). A broad coverage of constitutive models for soils and their finite element implementation is available in Potts and Zdravković (1999).

PROBLEMS

Problem 21.1: Show that (21.7) contains the Mohr–Coulomb criterion as a special case. Find the appropriate form of $r(\theta)$ and express the parameters c_1, c_2 and c_3 in terms of the cohesion, c, and the internal friction angle, ϕ.

Problem 21.2: Find the limits on the parameters K and e from (21.8) and (21.9), within which the resulting failure surface defined by (21.7) is convex.

Problem 21.3: Why is it not a good idea to define the function $r(\theta)$ from (21.7) as a cubic polynomial, even though such a formula would be very simple compared to (21.8) and (21.9) and could give a smooth deviatoric section?

Problem 21.4: Show that the plastic strain obtained by integration of the plastic strain rate (21.87) is equal to the permanent strain that remains in the material after unloading.

Problem 21.5: Construct an elastic potential $W(\varepsilon_m)$ such that the nonlinear elastic law (21.85) can be written in the hyperelastic format $\sigma_m = \mathrm{d}W/\mathrm{d}\varepsilon_m$.

Problem 21.6: Construct an elastic potential $W(\boldsymbol{\varepsilon})$ such that the nonlinear hypoelastic law with tangent stiffness given by (21.93) can be written in the hyperelastic format $\boldsymbol{\sigma} = \partial W/\partial \boldsymbol{\varepsilon}$.
Hint: Note that the volumetric and deviatoric parts of the hypoelastic law, $\dot{\sigma}_m = 3K_{et}\dot{\varepsilon}_m$ and $\boldsymbol{s}_d = 2G\dot{\boldsymbol{e}}$, are uncoupled. The elastic potential can be written as a sum of two terms, respectively, corresponding to the volumetric and deviatoric energy.

Problem 21.7: Show that the nonlinear hypoelastic law with tangent stiffness given by (21.92) cannot be written in the hyperelastic format.
Hint: Consider a stress cycle starting from a hydrostatic state $\boldsymbol{\sigma} = -p_1\boldsymbol{\delta}$ in which we first increase the hydrostatic pressure to p_2, then add a deviatoric component, $\boldsymbol{s}$, after that decrease the hydrostatic pressure back to p_1, and finally remove the deviatoric component.

22

Numerical Methods in Plasticity

22.1 FINITE ELEMENT METHOD

Closed-form solutions exist only for a narrow class of problems, usually those with a highly regular geometry and simple loading. Most of practical problems arising in engineering applications have to be solved in an approximate way by some numerical method. This is true not only for plastic analysis of structures, but for physical and engineering models in general. The most universal numerical technique applicable to problems with arbitrary geometry and boundary conditions is nowadays the *finite element method* (FEM). Even though in certain fields of application it competes with other techniques such as the finite difference method, boundary element method, finite volume method, spectral methods, or so-called meshless methods, the FEM retains its dominant position and is at the core of most of the commercial software packages for engineering computations.

A detailed description of the FEM is outside scope of the present work. Since the method has become a standard technique in computational mechanics, we assume that the reader is familiar with it. However, for convenience and to fix the notation, we briefly review the basic ideas. For a broad and systematic treatment of the subject we refer the reader to one of the classical textbooks (Bathe, 1996; Hughes, 1987; Zienkiewicz and Taylor, 2000). In this chapter we almost exclusively use the engineering notation, i.e. the strain and stress components are arranged into column matrices, and the stiffness tensors are represented by square matrices. This formalism is closer to the actual computer implementation than the tensorial notation. The correspondence between the two notations is explained in Appendix D.7.

22.1.1 Linear Elasticity

Let us start with a problem of linear elasticity, described by the basic equations (D.72)–(D.74), supplemented by boundary conditions (D.75)–(D.76). In the standard displacement version of the FEM, which is a generalization of the direct stiffness method of structural analysis, the displacement components are approximated as linear combinations of suitably chosen interpolation functions (in FEM called the *shape functions*) $N_I(\boldsymbol{x})$, $I = 1, 2, \ldots N_{nod}$. A typical property of the FE shape function is that each of them is associated with one of N_{nod} nodes of the model, and that the value of the Ith shape function is 1 at node number I and 0 at all the other nodes. The displacement approximation reads

$$u_i(\boldsymbol{x}) \approx \sum_{I=1}^{N_{nod}} N_I(\boldsymbol{x}) d_{Ii}, \qquad i = 1, \ldots N_{dim} \tag{22.1}$$

where d_{Ii} are unknown *nodal displacements*, and N_{dim} is the spatial dimension of the model. It is convenient to introduce matrix notation and collect the nodal displacements into a vector (column matrix), $\boldsymbol{d}$, and the shape functions into a matrix, $\boldsymbol{N}$. The approximation (22.1) is then rewritten as

$$\boldsymbol{u}(\boldsymbol{x}) \approx \boldsymbol{N}(\boldsymbol{x})\boldsymbol{d} \tag{22.2}$$

The kinematic (strain-displacement) equations (D.72) provide an approximation of the strains that can be written in the matrix notation as

$$\boldsymbol{\varepsilon}(\boldsymbol{x}) = \boldsymbol{\partial}\boldsymbol{u}(\boldsymbol{x}) \approx \boldsymbol{\partial}\boldsymbol{N}(\boldsymbol{x})\,\boldsymbol{d} = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{d} \tag{22.3}$$

where $\boldsymbol{\partial}$ is the kinematic operator matrix (D.112), $\boldsymbol{B} \equiv \boldsymbol{\partial}\boldsymbol{N}$ is the strain-displacement matrix (often called simply the B-matrix) containing the derivatives of the shape functions with respect to the spatial coordinates, and $\boldsymbol{\varepsilon}$ is a column matrix of the engineering strain components defined in (D.94).

Substituting the strain approximation (22.3) into the constitutive equations of linear elasticity (D.111), we obtain the stress approximation

$$\boldsymbol{\sigma}(\boldsymbol{x}) = \boldsymbol{D}_e(\boldsymbol{x})\boldsymbol{\varepsilon}(\boldsymbol{x}) \approx \boldsymbol{D}_e(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{x})\boldsymbol{d} \tag{22.4}$$

where the argument $\boldsymbol{x}$ at $\boldsymbol{D}_e$ marks explicitly that the elastic properties may vary in space.

By construction, the approximations of displacements, strains and stresses obviously satisfy the kinematic and constitutive equations exactly. It remains for us to use the equilibrium equations. However, since the adopted approximations depend only on a finite number of degrees of freedom (nodal displacements), the differential equations of equilibrium (D.71) cannot in general be satisfied exactly at every point of the body, i.e. in a strong sense. Therefore, they are replaced by the virtual work equality (D.114),

$$\int_V \boldsymbol{\sigma}^T \delta\boldsymbol{\varepsilon}\,\mathrm{d}V = \int_{S_t} \bar{\boldsymbol{t}}^T \delta\boldsymbol{u}\,\mathrm{d}S + \int_V \bar{\boldsymbol{b}}^T \delta\boldsymbol{u}\,\mathrm{d}V \tag{22.5}$$

This is the so-called *weak form* of the equilibrium equations. If the solution is sufficiently smooth, the strong form and the weak form are fully equivalent. However, to construct the equations for the nodal displacements $\boldsymbol{d}$, we test the virtual work equality only for those virtual displacement fields (and the corresponding virtual strain fields) that have the same form as the approximate solution. In other words, we consider the virtual displacements as a linear combination of the shape functions,

$$\delta\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{N}(\boldsymbol{x})\,\delta\boldsymbol{d} \tag{22.6}$$

where $\delta\boldsymbol{d}$ is a column matrix of (arbitrary) virtual nodal displacements. The corresponding virtual strain field derived from the kinematic equations (22.3) is then given by

$$\delta\boldsymbol{\varepsilon}(\boldsymbol{x}) = \boldsymbol{B}(\boldsymbol{x})\,\delta\boldsymbol{d} \tag{22.7}$$

Substituting (22.4), (22.6) and (22.7) into the virtual work equality (22.5), we obtain the *discretized weak form* of equilibrium equations,

$$\int_V \boldsymbol{d}^T\boldsymbol{B}^T(\boldsymbol{x})\boldsymbol{D}_e^T(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{x})\,\delta\boldsymbol{d}\,\mathrm{d}V = \int_{S_t} \bar{\boldsymbol{t}}^T\boldsymbol{N}(\boldsymbol{x})\,\delta\boldsymbol{d}\,\mathrm{d}S + \int_V \bar{\boldsymbol{b}}^T\boldsymbol{N}(\boldsymbol{x})\,\delta\boldsymbol{d}\,\mathrm{d}V \tag{22.8}$$

Taking into account that $\boldsymbol{d}$ and $\delta\boldsymbol{d}$ are not functions of the spatial coordinates and as such can be brought in front of the integrals, we can rewrite (22.8) as

$$\boldsymbol{d}^T\boldsymbol{K}_e^T\,\delta\boldsymbol{d} = \boldsymbol{f}_{\mathrm{ext}}^T\,\delta\boldsymbol{d} \tag{22.9}$$

where

$$\boldsymbol{K}_e = \int_V \boldsymbol{B}^T(\boldsymbol{x})\boldsymbol{D}_e(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{x})\,\mathrm{d}V \tag{22.10}$$

is the (global) *elastic stiffness matrix* and

$$\boldsymbol{f}_{\mathrm{ext}} = \int_{S_t} \boldsymbol{N}^T(\boldsymbol{x})\bar{\boldsymbol{t}}\,\mathrm{d}S + \int_V \boldsymbol{N}^T(\boldsymbol{x})\bar{\boldsymbol{b}}\,\mathrm{d}V \tag{22.11}$$

is the (equivalent) *external force vector*. Equation (22.9) must be satisfied for an arbitrary column matrix of virtual displacement parameters $\delta\boldsymbol{d}$, which is true only if

$$\boldsymbol{K}_e\boldsymbol{d} = \boldsymbol{f}_{\mathrm{ext}} \tag{22.12}$$

These are the *discretized equations of equilibrium* from which it is possible to compute the unknown nodal displacements $\boldsymbol{d}$.

22.1.2 Nonlinear Material Models

The preceding derivation can be extended to the more general case in which the material is not linearly elastic. The principal difference is that the constitutive equations are now nonlinear. Writing them formally as

$$\boldsymbol{\sigma} = \hat{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon}) \tag{22.13}$$

where $\hat{\boldsymbol{\sigma}}$ denotes the constitutive operator, we can replace the stress approximation (22.4) by

$$\boldsymbol{\sigma}(\boldsymbol{x}) \approx \hat{\boldsymbol{\sigma}}(\boldsymbol{B}(\boldsymbol{x})\boldsymbol{d}) \tag{22.14}$$

and the weak form of the equilibrium equations now reads

$$\int_V \hat{\boldsymbol{\sigma}}^T(\boldsymbol{B}(\boldsymbol{x})\boldsymbol{d})\,\boldsymbol{B}(\boldsymbol{x})\,\delta\boldsymbol{d}\,\mathrm{d}V = \int_{S_t} \bar{\boldsymbol{t}}^T\boldsymbol{N}(\boldsymbol{x})\,\delta\boldsymbol{d}\,\mathrm{d}S + \int_V \bar{\boldsymbol{b}}^T\boldsymbol{N}(\boldsymbol{x})\,\delta\boldsymbol{d}\,\mathrm{d}V \tag{22.15}$$

or, equivalently,

$$\boldsymbol{f}_{\mathrm{int}}^T(\boldsymbol{d})\,\delta\boldsymbol{d} = \boldsymbol{f}_{\mathrm{ext}}^T\,\delta\boldsymbol{d} \tag{22.16}$$

where

$$\boldsymbol{f}_{\mathrm{int}}(\boldsymbol{d}) = \int_V \boldsymbol{B}^T(\boldsymbol{x})\,\hat{\boldsymbol{\sigma}}(\boldsymbol{B}(\boldsymbol{x})\boldsymbol{d})\,\mathrm{d}V \tag{22.17}$$

is the column matrix of (equivalent) *internal forces* which now depend on the nodal displacements in a nonlinear way. The final form of the discretized equations of equilibrium is then

$$\boldsymbol{f}_{\mathrm{int}}(\boldsymbol{d}) = \boldsymbol{f}_{\mathrm{ext}} \tag{22.18}$$

which is a set of nonlinear algebraic equations. Equations (22.12) obtained for linear elasticity can be interpreted as a special case of (22.18), with internal forces given by a linear function $\boldsymbol{f}_{\mathrm{int}}(\boldsymbol{d}) = \boldsymbol{K}_e\boldsymbol{d}$.

The nonlinear constitutive equations (22.13) have been written in a form that is simple but not mathematically rigorous. With the exception of nonlinear elasticity, the stress cannot be expressed as a uniquely defined function of only the current strain. It must be described either by a functional that depends on the previous strain history or, which is more convenient for numerical applications, by a function depending on the current values of the strains and also of some *internal variables* characterizing the state of the material, such as the plastic strains $\boldsymbol{\varepsilon}_p$ or the isotropic hardening variable κ. Nevertheless, for the moment we will keep using the simple nonlinear stress-strain law (22.13), which is sufficient for the purpose of illustration. The role of internal variables and of the laws governing their evolution (such as the flow rule, the loading-unloading conditions, and the hardening law) will be discussed in Section 22.3.

22.1.3 Linearization

Numerical solution of nonlinear equations frequently exploits the concept of linearization. Suppose that $\bar{\boldsymbol{\varepsilon}}$ is a given strain state and $\bar{\boldsymbol{\sigma}} = \hat{\boldsymbol{\sigma}}(\bar{\boldsymbol{\varepsilon}})$ is the corresponding stress state evaluated from the constitutive law (22.13). The constitutive function $\hat{\boldsymbol{\sigma}}$ can be expanded into Taylor series in the vicinity of $\bar{\boldsymbol{\varepsilon}}$:

$$\hat{\boldsymbol{\sigma}}(\bar{\boldsymbol{\varepsilon}} + \Delta\boldsymbol{\varepsilon}) \quad \boldsymbol{\sigma}(\boldsymbol{\varepsilon}) + \left.\frac{\partial \hat{\boldsymbol{\sigma}}}{\partial \boldsymbol{\varepsilon}}\right|_{\boldsymbol{\varepsilon}=\bar{\boldsymbol{\varepsilon}}} \Delta\boldsymbol{\varepsilon} + \dots \tag{22.19}$$

Here, a general strain $\boldsymbol{\varepsilon}$ is expressed as the sum of the given initial strain $\bar{\boldsymbol{\varepsilon}}$ and an arbitrary increment $\Delta\boldsymbol{\varepsilon}$. For small $\Delta\boldsymbol{\varepsilon}$, the quadratic and higher-order terms (represented in (22.19) by dots) can be neglected, which leads to the linearized constitutive law

$$\boldsymbol{\sigma} \approx \boldsymbol{\sigma} + \boldsymbol{D}(\bar{\boldsymbol{\varepsilon}})\Delta\boldsymbol{\varepsilon} \tag{22.20}$$

that approximates (22.13) for strain states close to $\bar{\boldsymbol{\varepsilon}}$. In (22.20), $\boldsymbol{D} \equiv \partial\hat{\boldsymbol{\sigma}}/\partial\boldsymbol{\varepsilon}$ is the *tangent material stiffness matrix*, and the presence of $\bar{\boldsymbol{\varepsilon}}$ as an argument indicates that the tangent stiffness in general depends on the state at which it is evaluated. The material properties can also vary in space, and so we may add the spatial coordinate vector $\boldsymbol{x}$ as another argument and write $\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}})$.

With the linearized constitutive equations at hand, we can linearize the expression for internal forces. Suppose that $\boldsymbol{d}$ are given nodal displacements that generate the strain field $\bar{\boldsymbol{\varepsilon}}(\boldsymbol{x}) = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{d}$ and the stress field $\bar{\boldsymbol{\sigma}}(\boldsymbol{x}) = \hat{\boldsymbol{\sigma}}(\bar{\boldsymbol{\varepsilon}}(\boldsymbol{x})) = \hat{\boldsymbol{\sigma}}(\boldsymbol{B}(\boldsymbol{x})\bar{\boldsymbol{d}})$. Taylor expansion of the internal forces around $\bar{\boldsymbol{d}}$ leads to

$$\boldsymbol{f}_{\mathrm{int}}(\boldsymbol{d} + \Delta\boldsymbol{d}) = \boldsymbol{f}_{\mathrm{int}}(\bar{\boldsymbol{d}}) + \left.\frac{\partial \boldsymbol{f}_{\mathrm{int}}}{\partial \boldsymbol{d}}\right|_{\boldsymbol{d}=\bar{\boldsymbol{d}}} \Delta\boldsymbol{d} + \dots \tag{22.21}$$

For a small displacement increment, $\Delta\boldsymbol{d}$, the nonlinear terms represented by dots can be neglected, and then we obtain the linear approximation

$$\boldsymbol{f}_{\mathrm{int}}(\bar{\boldsymbol{d}} + \Delta\boldsymbol{d}) \approx \boldsymbol{f}_{\mathrm{int}}(\bar{\boldsymbol{d}}) + \boldsymbol{K}(\bar{\boldsymbol{d}})\Delta\boldsymbol{d} \tag{22.22}$$

in which $\boldsymbol{K} \equiv \partial\boldsymbol{f}_{\mathrm{int}}/\partial\boldsymbol{d}$ is the tangent stiffness matrix of the structural model, sometimes called the *global tangent stiffness matrix*. Differentiating (22.17), we arrive at

$$\boldsymbol{K} = \frac{\partial \boldsymbol{f}_{\mathrm{int}}}{\partial \boldsymbol{d}} = \frac{\partial}{\partial \boldsymbol{d}} \int_V \boldsymbol{B}^T \hat{\boldsymbol{\sigma}} \,\mathrm{d}V = \int_V \boldsymbol{B}^T \frac{\partial \hat{\boldsymbol{\sigma}}}{\partial \boldsymbol{d}} \,\mathrm{d}V$$

$$= \int_V \boldsymbol{B}^T \frac{\partial \hat{\boldsymbol{\sigma}}}{\partial \boldsymbol{\varepsilon}} \frac{\partial \boldsymbol{\varepsilon}}{\partial \boldsymbol{d}} \,\mathrm{d}V = \int_V \boldsymbol{B}^T \boldsymbol{D}\boldsymbol{B} \,\mathrm{d}V \tag{22.23}$$

This is the well-known formula for the evaluation of the global stiffness matrix. Compared to the linear elastic case, described by (22.10), the only difference is that the matrix of elastic moduli, $\boldsymbol{D}_e$, is replaced by the matrix of tangent moduli, $\boldsymbol{D}$. In a more rigorous notation, (22.23) could be written as

$$\boldsymbol{K}(\bar{\boldsymbol{d}}) = \int_V \boldsymbol{B}^T(\boldsymbol{x}) \boldsymbol{D}\left(\boldsymbol{x}, \boldsymbol{B}(\boldsymbol{x})\bar{\boldsymbol{d}}\right) \boldsymbol{B}(\boldsymbol{x}) \,\mathrm{d}V \tag{22.24}$$

where it is explicitly marked that, at a generic point $\boldsymbol{x}$, the matrix $\boldsymbol{D}$ is evaluated for the strain state $\bar{\boldsymbol{\varepsilon}} = \boldsymbol{B}(\boldsymbol{x})\bar{\boldsymbol{d}}$.

In actual numerical implementation, the domain V is divided into *finite elements* of simple geometrical shapes (e.g. triangles and quadrilaterals in two dimensions), and the shape function associated with a certain node is constructed such that it vanishes outside the patch of elements adjacent to that node. This means that in each element only a few shape functions are nonzero. The contributions of each finite element to the internal force vector and to the stiffness matrix are evaluated separately, and the global matrices are assembled from these contributions. Integration over the volume of each element, V_e, is performed numerically, using the approximations

$$\int_{V_e} \boldsymbol{B}^T(\boldsymbol{x}) \hat{\boldsymbol{\sigma}}\left(\boldsymbol{B}(\boldsymbol{x})\boldsymbol{d}\right) \mathrm{d}V \approx \sum_{k=1}^{N_{\mathrm{int}}^e} w_k^e \boldsymbol{B}^T(\boldsymbol{x}_k^e) \hat{\boldsymbol{\sigma}}\left(\boldsymbol{B}(\boldsymbol{x}_k^e)\boldsymbol{d}\right) \tag{22.25}$$

$$\int_{V_e} \boldsymbol{B}^T(\boldsymbol{x}) \boldsymbol{D}\left(\boldsymbol{x}, \boldsymbol{B}(\boldsymbol{x})\boldsymbol{d}\right) \boldsymbol{B}(\boldsymbol{x}) \,\mathrm{d}V \approx \sum_{k=1}^{N_{\mathrm{int}}^e} w_k^e \boldsymbol{B}^T(\boldsymbol{x}_k^e) \boldsymbol{D}\left(\boldsymbol{x}_k^e, \boldsymbol{B}(\boldsymbol{x}_k^e)\boldsymbol{d}\right) \boldsymbol{B}(\boldsymbol{x}_k^e)$$

where $\boldsymbol{x}_k^e$, $k = 1, 2, \dots N_{\mathrm{int}}^e$, are the integration points of element number e, and w_k^e, $k = 1, 2, \dots N_{\mathrm{int}}^e$, are the corresponding integration weights.

22.1.4 Finite Elements for Elastoplastic Analysis

Before we close the introductory section on finite elements, it is appropriate to mention one important issue which, due to space limitations, cannot be covered in detail. Routine finite element analysis is often performed with very simple types of elements, such as the linear triangle or bilinear isoparametric quadrilateral. Already in the elastic range, the model response can become far too stiff due to various types of locking effects. These effects typically arise due to the poor ability of the low-order displacement interpolation to capture certain types of kinematic fields.

For example, the bilinear quadrilateral element cannot reproduce pure bending without shear. If the standard 2×2 Gauss integration scheme is used, the element locks in bending-dominated situations, the deflections due to bending are grossly underpredicted, and the convergence to the exact solution upon mesh refinement becomes very slow. This is a typical case of the *shear locking*. Certain shell elements suffer by *membrane locking*, and elements for plane strain analysis can suffer by *volumetric locking* if the Poisson's ratio is close to 0.5, i.e. if the material is almost incompressible.

The danger of locking is even more threatening in the plastic range. After the onset of yielding, the plastic part of the strain rate often dominates, because changes of

the elastic strain are proportional to the changes of stress, which are limited since the stress must remain on the yield surface. In fact, if the material is perfectly elastoplastic, the stress in the limit state is constant, and so the elastic strain must be constant as well. All the kinematic changes are then due to the plastic flow. However, the plastic strain tensor at a given point can evolve only in a specific direction, dictated by the flow rule. For materials described by a pressure-independent yield condition and an associated flow rule, plastic flow takes place at constant volume. So, the material acts incrementally as if it were incompressible, and volumetric locking can appear. But even if the plastic flow is dilatant or contractive, it must still satisfy a rather strong kinematic constraint; only the physical interpretation of the constraint is then less straightforward. Depending on the element type and integration scheme, the solution can again be plagued by locking.

A classical example of volumetric locking during isochoric plastic flow in a finite element mesh consisting of linear triangular elements was given by Nagtegaal, Parks and Rice (1974). As a remedy, they proposed to arrange the elements into special cross-diagonal patterns. Many techniques that reduce or completely avoid the locking effects have been proposed since then. Their examples include strain projection methods (Hughes, 1980; Simo, Taylor and Pister, 1985), mixed displacement-pressure formulations, reduced and selective integration techniques (Malkus and Hughes, 1978), and assumed strain methods (Simo and Rifai, 1990). For a systematic exposition of this important subject, the reader is referred to fundamental textbooks on finite element analysis (Hughes, 1987, Chapter 4; Zienkiewicz and Taylor, 2000).

22.2 SOLUTION STRATEGIES IN NONLINEAR ANALYSIS

The objective of nonlinear FE analysis is to trace the response of the structural model subjected to the given loading history. This is usually done using an *incremental-iterative procedure*. The load is applied in a number of steps (increments), and the structural response after each step is computed from the equilibrium equations (22.18). As these equations are in general nonlinear, they have to be solved by an iterative method. This means that inside the top-level loop of load incrementation, we have an iterative loop that restores equilibrium. It is important to distinguish between the criterion that fixes the size of each step, and the iterative procedure that solves the equilibrium equations. Examples of step-size control techniques include the load control, direct or indirect displacement control, or various versions of the arc-length control. Examples of iterative methods include the standard or modified Newton–Raphson iteration, initial stiffness method, or quasi-Newton methods. Additional components of an incremental-iterative solution strategy are the convergence criteria and, optionally, convergence accelerators such as line-search techniques (Matthies and Strang, 1979; Crisfield, 1983) or step-size adjustment rules.

22.2.1 Newton–Raphson Iteration under Load Control

First of all, let us explain the notion of load control. The most natural way of specifying an incremental step is by prescribing the external loads that act on the structure at the end of the step. The loading can be either directly expressed in terms of the equivalent external force vector, $\boldsymbol{f}_{\text{ext}}$, or converted to this format by evaluating the integrals in (22.11). The discretized loading path is therefore described by the sequence $\boldsymbol{f}_{\text{ext}}^{(1)}$, $\boldsymbol{f}_{\text{ext}}^{(2)}$, $\boldsymbol{f}_{\text{ext}}^{(3)}$, etc. Our goal is to find the corresponding nodal displacement vectors $\boldsymbol{d}^{(1)}$, $\boldsymbol{d}^{(2)}$, $\boldsymbol{d}^{(3)}$, etc., such that each couple $(\boldsymbol{d}^{(n)}, \boldsymbol{f}_{\text{ext}}^{(n)})$ satisfies equations (22.18).

Let us now explain the iterative solution procedure applied within a generic loading step number n. At the beginning of the step, we start from the displacements $\boldsymbol{d}^{(n-1)}$ that were computed in the previous step from the condition of equilibrium between the corresponding internal forces $\boldsymbol{f}_{\text{int}}(\boldsymbol{d}^{(n-1)})$ and the external forces $\boldsymbol{f}_{\text{ext}}^{(n-1)}$ (at the beginning of the first step, we typically set $\boldsymbol{d}^{(0)} = \boldsymbol{0}$ and $\boldsymbol{f}_{\text{ext}}^{(0)} = \boldsymbol{0}$, i.e. we start from the initial undeformed state with zero applied loads). The external forces are now increased to $\boldsymbol{f}_{\text{ext}}^{(n)}$, and the resulting displacements $\boldsymbol{d}^{(n)}$ have to be found by solving the equilibrium equations

$$\boldsymbol{f}_{\text{int}}(\boldsymbol{d}^{(n)}) = \boldsymbol{f}_{\text{ext}}^{(n)} \tag{22.26}$$

As the initial guess for $\boldsymbol{d}^{(n)}$ we take $\boldsymbol{d}^{(n,0)} = \boldsymbol{d}^{(n-1)}$. After linearization of the left-hand side around $\boldsymbol{d}^{(n,0)}$, the nonlinear equations (22.26) are replaced by linear equations

$$\boldsymbol{f}_{\text{int}}(\boldsymbol{d}^{(n,0)}) + \boldsymbol{K}(\boldsymbol{d}^{(n,0)})\,\Delta\boldsymbol{d}^{(n,1)} = \boldsymbol{f}_{\text{ext}}^{(n)} \tag{22.27}$$

where $\Delta\boldsymbol{d}^{(n,1)}$ is the first approximation of the unknown displacement increments $\Delta\boldsymbol{d}^{(n)} = \boldsymbol{d}^{(n)} - \boldsymbol{d}^{(n-1)}$. Introducing the simplified notation $\boldsymbol{K}^{(n,0)} \equiv \boldsymbol{K}(\boldsymbol{d}^{(n,0)})$ and $\boldsymbol{f}_{\text{int}}^{(n,0)} = \boldsymbol{f}_{\text{int}}(\boldsymbol{d}^{(n,0)})$, we can rewrite (22.27) as

$$\boldsymbol{K}^{(n,0)}\Delta\boldsymbol{d}^{(n,1)} = \boldsymbol{f}_{\text{ext}}^{(n)} - \boldsymbol{f}_{\text{int}}^{(n,0)} \tag{22.28}$$

The displacement increments $\Delta\boldsymbol{d}^{(n,1)}$ solved from (22.28) satisfy the linearized equations (22.27) exactly, but the updated displacements $\boldsymbol{d}^{(n,1)} = \boldsymbol{d}^{(n,0)} + \Delta\boldsymbol{d}^{(n,1)}$ satisfy the original nonlinear equations (22.26) only approximately. Nevertheless, from the graphical presentation in Figure 22.1(a) it is clear that $\boldsymbol{d}^{(n,1)}$ is typically much closer to the exact solution $\boldsymbol{d}^{(n)}$ than was the initial vector $\boldsymbol{d}^{(n-1)}$. We can now linearize the governing equations around $\boldsymbol{d}^{(n,1)}$ and compute a correction of the displacement increment from equations similar to (22.28), of course with an updated stiffness matrix $\boldsymbol{K}^{(n,1)}$ and internal force vector $\boldsymbol{f}_{\text{int}}^{(n,1)}$. Such iterative correction can be repeated as many times as needed until the equations of equilibrium are satisfied with a sufficient accuracy. However, if the convergence is too slow, it is computationally more effective to reduce the step size than use an excessive number of iterations.

The iterative procedure is formally described by the recursive formulae

$$\left.\begin{aligned} \boldsymbol{K}^{(n,i-1)}\,\delta\boldsymbol{d}^{(n,i)} &= \boldsymbol{f}_{\text{ext}}^{(n)} - \boldsymbol{f}_{\text{int}}^{(n,i-1)} \\ \boldsymbol{d}^{(n,i)} &= \boldsymbol{d}^{(n,i-1)} + \delta\boldsymbol{d}^{(n,i)} \end{aligned}\right\} \quad i = 1, 2, 3, \ldots \tag{22.29}$$

where $\boldsymbol{d}^{(n,i)}$ is the displacement approximation after the ith iteration in step number n , and $\delta\boldsymbol{d}^{(n,i)}$ is the displacement correction applied in the ith iteration. Note that the symbol $\delta\boldsymbol{d}$ now has the meaning of a displacement correction and not of a virtual displacement. We denote the correction by lower case delta, so that $\Delta\boldsymbol{d}$ with capital delta can be reserved for the accumulated displacement increments, defined as

$$\Delta\boldsymbol{d}^{(n,i)} = \sum_{k=1}^{i} \delta\boldsymbol{d}^{(n,k)} \tag{22.30}$$

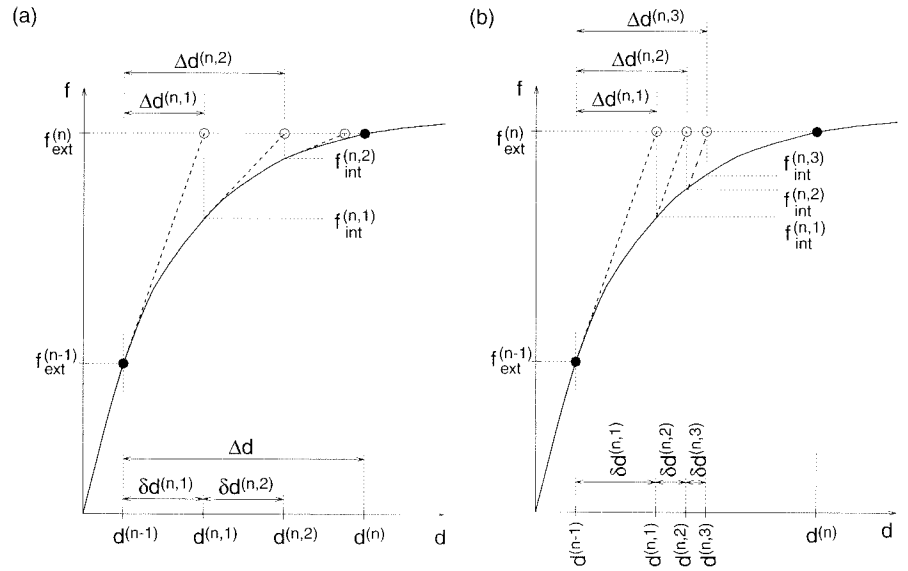


Figure 22.1 a) Standard Newton–Raphson iteration, b) modified Newton–Raphson iteration

Under 'normal' conditions, the sequence of approximations $\boldsymbol{d}^{(n,i)}$ converges to $\boldsymbol{d}^{(n)}$ as $i \to \infty$. Of course, in reality the iteration process is stopped after a finite number of cycles, and the last approximation is taken as the 'converged' value at the end of the step.

For a structure with a single degree of freedom, the iterative procedure just explained corresponds to the classical Newton method applied to the nonlinear equation $r(d) \equiv f_{\text{ext}}^{(n)} - f_{\text{int}}(d) = 0$; see Figure 22.1(a). Its generalization dealing with a set of nonlinear equations is referred to as the *Newton–Raphson method*. Each iteration involves three computationally expensive operations:

i) evaluation of the internal forces,
ii) evaluation of the global tangent stiffness matrix,
iii) solution of a system of linear equations.

For large-scale models, steps (ii) and especially (iii) may be excessively expensive. The computational cost can be reduced if the stiffness matrix is not updated in every iteration. The *modified Newton–Raphson method* evaluates the stiffness matrix only at the beginning of the incremental step and does not alter it in the subsequent iterations. i.e. the matrix $\boldsymbol{K}^{(n,i-1)}$ in (22.29) is replaced by $\boldsymbol{K}^{(n,0)} \equiv \boldsymbol{K}^{(n-1)}$. This means that $\boldsymbol{K}$ is not the exact tangent stiffness matrix anymore (except for the first iteration), and the rate of convergence slows down; see Figure 22.1(b). On the other hand, the increased number of iterations needed to achieve the same accuracy is often more than compensated by the reduced number of operations per iteration. The stiffness matrix is assembled and decomposed only once, so that, in subsequent iterations, step (ii) can be completely skipped, and step (iii) is limited to the forward reduction and backward substitution performed only on the right-hand side vector, without any manipulation of the coefficients stored in the system matrix.

It is also possible to use other modifications of the Newton–Raphson method, e.g. to update the stiffness matrix after every k iterations (where $k > 1$), or to use the elastic stiffness matrix $\boldsymbol{K}^{(0)} = \boldsymbol{K}(\boldsymbol{0}) = \boldsymbol{K}_e$ throughout the entire solution process (i.e. in all the load steps). The latter technique, called the *initial stiffness method*, is the simplest to program.

Another family of iterative algorithms, called the *quasi-Newton methods*, is related to the secant method of finding the root of a nonlinear function. The best-known member of this family is the BFGS method (Matthies and Strang, 1979; Dennis and Schnabel, 1983).

22.2.2 Direct Displacement Control

Independent of the choice of iterative algorithm, any solution strategy using load control fails if the prescribed loads cannot be equilibrated by the mobilized resistance of the structure. This typically happens if the load is monotonically increased until the load-carrying capacity of the structure is exhausted. Such a situation is illustrated in Figure 22.2, where the actual load-displacement diagram can only be traced up to point 5. After the sixth increment, the applied load exceeds the maximum load that can be transmitted by the structure, and the equilibrium equations have no solution. The iteration process then oscillates or diverges, and equilibrium cannot be restored. In many engineering analyses it is sufficient to determine the collapse load of the structure and the displacements at which collapse occurs. Divergence of the iterative process is often considered as an indicator of structural failure, and the last converged step is taken as an approximation of the collapse load. However, finite element simulations of complex engineering problems can diverge for a number of other reasons, many of which are purely numerical and have nothing to do with the real structural failure. It is therefore wise to use techniques that can follow the load-displacement diagram beyond its peak, so that the presence of a limit point can be clearly demonstrated. In addition, the post-peak branch of the load-displacement

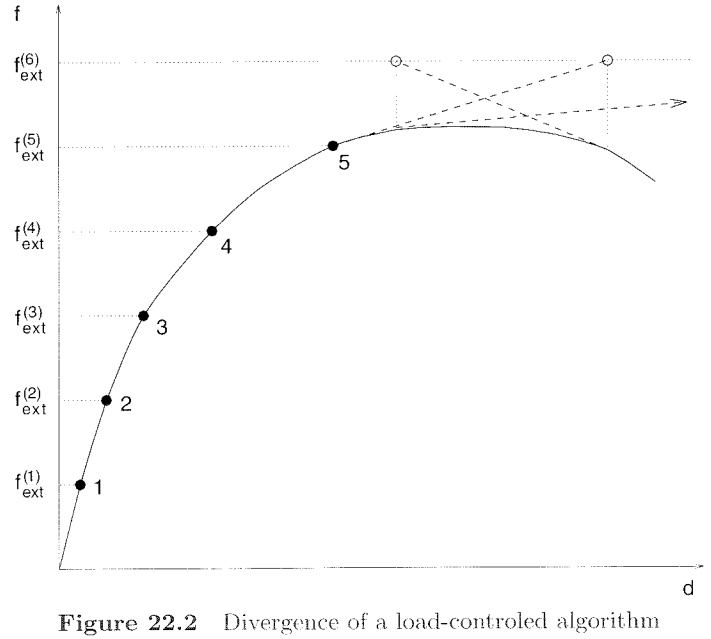


Figure 22.2 Divergence of a load-controled algorithm

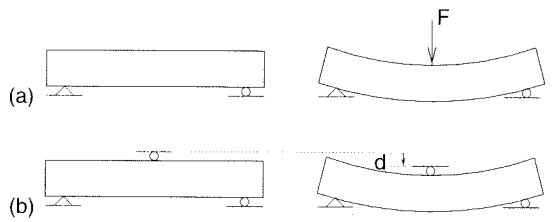


Figure 22.3 Experiment performed under a) load control, b) displacement control

diagram can be of interest in itself, because it provides useful information on the ductility and energy absorption capability of the structure.

The simplest incrementation control technique that can overcome limit points is motivated by physical considerations. Figure 22.3(a) shows a simply supported beam loaded by a concentrated force at midspan. Suppose that the material is perfectly elastoplastic or exhibits softening. Consequently, the load-displacement diagram is terminated by a horizontal plateau or has a peak followed by a descending branch. In an experiment performed under *load control*, the applied load is prescribed, and the resulting deflection is measured. When the applied load exceeds the collapse load of the structure, sudden failure occurs. During such failure, the structure is not in static equilibrium, and the displacements and strains grow in a dynamic process.

Alternatively, the experiment can be controled by the deflection at the point of load application. In this case, the load point can be considered as an additional support with prescribed displacement, and the force acting on the structure is measured as the reaction generated in this support; see Figure 22.3(b). In such an experiment, static equilibrium can be maintained even after the reaction has reached its maximum value, and the load-displacement diagram can be traced into the post-peak range. Since the values of displacements are directly prescribed at certain points by treating them as supports, this technique may be referred to as the *direct displacement control*.

Exactly the same idea can be exploited in numerical simulations. Let us divide the nodal displacements into two groups: The first one contains unknown displacements at nodes that are left 'free', and the second one consists of prescribed displacements at nodes that are controled. Accordingly, we partition the displacement vector into $\{\boldsymbol{d}_1, \boldsymbol{d}_2\}^T$ and the internal and external force vectors into $\{\boldsymbol{f}_{\text{int},1}, \boldsymbol{f}_{\text{int},2}\}^T$ and $\{\boldsymbol{f}_{\text{ext},1}, \boldsymbol{f}_{\text{ext},2}\}^T$, respectively. The external forces $\boldsymbol{f}_{\text{ext},1}$ corresponding to the unknown displacements $\boldsymbol{d}_1$ are prescribed, and for simplicity we will assume that they are equal to zero (treatment of the general case is left to the reader as an easy exercise). The only external forces acting on the structure are the reactions $\boldsymbol{f}_{\text{ext},2}$ at the supports with prescribed displacements $\boldsymbol{d}_2$. The equilibrium equations are partitioned as

$$\boldsymbol{f}_{\text{int},1}(\boldsymbol{d}_1, \boldsymbol{d}_2) = \boldsymbol{0} \tag{22.31}$$

$$\boldsymbol{f}_{\text{int},2}(\boldsymbol{d}_1, \boldsymbol{d}_2) = \boldsymbol{f}_{\text{ext},2} \tag{22.32}$$

For given $\boldsymbol{d}_2$, the unknown displacements $\boldsymbol{d}_1$ can be computed by solving (22.31). After that, the reactions $\boldsymbol{f}_{\text{ext},2}$ are obtained by simple evaluation of the left-hand side in (22.32).

In a typical incremental step number n, we start from the converged displacements $\boldsymbol{d}_1^{(n-1)}$ and $\boldsymbol{d}_2^{(n-1)}$ from the previous step, and we replace (22.31) by the linearized

equations

$$\boldsymbol{f}_{\text{int},1}^{(n-1)} + \boldsymbol{K}_{11}^{(n-1)} \Delta \boldsymbol{d}_1^{(n,1)} + \boldsymbol{K}_{12}^{(n-1)} \Delta \boldsymbol{d}_2^{(n,1)} = \boldsymbol{0} \tag{22.33}$$

where $\boldsymbol{K}_{11} \equiv \partial \boldsymbol{f}_{\text{int},1} / \partial \boldsymbol{d}_1$ and $\boldsymbol{K}_{12} \equiv \partial \boldsymbol{f}_{\text{int},1} / \partial \boldsymbol{d}_2$ are blocks of the global tangent stiffness matrix,

$$\boldsymbol{K} \equiv \frac{\partial \boldsymbol{f}_{\text{int}}}{\partial \boldsymbol{d}} = \begin{bmatrix} \dfrac{\partial \boldsymbol{f}_{\text{int},1}}{\partial \boldsymbol{d}_1} & \dfrac{\partial \boldsymbol{f}_{\text{int},1}}{\partial \boldsymbol{d}_2} \\ \dfrac{\partial \boldsymbol{f}_{\text{int},2}}{\partial \boldsymbol{d}_1} & \dfrac{\partial \boldsymbol{f}_{\text{int},2}}{\partial \boldsymbol{d}_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}_{11} & \boldsymbol{K}_{12} \\ \boldsymbol{K}_{21} & \boldsymbol{K}_{22} \end{bmatrix} \tag{22.34}$$

The increment of the prescribed displacements $\boldsymbol{d}_2$ is known in advance, and so we set $\Delta \boldsymbol{d}_2^{(n,1)} = \Delta \boldsymbol{d}_2^{(n)} \equiv \boldsymbol{d}_2^{(n)} - \boldsymbol{d}_2^{(n-1)}$ and rewrite (22.33) as

$$\boldsymbol{K}_{11}^{(n-1)} \Delta \boldsymbol{d}_1^{(n,1)} = -\boldsymbol{f}_{\text{int},1}^{(n-1)} - \boldsymbol{K}_{12}^{(n-1)} \Delta \boldsymbol{d}_2^{(n)} \tag{22.35}$$

Having solved for $\Delta \boldsymbol{d}_1^{(n,1)}$, we construct the first approximation $\boldsymbol{d}_1^{(n,1)} = \boldsymbol{d}_1^{(n-1)} + \Delta \boldsymbol{d}_1^{(n,1)}$ and $\boldsymbol{d}_2^{(n,1)} = \boldsymbol{d}_2^{(n-1)} + \Delta \boldsymbol{d}_2^{(n,1)} = \boldsymbol{d}_2^{(n)}$. Equations (22.31) are then linearized around $(\boldsymbol{d}_1^{(n,1)}, \boldsymbol{d}_2^{(n,1)})$, corrections of displacements $\boldsymbol{d}_1$ are computed, and the procedure is repeated until the convergence criteria are satisfied. The iterative process can be described by the recursive formulae

$$\left.\begin{aligned} \boldsymbol{K}_{11}^{(n,i-1)} \delta \boldsymbol{d}_1^{(n,i)} &= -\boldsymbol{f}_{\text{int}}^{(n,i-1)} - \boldsymbol{K}_{12}^{(n,i-1)} \delta \boldsymbol{d}_2^{(n,i)} \\ \boldsymbol{d}_1^{(n,i)} &= \boldsymbol{d}_1^{(n,i-1)} + \delta \boldsymbol{d}_1^{(n,i)} \end{aligned}\right\} \quad i = 1, 2, 3, \ldots \tag{22.36}$$

where

$$\boldsymbol{d}_1^{(n,0)} = \boldsymbol{d}_1^{(n-1)} \tag{22.37}$$

$$\boldsymbol{d}_2^{(n,0)} = \boldsymbol{d}_2^{(n-1)} \tag{22.38}$$

$$\delta \boldsymbol{d}_2^{(n,1)} = \boldsymbol{d}_2^{(n)} - \boldsymbol{d}_2^{(n-1)} \tag{22.39}$$

$$\delta \boldsymbol{d}_2^{(n,i)} = \boldsymbol{0} \qquad \text{for } i = 2, 3, \ldots \tag{22.40}$$

Note that, starting from the second iteration, the correction $\delta \boldsymbol{d}_2$ is zero, and so the term with $\boldsymbol{K}_{12}$ on the right-hand side of (22.36) vanishes. This term is present only in the first iteration. It might seem that one could start immediately from $\boldsymbol{d}_2^{(n,0)} = \boldsymbol{d}_2^{(n)}$ instead of $\boldsymbol{d}_2^{(n,0)} = \boldsymbol{d}_2^{(n-1)}$, and then the correction $\delta \boldsymbol{d}_2$ would vanish already in the first iteration and the matrix $\boldsymbol{K}_{12}$ would never have to be evaluated. However, this is, in general, not a good idea because such an initial approximation would be too far from the equilibrium path and the process might diverge or converge too slowly; see Problem 22.1.

22.2.3 Arc Length Control and Indirect Displacement Control

The direct displacement control can be applied in situations when the structure is loaded only at one point, or when the load is transmitted by a stiff platen so that all points on the loaded surface exhibit the same displacements. However, this is not always the case. As an example, consider a dam loaded by hydrostatic pressure due to

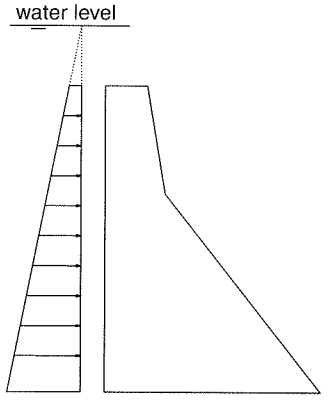


Figure 22.4 Gravity dam loaded by reservoir overflow

reservoir overflow; see Figure 22.4. Here, the load is applied along a large portion of the boundary, and the shape of the corresponding displacement profile is not known in advance. Another case in which direct displacement control fails is very brittle failure characterized by a load-displacement diagram with a snap-back; see Figure 22.5(c).

Advanced incrementation control techniques abandon the assumption that the values of external loads or displacements at supports after each step are prescribed in advance. Instead, the loading program is parameterized by a scalar load multiplier, μ, and the external force vector (or the vector of support displacements) becomes a function of μ. The simplest example is *proportional loading*, characterized by

$$\boldsymbol{f}_{\text{ext}}(\mu) = \mu \bar{\boldsymbol{f}} \tag{22.41}$$

where $\bar{\boldsymbol{f}}$ is a given reference load vector. Equation (22.41) means that all components of the external force vector increase (or decrease) proportionally. Of course, it is possible to prescribe more general loading programs, such as the *linear loading*

$$\boldsymbol{f}_{\text{ext}}(\mu) = \bar{\boldsymbol{f}}_0 + \mu \bar{\boldsymbol{f}} \tag{22.42}$$

with a constant component, $\boldsymbol{f}_0$, and a proportionally increasing component, $\mu\bar{\boldsymbol{f}}$. For example, for the dam problem, $\bar{\boldsymbol{f}}_0$ reflects the effect of gravity and full reservoir pressure, and $\bar{\boldsymbol{f}}$ corresponds to the additional hydrostatic pressure due to a unit overflow.

The basic idea of a flexible incrementation control technique is that the step size is specified by a constraint equation that involves the unknown displacements as well as the load multiplier. The original motivation was provided by the requirement that the size of the step, measured as the geometric distance between the initial and final state in the load-displacement space, should be equal to a prescribed constant. Figure 22.5 illustrates the fact that such a condition combines load control with displacement control in an optimal manner. The load-displacement diagram exhibits a *limit point* (snap-through) and later a *turning point* (snap-back). Figure 22.5(b) shows that load control fails around the limit point, where the diagram has a horizontal tangent. As shown in Figure 22.5(c), displacement control can overcome this point but it fails at snapback, where the diagram has a vertical tangent. The criterion of a constant arc length automatically emphasizes the contribution of the variable that changes faster – the displacement around the limit point and the load parameter around the point of snapback; see Figure 22.5(d).

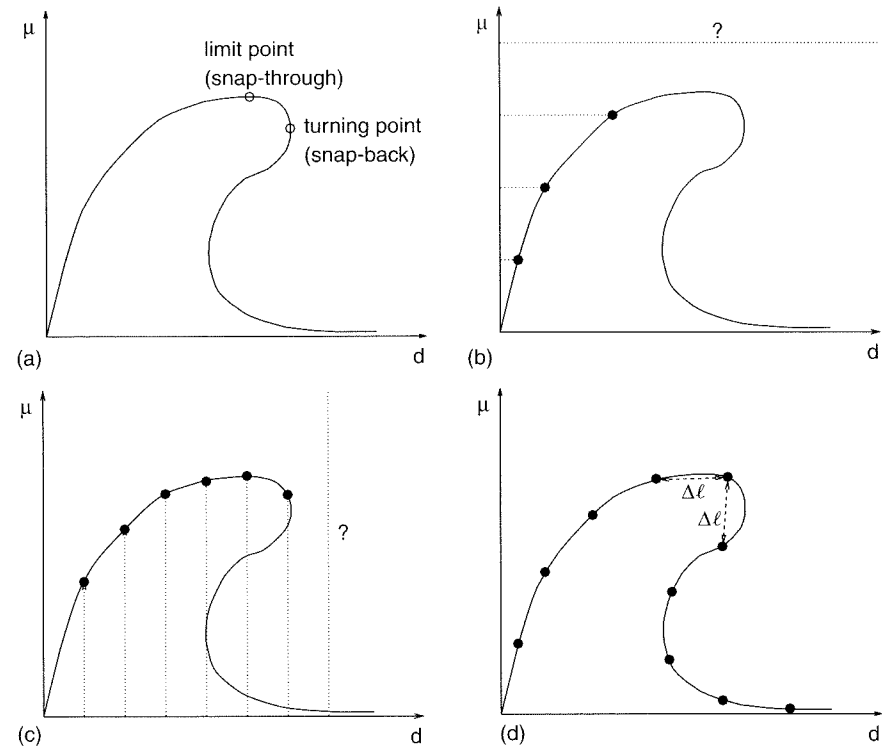


Figure 22.5 a) Load-displacement diagram with snap-back, b) load control, c) direct displacement control, d) arc length control

Despite the apparent simplicity of the condition of a constant arc length, it must be used with caution. First of all, it is important to realize that forces and displacements have completely different units, and so the purely geometrical measure of length in the load-displacement space does not make sense. It is necessary to specify a suitable metric. The easiest way is to introduce a scaling factor, denoted as c, that multiplies the load parameter and converts it into a quantity with the physical dimension of displacement. The 'length' of a step during which the load parameter changes by $\Delta\mu$ and the displacements change by $\Delta\boldsymbol{d}$ is then defined as

$$\Delta\ell = \sqrt{\Delta\boldsymbol{d}^T \Delta\boldsymbol{d} + (c\,\Delta\mu)^2} \tag{22.43}$$

By adjusting the scaling factor, we can amplify or suppress the relative contribution of loads and displacements. One reasonable choice is derived from the condition that the contributions should be equal as long as the response remains linear elastic, which leads to $c = \sqrt{\boldsymbol{d}_e^T \boldsymbol{d}_e}$, where $\boldsymbol{d}_e$ is the solution of $\boldsymbol{K}_e \boldsymbol{d}_e = \bar{\boldsymbol{f}}$. In some cases (e.g. for frame, plate and shell models that use both translational and rotational degrees of freedom), the components of the generalized displacement vector $\boldsymbol{d}$ do not have the same physical dimension. It is then necessary to apply scaling also to the vector $\Delta\boldsymbol{d}$.

Consider an incremental solution process controled by the arc length method. In a typical step number n, we start with displacements $\boldsymbol{d}^{(n-1)}$ and load parameter $\mu^{(n-1)}$ computed in the previous step, and we search for displacements $\boldsymbol{d}^{(n)}$ and load parameter $\mu^{(n)}$. The state at the end of the step must satisfy the equations of equilibrium between the internal forces $\boldsymbol{f}_{\text{int}}(\boldsymbol{d}^{(n)})$ and external forces $\boldsymbol{f}_{\text{ext}}(\mu^{(n)})$.

Compared to the load control or direct displacement control, the load parameter is now an additional unknown. The corresponding additional equation is provided by the constraint that fixes the size of the step.

For example, we can require that the length of the step evaluated from formula (22.43) be equal to a prescribed value, $\overline{\Delta\ell}$. We could treat the problem as a system of $N_{df}+1$ nonlinear equations, where N_{df} is the number of unknown displacement components (degrees of freedom), and solve it by Newton–Raphson iteration. However, a more elegant and computationally more efficient procedure treats the equilibrium equations and the constraint equation to a certain extent separately. Assume for simplicity that the loading program is described by (22.42). The linearized equations of equilibrium in the ith iteration of the nth step read

$$\boldsymbol{K}^{(n,i-1)}\delta\boldsymbol{d}^{(n,i)} = \bar{\boldsymbol{f}}_0 + \mu^{(n,i-1)}\bar{\boldsymbol{f}} - \boldsymbol{f}_{\text{int}}^{(n,i-1)} + \delta\mu^{(n,i)}\bar{\boldsymbol{f}} \tag{22.44}$$

where $\delta\boldsymbol{d}^{(n,i)}$ is the unknown displacement correction, and $\delta\mu^{(n,i)}$ is the unknown correction of the load parameter. The first three terms on the right-hand side are known, and the last term is an unknown scalar multiple of a given vector $\bar{\boldsymbol{f}}$. We can therefore separately solve equations

$$\boldsymbol{K}^{(n,i-1)}\delta\boldsymbol{d}_0 = \boldsymbol{f}_0 + \mu^{(n,i-1)}\bar{\boldsymbol{f}} - \boldsymbol{f}_{\text{int}}^{(n,i-1)} \tag{22.45}$$

$$\boldsymbol{K}^{(n,i-1)}\delta\boldsymbol{d}_f = \boldsymbol{f} \tag{22.46}$$

and then express the displacement correction as

$$\delta\boldsymbol{d}^{(n,i)} = \delta\boldsymbol{d}_0 + \delta\mu^{(n,i)}\,\delta\boldsymbol{d}_f \tag{22.47}$$

When this expression is substituted into the constraint condition,

$$(\Delta\boldsymbol{d}^{(n,i-1)}+\delta\boldsymbol{d}^{(n,i)})^T(\Delta\boldsymbol{d}^{(n,i-1)}+\delta\boldsymbol{d}^{(n,i)})+c^2(\Delta\mu^{(n,i-1)}+\delta\mu^{(n,i)})^2 = (\overline{\Delta\ell})^2 \tag{22.48}$$

we obtain a quadratic equation for a single unknown, $\delta\mu^{(n,i)}$. This equation usually has two real roots, corresponding to the two points of the equilibrium path that have the prescribed distance $\overline{\Delta\ell}$ from point $(\boldsymbol{d}^{(n-1)}, \mu^{(n-1)})$; see Figure 22.6. The correct root is selected depending on the sense in which we march on the equilibrium path (Crisfield, 1981), and the displacement correction is determined from (22.47). After standard updates of the displacement vector and the load parameter, the iteration cycle is repeated until the convergence criteria are satisfied.

There is a vast body of literature related to arc length methods and addressing the criteria for proper selection of the root of (22.48), escape procedures in the case when the roots are imaginary, various modifications of the constraint equation, initial choice and adjustment of the step size $\overline{\Delta\ell}$, and other issues (Riks, 1979; Mathies and Strang, 1979; Bergan, 1980; Ramm, 1981; Crisfield, 1981, 1983; Fried, 1984; Riks, 1984). The purpose of this short summary was to explain the basic idea behind this class of techniques.

The standard arc length control performs well if the entire structure or its large portion participates in the failure mechanism. In the cases where the failure pattern is highly localized, robustness of the technique may deteriorate. The remedy is to adapt the constraint equation to the particular problem and control the incrementation process by a few carefully selected displacement components. Motivation is again provided by the physical behavior of a real structure or specimen.

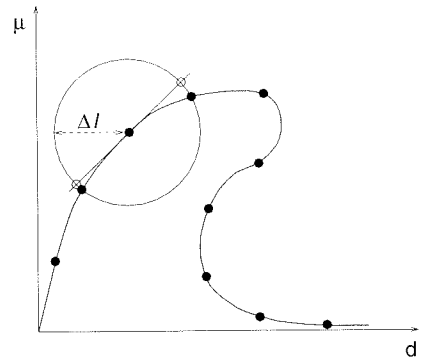


Figure 22.6 Two points on the load-displacement curve satisfying the arc-length constraint

If the load-displacement diagram of a brittle structure exhibits snap-back, direct displacement control applied in an experiment leads to sudden catastrophic failure. When the displacement imposed by the loading device reaches a critical value, fracture starts propagating spontaneously even though the imposed displacement at the load point is kept fixed. However, the crack opening monotonically increases during the entire failure process, and so it can be used as a control variable. If the experimental setup is arranged in a closed loop such that the applied force is continuously adjusted depending on the currently measured value of the crack opening, the response can be traced in a stable manner even after the point at which the load-displacement diagram snaps back.

The same idea can be exploited in a numerical simulation. It is sufficient to select a suitable linear combination of nodal displacements that increases monotonically during the entire failure process, and to use this combination as the control variable. The constraint equation is then formulated as

$$\boldsymbol{c}^T\Delta\boldsymbol{d} = \overline{\Delta\ell} \tag{22.49}$$

where $\boldsymbol{c}$ is a column matrix containing coefficients at individual nodal displacements in the selected linear combination.

In mathematical terms, (22.49) prescribes the projection of the displacement vector onto the direction given by $\boldsymbol{c}$. If we plot the load-displacement diagram with this projected value on the horizontal axis, the incrementation process looks very similar to that driven by direct displacement control. However, the difference lies in the fact that the desired value of the projected displacement cannot be imposed directly (in contrast to the displacement in a support) but the load parameter must be adjusted such that the desired displacement projection is produced. It is therefore suggested this technique is called an *indirect displacement control*. The algorithmic treatment of the problem is very similar to the standard arc length approach, with the quadratic constraint condition (22.48) replaced by a linear one,

$$\boldsymbol{c}^T(\Delta\boldsymbol{d}^{(n,i-1)} + \delta\boldsymbol{d}^{(n,i)}) = \overline{\Delta\ell} \tag{22.50}$$

After substitution from (22.47) we obtain a linear equation, from which the correction of the load multiplier is easily expressed as

$$\delta\mu^{(n,i)} = \frac{\overline{\Delta\ell} - \boldsymbol{c}^T\Delta\boldsymbol{d}^{(n,i-1)} - \boldsymbol{c}^T\delta\boldsymbol{d}_0}{\boldsymbol{c}^T\delta\boldsymbol{d}_f} \tag{22.51}$$

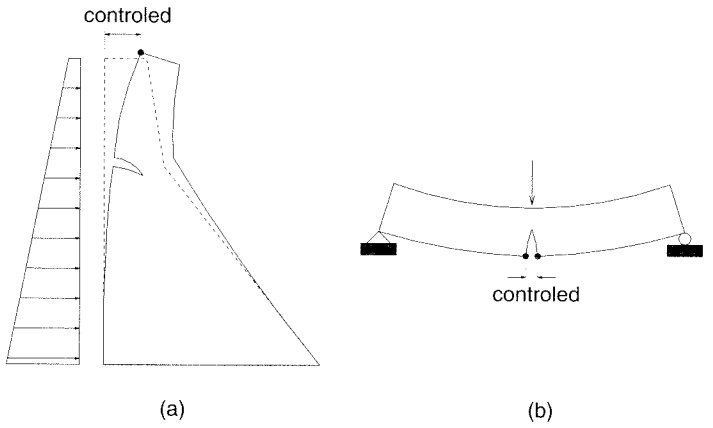


Figure 22.7 Indirect displacement control of a) gravity dam, b) fracture specimen

This approach is simpler than the arc length technique, and it does not require special rules for the selection of the proper root. The key point is, of course, a good choice of the projection vector, $\boldsymbol{c}$. In many practical analyses, the choice can be guided by the expected failure mechanism. It is often sufficient to control the displacement of one specific point of the structure, e.g. horizontal displacement at the crest of the dam in Figure 22.7(a). In simulated fracture experiments, the control variable can be the opening at the mouth of a pre-existing crack (notch), which is expressed as the difference between the displacements of two points located at opposite sides of the crack; see Figure 22.7(b). Procedures leading to an automatic choice of the projection vector have been proposed by Geers (1997).

22.3 STRESS RETURN ALGORITHMS

As explained in the preceding section, nonlinear structural analysis usually proceeds in loading steps, and the equilibrated state at the end of the step is found by an iterative process. In every iteration it is necessary to evaluate the internal forces $\boldsymbol{f}_{\text{int}}^{(n,i)}$ that are produced by a change of displacement parameters from $\boldsymbol{d}^{(n-1)}$ (the last converged state) to $\boldsymbol{d}^{(n,i)}$. This is done according to the formula

$$\boldsymbol{f}_{\text{int}}^{(n,i)} = \sum_{k=1}^{N_{\text{int}}} w_k \boldsymbol{B}^T(\boldsymbol{x}_k)\boldsymbol{\sigma}^{(n,i)}(\boldsymbol{x}_k) \tag{22.52}$$

in which $\boldsymbol{\sigma}^{(n,i)}(\boldsymbol{x}_k)$ are the stresses produced at integration points $\boldsymbol{x}_k$ ($k = 1, 2, \ldots N_{\text{int}}$) by a change of strains from $\boldsymbol{\varepsilon}^{(n-1)}(\boldsymbol{x}_k) = \boldsymbol{B}(\boldsymbol{x}_k)\boldsymbol{d}^{(n-1)}$ to $\boldsymbol{\varepsilon}^{(n,i)}(\boldsymbol{x}_k) = \boldsymbol{B}(\boldsymbol{x}_k)\boldsymbol{d}^{(n,i)}$. The evaluation of $\boldsymbol{\sigma}^{(n,i)}$ is straightforward only if the constitutive equations are formulated as an explicit dependence of stress on strain, which is the case of elasticity. Material models taking into account irreversibility usually work with some *internal variables*, which appear as additional arguments of the constitutive operator that maps strains on stresses. Denoting the set of internal variables by $\boldsymbol{\alpha}$, we can write the constitutive equations as

$$\boldsymbol{\sigma} = \hat{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \tag{22.53}$$

Of course, these equations for stress evaluation must be supplemented by suitable *evolution laws* for the internal variables, such as the flow rule and the loading-unloading conditions in plasticity. The evolution laws often have the form of differential equations, and their integration in a closed form is possible only in very special cases. Most material models require special numerical integration schemes providing approximate solutions of the constitutive equations. In plasticity, such procedures are often referred to as the *stress return algorithms*, because the stresses must be 'returned' to the yield surface.

For simplicity, consider a perfectly elastoplastic material. The role of internal variables is played by the plastic strain, $\boldsymbol{\varepsilon}_p$, and the stresses can be evaluated from the elastic law

$$\boldsymbol{\sigma} = \boldsymbol{D}_e(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) \tag{22.54}$$

which represents a special form of the general relation (22.53). Evolution of the plastic strain is governed by the flow rule, which we state here in the nonassociated form

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda}\boldsymbol{g}_\sigma(\boldsymbol{\sigma}) \tag{22.55}$$

To simplify the notation, we have denoted the gradient of the plastic potential as $\boldsymbol{g}_\sigma \equiv \partial g/\partial\boldsymbol{\sigma}$; cf. (20.109). Equation (22.55), defining only the direction of plastic flow but not its magnitude, must be supplemented by the loading-unloading conditions

$$f(\boldsymbol{\sigma}) \le 0, \qquad \dot{\lambda} \ge 0, \qquad \dot{\lambda} f(\boldsymbol{\sigma}) = 0 \tag{22.56}$$

from which it is possible to compute the rate of the plastic multiplier, $\dot{\lambda}$.

Suppose that we know all the variables (stresses, strains and internal variables) at the beginning of a generic step number n, denoted by superscripts $^{(n-1)}$. The basic task of computational plasticity is to find the values of all the variables at the end of the step, provided that the increment of total strain, $\Delta\boldsymbol{\varepsilon}^{(n)}$, is given. To keep the notation simple, we do not explicitly mark the increment that corresponds to a certain equilibrium iteration number i, because this has no influence on the stress evaluation procedure. In Sections 22.3.1–22.3.4, the superscript $^{(n,i)}$ is replaced by $^{(n)}$.

It is important to bear in mind that the stress at the end of the step should be evaluated for the entire strain increment $\Delta\boldsymbol{\varepsilon}^{(n)}$ applied at once, and not for a sequence of superimposed corrections $\delta\boldsymbol{\varepsilon}^{(n,1)}$, $\delta\boldsymbol{\varepsilon}^{(n,2)}$, etc. The reason is that the intermediate strain approximations $\boldsymbol{\varepsilon}^{(n,1)} = \boldsymbol{\varepsilon}^{(n-1)} + \delta\boldsymbol{\varepsilon}^{(n,1)}$, $\boldsymbol{\varepsilon}^{(n,2)} = \boldsymbol{\varepsilon}^{(n,1)} + \delta\boldsymbol{\varepsilon}^{(n,2)}$, etc., correspond to non-equilibrated states of the structure and they do not represent the actual strain path during the nth step (they strongly depend on the selected iterative technique. while the final converged state should be independent of that technique). Since the elastoplastic constitutive law is path-dependent, approximation of the actual strain path by a series of segments connecting points $\boldsymbol{\varepsilon}^{(n-1)}$, $\boldsymbol{\varepsilon}^{(n,1)}$, $\boldsymbol{\varepsilon}^{(n,2)}$, $\ldots\boldsymbol{\varepsilon}^{(n)}$ can lead to a much larger error than the approximation by a single straight segment connecting $\boldsymbol{\varepsilon}^{(n-1)}$ and $\boldsymbol{\varepsilon}^{(n)}$.

This is illustrated in a uniaxial setting in Figure 22.8. For simplicity, we assume that the global equilibrium is reached in just two iterations. It can easily happen that the first equilibrium iteration overpredicts the actual strain increment, and the subsequent correction has an opposite sign. If the stress is evaluated for a sequence of partial increments and if the material is yielding during the first partial increment from $\varepsilon^{(n-1)}$ to $\varepsilon^{(n,1)}$, the subsequent correction leading from $\varepsilon^{(n,1)}$ to $\varepsilon^{(n,2)} = \varepsilon^{(n)}$ will be considered as elastic unloading and the stress will drop below the yield limit (Figure

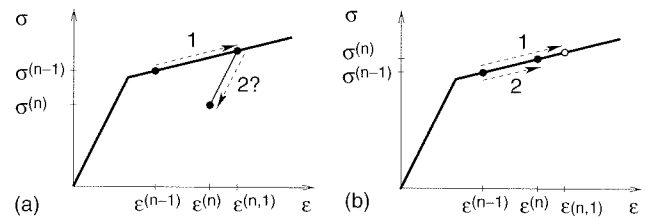


Figure 22.8 Stress evaluation a) from successive corrections (wrong), b) from the entire strain increment applied at once (correct)

22.8(a)). In reality, the strain increases from $\varepsilon^{(n-1)}$ to $\varepsilon^{(n)}$, and the nonconverged state $\varepsilon^{(n,1)}$ is just an inaccurate numerical approximation of $\varepsilon^{(n)}$. The material is therefore yielding during the entire increment. To avoid spurious unloading, it is necessary to discard the information about the previous nonconverged state and evaluate the stress in each iteration starting from $\varepsilon^{(n-1)}$ and applying the entire increment at once (Figure 22.8(b)).

22.3.1 Radial Return Mapping

The stress-strain law (22.54) written at the end of the step reads

$$\boldsymbol{\sigma}^{(n)} = \boldsymbol{D}_e \left(\boldsymbol{\varepsilon}^{(n)} - \boldsymbol{\varepsilon}_p^{(n)} \right) \tag{22.57}$$

The flow rule (22.55) has the character of a differential equation that must be integrated by a numerical technique. The simplest approach is to replace the rates by finite differences and write the discretized form of (22.55) as

$$\boldsymbol{\varepsilon}_p^{(n)} - \boldsymbol{\varepsilon}_p^{(n-1)} = \Delta\lambda^{(n)} \boldsymbol{g}_\sigma(\boldsymbol{\sigma}^{(n)}) \tag{22.58}$$

Note that the flow direction $\boldsymbol{g}_\sigma$ is evaluated for the stress state at the end of the step, $\boldsymbol{\sigma}^{(n)}$, which means that (22.58) corresponds to the *backward Euler scheme*. Finally, the loading-unloading conditions (22.56) are replaced by

$$f(\boldsymbol{\sigma}^{(n)}) \le 0, \qquad \Delta\lambda^{(n)} \ge 0, \qquad \Delta\lambda^{(n)} f(\boldsymbol{\sigma}^{(n)}) = 0 \tag{22.59}$$

The given strain increment directly determines the total strain at the end of the step, $\boldsymbol{\varepsilon}^{(n)} = \boldsymbol{\varepsilon}^{(n-1)} + \Delta\boldsymbol{\varepsilon}^{(n)}$. The values of $\boldsymbol{\sigma}^{(n)}$, $\boldsymbol{\varepsilon}_p^{(n)}$ and $\Delta\lambda^{(n)}$ have to be calculated from (22.57)–(22.59).

Recall the analysis of the constitutive equations in the rate form (i.e. for an infinitely small increment) presented in Section 15.2.3. Now we carry out a similar analysis for a finite increment. According to the loading-unloading conditions (22.59), two basic cases can be distinguished:

1. An *elastic step* is characterized by $\Delta\lambda^{(n)} = 0$. From (22.58) we get $\boldsymbol{\varepsilon}_p^{(n)} = \boldsymbol{\varepsilon}_p^{(n-1)}$, and (22.57) yields

$$\boldsymbol{\sigma}^{(n)} = \boldsymbol{D}_e \left(\boldsymbol{\varepsilon}^{(n)} - \boldsymbol{\varepsilon}_p^{(n-1)} \right) \tag{22.60}$$

   The solution is admissible if the resulting stress is plastically admissible, i.e. if $f(\boldsymbol{\sigma}^{(n)}) \le 0$; see the first loading-unloading condition (22.59). The other two loading-unloading conditions are satisfied automatically.

2. An *elastoplastic step* is characterized by

$$f(\boldsymbol{\sigma}^{(n)}) = 0 \tag{22.61}$$

   Substituting (22.58) into (22.57), we eliminate the plastic strains. The resulting equation,

$$\boldsymbol{\sigma}^{(n)} = \boldsymbol{D}_e \left(\boldsymbol{\varepsilon}^{(n)} - \boldsymbol{\varepsilon}_p^{(n-1)} \right) - \Delta\lambda^{(n)} \boldsymbol{D}_e \boldsymbol{g}_\sigma(\boldsymbol{\sigma}^{(n)}) \tag{22.62}$$

   must be combined with (22.61) and solved for $\boldsymbol{\sigma}^{(n)}$ and $\Delta\lambda^{(n)}$. The solution is admissible if the plastic multiplier does not decrease, i.e. if $\Delta\lambda^{(n)} \ge 0$; see the second loading-unloading condition (22.59).

On the theoretical level, the foregoing two cases have been discussed separately, but their numerical treatment can be integrated into one common algorithm. Its first step consists in the evaluation of the *trial stress*,

$$\boldsymbol{\sigma}^{(tr)} = \boldsymbol{D}_e \left(\boldsymbol{\varepsilon}^{(n)} - \boldsymbol{\varepsilon}_p^{(n-1)} \right) = \boldsymbol{\sigma}^{(n-1)} + \boldsymbol{D}_e \Delta\boldsymbol{\varepsilon}^{(n)} \tag{22.63}$$

The adjective 'trial' refers to the fact that, not knowing whether plastic flow takes place or not, we first try a purely elastic step; cf. (22.60). If $\boldsymbol{\sigma}^{(tr)}$ is plastically admissible, it is accepted as the actual stress, $\boldsymbol{\sigma}^{(n)}$. Since $\boldsymbol{\sigma}^{(tr)}$ appears as the first term on the right-hand side of (22.62), we can rewrite this equation as

$$\boldsymbol{\sigma}^{(n)} + \Delta\lambda^{(n)} \boldsymbol{D}_e \boldsymbol{g}_\sigma(\boldsymbol{\sigma}^{(n)}) = \boldsymbol{\sigma}^{(tr)} \tag{22.64}$$

If the trial stress is not plastically admissible, the trial state $\boldsymbol{\sigma}^{(n)} = \boldsymbol{\sigma}^{(tr)}$ and $\Delta\lambda^{(n)} = 0$ serves as the initial approximation of the solution of (22.61) and (22.64). This initial guess (satisfying (22.64) exactly but violating (22.61)) is then corrected. In general, this requires an iterative solution procedure, which can be based, for example, on the Newton–Raphson technique. Evaluation of the trial stress is sometimes referred to as the *elastic predictor*, and the procedure that returns the stress to the yield surface is then called the *plastic corrector*. The special case presented in the following example can be handled analytically.

Example 22.1: Develop a stress return algorithm based on the backward Euler scheme for von Mises elastoplasticity with an associated flow rule.

Solution: The von Mises condition, $\sqrt{J_2(\boldsymbol{\sigma})} = \tau_0$, in the engineering notation reads $\sqrt{\boldsymbol{s}^T \boldsymbol{P} \boldsymbol{s}/2} = \tau_0$ where $\boldsymbol{s}$ is the deviatoric stress and $\boldsymbol{P}$ is the diagonal scaling matrix defined in (D.97). For associated flow, the flow direction $\boldsymbol{g}_\sigma$ is given by the gradient of the yield function, $\boldsymbol{f}_\sigma = \partial f/\partial\boldsymbol{\sigma} = \boldsymbol{P}\boldsymbol{s}/2\sqrt{J_2} = \boldsymbol{P}\boldsymbol{s}/2\tau_0$. The volumetric response is purely elastic, i.e. $\sigma_V^{(n)} = 3K\varepsilon_V^{(n)}$. For the deviatoric part, equations (22.61) and (22.64) can be rewritten as

$$\sqrt{\boldsymbol{s}^{(n)T} \boldsymbol{P} \boldsymbol{s}^{(n)}} = \sqrt{2}\,\tau_0 \tag{22.65}$$

$$\boldsymbol{s}^{(n)} + \Delta\lambda^{(n)} 2G\boldsymbol{P}^{-1} \frac{\boldsymbol{P}\boldsymbol{s}^{(n)}}{2\tau_0} = \boldsymbol{s}^{(tr)} \tag{22.66}$$

The trial stress deviator is evaluated as

$$\boldsymbol{s}^{(tr)} = 2G\boldsymbol{P}^{-1}(\boldsymbol{e}^{(n)} - \boldsymbol{e}_p^{(n-1)}) = \boldsymbol{s}^{(n-1)} + 2G\boldsymbol{P}^{-1}\Delta\boldsymbol{e}^{(n)} \tag{22.67}$$

where $\boldsymbol{e}$ is the deviatoric strain, $\boldsymbol{e}_p$ is its plastic part (equal to the plastic strain, $\boldsymbol{\varepsilon}_p$, because the flow is purely deviatoric), and G is the elastic shear modulus. Since the product $\boldsymbol{P}^{-1}\boldsymbol{P}$ cancels out from (22.66), it is clear that $\boldsymbol{s}^{(n)}$ must be colinear with $\boldsymbol{s}^{(tr)}$. Then, (22.65) makes it possible to determine the corresponding proportionality factor. The resulting formula is

$$\boldsymbol{s}^{(n)} = \frac{\sqrt{2}\,\tau_0}{\sqrt{\boldsymbol{s}^{(tr)T}\boldsymbol{P}\boldsymbol{s}^{(tr)}}}\boldsymbol{s}^{(tr)} = \frac{\sqrt{2}\,\tau_0}{\|\boldsymbol{s}^{(tr)}\|_\sigma}\boldsymbol{s}^{(tr)} \tag{22.68}$$

where $\|.\|_\sigma$ is the norm defined in (D.98). □

The stress evaluation procedure developed in Example 22.1 has an interesting geometrical interpretation. If the elastically computed trial stress deviator is inside the elastic domain, it is accepted as the actual stress deviator, $\boldsymbol{s}^{(n)}$; see Figure 22.9(a). However, if the trial stress is not plastically admissible, $\boldsymbol{s}^{(n)}$ is obtained by radial scaling of $\boldsymbol{s}^{(tr)}$; see Figure 22.9(b). This technique was proposed by Wilkins (1964) and became known as the *radial return mapping*. The name comes from the fact that the trial stress 'returns' to the yield surface in the radial direction. The algorithm, originally devised for perfect plasticity, was extended to linear hardening by Krieg and Key (1976).

22.3.2 Closest-Point Projection

For associated von Mises elastoplasticity, the radial direction is normal to the yield surface. In a general case, the direction of plastic flow does not have to coincide with the radial direction. However, if the flow rule is associated and the yield surface is convex, the stress $\boldsymbol{\sigma}^{(n)}$ computed from (22.61) and (22.64) corresponds to the point on the yield surface that minimizes the distance (in the metric induced by the elastic stiffness tensor) from the trial stress, $\boldsymbol{\sigma}^{(tr)}$; see Figure 22.9(c). The technique based on the backward Euler approximation of the flow rule is therefore called the *closest-point projection* (CPP) method (Ortiz and Pinsky, 1981; Simo and Taylor, 1985).

As already mentioned, equations (22.64) and (22.61) must, in general, be solved iteratively. For a while, we restrict our attention to the associated case and rewrite these equations as

$$\boldsymbol{\sigma}^{(n)} + \Delta\lambda^{(n)}\boldsymbol{D}_e\,\boldsymbol{f}_\sigma(\boldsymbol{\sigma}^{(n)}) = \boldsymbol{\sigma}^{(tr)} \tag{22.69}$$

$$f(\boldsymbol{\sigma}^{(n)}) = 0 \tag{22.70}$$

Recall that the unknowns are $\boldsymbol{\sigma}^{(n)}$ and $\Delta\lambda^{(n)}$. The initial approximation is taken as $\Delta\lambda^{(n,0)} = 0$ and $\boldsymbol{\sigma}^{(n,0)} = \boldsymbol{\sigma}^{(tr)}$. In a generic iteration cycle number $k+1$ (where $k = 0, 1, 2, \ldots$), corrections $\delta\boldsymbol{\sigma}$ and $\delta\lambda$ are computed from equations (22.69)–(22.70)

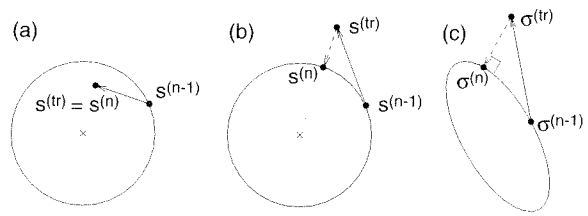


Figure 22.9 a) Elastic step, b) radial return, c) closest-point projection

linearized around the current approximation $\boldsymbol{\sigma}^{(n,k)}$ and $\Delta\lambda^{(n,k)}$. Linearization of the gradient $\boldsymbol{f}_\sigma = \partial f/\partial\boldsymbol{\sigma}$ requires evaluation of the so-called Hessian matrix of the yield function, $\boldsymbol{F}_{\sigma\sigma} = \partial\boldsymbol{f}_\sigma/\partial\boldsymbol{\sigma} = \partial^2 f/\partial\boldsymbol{\sigma}^2$. Denoting $f^{(n,k)} \equiv f(\boldsymbol{\sigma}^{(n,k)})$, $\boldsymbol{f}_\sigma^{(n,k)} \equiv \boldsymbol{f}_\sigma(\boldsymbol{\sigma}^{(n,k)})$, and $\boldsymbol{F}_{\sigma\sigma}^{(n,k)} \equiv \boldsymbol{F}_{\sigma\sigma}(\boldsymbol{\sigma}^{(n,k)})$, we can write the linearized equations as

$$\boldsymbol{\sigma}^{(n,k)} + \delta\boldsymbol{\sigma} + \Delta\lambda^{(n,k)}\boldsymbol{D}_e\,\boldsymbol{f}_\sigma^{(n,k)} + \delta\lambda\boldsymbol{D}_e\,\boldsymbol{f}_\sigma^{(n,k)} + \Delta\lambda^{(n,k)}\boldsymbol{D}_e\,\boldsymbol{F}_{\sigma\sigma}^{(n,k)}\delta\boldsymbol{\sigma} = \boldsymbol{\sigma}^{(tr)} \tag{22.71}$$

$$f^{(n,k)} + \boldsymbol{f}_\sigma^{(n,k)}\delta\boldsymbol{\sigma} = 0 \tag{22.72}$$

To recover a symmetric form of the resulting set of linear equations, we premultiply (22.71) by the elastic compliance matrix $\boldsymbol{C}_e$ and transfer all the known terms to the right-hand side. The resulting linear equations read

$$\begin{bmatrix} \boldsymbol{C}_e + \Delta\lambda^{(n,k)}\boldsymbol{F}_{\sigma\sigma}^{(n,k)} & \boldsymbol{f}_\sigma^{(n,k)} \\ \boldsymbol{f}_\sigma^{(n,k)T} & 0 \end{bmatrix} \begin{Bmatrix} \delta\boldsymbol{\sigma} \\ \delta\lambda \end{Bmatrix} = \begin{Bmatrix} \Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)} - \Delta\lambda^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)} \\ -f^{(n,k)} \end{Bmatrix} \tag{22.73}$$

where $\Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)} = \boldsymbol{C}_e(\boldsymbol{\sigma}^{(tr)} - \boldsymbol{\sigma}^{(n,k)})$. The tilde over $\boldsymbol{\varepsilon}$ emphasizes that $\Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)}$ is the plastic strain increment computed from the difference between the trial stress and the current stress approximation, which is in general different from the plastic strain increment computed from the flow rule, $\Delta\boldsymbol{\varepsilon}_p^{(n,k)} = \Delta\lambda^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)}$. In fact, equation (22.69) could be rewritten as

$$\Delta\lambda^{(n)}\boldsymbol{f}_\sigma(\boldsymbol{\sigma}^{(n)}) = \boldsymbol{C}_e(\boldsymbol{\sigma}^{(tr)} - \boldsymbol{\sigma}^{(n)}) \tag{22.74}$$

and interpreted as the condition that $\Delta\boldsymbol{\varepsilon}_p^{(n,k)}$ and $\Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)}$ be equal. For the converged values $\boldsymbol{\sigma}^{(n)}$ and $\Delta\lambda^{(n)}$, this condition is satisfied exactly and both ways of computing the plastic strain increment give the same result, $\Delta\boldsymbol{\varepsilon}_p^{(n)}$.

If the yield function f is convex, its Hessian matrix $\boldsymbol{F}_{\sigma\sigma}$ is at least positive semi-definite, and the left upper block of the system matrix in (22.73) is positive definite and thus regular. Its inverse will be denoted as

$$\boldsymbol{\Xi}^{(n,k)} = \left(\boldsymbol{C}_e + \Delta\lambda^{(n,k)}\boldsymbol{F}_{\sigma\sigma}^{(n,k)}\right)^{-1} \tag{22.75}$$

Using the first block of equations from (22.73), the unknown

$$\delta\boldsymbol{\sigma} = \boldsymbol{\Xi}^{(n,k)}\left(\Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)} - \boldsymbol{f}_\sigma^{(n,k)}\Delta\lambda^{(n,k)} - \boldsymbol{f}_\sigma^{(n,k)}\delta\lambda\right) \tag{22.76}$$

can be expressed in terms of $\delta\lambda$ and subsequently eliminated from the last equation in (22.73). The resulting condition reads

$$\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{\Xi}^{(n,k)}\left(\Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)} - \boldsymbol{f}_\sigma^{(n,k)}\Delta\lambda^{(n,k)} - \boldsymbol{f}_\sigma^{(n,k)}\delta\lambda\right) = -f^{(n,k)} \tag{22.77}$$

Since $\delta\lambda$ is multiplied by the same factor as $\Delta\lambda^{(n,k)}$, we can directly express their sum.

$$\Delta\lambda^{(n,k+1)} = \Delta\lambda^{(n,k)} + \delta\lambda = \frac{f^{(n,k)} + \boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{\Xi}^{(n,k)}\Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)}}{\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{\Xi}^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)}} \tag{22.78}$$

which represents the improved approximation of $\Delta\lambda^{(n)}$ after iteration number $k+1$. Due to the positive definiteness of $\boldsymbol{\Xi}$, the denominator in (22.78) is always positive. Finally, substituting the result back into (22.76), we obtain the stress correction

$$\delta\boldsymbol{\sigma} = \left[\boldsymbol{\Xi}^{(n,k)} - \frac{\boldsymbol{\Xi}^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{\Xi}^{(n,k)}}{\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{\Xi}^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)}}\right]\Delta\tilde{\boldsymbol{\varepsilon}}_p^{(n,k)} - \frac{f^{(n,k)}\boldsymbol{\Xi}^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)}}{\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{\Xi}^{(n,k)}\boldsymbol{f}_\sigma^{(n,k)}} \tag{22.79}$$

For the updated approximation $\boldsymbol{\sigma}^{(n,k+1)} = \boldsymbol{\sigma}^{(n,k)} + \delta\boldsymbol{\sigma}$ and $\Delta\lambda^{(n,k+1)}$, new equations (22.73) are set up and new corrections are computed. The iterative process is terminated when the right-hand side of (22.73) becomes negligible, i.e. when the original nonlinear equations (22.69)–(22.70) are satisfied with a prescribed tolerance. As we have seen, for smooth, convex, associated plasticity there are no problems with the inversion of a singular matrix in (22.75) or with division by zero in (22.78). Moreover, the iterative process always converges, and it does so at a quadratic rate.

If the flow rule is nonassociated, equations (22.61) and (22.64) can still be solved by the Newton–Raphson method but the interpretation of the algorithm as the closest point projection is lost. In (22.71), the gradient of the yield function, $\boldsymbol{f}_\sigma$, must be replaced by the flow direction, $\boldsymbol{g}_\sigma$, and the Hessian matrix of the yield function, $\boldsymbol{F}_{\sigma\sigma}$, must be replaced by the Hessian matrix of the plastic potential, $\boldsymbol{G}_{\sigma\sigma} = \partial\boldsymbol{g}_\sigma/\partial\boldsymbol{\sigma} = \partial^2 g/\partial\boldsymbol{\sigma}^2$. The system matrix in (22.73) is no longer symmetric but the inverse of the upper left block still exists, provided that g is convex. However, the denominator in (22.78) is replaced by $\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{\Xi}^{(n,k)}\boldsymbol{g}_\sigma^{(n,k)}$ which is not guaranteed to be positive.

22.3.3 Generalized Trapezoidal and Midpoint Rules

Formal integration of the flow rule (22.55) leads to the formula

$$\Delta\boldsymbol{\varepsilon}_p^{(n)} = \int_{\lambda^{(n-1)}}^{\lambda^{(n)}} \boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda)\right)\,\mathrm{d}\lambda \tag{22.80}$$

where the plastic multiplier λ parameterizes the process of plastic yielding. As is clear from (22.58), the CPP method approximates the integral in (22.80) by the product $\Delta\lambda^{(n)}\ \boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda^{(n)})\right)$. This is schematically shown in Figure 22.10(a) where the area under the thick curve corresponds to the exact integral and the shaded rectangle corresponds to the approximate expression. More accurate results can be expected if we use the trapezoidal rule illustrated in Figure 22.10(b),

$$\int_{\lambda^{(n-1)}}^{\lambda^{(n)}} \boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda)\right)\,\mathrm{d}\lambda \approx \Delta\lambda^{(n)}\,\frac{\boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda^{(n-1)})\right) + \boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda^{(n)})\right)}{2} \tag{22.81}$$

or the midpoint rule illustrated in Figure 22.10(c),

$$\int_{\lambda^{(n-1)}}^{\lambda^{(n)}} \boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda)\right)\,\mathrm{d}\lambda \approx \Delta\lambda^{(n)}\ \boldsymbol{g}_\sigma\left(\frac{\boldsymbol{\sigma}(\lambda^{(n-1)}) + \boldsymbol{\sigma}(\lambda^{(n)})}{2}\right) \tag{22.82}$$

Figure 22.10(c) is only schematic because λ is taken as the horizontal coordinate, but the midpoint is computed by averaging $\boldsymbol{\sigma}^{(n-1)}$ and $\boldsymbol{\sigma}^{(n)}$.

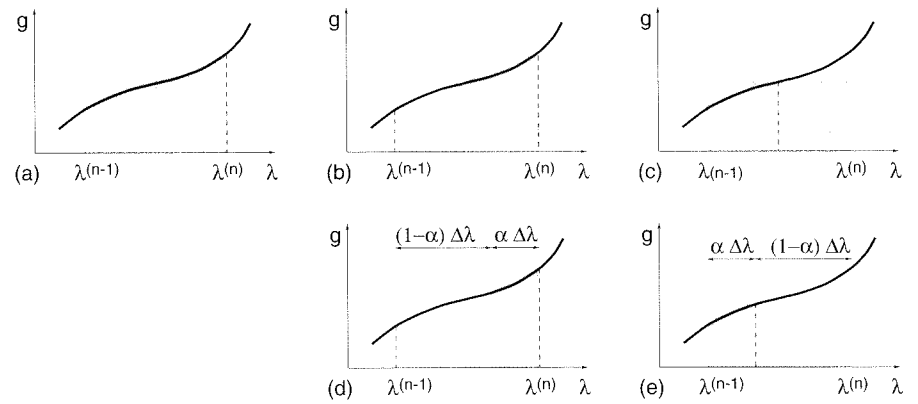


Figure 22.10 Integration schemes: a) backward Euler, b) trapezoidal rule, c) midpoint rule, d) generalized trapezoidal rule, e) generalized midpoint rule

Ortiz and Popov (1985) considered the generalized forms of these integration rules with adjustable weight factors $(1-\alpha)$ and α at the terms corresponding to the initial and final state, respectively. The *generalized trapezoidal rule* (GTR) employs the discretized flow rule in the form

$$\Delta\boldsymbol{\varepsilon}_p^{(n)} = \Delta\lambda^{(n)}\left[(1-\alpha)\boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda^{(n-1)})\right) + \alpha\boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}(\lambda^{(n)})\right)\right] \tag{22.83}$$

and the *generalized midpoint rule* (GMR) uses

$$\Delta\boldsymbol{\varepsilon}_p^{(n)} = \Delta\lambda^{(n)}\ \boldsymbol{g}_\sigma\left([1-\alpha]\boldsymbol{\sigma}(\lambda^{(n-1)}) + \alpha\boldsymbol{\sigma}(\lambda^{(n)})\right) \tag{22.84}$$

A graphical illustration is provided in Figure 22.10(d), (e).

Formula (22.83) or (22.84) replaces (22.58) while equations (22.57) and (22.59) remain unchanged. For $\alpha = 1$, we recover the CPP method as a special case. Ortiz and Popov (1985) evaluated the performance of the GTR and GMR and showed that both techniques are second-order accurate for $\alpha = 0.5$ and first-order accurate for any other $\alpha \in [0,1]$. Second-order accuracy assures fast convergence to the exact solution as the step size tends to zero. In practical applications with large strain increments, the accuracy can nevertheless be better for some other values of α. Ortiz and Popov (1985) constructed the so-called iso-error maps for associated von Mises plasticity, and they concluded that when large strain increments are anticipated, the radial return mapping ($\alpha = 1$) is probably optimal.

The generalized algorithms (GTR and GMR) again lead to a system of nonlinear equations that must be solved iteratively, usually by the Newton Raphson method. To illustrate this procedure, consider the GTR, characterized by equation (22.83). Substituting this into (22.57), we obtain a generalization of (22.62),

$$\boldsymbol{\sigma}^{(n)} = \boldsymbol{D}_e\left(\boldsymbol{\varepsilon}^{(n)} - \boldsymbol{\varepsilon}_p^{(n-1)}\right) - \Delta\lambda^{(n)}\left[(1-\alpha)\boldsymbol{D}_e\boldsymbol{g}_\sigma^{(n-1)} + \alpha\boldsymbol{D}_e\boldsymbol{g}_\sigma\left(\boldsymbol{\sigma}^{(n)}\right)\right] \tag{22.85}$$

Note that the flow direction $\boldsymbol{g}_\sigma^{(n-1)} \equiv \boldsymbol{g}_\sigma(\boldsymbol{\sigma}^{(n-1)})$ can be evaluated from the known stress at the beginning of the step, while the flow direction $\boldsymbol{g}_\sigma(\boldsymbol{\sigma}^{(n)})$ depends on the unknown final stress, $\boldsymbol{\sigma}^{(n)}$. Another unknown variable in (22.85) is the increment of the plastic multiplier, $\Delta\lambda^{(n)}$. For convenience, we define the function

$$\tilde{\boldsymbol{\sigma}}(\boldsymbol{\sigma}, \Delta\lambda) \equiv \boldsymbol{\sigma}^{(tr)} - \Delta\lambda\left[(1-\alpha)\boldsymbol{D}_e\boldsymbol{g}_\sigma^{(n-1)} + \alpha\boldsymbol{D}_e\boldsymbol{g}_\sigma(\boldsymbol{\sigma})\right] \tag{22.86}$$

and we rewrite (22.85) in the form

$$\boldsymbol{\sigma}^{(n)} = \tilde{\boldsymbol{\sigma}}(\boldsymbol{\sigma}^{(n)}, \Delta\lambda^{(n)}) \tag{22.87}$$

Equations (22.87) combined with the yield condition

$$f(\boldsymbol{\sigma}^{(n)}) = 0 \tag{22.88}$$

(which is in fact the discrete counterpart of the consistency condition) represent a system of nonlinear equations from which the unknown variables can be computed. After premultiplying (22.87) by the elastic compliance matrix $\boldsymbol{C}_e$ and linearizing around approximations $\Delta\lambda^{(n,k)}$ and $\boldsymbol{\sigma}^{(n,k)}$, we obtain a set of linear equations

$$\begin{bmatrix} \boldsymbol{C}_e + \alpha\,\Delta\lambda^{(n,k)}\boldsymbol{G}_{\sigma\sigma}^{(n,k)} & (1-\alpha)\boldsymbol{g}_\sigma^{(n-1)} + \alpha\boldsymbol{g}_\sigma^{(n,k)} \\ \boldsymbol{f}_\sigma^{(n,k)T} & 0 \end{bmatrix} \begin{Bmatrix} \delta\boldsymbol{\sigma} \\ \delta\lambda \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{C}_e\left(\hat{\boldsymbol{\sigma}}^{(n,k)} - \boldsymbol{\sigma}^{(n,k)}\right) \\ -f^{(n,k)} \end{Bmatrix} \tag{22.89}$$

in which $\boldsymbol{g}_\sigma^{(n,k)} \equiv \boldsymbol{g}_\sigma(\boldsymbol{\sigma}^{(n,k)})$, $\boldsymbol{G}_{\sigma\sigma}^{(n,k)} \equiv \boldsymbol{G}_{\sigma\sigma}(\boldsymbol{\sigma}^{(n,k)})$, $\tilde{\boldsymbol{\sigma}}^{(n,k)} \equiv \tilde{\boldsymbol{\sigma}}(\boldsymbol{\sigma}^{(n,k)}, \Delta\lambda^{(n,k)})$, $\boldsymbol{f}_\sigma^{(n,k)} \equiv \boldsymbol{f}_\sigma(\boldsymbol{\sigma}^{(n,k)})$ and $f^{(n,k)} \equiv f(\boldsymbol{\sigma}^{(n,k)})$. After $\delta\boldsymbol{\sigma}$ and $\delta\lambda$ are solved, improved estimates $\boldsymbol{\sigma}^{(n,k+1)} = \boldsymbol{\sigma}^{(n,k)} + \delta\boldsymbol{\sigma}$ and $\Delta\lambda^{(n,k+1)} = \Delta\lambda^{(n,k)} + \delta\lambda$ are evaluated. The iterative process is terminated once the residuals $\tilde{\boldsymbol{\sigma}}^{(n,k+1)} - \boldsymbol{\sigma}^{(n,k+1)}$ and $f^{(n,k+1)}$ become negligible.

22.3.4 Generalized Cutting Plane Algorithm

The disadvantage of the procedure described in the preceding subsection is that it requires evaluating the Hessian matrix $\boldsymbol{G}_{\sigma\sigma} = \partial^2 g/\partial\boldsymbol{\sigma}^2$, which can be (for complex plasticity models) extremely tedious. An alternative approach is the *generalized cutting plane* (GCP) algorithm (Simo and Ortiz, 1985; Ortiz and Simo, 1986), which is fully explicit, i.e. it does not require the solution of any system of equations. Of course, such an explicit procedure can be only conditionally stable.[1]

To motivate the GCP algorithm, let us start from the GTR (or GMR) with $\alpha = 0$, which corresponds to the *forward Euler formula*

$$\Delta\boldsymbol{\varepsilon}_p^{(n)} = \Delta\lambda^{(n)}\boldsymbol{g}_\sigma^{(n-1)} \tag{22.90}$$

In this case, the function $\tilde{\boldsymbol{\sigma}}$ defined in (22.86) does not depend on $\boldsymbol{\sigma}$, and then (22.87) provides a formula expressing $\boldsymbol{\sigma}^{(n)}$ in terms of $\Delta\lambda^{(n)}$. Substituting this formula into the discrete consistency condition (22.88), we obtain a single nonlinear equation for $\Delta\lambda^{(n)}$,

$$f\left(\boldsymbol{\sigma}^{(tr)} - \Delta\lambda^{(n)}\boldsymbol{D}_e\boldsymbol{g}_\sigma^{(n-1)}\right) = 0 \tag{22.91}$$

[1] The basic notions of numerical stability are explained in courses and textbooks on numerical solution of differential equations or on structural dynamics. Roughly speaking, a numerical solution is stable if a small perturbation of the initial conditions is not amplified by the integration scheme. If this is true regardless of the step size, the method is called unconditionally stable. If stability can be guaranteed only for step sizes below a certain critical value, the method is conditionally stable.

An iterative solution by the Newton method leads to a sequence of approximations

$$\Delta\lambda^{(n,k+1)} = \Delta\lambda^{(n,k)} + \delta\lambda \tag{22.92}$$

$$\boldsymbol{\sigma}^{(n,k+1)} = \boldsymbol{\sigma}^{(n,k)} - \delta\lambda\,\boldsymbol{D}_e\boldsymbol{g}_\sigma^{(n-1)} \tag{22.93}$$

in which

$$\delta\lambda = \frac{f^{(n,k)}}{\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{D}_e\boldsymbol{g}_\sigma^{(n-1)}} \tag{22.94}$$

Geometrically, we are looking for the intersection of the yield surface with a line constructed from $\boldsymbol{\sigma}^{(n,0)} = \boldsymbol{\sigma}^{(tr)}$ in the direction corresponding to the plastic flow direction at $\boldsymbol{\sigma}^{(n-1)}$ transformed by the elastic operator; this is illustrated for the case of an associated flow in Figure 22.11(a). As shown in Figure 22.11(b), the solution of (22.91) does not always exist. The procedure can be improved by updating the direction of plastic flow in each iteration. This leads to the formulae of the GCP algorithm,

$$\delta\lambda = \frac{f_\sigma^{(n,k)}}{\boldsymbol{f}_\sigma^{(n,k)T}\boldsymbol{D}_e\boldsymbol{g}_\sigma^{(n,k)}} \tag{22.95}$$

$$\Delta\lambda^{(n,k+1)} = \Delta\lambda^{(n,k)} + \delta\lambda \tag{22.96}$$

$$\boldsymbol{\sigma}^{(n,k+1)} = \boldsymbol{\sigma}^{(n,k)} - \delta\lambda\,\boldsymbol{D}_e\boldsymbol{g}_\sigma^{(n,k)} \tag{22.97}$$

The geometric interpretation is given in Figure 22.11(c). The important advantage is that, for a convex yield function and an associated flow, the iterative process always converges (and that it does so at a quadratic rate). The algorithm is explicit, and requires only repeated evaluations of the yield function and its gradient.

22.3.5 Algorithms for Plane Stress Problems

So far, we have presented all the stress return algorithms having in mind general three-dimensional elastoplastic models. Many engineering analyses are performed under the simplifying assumptions of axial symmetry, plane strain or plane stress. The first two cases can be handled by slight modifications of the general procedures (the terms containing those shear components that are known to be zero can be deleted).

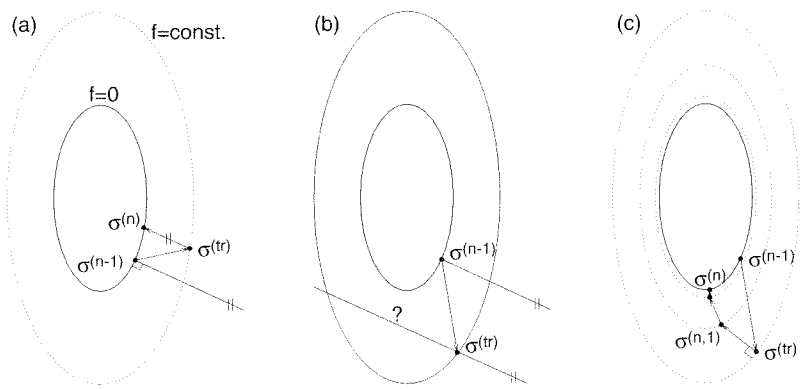


Figure 22.11 a) Stress return corresponding to the forward Euler scheme, b) case when no solution exists, c) generalized cutting plane algorithm

In contrast to that, the case of plane stress needs special attention, since the resulting stress tensor must have a prescribed (zero) value of the out-of-plane normal component, σ_z. The stress return algorithm must lead to a result that is contained in a certain three-dimensional subspace of the six-dimensional space of symmetric second-order tensors. The difficulty stems from the fact that the direction of stress relaxation by plastic flow (given by $\boldsymbol{D}_e\boldsymbol{g}_\sigma$) is not generally contained in this subspace. For this reason, it is not possible to simply reduce the column matrices of stress and strain components to $\boldsymbol{\sigma} = \{\sigma_x, \sigma_y, \sigma_{xy}\}^T$ and $\boldsymbol{\varepsilon} = \{\varepsilon_x, \varepsilon_y, \gamma_{xy}\}^T$, and keep using the general formulae with $\boldsymbol{D}_e$ replaced by the elastic stiffness matrix corresponding to plane stress.

A rather general remedy is to include the out-of-plane normal components σ_z and ε_z in $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$, and to treat ε_z as an unknown that can be determined from the condition $\sigma_z = 0$. An alternative, elegant procedure was proposed by Simo and Taylor (1986). The basic idea is to perform the return mapping directly in the constrained plane stress subspace.

22.4 ALGORITHMIC STIFFNESS

It is important to correctly understand the role of the stress evaluation algorithm in the context of the incremental-iterative solution procedure. Every incremental step n starts from the converged state obtained in the previous step, and the state at the end of the step is determined iteratively. In every iteration i (and at every Gauss integration point), it is necessary to evaluate the stress that would be produced by a change of the strain field from $\boldsymbol{\varepsilon}^{(n-1)}$ to $\boldsymbol{\varepsilon}^{(n,i)}$. The stress return algorithm may involve another iteration loop, nested in the global equilibrium iteration loop. The approximate solution of (22.69) (22.70) or of (22.87)–(22.88) after the kth cycle of the inner loop can be denoted as $\boldsymbol{\sigma}^{(n,i,k)}$ (for explaining the iterative stress return algorithms in the previous section, we simplified the notation by omitting the superscript i that refers to the global equilibrium iteration). As $k \to \infty$, the approximation usually converges to the exact solution of (22.87)–(22.88), now denoted as $\boldsymbol{\sigma}^{(n,i)}$. In the next cycle of the equilibrium iteration process, the stress evaluation is repeated with an updated value of the final strain, $\boldsymbol{\varepsilon}^{(n,i+1)}$, but with the same initial strain, $\boldsymbol{\varepsilon}^{(n-1)}$.

The stress evaluation algorithm can be conveniently described by a function, $\boldsymbol{\theta}^{(n)}$, mapping an arbitrary strain $\boldsymbol{\varepsilon}$ on the stress that would be produced by a change of strain from the given value $\boldsymbol{\varepsilon}^{(n-1)}$ to $\boldsymbol{\varepsilon}$. The resulting stress, $\boldsymbol{\theta}^{(n)}(\boldsymbol{\varepsilon})$, in general depends not only on the final strain, $\boldsymbol{\varepsilon}$, but also on the state variables (strain, plastic strain and hardening variables) at the beginning of the step – this is reflected by the superscript $^{(n)}$. The specific form of $\boldsymbol{\theta}^{(n)}$ depends on the choice of a particular stress return algorithm. For example, for the radial return algorithm applied to associated von Mises elastoplasticity (without hardening), we have

$$\boldsymbol{\theta}^{(n)}(\boldsymbol{\varepsilon}) = \sqrt{2}\tau_0 \frac{\boldsymbol{s}^{(tr)}(\boldsymbol{e})}{\|\boldsymbol{s}^{(tr)}(\boldsymbol{e})\|_\sigma} + K\boldsymbol{\delta}\boldsymbol{\delta}^T\boldsymbol{\varepsilon} \tag{22.98}$$

where $\boldsymbol{e} = \boldsymbol{I}_D\boldsymbol{\varepsilon}$ is the deviatoric part of $\boldsymbol{\varepsilon}$, $\boldsymbol{s}^{(tr)}(\boldsymbol{e}) = 2G\boldsymbol{P}^{-1}(\boldsymbol{e} - \boldsymbol{e}_p^{(n-1)})$ is the trial stress deviator considered as a function of $\boldsymbol{e}$, $\|.\|_\sigma$ is the norm defined in (D.98), K is the bulk modulus of elasticity, and $\boldsymbol{\delta}$ is the column matrix defined in (D.100) and representing Kronecker delta in the engineering notation. The first term on the right-hand side of (22.98) is the deviatoric stress evaluated by the radial return (22.68) while the second term is the elastically evaluated volumetric stress.

Linearization of the internal forces around the (nonequilibrated) state after the ith global equilibrium iteration leads to

$$\begin{aligned}
\boldsymbol{f}_{\text{int}}(\boldsymbol{d}^{(n,i)} + \delta\boldsymbol{d}) &= \int_V \boldsymbol{B}^T\boldsymbol{\theta}^{(n)}(\boldsymbol{B}\boldsymbol{d}^{(n,i)} + \boldsymbol{B}\,\delta\boldsymbol{d})\,\mathrm{d}V \\
&\approx \int_V \boldsymbol{B}^T\boldsymbol{\theta}^{(n)}(\boldsymbol{\varepsilon}^{(n,i)})\,\mathrm{d}V + \int_V \boldsymbol{B}^T \left.\frac{\partial\boldsymbol{\theta}^{(n)}}{\partial\boldsymbol{\varepsilon}}\right|_{\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}^{(n,i)}} \boldsymbol{B}\,\delta\boldsymbol{d}\,\mathrm{d}V \\
&= \int_V \boldsymbol{B}^T\boldsymbol{\sigma}^{(n,i)}\,\mathrm{d}V + \int_V \boldsymbol{B}^T\boldsymbol{\Theta}^{(n)}(\boldsymbol{\varepsilon}^{(n,i)})\boldsymbol{B}\,\mathrm{d}V\;\delta\boldsymbol{d}
\end{aligned} \tag{22.99}$$

where we have denoted

$$\boldsymbol{\Theta}^{(n)} \equiv \frac{\partial\boldsymbol{\theta}^{(n)}}{\partial\boldsymbol{\varepsilon}} \tag{22.100}$$

If we want to write (22.99) in the form

$$\boldsymbol{f}_{\text{int}}(\boldsymbol{d}^{(n,i)} + \delta\boldsymbol{d}) \approx \boldsymbol{f}_{\text{int}}^{(n,i)} + \boldsymbol{K}^{(n,i)}\delta\boldsymbol{d} \tag{22.101}$$

we must define the global stiffness matrix as

$$\boldsymbol{K}^{(n,i)} = \int_V \boldsymbol{B}^T\boldsymbol{\Theta}^{(n)}(\boldsymbol{\varepsilon}^{(n,i)})\boldsymbol{B}\,\mathrm{d}V \tag{22.102}$$

This is similar to the standard formula (22.23), with the only difference that the tangent material stiffness matrix is replaced by the Jacobi matrix of the mapping $\boldsymbol{\theta}^{(n)}$. Now, an important point is that in general

$$\boldsymbol{\Theta}^{(n)}(\boldsymbol{\varepsilon}) \neq \boldsymbol{D}(\boldsymbol{\varepsilon}) \tag{22.103}$$

i.e. the matrix obtained by consistent linearization of the stress return algorithm is different from the tangent material stiffness matrix.

Example 22.2: Perform consistent linearization of the radial return mapping and compare the resulting matrix $\boldsymbol{\Theta}$ with the elastoplastic tangent stiffness derived in Example 15.3.

Solution: Radial return mapping is described by formula (22.98). Applying the chain rule of differentiation we obtain

$$\boldsymbol{\Theta}^{(n)} \equiv \frac{\partial\boldsymbol{\theta}^{(n)}}{\partial\boldsymbol{\varepsilon}} = \sqrt{2}\tau_0 \frac{\partial}{\partial\boldsymbol{s}^{(tr)}}\left(\frac{\boldsymbol{s}^{(tr)}}{\|\boldsymbol{s}^{(tr)}\|_\sigma}\right)\frac{\partial\boldsymbol{s}^{(tr)}}{\partial\boldsymbol{e}}\frac{\partial\boldsymbol{e}}{\partial\boldsymbol{\varepsilon}} + K\boldsymbol{\delta}\boldsymbol{\delta}^T \tag{22.104}$$

It is useful to derive first the formula

$$\frac{\partial\|\boldsymbol{s}\|_\sigma}{\partial\boldsymbol{s}} = \frac{\partial\sqrt{\boldsymbol{s}^T\boldsymbol{P}\boldsymbol{s}}}{\partial\boldsymbol{s}} = \frac{2\boldsymbol{P}\boldsymbol{s}}{2\sqrt{\boldsymbol{s}^T\boldsymbol{P}\boldsymbol{s}}} = \frac{\boldsymbol{P}\boldsymbol{s}}{\|\boldsymbol{s}\|_\sigma} \tag{22.105}$$

Now we can evaluate the individual terms in (22.104):

$$\frac{\partial}{\partial \boldsymbol{s}^{(tr)}}\left(\frac{\boldsymbol{s}^{(tr)}}{\|\boldsymbol{s}^{(tr)}\|_\sigma}\right) = \frac{\boldsymbol{I}\|\boldsymbol{s}^{(tr)}\|_\sigma - \boldsymbol{s}^{(tr)}\left(\dfrac{\partial\|\boldsymbol{s}^{(tr)}\|_\sigma}{\partial \boldsymbol{s}^{(tr)}}\right)^T}{\|\boldsymbol{s}^{(tr)}\|_\sigma^2}$$

$$= \frac{1}{\|\boldsymbol{s}^{(tr)}\|_\sigma}\left(\boldsymbol{I} - \frac{\boldsymbol{s}^{(tr)}\boldsymbol{s}^{(tr)T}\boldsymbol{P}}{\|\boldsymbol{s}^{(tr)}\|_\sigma^2}\right) \tag{22.106}$$

$$\frac{\partial \boldsymbol{s}^{(tr)}}{\partial \boldsymbol{e}} = \frac{\partial}{\partial \boldsymbol{e}}\left[2G\boldsymbol{P}^{-1}\left(\boldsymbol{e} - \boldsymbol{e}_p^{(n-1)}\right)\right] = 2G\boldsymbol{P}^{-1} \tag{22.107}$$

$$\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{\varepsilon}} = \frac{\partial}{\partial \boldsymbol{\varepsilon}}\left(\boldsymbol{I}_D\boldsymbol{\varepsilon}\right) = \boldsymbol{I}_D = \boldsymbol{I} - \frac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T \tag{22.108}$$

Substituting this back into (22.104) and taking into account that $\boldsymbol{s}^{(tr)T}\boldsymbol{\delta} = 0$, we obtain the Jacobi matrix of the radial return mapping,

$$\boldsymbol{\Theta}^{(n)} = \sqrt{2}\tau_0\frac{1}{\|\boldsymbol{s}^{(tr)}\|_\sigma}\left(\boldsymbol{I} - \frac{\boldsymbol{s}^{(tr)}\boldsymbol{s}^{(tr)T}\boldsymbol{P}}{\|\boldsymbol{s}^{(tr)}\|_\sigma^2}\right)2G\boldsymbol{P}^{-1}\left(\boldsymbol{I} - \frac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T\right) + K\boldsymbol{\delta}\boldsymbol{\delta}^T$$

$$= \frac{2\sqrt{2}G\tau_0}{\|\boldsymbol{s}^{(tr)}\|_\sigma}\left(\boldsymbol{P}^{-1} - \frac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T - \frac{\boldsymbol{s}^{(tr)}\boldsymbol{s}^{(tr)T}}{\|\boldsymbol{s}^{(tr)}\|_\sigma^2}\right) + K\boldsymbol{\delta}\boldsymbol{\delta}^T \tag{22.109}$$

For the present simple model, the elastoplastic stiffness matrix from (15.62) would be in the engineering notation given by

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{G}{\tau_0^2}\boldsymbol{s}\boldsymbol{s}^T = 2G\left(\boldsymbol{P}^{-1} - \frac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T - \frac{\boldsymbol{s}\boldsymbol{s}^T}{2\tau_0^2}\right) + K\boldsymbol{\delta}\boldsymbol{\delta}^T \tag{22.110}$$

Comparison of (22.109) and (22.110) reveals that both expressions only coincide if $\boldsymbol{s}^{(tr)} = \boldsymbol{s}$ and $\|\boldsymbol{s}^{(tr)}\|_\sigma = \sqrt{2}\tau_0$. □

The difference between the stiffness matrices in (22.109) and (22.110) can be explained as follows:

- The elastoplastic material stiffness matrix $\boldsymbol{D}^{(n)}$ is related to a given *state*, characterized by a strain $\boldsymbol{\varepsilon}^{(n)}$ and a stress $\boldsymbol{\sigma}^{(n)}$ satisfying the yield condition $f(\boldsymbol{\sigma}^{(n)}) = 0$. It describes the infinitesimal change of stress $\mathrm{d}\boldsymbol{\sigma} = \boldsymbol{D}^{(n)}\mathrm{d}\boldsymbol{\varepsilon}$ produced by an infinitesimal change of the current strain, $\mathrm{d}\boldsymbol{\varepsilon}$; see Figure 22.12(a).
- To construct the matrix $\boldsymbol{\Theta}^{(n)}(\boldsymbol{\varepsilon})$, we consider the stress evaluated by a specific return mapping algorithm for a finite increment of strain from $\boldsymbol{\varepsilon}^{(n-1)}$ to $\boldsymbol{\varepsilon}$, and we look at the change of this stress corresponding to an infinitesimal change of $\boldsymbol{\varepsilon}$, i.e. to an infinitesimal change of the finite increment of strain. This is illustrated in Figure 22.12(b). Note that, due to the finite distance of the trial stress state from the yield surface, the stress obtained by changing the trial state and then projecting it on the yield surface (Figure 22.12(b)) is different from the stress obtained by changing the projected trial state and then projecting again (Figure 22.12(c)).

In summary, the elastoplastic material stiffness $\boldsymbol{D}_{ep}$ is obtained by differentiation of the constitutive law while $\boldsymbol{\Theta}^{(n)}$ is obtained by differentiation of the numerical algorithm used for stress evaluation. For this reason, $\boldsymbol{\Theta}^{(n)}$ is called the *algorithmic stiffness*. Some authors use the expression 'consistent stiffness', but this is not the best choice since the tangent material stiffness is also derived consistently. The concept of

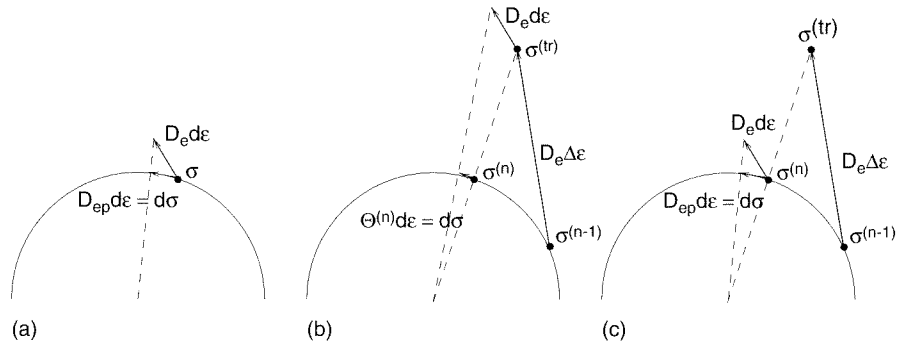


Figure 22.12 Radial return mapping applied to (a) an infinitesimal increment, (b) a finite increment modified by an infinitesimal change, (c) a finite increment followed by an infinitesimal increment

consistent linearization was introduced by Hughes and Pister (1978) and consistent algorithmic moduli were first derived for viscoplasticity by Hughes and Taylor (1978). Pioneering contributions to this area also include Nagtegaal (1982) and Simo and Taylor (1985).

Many authors have shown that only the algorithmic stiffness preserves the quadratic rate of convergence of the Newton–Raphson equilibrium iteration. At the beginning of an incremental loading step, we only have information about the last converged state, so the only tangent stiffness that can be constructed is the elastoplastic material stiffness, $\boldsymbol{D}_{ep}$. Suppose that we decide to use the full Newton–Raphson method with a stiffness update after each equilibrium iteration. Starting from the second iteration, it is important to construct the global stiffness matrix (to be used for the calculation of displacement corrections) from the algorithmic stiffness (22.100) rather than from the elastoplastic stiffness corresponding to the most recent (nonequilibrated) stress state. In a typical equilibrium iteration number $i+1$, the algorithmic stiffness is evaluated for the increment from $\boldsymbol{\varepsilon}^{(n-1)}$ to $\boldsymbol{\varepsilon}^{(n,i)}$. In (22.109), the dependence on the strain at the end of the step is hidden in the trial deviatoric stress, $\boldsymbol{s}^{(tr)} = 2G(\boldsymbol{e} - \boldsymbol{e}_p^{(n-1)})$. In the limit for an infinitely small step from $\boldsymbol{\varepsilon}^{(n-1)}$ to $\boldsymbol{\varepsilon}^{(n)} = \boldsymbol{\varepsilon}^{(n-1)}$, the deviatoric stress $\boldsymbol{s}^{(tr)}$ is equal to the current deviatoric stress $\boldsymbol{s}^{(n)} = \boldsymbol{s}^{(n-1)}$, which is located on the yield surface. In this limit case, the algorithmic stiffness is equal to the elastoplastic stiffness. This is in agreement with the natural requirement that, as the step size tends to zero, the numerical solution must tend to the exact solution.

In Example 22.2, we performed consistent linearization of the radial return algorithm. In this special case, the stress return mapping $\boldsymbol{\theta}(\boldsymbol{\varepsilon})$ was defined explicitly by formula (22.98) and it was relatively easy to obtain the algorithmic stiffness by differentiating that formula. However, such a case is rare. Stress return procedures usually lead to systems of nonlinear equations that must be solved iteratively, and the stress return mapping is defined implicitly as the solution of such a nonlinear system. In order to derive the algorithmic stiffness, we have to study how the solution changes due to an infinitesimal change of the prescribed strain increment.

Example 22.3: Derive the algorithmic stiffness matrix for the generalized trapezoidal rule.

Solution: The result of the mapping algorithm based on GTR is defined as the solution of nonlinear equations (22.85) and (22.88). For a given state at the beginning of the step, characterized by $\boldsymbol{\varepsilon}_p^{(n-1)}$ and $\boldsymbol{g}_\sigma^{(n-1)} \equiv \boldsymbol{g}_\sigma(\boldsymbol{\sigma}^{(n-1)})$, the final stress $\boldsymbol{\sigma}^{(n)}$ only depends on the final strain $\boldsymbol{\varepsilon}^{(n)}$. Suppose that, for a certain value of $\boldsymbol{\varepsilon}^{(n)}$, we know the corresponding $\boldsymbol{\sigma}^{(n)}$. Our goal is to find the change $\mathrm{d}\boldsymbol{\sigma}$ caused by a change of $\boldsymbol{\varepsilon}^{(n)}$ by $\mathrm{d}\boldsymbol{\varepsilon}$. This can be deduced from the governing equations linearized around the known solution, taking into account that the plastic multiplier is also subjected to an infinitesimal change $\mathrm{d}\lambda$. Since all the changes are infinitesimal, second- and higher-order terms are negligible and the linearized equations are 'exact'. Replacing $\boldsymbol{\sigma}^{(n)}$ by $\boldsymbol{\sigma}^{(n)} + \mathrm{d}\boldsymbol{\sigma}$, $\boldsymbol{\varepsilon}^{(n)}$ by $\boldsymbol{\varepsilon}^{(n)} + \mathrm{d}\boldsymbol{\varepsilon}$, and $\Delta\lambda^{(n)}$ by $\Delta\lambda^{(n)} + \mathrm{d}\lambda$ in (22.85) and (22.88), performing linearization and separating the unknown and known terms, we obtain a set of linear equations

$$\begin{bmatrix} \boldsymbol{C}_e + \alpha\,\Delta\lambda^{(n)}\boldsymbol{G}_{\sigma\sigma}^{(n)} & (1-\alpha)\boldsymbol{g}_\sigma^{(n-1)} + \alpha\boldsymbol{g}_\sigma^{(n)} \\ \boldsymbol{f}_\sigma^{(n)T} & 0 \end{bmatrix} \begin{Bmatrix} \mathrm{d}\boldsymbol{\sigma} \\ \mathrm{d}\lambda \end{Bmatrix} = \begin{Bmatrix} \mathrm{d}\boldsymbol{\varepsilon} \\ 0 \end{Bmatrix} \tag{22.111}$$

Not surprisingly, the resulting equations are closely related to equations (22.89), which are used by the GTR algorithm to determine the iterative corrections of the stress and plastic multiplier. However, some important differences exist. Equations (22.89) were obtained by linearization around the state $\boldsymbol{\sigma}^{(n,k)}$ and $\Delta\lambda^{(n,k)}$ that does not satisfy the governing equations (22.85) and (22.88). Since the corrections $\delta\boldsymbol{\sigma}$ and $\delta\lambda$ were 'small' but finite, the linearized equations were not exactly equivalent to the original ones, and a sequence of iterations was needed to bring the residuals to (almost) zero. On the other hand, equations (22.111) are obtained by linearization around the state $\boldsymbol{\sigma}^{(n)}$ and $\Delta\lambda^{(n)}$ that solves the nonlinear equations. Consequently, the residuals on the right-hand side are zero, except for the term $\mathrm{d}\boldsymbol{\varepsilon}$ that appears due to a perturbation of the converged state. Since the assumed perturbation is infinitesimal, the corresponding changes of the exact solution of (22.85) and (22.88) are solved directly from the linearized equations, without any need for iteration. Denoting

$$\boldsymbol{\Xi}^{(n)} = \left(\boldsymbol{C}_e + \alpha\,\Delta\lambda^{(n)}\boldsymbol{G}_{\sigma\sigma}^{(n)}\right)^{-1} \tag{22.112}$$

and proceeding in analogy to the treatment of equation (22.73), we can eliminate $\mathrm{d}\lambda$ and find the relationship between $\mathrm{d}\boldsymbol{\sigma}$ and $\mathrm{d}\boldsymbol{\varepsilon}$ in the form

$$\mathrm{d}\boldsymbol{\sigma} = \boldsymbol{\Theta}^{(n)}\mathrm{d}\boldsymbol{\varepsilon} \tag{22.113}$$

where

$$\boldsymbol{\Theta}^{(n)} = \boldsymbol{\Theta} - \frac{\boldsymbol{\Theta}\boldsymbol{g}_\sigma^{(n-1+\alpha)}\boldsymbol{f}_\sigma^{(n)T}\boldsymbol{\Theta}}{\boldsymbol{f}_\sigma^{(n)T}\boldsymbol{\Theta}\boldsymbol{g}_\sigma^{(n-1+\alpha)}} \tag{22.114}$$

is the algorithmic stiffness matrix of the GTR method. For convenience, we have denoted

$$\boldsymbol{g}_\sigma^{(n-1+\alpha)} = (1-\alpha)\boldsymbol{g}_\sigma^{(n-1)} + \alpha\boldsymbol{g}_\sigma^{(n)} \tag{22.115}$$

Note that the algorithmic stiffness (22.114) has the same structure as the elastoplastic stiffness of a perfectly plastic material with a nonassociated flow rule,

$$\boldsymbol{D}_{ep} = \boldsymbol{D}_e - \frac{\boldsymbol{D}_e : \boldsymbol{g}_\sigma \otimes \boldsymbol{f}_\sigma : \boldsymbol{D}_e}{\boldsymbol{f}_\sigma : \boldsymbol{D}_e : \boldsymbol{g}_\sigma} \tag{22.116}$$

which easily follows from(20.112) by deleting the terms related to hardening. The difference is that $\boldsymbol{D}_e$ from (22.116) is replaced by $\boldsymbol{\Theta}$ in (22.114), and the yield function gradient $\boldsymbol{f}_\sigma^{(n)}$ is computed at $\boldsymbol{\sigma}^{(n)}$ while the flow direction $\boldsymbol{g}_\sigma^{(n-1+\alpha)}$ is determined as a weighted average of the directions at $\boldsymbol{\sigma}^{(n-1)}$ and $\boldsymbol{\sigma}^{(n)}$.

PROBLEMS

Problem 22.1: Consider a one-dimensional model of a bar of length 1 m, consisting of 10 equally sized linear finite elements. The left end of the bar is fixed and the right end is subjected to a prescribed displacement 1 mm, applied in one single step. The material is perfectly elastoplastic, with Young's modulus of 200 Gpa and uniaxial yield stress 400 MPa. How many Newton–Raphson iterations are needed to restore equilibrium using the algorithm described in Section 22.2.2? How many iterations would be needed with $\boldsymbol{d}_2^{(n,0)} = \boldsymbol{d}_2^{(n)}$ instead of $\boldsymbol{d}_2^{(n,0)} = \boldsymbol{d}_2^{(n-1)}$, as described in the last paragraph of Section 22.2.2?

Problem 22.2: Develop the radial return algorithm for associated von Mises plasticity with linear isotropic hardening.

Problem 22.3: Under which conditions is the algorithmic stiffness matrix of the GTR method symmetric?

Problem 22.4*: Develop a stress return algorithm for J_2-plasticity under plane stress conditions.

23

Thermodynamic Approach to Constitutive Modeling

There are two aspects of thermodynamics relevant to plasticity: (1) thermodynamics of a theoretically infinitesimal material volume, in practice applicable to the so-called representative volume of material, and (2) thermodynamics of structures, important for stability of, and bifurcations in, elastoplastic structures. In this book we deal only with the former. For a detailed exposition of the latter, see Bažant and Cedolin (1991, Chapter 10).

Thermodynamics provides important restrictions on the form of constitutive relations, but no more than restrictions. It reveals symmetries due to interaction of different variables, and shows what cross-couplings between different variables to expect and how to treat them. The restrictions of thermodynamics are particularly important and helpful in the treatment of 'switch' conditions, such as the loading-unloading conditions in plasticity. However, keep in mind that a constitutive law cannot be derived from thermodynamics; only its general admissible form can.

23.1 BASIC CONCEPTS AND LAWS OF THERMODYNAMICS

23.1.1 Internal Energy and Entropy

The thermodynamic approach to the description of a continuum starts from the assumption that the current state of the material can be uniquely characterized by a suitably selected set of *state variables*. The state variables can always be selected such that they are independent, i.e. neither of them can be expressed as a unique function of the others. Any quantity that uniquely depends on the state variables is called a *state function*.

For the sake of simplicity, we assume that the distribution of the state variables in space and their evolution in time are sufficiently regular, i.e. we exclude discontinuities such as shock waves and moving phase boundaries from our consideration. We will only work with closed systems, in which no mass exchange and no phase changes take place. Similar to all the topics covered so far, in this chapter we adhere to the small-strain theory. In particular, the variation of mass density in time is considered as negligible. For a generalization to large strain, see Section 24.2.

All thermodynamics rests on two basic laws – the First and Second Law, which result from overwhelming experience but cannot be theoretically proven. The First

Law (Mayer, 1845; Joule, 1845; Clausius, 1865) in one way or another expresses the principle of conservation of energy. The Second Law (Thomson, 1851; Clausius, 1854, 1865; Duhem, 1892) is an inequality that can be seen either as a restriction on the direction in which various irreversible processes can proceed, vital for our purpose, or as a condition of stability limit, equivalent to the Lyapunov concept of stability applied to equilibrium (Bažant and Cedolin, 1991, Chapter 10). These laws, particularly the second, can be stated in many different forms, ranging from intuitive physical statements (such as 'heat cannot spontaneously flow from a colder body to a hotter body') to very abstract mathematical postulates. Here we present them as statements postulating the existence of certain special state functions.

First Law of Thermodynamics
There exists a state function $\mathcal{U}$, called the **internal energy**, *such that its sum with the kinetic energy changes at a rate equal to the sum of the external power and the rate of heat supply.*

Second Law of Thermodynamics
There exists a state function $\mathcal{S}$, called the **entropy**, *such that its rate of change is never smaller than the rate of heat supply divided by the absolute temperature.*

The fundamental laws of thermodynamics introduce the notions of internal energy and entropy, provided that we know what is meant by the kinetic energy, external power, heat supply and absolute temperature. Consider a solid body occupying a domain V, which is bounded by a surface S. The supply of heat to the body is provided by the heat flux $\boldsymbol{q}$ on the boundary (rate of heat per unit area) and by a distributed heat source of intensity r (per unit mass) coming, for example, from absorption of radiation or from ohmic resistance to electrical current. The received *heat power* (rate of heat supply to the body) is expressed as

$$\mathcal{P}_{\text{cal}} \qquad \int_V \rho r \,\mathrm{d}V - \int_S \boldsymbol{q}\cdot\boldsymbol{n}\,\mathrm{d}S \tag{23.1}$$

where ρ is the density and $\boldsymbol{n}$ is the unit vector normal to the surface and oriented towards the exterior of V. The negative sign in front of the second integral means that, if the heat flux $\boldsymbol{q}$ is oriented in the direction of the outward normal $\boldsymbol{n}$, the heat is flowing out.

The *external power*

$$\mathcal{P}_{\text{ext}} = \int_V \bar{\boldsymbol{b}}\cdot\boldsymbol{v}\,\mathrm{d}V + \int_S \bar{\boldsymbol{t}}\cdot\boldsymbol{v}\,\mathrm{d}S \tag{23.2}$$

is supplied by tractions $\bar{\boldsymbol{t}}$ acting on the boundary, and by body forces $\bar{\boldsymbol{b}}$ due to an external force field acting directly on the material particles that compose the body (e.g. a gravity field). The surface tractions are taken per unit surface area and the body forces per unit volume. The velocity vector $\boldsymbol{v} = \dot{\boldsymbol{u}}$ is the (material) time derivative of the displacement vector $\boldsymbol{u}$.

The *kinetic energy* is given by

$$\mathcal{K} = \frac{1}{2}\int_V \rho\,\|\boldsymbol{v}\|^2\,\mathrm{d}V = \frac{1}{2}\int_V \rho\,\boldsymbol{v}\cdot\boldsymbol{v}\,\mathrm{d}V \tag{23.3}$$

and thus its rate is

$$\dot{\mathcal{K}} = \int_V \rho\,\dot{\boldsymbol{v}}\cdot\boldsymbol{v}\,\mathrm{d}V \tag{23.4}$$

where $\dot{\boldsymbol{v}} = \ddot{\boldsymbol{u}}$ is the acceleration.

23.1.2 First Law of Thermodynamics

The First Law represents the *principle of conservation of energy*, written mathematically as

$$\dot{\mathcal{U}} + \dot{\mathcal{K}} = \mathcal{P}_{\text{ext}} + \mathcal{P}_{\text{cal}} \tag{23.5}$$

The sum of the internal and kinetic energies is the *total energy* of the body, and the energy balance equation (23.5) integrated over a finite time interval states that the change of the total energy is equal to the sum of the work and heat supplied to the body from its environment.

In equation (23.5), the First Law is written in the *global form*, referred to as a finite volume. A *local form* is obtained by taking the limit for an infinitely small volume. To this end, it is necessary to convert all the surface integrals into volume integrals. Writing the tractions as stress projections, $\bar{\boldsymbol{t}} = \boldsymbol{\sigma}\cdot\boldsymbol{n}$, and applying the Gauss–Green theorem (integration by parts in multiple dimensions), we can express the power of surface tractions as

$$\begin{aligned}\int_S \bar{\boldsymbol{t}}\cdot\boldsymbol{v}\,\mathrm{d}S &= \int_S \bar{t}_i v_i\,\mathrm{d}S = \int_S v_i\sigma_{ij}n_j\,\mathrm{d}S = \int_V \frac{\partial(v_i\sigma_{ij})}{\partial x_j}\,\mathrm{d}V \\ &= \int_V \sigma_{ij}\frac{\partial v_i}{\partial x_j}\,\mathrm{d}V + \int_V v_i\frac{\partial\sigma_{ij}}{\partial x_j}\,\mathrm{d}V = \int_V \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V + \int_V \boldsymbol{v}\cdot(\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,\mathrm{d}V\end{aligned} \tag{23.6}$$

where $\boldsymbol{\nabla}$ is the differential operator defined in Section D.1. Since $\boldsymbol{\sigma}$ is a symmetric tensor, it was possible to replace the tensor $\partial\boldsymbol{v}/\partial\boldsymbol{x}$ by its symmetric part, $(\partial\boldsymbol{v}/\partial\boldsymbol{x})_{\text{sym}} = (\partial\dot{\boldsymbol{u}}/\partial\boldsymbol{x})_{\text{sym}} = \dot{\boldsymbol{\varepsilon}}$, without affecting the value of the product $\sigma_{ij}\partial v_i/\partial x_j = \sigma_{ij}\dot{\varepsilon}_{ij} = \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}$.

Substituting (23.6) into (23.2) and making use of the momentum equations (D.86) (i.e. of equilibrium equations enriched by the inertial forces), we obtain

$$\begin{aligned}\mathcal{P}_{\text{ext}} &= \int_V \bar{\boldsymbol{b}}\cdot\boldsymbol{v}\,\mathrm{d}V + \int_S \bar{\boldsymbol{t}}\cdot\boldsymbol{v}\,\mathrm{d}S = \int_V \boldsymbol{v}\cdot(\bar{\boldsymbol{b}} + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,\mathrm{d}V + \int_V \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V \\ &= \int_V \boldsymbol{v}\cdot\rho\dot{\boldsymbol{v}}\,\mathrm{d}V + \int_V \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V = \dot{\mathcal{K}} + \mathcal{P}_{\text{int}}\end{aligned} \tag{23.7}$$

where

$$\mathcal{P}_{\text{int}} = \int_V \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}\,\mathrm{d}V \tag{23.8}$$

is the *internal power* (or *stress power*). The surface integral in the heat power expression (23.1) is also easily converted into a volume integral by a direct application of the Gauss–Green theorem,

$$\int_S \boldsymbol{q}\cdot\boldsymbol{n}\,\mathrm{d}S = \int_V \boldsymbol{\nabla}\cdot\boldsymbol{q}\,\mathrm{d}V \tag{23.9}$$

where $\boldsymbol{\nabla} \cdot \boldsymbol{q} = \operatorname{div} \boldsymbol{q}$ is the divergence of the heat flux. Consequently, the energy balance equation (23.5) can be simplified to

$$\dot{\mathcal{U}} = \mathcal{P}_{\text{int}} + \mathcal{P}_{\text{cal}} = \int_V \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} \, \mathrm{d}V + \int_V (\rho r - \boldsymbol{\nabla} \cdot \boldsymbol{q}) \, \mathrm{d}V \tag{23.10}$$

Introducing the specific internal energy u (taken per unit mass), we can express the internal energy of the finite body as $\mathcal{U} = \int_V \rho u \, \mathrm{d}V$ and rewrite (23.10) as

$$\int_V (\rho \dot{u} - \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \rho r + \boldsymbol{\nabla} \cdot \boldsymbol{q}) \, \mathrm{d}V = 0 \tag{23.11}$$

This equation must hold for any partial volume V within the material body. Therefore, if the integrand is continuous, it must vanish identically. This argument leads to the local form of the First Law,

$$\rho \dot{u} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} + \rho r - \boldsymbol{\nabla} \cdot \boldsymbol{q} \tag{23.12}$$

If the integrand in (23.11) is discontinuous on certain surfaces inside the body, equation (23.12) must be replaced at the points of discontinuity by certain jump conditions. As already mentioned, this case is not considered here.

Equation (23.12) describes the local energy balance: The rate of increase of internal energy in an elementary material volume is equal to the sum of: (1) the power of stress $\boldsymbol{\sigma}$ working on the strain rate $\dot{\boldsymbol{\varepsilon}}$; (2) the heat supplied by an internal source of intensity r; and (3) the negative divergence of the heat flux, which represents the net rate of heat entering the elementary volume through its boundary. The fact that $\boldsymbol{\nabla} \cdot \boldsymbol{q} \, \mathrm{d}V = \operatorname{div} \boldsymbol{q} \, \mathrm{d}V$ represents the rate of heat flow out of the infinitesimal volume $\mathrm{d}V = \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$ becomes clear upon summing the fluxes shown in Figure 23.1. The fluxes in are $q_x \, \mathrm{d}y \, \mathrm{d}z$, $q_y \, \mathrm{d}x \, \mathrm{d}z$ and $q_z \, \mathrm{d}x \, \mathrm{d}y$, and the fluxes out are $[q_x + (\partial q_x / \partial x) \mathrm{d}x] \, \mathrm{d}y \, \mathrm{d}z$, etc. The sum gives the net flux out as

$$\begin{aligned} \left(\frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z} \right) \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z &= \left\{ \begin{matrix} \partial / \partial x & \partial / \partial y & \partial / \partial z \end{matrix} \right\} \left\{ \begin{matrix} q_x \\ q_y \\ q_z \end{matrix} \right\} \mathrm{d}V \\ &= \boldsymbol{\nabla} \cdot \boldsymbol{q} \, \mathrm{d}V = \operatorname{div} \boldsymbol{q} \, \mathrm{d}V \end{aligned} \tag{23.13}$$

23.1.3 Second Law of Thermodynamics

The Second Law is stated in its global form as

$$\dot{\mathcal{S}} \geq \int_V \frac{\rho r}{T} \, \mathrm{d}V - \int_S \frac{\boldsymbol{q} \cdot \boldsymbol{n}}{T} \, \mathrm{d}S \tag{23.14}$$

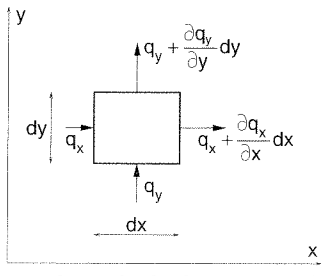


Figure 23.1 Heat fluxes through the boundary of an infinitesimal volume

where T is the absolute temperature. The script $\mathcal{S}$, denoting the entropy of the body, should not be confused with the italic S, denoting the surface. Applying the Gauss–Green theorem again, we obtain

$$\int_S \frac{\boldsymbol{q} \cdot \boldsymbol{n}}{T} \, \mathrm{d}S = \int_V \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{q}}{T} \right) \mathrm{d}V = \int_V \left(\frac{\boldsymbol{\nabla} \cdot \boldsymbol{q}}{T} - \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T^2} \right) \mathrm{d}V \tag{23.15}$$

Introducing the specific entropy s (taken per unit mass) and using (23.15), we can rewrite (23.14) as

$$\int_V \rho \dot{s} \, \mathrm{d}V \geq \int_V \left(\frac{\rho r}{T} - \frac{\boldsymbol{\nabla} \cdot \boldsymbol{q}}{T} + \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T^2} \right) \mathrm{d}V \tag{23.16}$$

Requiring this to hold for any partial volume V of the structure (and assuming again that the integrands are continuous), we get the local form of the Second Law,

$$\rho \dot{s} \geq \frac{\rho r - \boldsymbol{\nabla} \cdot \boldsymbol{q}}{T} + \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T^2} \tag{23.17}$$

The left-hand side of (23.17) is the rate of entropy, the right-hand side is called the rate of *external entropy supply*, and the difference between the left-hand side and the right-hand side is called the rate of *internal entropy production* (all considered here per unit volume). Physically, the entropy is a measure of the complexity and disorder of the internal state.

It is convenient to introduce the density of energy dissipation rate (henceforth termed just *dissipation*), $\mathcal{D}$, defined as the rate of internal entropy production per unit volume multiplied by the absolute temperature. Since the absolute temperature is always positive, the Second Law is equivalent to the condition of non-negative dissipation,

$$\mathcal{D} \equiv T \rho \dot{s} - \rho r + \boldsymbol{\nabla} \cdot \boldsymbol{q} - \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T} \geq 0 \tag{23.18}$$

Based on the local energy balance equation (23.12), we can replace $\boldsymbol{\nabla} \cdot \boldsymbol{q} - \rho r$ by $\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \rho \dot{u}$ and derive from (23.18) the *Clausius–Duhem inequality*,

$$\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \rho(\dot{u} - T \dot{s}) - \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T} \geq 0 \tag{23.19}$$

The name 'Clausius–Duhem inequality' is in the literature frequently used for other forms of the inequality originating from the Second Law, e.g. for (23.14), (23.16), (23.17), (23.18), or for inequalities (23.22) and (23.29), to be derived later. The general Clausius-Duhem inequality first appeared in Truesdell and Toupin (1960), although particular cases were treated much earlier by Clausius (1865) and Duhem (1903). A clear interpretation of this inequality as a restriction that must be satisfied in every smooth process consistent with the balance and constitutive equations was first clearly given by Coleman and Noll (1963).

23.1.4 Free Energy and Dissipation

So far, we have not specified the nature of the state variables. Some of them, e.g. the temperature or displacement, can be observed on the macroscopic level and controled

from the exterior; they are usually called the *observable variables*. The remaining state variables, e.g. plastic strain or damage, characterize the internal changes of the material and are called the *internal variables*. They can sometimes be measured but they cannot be controled from the outside. In solid mechanics, the observable local state variables usually include the strain tensor, $\boldsymbol{\varepsilon}$, and the absolute temperature, T. Alternatively, the temperature can be replaced by the specific entropy, s. The choice of internal variables depends on the type of constitutive behavior to be described, and at this stage we will denote them collectively by the symbol $\boldsymbol{\alpha}$. Since $\boldsymbol{\alpha}$ can be a scalar, tensor, or a collection of scalars and tensors of various orders, we introduce the symbol $\bullet$ for the scalar product of $\boldsymbol{\alpha}$ with a similarly structured object (e.g. with the partial gradient $\partial\psi/\partial\boldsymbol{\alpha}$ of a state function ψ). Once the structure of $\boldsymbol{\alpha}$ is known, the operator $\bullet$ may be interpreted accordingly (as a scalar product of vectors, doubly contracted product of second-order tensors, etc.).

Suppose that $(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}, T)$ are the state variables. It turns out to be useful to transform the internal energy density u into a new state function

$$\psi(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}, T) = u(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}, T) - Ts(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}, T) \tag{23.20}$$

called the *Helmholtz free energy density*. Since, from now on, we will work exclusively on the level of an infinitesimal material volume, we will often leave out the word 'density' and speak simply of 'free energy', 'internal energy', etc., understood as taken per unit mass. Also, we will abbreviate the 'density of energy dissipation rate' (per unit volume) as 'dissipation'.

As we will see later, the free energy ψ is the portion of the internal energy u available under constant temperature. Its rate is given by

$$\dot{\psi} = \dot{u} - \dot{T}s - T\dot{s} \tag{23.21}$$

and substituting this into (23.19) leads to an alternative form of the Clausius–Duhem inequality,

$$\mathcal{D} \quad \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \rho(\dot{\psi} + \dot{T}s) - \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T} \geq 0 \tag{23.22}$$

often called the *dissipation inequality*.

Now consider a process of pure heating or cooling, during which the strain and the internal variables remain constant and only the temperature varies in time but remains constant in space. The dissipation, in this case given by

$$\mathcal{D} = -\rho \left(\frac{\partial \psi}{\partial T} + s \right) \dot{T} \tag{23.23}$$

must not be negative, independent of the sign of $\dot{T}$. Consequently, we have

$$s = -\frac{\partial \psi}{\partial T} \tag{23.24}$$

As both sides of (23.24) are state functions, this equality must hold in general, independent of the type of process (i.e. not only for pure heating). Therefore, the free energy plays the role of a *thermodynamic potential* from which it is possible to deduce the state function describing the dependence of entropy on the state variables.

For a general process with variable $\boldsymbol{\varepsilon}$ and $\boldsymbol{\alpha}$, the free energy rate reads

$$\dot{\psi} = \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}} : \dot{\boldsymbol{\varepsilon}} + \frac{\partial \psi}{\partial \boldsymbol{\alpha}} \bullet \dot{\boldsymbol{\alpha}} + \frac{\partial \psi}{\partial T} \dot{T} \tag{23.25}$$

Substituting this into (23.22) and taking into account (23.24), we get

$$\mathcal{D} = \left(\boldsymbol{\sigma} - \rho \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}} \right) : \dot{\boldsymbol{\varepsilon}} - \rho \frac{\partial \psi}{\partial \boldsymbol{\alpha}} \bullet \dot{\boldsymbol{\alpha}} - \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T} \tag{23.26}$$

Defining the *quasi-conservative stress* (e.g. Ziegler, 1983),

$$\boldsymbol{\sigma}_Q = \rho \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}} \tag{23.27}$$

and the *quasi-conservative thermodynamic forces* conjugate to the internal variables (for the sake of brevity referred to as the quasi-conservative forces),

$$\boldsymbol{\beta}_Q = \rho \frac{\partial \psi}{\partial \boldsymbol{\alpha}} \tag{23.28}$$

we can finally write the condition of non-negative dissipation as

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}_Q) : \dot{\boldsymbol{\varepsilon}} - \boldsymbol{\beta}_Q \bullet \dot{\boldsymbol{\alpha}} - \frac{\boldsymbol{q} \cdot \boldsymbol{\nabla} T}{T} \geq 0 \tag{23.29}$$

The quasi-conservative stresses and forces are also called *reversible* or *non-dissipative*.

Condition (23.29) is one of the conditions often called the Clausius–Duhem inequality. The first two terms on its left-hand side represent the *mechanical dissipation*, mechanical the last term is the *thermal dissipation*. The processes for which the dissipation is zero are called *reversible*, and all the other processes *irreversible*. Relations of the type of (23.24), (23.27) and (23.28) are called the *state laws*.

As is clear from (23.29), the Second Law requires only that the sum of the mechanical and thermal dissipation be nonnegative. This condition is usually replaced by a stronger one, namely that each of these two dissipation terms taken separately must be non-negative (e.g. Truesdell and Noll, 1960). For the thermal part it means that

$$\boldsymbol{q} \cdot \boldsymbol{\nabla} T \leq 0 \tag{23.30}$$

In other words, the heat flux cannot be oriented in a direction in which temperature increases. Stated in a simple way, heat cannot spontaneously flow from a colder body to a warmer one. For the mechanical dissipation we obtain the *Kelvin inequality*

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}_Q) : \dot{\boldsymbol{\varepsilon}} - \boldsymbol{\beta}_Q \bullet \dot{\boldsymbol{\alpha}} \geq 0 \tag{23.31}$$

23.1.5 Historical Remarks

Before proceeding to the specific thermodynamic formulations of diverse constitutive theories, let us insert a few historical remarks (which may be skipped by a reader interested mainly in the practical aspects). The classical thermodynamics, which was founded already in the 19th century and is lucidly but compactly presented by Guggenheim (1959), is concerned with systems in or very near thermodynamic equilibrium – more precisely, with states of both mechanical and thermal equilibrium and with infinitesimal deviations from equilibrium. Therefore, the classical thermodynamics is sometimes called the *thermostatics*. In this well-established theory, it is possible to prove the existence of an absolute temperature and

a universal entropy based on a few fundamental postulates. A rigorous mathematical formulation of classical thermodynamics was given by Carathéodory (1909).

Extensions of the classical theory to continua a finite distance away from equilibrium were proposed much later and created serious controversies. There exist several competing schools and methodologies that differ by the levels of generality and abstraction. One prominent class of theories grew out of the *rational thermodynamics*, pioneered by Coleman, Noll and Truesdell (Coleman and Noll, 1963; Coleman, 1964; Truesdell, 1969) and their coworkers, and further developed, for example, by Müller (1975).

For our present purpose, the most suitable framework is provided by the *thermodynamics with internal variables*, based on the hypothesis of a local accompanying state. The main assumption is that each particle (elementary material volume) of a continuum is associated with an accompanying thermostatic system (in, or infinitesimally close to, thermodynamic equilibrium). For such a system, the absolute temperature and entropy can be defined in the standard manner.

The aforementioned class of theories takes inspiration from the *thermodynamics of irreversible processes* (de Groot and Mazur, 1962; Woods, 1975), which originated from the contributions of Onsager (1931) and Casimir (1945). The famous Onsager-Casimir reciprocity relations were generalized to the orthogonality principle (Ziegler, 1963), and finally led to the concept of standard dissipative materials (Moreau, 1970), which make use of the so-called dissipation potentials. A succint comparison of various thermodynamic approaches can be found, for instance, in Germain, Nguyen and Suquet (1983).

23.2 LINEAR THERMOELASTICITY

As the simplest example, consider a material model with no dissipation. In this case, all the processes are reversible and no internal variables are needed. The state of the material is uniquely characterized by the strain and temperature. The second term in the expression for mechanical dissipation (23.31) disappears, and the first term, $(\boldsymbol{\sigma} - \boldsymbol{\sigma}_Q) : \dot{\boldsymbol{\varepsilon}}$, must vanish for any strain rate $\dot{\boldsymbol{\varepsilon}}$. Consequently, the state law (23.27) provides a constitutive equation for the stress. This means that the specific choice of the free-energy potential and the assumption of zero dissipation completely define the material model.

Example 23.1: Construct the state laws corresponding to a quadratic free-energy potential

$$\rho\psi(\boldsymbol{\varepsilon}, T) = \rho\psi_0 - \rho s_0(T - T_0) + \frac{1}{2}\boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{a}(T - T_0) - \frac{C\rho}{2T_0}(T - T_0)^2 \tag{23.32}$$

and discuss the physical meaning of parameters ψ_0, s_0, T_0, $\boldsymbol{a}$ and C.

Solution: Differentiation of the free energy leads to the state laws

$$\boldsymbol{\sigma} = \rho\frac{\partial\psi}{\partial\boldsymbol{\varepsilon}} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \boldsymbol{D}_e : \boldsymbol{a}(T - T_0) = \boldsymbol{D}_e : [\boldsymbol{\varepsilon} - (T - T_0)\boldsymbol{a}] \tag{23.33}$$

$$s = -\frac{\partial\psi}{\partial T} = s_0 + \frac{1}{\rho}\boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{a} + \frac{C}{T_0}(T - T_0) \tag{23.34}$$

We identify ψ_0 as the initial specific free energy in the reference state (at zero strain and at reference temperature T_0), s_0 as the specific entropy in the reference state, $\boldsymbol{D}_e$ as the elastic stiffness tensor, $\boldsymbol{a}$ as the thermal expansion tensor, and C as the specific heat (per unit mass); $\boldsymbol{D}_e$, $\boldsymbol{a}$ and C are material properties.

The elastic stiffness $\boldsymbol{D}_e$ from (23.33)–(23.34) corresponds to *isothermal conditions*, i.e. to slow loading during which the temperature of the sample is kept constant. The other extreme is an *adiabatic process*, which is so fast that no heat exchange can take place. In the absence of dissipation, adiabatic processes are at the same time isentropic. This is clear upon recalling that the rate of entropy $\rho\dot{s}$ is equal to the sum of the right-hand side of (23.17) with the dissipation multiplied by the temperature. If there is no internal heat source r and the process is adiabatic, the right-hand side of (23.17) is zero. In the absence of dissipation, this means that the entropy remains constant, which is called an *isentropic process*.

The evolution of temperature as a function of strain during an isentropic process can be obtained from (23.34) with $s = \text{const.} = s_0$:

$$T = T_0 - \frac{T_0}{C\rho}\boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{a} = T_0 - \frac{T_0}{C\rho}\boldsymbol{a} : \boldsymbol{D}_e : \boldsymbol{\varepsilon} \tag{23.35}$$

Now it is possible to eliminate the temperature from (23.33) and derive the stress-strain law

$$\boldsymbol{\sigma} = \left(\boldsymbol{D}_e + \frac{T_0}{C\rho}\boldsymbol{D}_e : \boldsymbol{a} \otimes \boldsymbol{a} : \boldsymbol{D}_e\right) : \boldsymbol{\varepsilon} \tag{23.36}$$

The expression in parentheses represents the elastic stiffness tensor valid under adiabatic conditions. It turns out that the adiabatic stiffness is larger than the isothermal one, but the difference is often negligible.

For isotropic materials we have

$$\boldsymbol{D}_e = \left(K - \frac{2}{3}G\right)\boldsymbol{\delta} \otimes \boldsymbol{\delta} + 2G\boldsymbol{I}_s, \qquad \boldsymbol{a} = \alpha\boldsymbol{\delta} \tag{23.37}$$

$$\boldsymbol{D}_e : \boldsymbol{a} \otimes \boldsymbol{a} : \boldsymbol{D}_e = 9K^2\alpha^2\boldsymbol{\delta} \otimes \boldsymbol{\delta} \tag{23.38}$$

where K is the (isothermal) bulk modulus, G is the shear modulus, α is the coefficient of thermal expansion, $\boldsymbol{\delta}$ is the unit second-order tensor (Kronecker delta), and $\boldsymbol{I}_s$ is the unit symmetric fourth-order tensor. In this case, the adiabatic stiffness can be evaluated as

$$\boldsymbol{D}_e^{\text{adb}} = \boldsymbol{D}_e + \frac{T_0}{C\rho}9K^2\alpha^2\boldsymbol{\delta} \otimes \boldsymbol{\delta} = \left(K + \frac{9K^2\alpha^2T_0}{C\rho} - \frac{2}{3}G\right)\boldsymbol{\delta} \otimes \boldsymbol{\delta} + 2G\boldsymbol{I}_s \tag{23.39}$$

It has the same structure as the isothermal stiffness, with the same shear modulus G but with the bulk modulus K replaced by

$$K^{\text{adb}} = K\left(1 + \frac{9K\alpha^2T_0}{C\rho}\right) \tag{23.40}$$

The dimensionless factor $9K\alpha^2T_0/C\rho$ is usually small compared to 1. For example, the typical values of material parameters for concrete at $T_0 = 300\,\text{K}$ are $K = 18 \times 10^9$ Nm^{-2}, $\alpha = 10 \times 10^{-6}\text{K}^{-1}$, $C = 1000\,\text{J}\,\text{kg}^{-1}\text{K}^{-1}$, $\rho = 2400\,\text{kg}\,\text{m}^{-3}$, and so

$$\frac{9K\alpha^2T_0}{C\rho} = \frac{9 \times 18 \times 10^9 \times 100 \times 10^{-12} \times 300}{1000 \times 2400} = 2.025 \times 10^{-3} \tag{23.41}$$

Thus the difference between the adiabatic and isothermal bulk moduli is only about 0.2%, barely within the resolution of normal experiments.

Similarly, it can be shown that the specific heat C from (23.34) corresponds to heating under constant strain. If the sample is heated under constant stress, the specific heat increases to

$$C_\sigma = C\left(1 + \frac{T_0}{C\rho}\boldsymbol{a} : \boldsymbol{D}_e : \boldsymbol{a}\right) \tag{23.42}$$

For isotropic materials we have $\boldsymbol{a} : \boldsymbol{D}_e : \boldsymbol{a} = 9K\alpha^2$ and

$$C_\sigma = C\left(1 + \frac{9K\alpha^2 T_0}{C\rho}\right) \tag{23.43}$$

The relative increase of specific heat is again determined by the factor (23.41), which is negligible for common structural materials. □

23.3 DISSIPATION AND EVOLUTION LAWS

From now on we will assume that all the processes are isothermal, which gives us the right to exclude temperature from the list of state variables. We focus on the purely mechanical description of materials. The temperature is assumed to be constant in space and time, and expression (23.22) for the dissipation reduces to

$$\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \rho\dot{\psi} \tag{23.44}$$

When the material can behave in an irreversible manner, e.g. undergo plastic yielding, its state must be characterized not only by the strain but also by some additional internal variables that reflect the changes in the microstructure. In that case, the free-energy function alone does not uniquely specify the material model. The reason is that, while the evolution of strain can be controled externally,[1] the internal variables evolve spontaneously, governed by some additional laws which complement the state laws derived from the free-energy potential. These *evolution laws* can be approached in several different ways.

Perhaps the most general approach is to specify the dissipation $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ as a function of the state variables, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\alpha}$, and their rates, $\dot{\boldsymbol{\varepsilon}}$ and $\dot{\boldsymbol{\alpha}}$, which are often called the *thermodynamic fluxes*.

Note that expression (23.44) is in fact a transformed version of the First Law, stating that the (mechanical) dissipation must be equal to the difference between the external power and the rate of free energy. To make use of this energy balance equation, we need an independent, intrinsic specification of the dissipation for an arbitrary admissible evolution of the material state. Then it is possible to find the rates of internal variables induced at a given state by a certain applied strain rate, and

[1] The strain can be directly controled only if we deal with a very small material volume (theoretically infinitesimal, practically the smallest possible volume that can still be considered as a continuum). In experiments on laboratory specimens or structures, we impose displacements on certain parts of the boundary, and the resulting strain distribution is governed by the equations of motion combined with compatibility conditions and the constitutive equations. Conceptually, however, we may think of the evolution of strain at a given material 'point' being controled by the deformation of the surrounding material, i.e. from the exterior.

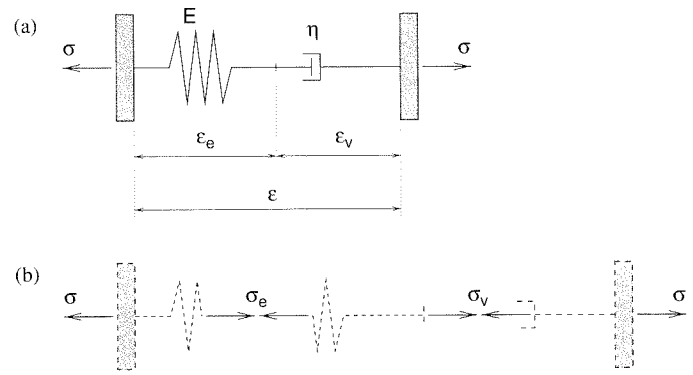


Figure 23.2 a) Maxwell model, b) free-body diagram

subsequently eliminate them from the rate form of the constitutive equation (state law for stress). This leads to a relation directly linking the rates of strain and stress. The procedure will now be illustrated by an example of a very simple viscoelastic model.

23.3.1 Maxwell Viscoelasticity

Example 23.2: Develop the rate form of the stress-strain law for the Maxwell model, consisting of a linearly elastic unit coupled in series with a linearly viscous unit. For the sake of simplicity, restrict attention to the uniaxial case.

Solution: It is a common practice to visualize material models by means of rheological models consisting, for example, of elastic units represented by springs and viscous units represented by dashpots (also called 'dampers'). For our case, such a model is shown in Figure 23.2(a). It should be kept in mind that this is only a conceptual scheme and the entire spring-dashpot system corresponds to a single material point. The elastic unit is an idealized representation of the free energy stored in the material and the viscous unit is an idealized representation of energy dissipation. The serial coupling of these units corresponds to the assumption that the total strain ε can be additively decomposed[2] into an elastic part, ε_e, and a viscous part, ε_v.

Only the total strain is directly observable, but it does not uniquely characterize the internal details of the material state. Knowing only the total strain, we cannot say how much free energy is stored in the material, and knowing only the total strain rate we cannot say how much energy is being dissipated. It is therefore necessary to include either the elastic strain or the viscous strain in the list of state variables (including both would be redundant). Here we choose ε_v as the internal variable, following a practice initiated in general by Biot (1955), who called ε_v the 'hidden strain'. The other choice is left to be considered in Problem 23.1.

[2] In the rheological models, the lengths of individual units represent the differences between the current state and the initial stress-free state. Consequently, we label the length of the spring as the elastic strain, ε_e, and the length of the dashpot as the viscous strain, ε_v. This is only a simplified illustration of the additive decomposition. Physically, it might seem more reasonable to denote the total length of the model as $1 + \varepsilon$, but then we would have to denote the lengths of the spring and dashpot, respectively, as $0.5 + \varepsilon_e$ and $0.5 + \varepsilon_v$, which would be unappealing and unnecessarily contrived.

Motivated by the previous considerations, we specify the free-energy potential and the dissipation function as

$$\rho\psi(\varepsilon,\varepsilon_v) = \tfrac{1}{2}E(\varepsilon-\varepsilon_v)^2 \tag{23.45}$$

$$\mathcal{D}(\dot\varepsilon,\dot\varepsilon_v,\varepsilon,\varepsilon_v) = \eta\dot\varepsilon_v^2 \tag{23.46}$$

where E and η are material parameters, later identified as the elastic modulus and the viscosity parameter, respectively. The dissipation in this particular case depends only on the viscous strain rate; in (23.46) we nevertheless list all the arguments that could be involved in the general case. The particular form of ψ and $\mathcal{D}$ has been deduced from the selected rheological model, but one could also bypass the mechanical analogy and explore a theory based on simple quadratic expressions postulated *ad hoc*. A quadratic form of the free energy will always lead to linear state laws, and the quadratic function (23.46), defining the dissipation, is guaranteed to always have a non-negative value, which means that any possible evolution of state variables is thermodynamically admissible.

The choice of state variables $(\varepsilon,\varepsilon_v)$, free energy (23.45) and dissipation function (23.46) fully defines the constitutive model. To derive the corresponding set of equations, we differentiate the free energy and substitute the result into the energy balance law (23.44) rewritten as

$$\sigma\dot\varepsilon = \rho\dot\psi + \mathcal{D} \tag{23.47}$$

The resulting equation

$$\sigma\dot\varepsilon = E(\varepsilon-\varepsilon_v)(\dot\varepsilon-\dot\varepsilon_v) + \eta\dot\varepsilon_v^2 \tag{23.48}$$

has an interesting interpretation as the material-point analogy of the principle of virtual power (i.e. the principle of virtual work with virtual increments formally replaced by rates). Indeed, if we denote

$$\sigma_e = E\varepsilon_e = E(\varepsilon-\varepsilon_v) \tag{23.49}$$

$$\sigma_v = \eta\dot\varepsilon_v \tag{23.50}$$

then (23.48) can be rewritten as

$$\sigma\dot\varepsilon = \sigma_e\dot\varepsilon_e + \sigma_v\dot\varepsilon_v \tag{23.51}$$

Variables σ_e and σ_v can be interpreted as the stresses respectively carried by the elastic spring and the viscous dashpot (Figure 23.2(b)), which also ellucidates the physical meaning of constants E and η.

Each energy rate is thus expressed as the product of a stress and a strain rate (in general terms, of a thermodynamic force and the conjugate flux). On the material-point level, the actual stress σ plays the role of 'loading' while the stresses σ_e and σ_v play the role of 'internal forces'. The total strain ε is intimately related to the displacements applied at the boundary of the infinitesimal material volume. Partial strains ε_e and ε_v characterize the internal details of deformation 'inside the material point'; they are not completely independent but are tied by the compatibility condition $\varepsilon_e+\varepsilon_v=\varepsilon$. Replacing $\dot\varepsilon_e$ by $\dot\varepsilon-\dot\varepsilon_v$ again, we can write (23.51) as

$$\sigma\dot\varepsilon = \sigma_e(\dot\varepsilon-\dot\varepsilon_v) + \sigma_v\dot\varepsilon_v \tag{23.52}$$

The First Law expressing energy conservation refers to the actual processes. When, at a given state, we impose a certain strain rate $\dot\varepsilon$, the corresponding internal variable rate $\dot\varepsilon_v$ is a result of processes in the material microstructure that can hardly be affected by an external agent. This means that, to guarantee energy conservation, it is sufficient if equation (23.52) holds only for those pairs $(\dot\varepsilon,\dot\varepsilon_v)$ for which $\dot\varepsilon_v$ is the viscous strain rate produced by the applied strain rate $\dot\varepsilon$. However, it is natural to require (23.52) to also remain valid for fictitious processes described by arbitrary combinations of $\dot\varepsilon$ and $\dot\varepsilon_v$.

This is in complete analogy to structural analysis, in which the external and internal powers remain equal not only for the actual displacement and strain rates, as required by the energy conservation principle, but also for any virtual (fictitious) displacement and strain rates satisfying the compatibility conditions, as required by the principle of virtual work. The latter principle provides a test of equilibrium inside the structure and on its boundary. Similarly, equation (23.52), if regarded as an identity valid for an arbitrary evolution of state variables, automatically provides the conditions of 'internal equilibrium' for the thermodynamic forces conjugate to the state variables and also between these forces and the stress, understood here as an external force.

If we accept this reasoning, we can test (23.52) for two fictitious evolutions – one for which $\dot\varepsilon_v$ vanishes and $\dot\varepsilon$ is arbitrary, and the other for which $\dot\varepsilon$ vanishes and $\dot\varepsilon_v$ is arbitrary. This leads to equations

$$\sigma = \sigma_e \tag{23.53}$$

$$0 = -\sigma_e + \sigma_v \tag{23.54}$$

which describe equilibrium between the stress in the elastic unit and the 'applied' stress, and between the stresses in the elastic and viscous units. Of course, such equations are trivial and immediately obvious from the free-body diagram in Figure 23.2(b).

The simplicity of the present example has helped to clarify the basic concepts. The rheological models may of course be quite complex or might not exist at all. After a proper generalization, the thermodynamic formalism is universally applicable and provides a tool for a systematic construction of the governing equations.

Returning to our example, we note that the thermodynamic forces that appear in (23.53)–(23.54) are functions of the internal variables and their rates, as follows from relations (23.49)–(23.50). Combining these two sets of equations leads to

$$\sigma = E(\varepsilon-\varepsilon_v) \tag{23.55}$$

$$0 = -E(\varepsilon-\varepsilon_v) + \eta\dot\varepsilon_v \tag{23.56}$$

The first equation is the constitutive law, expressing the stress in terms of the state variables. The second equation is the desired evolution law that governs the evolution of the internal variable. It could be rephrased as

$$\dot\varepsilon_v = \frac{E}{\eta}(\varepsilon-\varepsilon_v) \tag{23.57}$$

For a given state $(\varepsilon,\varepsilon_v)$, we can evaluate the stress from (23.55) and solve for the rate of the internal variable from (23.56). Substituting the solution (23.57) into the rate form of (23.55), we get an expression for the stress rate

$$\dot\sigma = E(\dot\varepsilon-\dot\varepsilon_v) = E\dot\varepsilon - \frac{E^2}{\eta}(\varepsilon-\varepsilon_v) \tag{23.58}$$

in terms of the current values of internal variables and the rate of strain. This means that, for a given state, the stress rate is a unique function of the strain rate. In the present example, this function is even invertible, i.e. the strain rate is a unique function of the stress rate, but this is not the case in general. Finally, we can see that if the strain history during a certain time interval is prescribed, the corresponding history of the internal variable ε_v can be computed by integrating the differential equation (23.57). The stress history is then evaluated simply by substitution into (23.55). This shows that the two fundamental thermodynamic functions (23.45)–(23.46), i.e. the free energy potential and the dissipation function, uniquely define the constitutive model, and all the necessary equations can be derived from them. □

The reason why we have paid so much attention to such a trivial example is that it nicely illustrates the basic structure of the thermodynamic formulation of a constitutive model. Now we can proceed to its generalization, which would be, without the example, perhaps too abstract.

23.3.2 Dissipative and Quasiconservative Thermodynamic Forces

Suppose that a model is defined by particular functions describing the free energy and the dissipation, $\psi(\boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ and $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$. It is useful to write the dissipation function in the form

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \equiv \boldsymbol{\sigma}_D(\dot{\boldsymbol{\varepsilon}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) : \dot{\boldsymbol{\varepsilon}} + \boldsymbol{\beta}_D(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \bullet \dot{\boldsymbol{\alpha}} \tag{23.59}$$

where $\boldsymbol{\sigma}_D(\dot{\boldsymbol{\varepsilon}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ is the *dissipative stress* (also called the *viscous stress*) and $\boldsymbol{\beta}_D(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ are the *dissipative thermodynamic forces* conjugate to the internal variables (for the sake of brevity referred to as the dissipative forces). The first term on the right-hand side of (23.59), $\boldsymbol{\sigma}_D : \dot{\boldsymbol{\varepsilon}}$, represents the part of dissipation that does not disappear even when the internal variables are kept fixed (such as the dissipation in a viscous unit coupled in parallel to the Maxwell model, see Problem 23.2). That is why we can consider $\boldsymbol{\sigma}_D$ as independent of $\dot{\boldsymbol{\alpha}}$.

All the relevant equations now follow from the energy balance equation (a transformed version of the First Law), $\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} = \rho\dot{\psi} + \mathcal{D}$, interpreted as an identity that must be valid for an arbitrary (fictitious) history of state variables, which are considered as completely independent. Differentiating the free energy, we can expand the energy balance equation as

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} = \rho\frac{\partial\psi(\boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial\boldsymbol{\varepsilon}} : \dot{\boldsymbol{\varepsilon}} + \rho\frac{\partial\psi(\boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial\boldsymbol{\alpha}} \bullet \dot{\boldsymbol{\alpha}} + \mathcal{D}(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \tag{23.60}$$

Recall that, according to the state laws (23.27) and (23.28), the gradients $\partial\psi/\partial\boldsymbol{\varepsilon}$ and $\partial\psi/\partial\boldsymbol{\alpha}$ multiplied by the density ρ are the quasi-conservative stress $\boldsymbol{\sigma}_Q$ and the quasi-conservative forces $\boldsymbol{\beta}_Q$. Taking this into account, and substituting from (23.59), we can rewrite (23.60) as

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} = [\boldsymbol{\sigma}_Q(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) + \boldsymbol{\sigma}_D(\dot{\boldsymbol{\varepsilon}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})] : \dot{\boldsymbol{\varepsilon}} + [\boldsymbol{\beta}_Q(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) + \boldsymbol{\beta}_D(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})] \bullet \dot{\boldsymbol{\alpha}} \tag{23.61}$$

This is identically satisfied, for any combination of fluxes $\dot{\boldsymbol{\varepsilon}}$ and $\dot{\boldsymbol{\alpha}}$, if

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_Q(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) + \boldsymbol{\sigma}_D(\dot{\boldsymbol{\varepsilon}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \tag{23.62}$$

$$\mathbf{0} = \boldsymbol{\beta}_Q(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) + \boldsymbol{\beta}_D(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \tag{23.63}$$

These equations are sufficient conditions for satisfying (23.61) in general but, strictly speaking, they are not necessary. Indeed, $\boldsymbol{\beta}_Q + \boldsymbol{\beta}_D$ could be equal to some nonzero gyroscopic forces that always remain orthogonal to $\dot{\boldsymbol{\alpha}}$ and thus do not affect the energy balance. Since the present chapter is meant only as an introduction to the thermodynamic approach, we will not go deeper into such subtle details.

Equation (23.62) is a *constitutive law* specifying the stress as a function of the internal variables and the strain rate; equation (23.63) is an *evolution law* implicitly specifying the evolution of internal variables. For a given state (and provided that the model is 'well-posed'), the rates $\dot{\boldsymbol{\alpha}}$ can be determined from (23.63) and subsequently eliminated from the rate form of (23.62). For a given strain history $\boldsymbol{\varepsilon}(t)$, equations (23.63) usually have the character of a set of ordinary differential equations, from which the corresponding history of internal variables $\boldsymbol{\alpha}(t)$ can be solved. Substituting the solution into (23.62) we obtain the stress history $\boldsymbol{\sigma}(t)$. However, the existence and uniqueness of the solution depend on the specific formulation and in general cannot be guaranteed. Furthermore, the model is only thermodynamically admissible if the dissipation function $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}, \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ is nonnegative.

Let us now proceed to the next example, which illustrates certain difficulties with applying the foregoing formulation to models with 'switches', such as those encountered in plasticity.

Example 23.3: Develop a thermodynamic formulation of uniaxial plasticity with linear kinematic hardening.

Solution: The basic equations can be derived from the rheological model in Figure 23.3. Beside two springs of stiffnesses E and H, the model contains a purely dissipative unit (the so-called *Saint–Venant body* or *friction unit*) that generates the dissipation $\sigma_0|\dot{\varepsilon}_p|$, in which σ_0 is a material parameter characterizing that unit, and ε_p is the corresponding strain. Intuitively it is clear that σ_0 is the (initial) yield stress and ε_p is the plastic strain, but the governing equations can be set up following a purely formal procedure, without knowing the physical meaning of the individual parameters and variables (which can be identified later).

Similar to the preceding example, either ε_e or ε_p must be considered as the internal variable. Selecting ε_p, we can write the free energy as

$$\rho\psi(\varepsilon, \varepsilon_p) = \frac{1}{2}E(\varepsilon - \varepsilon_p)^2 + \frac{1}{2}H\varepsilon_p^2 \tag{23.64}$$

and, as stated before, the dissipation is given by

$$\mathcal{D}(\dot{\varepsilon}, \dot{\varepsilon}_p, \varepsilon, \varepsilon_p) = \sigma_0|\dot{\varepsilon}_p| \tag{23.65}$$

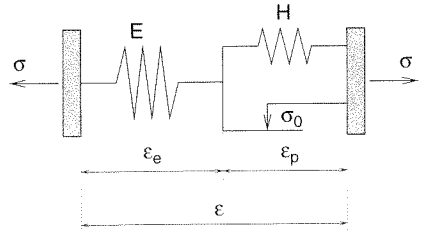


Figure 23.3 Rheological model for plasticity with linear kinematic hardening

The state laws defining the quasi-conservative forces are easy to derive;

$$\sigma_Q(\varepsilon,\varepsilon_p)=\rho\frac{\partial\psi(\varepsilon,\varepsilon_p)}{\partial\varepsilon}=E(\varepsilon-\varepsilon_p) \tag{23.66}$$

$$\beta_Q(\varepsilon,\varepsilon_p)=\rho\frac{\partial\psi(\varepsilon,\varepsilon_p)}{\partial\varepsilon_p}=-E(\varepsilon-\varepsilon_p)+H\varepsilon_p=(E+H)\varepsilon_p-E\varepsilon \tag{23.67}$$

The dissipative stress σ_D is obviously zero. The dissipative force β_D requires special attention. One is tempted to write

$$\beta_D(\dot\varepsilon,\dot\varepsilon_p,\varepsilon,\varepsilon_p)=\sigma_0\,\mathrm{sgn}\,\dot\varepsilon_p \tag{23.68}$$

because this is consistent with the relation $\mathcal{D}=\beta_D\dot\varepsilon_p$. However, while for $\dot\varepsilon_p>0$ we must certainly have $\beta_D=\sigma_0$ and for $\dot\varepsilon_p<0$ we must have $\beta_D=-\sigma_0$, for $\dot\varepsilon_p=0$ the dissipative force does not necessarily vanish. In this last case, the relation $\mathcal{D}=\beta_D\dot\varepsilon_p$ holds for any β_D. This was, in fact, already true in Example 23.2, where we set $\beta_D\equiv\sigma_v=\eta\dot\varepsilon_v$ but for $\dot\varepsilon_v=0$ we could have used an arbitrary value of β_D without affecting the dissipation, $\beta_D\dot\varepsilon_v=0$. In that example, β_D was a continuous function of $\dot\varepsilon_v$, and it was natural to preserve continuity even when passing through a zero viscous strain rate. In the present example, the dissipative force β_D jumps from $-\sigma_0$ to σ_0 when the plastic strain rate changes sign from negative to positive, and it is reasonable to admit that, at $\dot\varepsilon_p=0$, the force β_D can attain any value from the interval $[-\sigma_0,\sigma_0]$. This is clear when we consider the physical meaning of β_D, which represents the stress transmitted by the Saint–Venant unit. If the stress carried by the unit is smaller in magnitude than σ_0, the unit is locked and the corresponding plastic strain rate is zero. Stresses larger in magnitude than σ_0 cannot be transmitted. In contrast to the viscous damper, we have here a non-invertible (and even non-unique) relationship between the thermodynamic flux and the conjugate dissipative force.

Having accepted the above reasoning, we can proceed to the basic equations. The constitutive law (23.62) assumes a simple form

$$\sigma=E(\varepsilon-\varepsilon_p) \tag{23.69}$$

which can be interpreted as the elastic stress-strain law, $\sigma=E\varepsilon_e$. The evolution law (23.63) must be written separately for the cases where the plastic strain rate is zero and nonzero:

$$0=(E+H)\varepsilon_p-E\varepsilon+\beta_D,\qquad -\sigma_0\le\beta_D\le\sigma_0\qquad\text{if }\dot\varepsilon_p=0 \tag{23.70}$$

$$0=(E+H)\varepsilon_p-E\varepsilon+\sigma_0\,\mathrm{sgn}\,\dot\varepsilon_p\qquad\text{if }\dot\varepsilon_p\neq0 \tag{23.71}$$

In the former case, we do not really have an equation but an inequality, which can be written as

$$-\sigma_0\le E\varepsilon-(E+H)\varepsilon_p\le\sigma_0 \tag{23.72}$$

Still, this plays the role of an evolution law, because for those states $(\varepsilon,\varepsilon_p)$ that satisfy the inequalities in (23.72) in the strict sense,

$$-\sigma_0<E\varepsilon-(E+H)\varepsilon_p<\sigma_0 \tag{23.73}$$

the rate $\dot\varepsilon_p$ is zero. The reason is that, for states satisfying (23.73), $\dot\varepsilon_p=0$ is the only admissible solution, because the equality in (23.71) can be satified neither with $\dot\varepsilon_p>0$ nor $\dot\varepsilon_p<0$. In the limit case, in which one of the inequalities in (23.72) is satisfied as an equality, the plastic strain rate can be zero or nonzero, depending on the total strain rate.

On the other hand, (23.71) contains $\dot\varepsilon_p$, but this rate cannot be uniquely determined because it enters only through its sign, which leads to a non-invertible relation. To obtain the desired evolution law, we note that if the rate $\dot\varepsilon_p$ varies continuously and is nonzero at a given time instant $\bar t$, it must remain nonzero (and keep the same sign) during a finite (though possibly short) time interval around $\bar t$. Consequently, equation (23.71) must remain valid in the vicinity of $\bar t$ and it can be differentiated with respect to time. Since $\dot\varepsilon_p$ keeps the same sign, the last term on the right-hand side is constant and the differentiation leads to

$$0=(E+H)\dot\varepsilon_p-E\dot\varepsilon \tag{23.74}$$

from which

$$\dot\varepsilon_p=\frac{E}{E+H}\dot\varepsilon \tag{23.75}$$

In summary, if the current state $(\varepsilon,\varepsilon_p)$ satisfies condition (23.73), then $\dot\varepsilon_p=0$ and $\dot\sigma=E(\dot\varepsilon-\dot\varepsilon_p)=E\dot\varepsilon$, and if (23.73) is violated, then $\dot\varepsilon_p$ is given by (23.75) until $\dot\varepsilon$ changes its sign. In this latter case, the elimination of the thermodynamic flux $\dot\varepsilon_p$ from the rate form of (23.69) leads to the rate form of the stress-strain law,

$$\dot\sigma=E(\dot\varepsilon-\dot\varepsilon_p)=E\left(\dot\varepsilon-\frac{E}{E+H}\dot\varepsilon\right)=\frac{EH}{E+H}\dot\varepsilon \tag{23.76}$$

The reader certainly recognizes (23.72) as the condition of plastic admissibility in terms of the state variables, i.e. in the 'strain space'. Substituting the constitutive law (23.69) into that inequality, we could obtain an equivalent description in the stress space,

$$-\sigma_0\le\sigma-H\varepsilon_p\le\sigma_0 \tag{23.77}$$

When $H\varepsilon_p$ is identified with the back stress, σ_b, (23.77) can be written as

$$-\sigma_0+\sigma_b\le\sigma\le\sigma_0+\sigma_b \tag{23.78}$$

which indicates that the yield limits in tension and in compression are shifted by the same amount and in the same direction (kinematic hardening). □

23.4 DISSIPATION POTENTIAL

In the previous section, we have demonstrated the derivation of the evolution law starting from the description of the dissipation as a function of the state variables and their rates. It has been shown that certain difficulties can arise when the expression for the dissipative forces $\boldsymbol{\beta}_D$ needs to be identified from a given expression for the dissipation. This task is even more delicate for models that use a set of tensorial internal variables. On top of that, in a general case it is not trivial to verify the thermodynamic admissibility of the model, i.e. to check that the dissipation cannot become negative.

All these problems can be avoided if the dissipative forces are defined as the gradient of a suitably chosen scalar function ϕ, called the *dissipation potential* (or the

dissipation pseudo-potential). This potential in general depends on the state variables and their rates, and the dissipative forces are given as

$$\boldsymbol{\sigma}_D = \frac{\partial \phi(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial \dot{\boldsymbol{\varepsilon}}} \tag{23.79}$$

$$\boldsymbol{\beta}_D = \frac{\partial \phi(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial \dot{\boldsymbol{\alpha}}} \tag{23.80}$$

The dependence of the dissipation potential on the *rates* of state variables is essential, while the *values* of state variables play only the role of parameters and often can be omitted (this is emphasized by a semicolon separating the 'essential' variables $\dot{\boldsymbol{\varepsilon}}$ and $\dot{\boldsymbol{\alpha}}$ from the parameters $\boldsymbol{\varepsilon}$ and $\boldsymbol{\alpha}$). Equations (23.79) and (23.80) will be called the *complementary laws*, because they are in some sense the counterpart of the state laws (23.27) and (23.28).

For our purpose, it is sufficient to restrict attention to models with zero dissipative stresses $\boldsymbol{\sigma}_D$. In this case, the dissipation potential is independent of the strain rate, and the complementary laws (23.79) (23.80) are simplified to

$$\boldsymbol{\sigma}_D = \mathbf{0} \tag{23.81}$$

$$\boldsymbol{\beta}_D = \frac{\partial \phi(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial \dot{\boldsymbol{\alpha}}} \tag{23.82}$$

For example, for the linear viscous damper, the dissipative force $\sigma_v = \eta \dot{\varepsilon}_v$ can be derived from the potential

$$\phi(\dot{\varepsilon}_v) = \tfrac{1}{2} \eta \dot{\varepsilon}_v^2 \tag{23.83}$$

Here, σ_v is the only component of $\boldsymbol{\beta}_D$, ε_v is the only component of $\boldsymbol{\alpha}$, and the dissipation potential in this special case only depends on the thermodynamic flux $\dot{\boldsymbol{\alpha}} = \{\dot{\varepsilon}_v\}$. Note that the quadratic form of (23.83) is formally similar to the expression for the free energy of a linear elastic spring, $\psi(\varepsilon_e) = E\varepsilon_e^2/2$. The difference is that ϕ is not a state function, because it depends not only on the state variables, but also on the fluxes (this latter dependence is essential, since without it the dissipative forces derived from the dissipation potential would vanish).

A dissipation potential quadratic in terms of thermodynamic fluxes $\dot{\boldsymbol{\alpha}}$ leads to linear expressions for the dissipative forces $\boldsymbol{\beta}_D$. The coefficients in these expressions can be functions of the state variables, but do not depend on the fluxes. The complementary law (23.82) in this case reads

$$\boldsymbol{\beta}_D = \boldsymbol{A}(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \bullet \dot{\boldsymbol{\alpha}} \tag{23.84}$$

where

$$\boldsymbol{A} = \frac{\partial \boldsymbol{\beta}_D}{\partial \dot{\boldsymbol{\alpha}}} = \frac{\partial^2 \phi}{\partial \dot{\boldsymbol{\alpha}}\, \partial \dot{\boldsymbol{\alpha}}} \tag{23.85}$$

is a symmetric tensor (or collection of tensors), because the second derivatives are interchangable. For instance, if $\boldsymbol{\alpha}$ is a first-order tensor, we obtain

$$A_{ij} = \frac{\partial \beta_i^D}{\partial \dot{\alpha}_j} = \frac{\partial^2 \phi}{\partial \dot{\alpha}_i \partial \dot{\alpha}_j} = \frac{\partial \beta_j^D}{\partial \dot{\alpha}_i} = A_{ji} \tag{23.86}$$

These are the *Onsager–Casimir reciprocity relations*, deduced by Onsager (1931) from the principle of microscopic reversibility, using statistical arguments (his work was

not directly concerned with plasticity of solids). The reciprocity relations have to be considered as an additional postulate that does not follow from the Second Law (and is not necessarily realistic, in general). Note that the dissipation

$$\mathcal{D} = \boldsymbol{\beta}_D \bullet \dot{\boldsymbol{\alpha}} = \dot{\boldsymbol{\alpha}} \bullet \boldsymbol{A} \bullet \dot{\boldsymbol{\alpha}} \tag{23.87}$$

is non-negative if and only if the symmetric part of $\boldsymbol{A}$ is positive semi-definite. The skew-symmetric part of $\boldsymbol{A}$ has no effect on the dissipation, and so its existence does not violate the Second Law. If the complementary law has the linear form (23.84), the symmetry of $\boldsymbol{A}$ is a necessary and sufficient condition for the existence of a dissipation potential

$$\phi(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) = \tfrac{1}{2} \dot{\boldsymbol{\alpha}} \bullet \boldsymbol{A}(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \bullet \dot{\boldsymbol{\alpha}} \tag{23.88}$$

Replacing the quadratic expression (23.88) by a more general one, we can cover nonlinear relations between the fluxes and dissipative forces. The Onsager relations, originally formulated for linear complementary laws, can be generalized in the sense that $\partial \boldsymbol{\beta}_D / \partial \dot{\boldsymbol{\alpha}}$ is a symmetric tensor (or collection of tensors). This fact can be used to explore various couplings. For example, if a certain force β_{ij}^D does not depend on the flux $\dot{\alpha}_{kl}$, then the force β_{kl}^D must not depend on the flux $\dot{\alpha}_{ij}$ (provided that the complementary laws are derived from a dissipation potential).

A special case arises when the dissipation potential is *positively homogeneous of degree one* with respect to the fluxes, i.e. if

$$\phi(\lambda \dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) = \lambda \phi(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \qquad \forall \lambda \geq 0 \tag{23.89}$$

This particular form of ϕ corresponds to time-independent models. Recall that, in general, $f(\boldsymbol{x})$ is called a *homogeneous function* of degree k if $f(\lambda \boldsymbol{x}) = \lambda^k f(\boldsymbol{x})$ for all $\lambda \in \mathcal{R}$. Differentiating this with respect to λ and then setting $\lambda = 1$, we can derive the Euler theorem, $\boldsymbol{x} \bullet \partial f(\boldsymbol{x}) / \partial \boldsymbol{x} = k\, f(\boldsymbol{x})$. If $f(\boldsymbol{x})$ is positively homogeneous of degree k, the theorem remains valid for all $\boldsymbol{x} \neq \mathbf{0}$. In our case, when $f \equiv \phi$ and $k = 1$, we have

$$\phi = \frac{\partial \phi}{\partial \dot{\boldsymbol{\alpha}}} \bullet \dot{\boldsymbol{\alpha}} = \boldsymbol{\beta}_D \bullet \dot{\boldsymbol{\alpha}} = \mathcal{D} \qquad \forall \dot{\boldsymbol{\alpha}} \neq \mathbf{0} \tag{23.90}$$

For $\dot{\boldsymbol{\alpha}} = \mathbf{0}$, we cannot write $\phi = \partial \phi / \partial \dot{\boldsymbol{\alpha}} \bullet \dot{\boldsymbol{\alpha}}$, because the dissipation potential is not differentiable at the origin. Even then the identity $\phi = \mathcal{D}$ holds, because both ϕ and $\mathcal{D}$ are zero (ϕ must be zero due to (23.89), and the dissipation $\mathcal{D}$ vanishes due to its physical meaning). This means that, for this class of models, the dissipation potential is identical to the dissipation function $\mathcal{D}$. The simplest example is provided by the Saint-Venant unit from Example 23.3, for which $\phi(\dot{\boldsymbol{\varepsilon}}_p) = \sigma_0 |\dot{\boldsymbol{\varepsilon}}_p|$.

Let us return to the general case, not necessarily satisfying (23.89). An important advantage of the formulation using a dissipation potential is that the dissipation is guaranteed to be non-negative if the potential has the following properties:[3]

---

[3] The third property is easy to achieve. Indeed, since only the gradient of ϕ with respect to $\dot{\boldsymbol{\alpha}}$ affects the dissipative forces, an arbitrary $\tilde{\phi}(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ can be replaced by $\phi(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) = \tilde{\phi}(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) - \tilde{\phi}(\mathbf{0}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$, which has the same gradient with respect to $\dot{\boldsymbol{\alpha}}$ and satisfies condition 3. Thus, the third property only simplifies the description of the second property, which would otherwise be replaced by the requirement that, for any fixed state $(\boldsymbol{\varepsilon}, \boldsymbol{\alpha})$, ϕ must attain its minimal value at the origin of the space of fluxes $\dot{\boldsymbol{\alpha}}$.

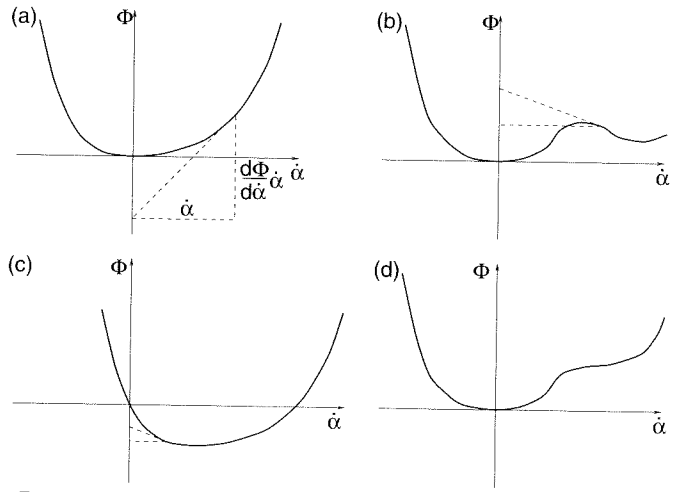


Figure 23.4 Dissipation potential: a) satisfying conditions 1.-3., b) violating convexity, c) violating condition of non-negativity, d) violating convexity but satisfying monotonicity

1. ϕ is convex with respect to $\dot{\boldsymbol{\alpha}}$,
2. ϕ is non-negative for any set of arguments,
3. $\phi(\mathbf{0}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) = 0$ for any combination of state variables $\boldsymbol{\varepsilon}$ and $\boldsymbol{\alpha}$.

The geometrical interpretation of these conditions is shown in Figure 23.4(a) for a smooth dissipation potential depending only on the rate of one scalar internal variable. All the differentiable convex functions ϕ satisfy the inequality

$$\phi(\dot{\boldsymbol{\alpha}} + \Delta\dot{\boldsymbol{\alpha}}) \geq \phi(\dot{\boldsymbol{\alpha}}) + \frac{\partial \phi(\dot{\boldsymbol{\alpha}})}{\partial \dot{\boldsymbol{\alpha}}} \bullet \Delta\dot{\boldsymbol{\alpha}} \qquad \forall \dot{\boldsymbol{\alpha}}, \Delta\dot{\boldsymbol{\alpha}} \tag{23.91}$$

Geometrically this means that the entire graph of $\phi(\dot{\boldsymbol{\alpha}})$ is on or above the tangent constructed at any of its points. Exploiting the special case $\Delta\dot{\boldsymbol{\alpha}} = -\dot{\boldsymbol{\alpha}}$, we obtain

$$\phi(\mathbf{0}) \geq \phi(\dot{\boldsymbol{\alpha}}) - \frac{\partial \phi(\dot{\boldsymbol{\alpha}})}{\partial \dot{\boldsymbol{\alpha}}} \bullet \dot{\boldsymbol{\alpha}} \tag{23.92}$$

Using properties 2. and 3., we finally get

$$\mathcal{D} = \boldsymbol{\beta}_D \bullet \dot{\boldsymbol{\alpha}} = \frac{\partial \phi}{\partial \dot{\boldsymbol{\alpha}}} \bullet \dot{\boldsymbol{\alpha}} \geq \phi(\dot{\boldsymbol{\alpha}}) - \phi(\mathbf{0}) \geq 0 \tag{23.93}$$

The consequences of the lack of convexity and of the violation of the second condition are illustrated in Figures 23.4(b),(c). In fact, the condition of convexity is unnecessarily strong. From the Second Law it only follows that, for any fixed state $(\boldsymbol{\varepsilon}, \boldsymbol{\alpha})$, the dissipation potential must be monotonically increasing on any semi-infinite ray emanating from the origin of the space of fluxes $\dot{\boldsymbol{\alpha}}$. A nonconvex potential with this property is shown in Figure 23.4(d).

The evaluation of the dissipative forces from the gradient of the dissipation potential is not straightforward if the potential is not smooth. This is obviously the case for potentials corresponding to rate-independent models. If the potential is a positively homogeneous function of degree 1, it is not differentiable (in the standard sense) at the origin. Figure 23.5(a) shows a typical example with $\phi(\dot{\alpha}) = \sigma_0|\dot{\alpha}|$ where σ_0 is a

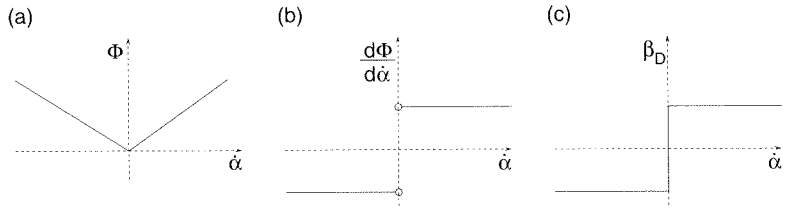


Figure 23.5 Rigid-plastic unit: a) dissipation potential, b) derivative of the dissipation potential, c) dissipative force

positive constant. The classical derivative $\mathrm{d}\phi/\mathrm{d}\dot{\alpha}$ is plotted in Figure 23.5(b). It is defined everywhere except at the origin. From the physical meaning of the dissipative forces (see Example 23.3), we know that the actual relationship between $\dot{\alpha}$ and β_D should correspond to Figure 23.5(c). For $\dot{\alpha} = 0$, the force β_D can assume any value between $-\sigma_0$ and σ_0.

An elegant mathematical description may be based on the concept of the subdifferential, which is a generalization of the differential of smooth functions (Moreau, 1963, 1966, 1970). Suppose that $f(\boldsymbol{x})$ is a scalar function defined on a vector space X. The *subgradient*[4] of f at point $\boldsymbol{x}_0 \in X$ is any vector $\boldsymbol{u} \in X$ for which

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0) + \boldsymbol{u} \bullet (\boldsymbol{x} - \boldsymbol{x}_0) \qquad \forall \boldsymbol{x} \in X \tag{23.94}$$

where $\bullet$ is the scalar product in X. Note the similarity between (23.94) and the inequality (23.91), valid for smooth convex functions. The set of all vectors $\boldsymbol{u}$ satisfying (23.94), denoted as $\partial f(\boldsymbol{x}_0)$, is called the *subdifferential* of f at point $\boldsymbol{x}_0$.

In a one-dimensional space $X = \mathcal{R}$ = set of all real numbers, the subdifferential is the set of all the real numbers u that satisfy the inequality

$$u \leq \frac{f(x) - f(x_0)}{x - x_0} \tag{23.95}$$

for any $x \neq x_0$. If the function f is convex and differentiable at x_0, the subdifferential is a set containing a single element – the usual derivative $f'(x_0)$. If the function is convex and has a 'kink' at x_0, its subdifferential at that point is the closed interval $[f'_-(x_0), f'_+(x_0)]$, where $f'_-(x_0)$ and $f'_+(x_0)$ are the derivatives from the left and from the right, respectively. According to this definition, the subdifferential of $\phi(\dot{\varepsilon}) = \sigma_0|\dot{\varepsilon}|$ at $\dot{\varepsilon} = 0$ is indeed the interval $[-\sigma_0, \sigma_0]$. Expression (23.82) for the dissipative forces is then generalized to

$$\boldsymbol{\beta}_D(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \in \partial_{\dot{\boldsymbol{\alpha}}} \phi(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \tag{23.96}$$

where $\partial_{\dot{\boldsymbol{\alpha}}} \phi$ denotes the subdifferential of ϕ with respect to $\dot{\boldsymbol{\alpha}}$. Recall that the subdifferential is a set and so, compared to (23.82), the symbol $=$ must be replaced by $\in$.

---

[4] In rigorous mathematical texts, subgradients are considered as elements of the dual space X^*, i.e. of the space of all linear forms on the Banach space X, and $\bullet$ is the inner product defined by the duality. Early work on subgradients, mainly by Rockafellar and Moreau, dealt with convex functions. Clarke (1975) extended this concept to general nonconvex functions. Instead of discussing all the mathematical details, we prefer to follow a more intuitive path and focus on the physical interpretation. Precise definitions can be found in textbooks on convex analysis, e.g. Rockafellar (1970) and Ekeland and Temam (1976).

23.5 DUAL DISSIPATION POTENTIAL

23.5.1 Legendre–Fenchel Transform

It has already been mentioned that the rates of internal variables $\dot{\boldsymbol{\alpha}}$ can be eliminated from the rate form of the constitutive equations (23.62), provided that the evolution law (23.63) has a unique solution $\dot{\boldsymbol{\alpha}}$ for any admissible combination of state variables $\boldsymbol{\varepsilon}$ and $\boldsymbol{\alpha}$ (similar to the previous section, we restrict our attention to models for which the dissipative forces $\boldsymbol{\beta}_D$ do not depend on the strain rate $\dot{\boldsymbol{\varepsilon}}$). This is the case if the function $\boldsymbol{\beta}_D$ is invertible with respect to $\dot{\boldsymbol{\alpha}}$. The evolution law (23.63) can then be converted to the *explicit form*

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{\beta}_D^{-1}(-\boldsymbol{\beta}_Q; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}) = \boldsymbol{\beta}_D^{-1}\left(-\rho\frac{\partial\psi(\boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha}\right) \equiv \boldsymbol{k}(\boldsymbol{\varepsilon}, \boldsymbol{\alpha}) \tag{23.97}$$

which is amenable to numerical integration. In this equation, $\boldsymbol{\beta}_D^{-1}$ denotes a partial inverse of the function $\boldsymbol{\beta}_D$ with respect to $\dot{\boldsymbol{\alpha}}$, with the state variables $\boldsymbol{\varepsilon}$ and $\boldsymbol{\alpha}$ taken as parameters. In other words, the function $\boldsymbol{\beta}_D^{-1}(\boldsymbol{\beta}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ assigns to given thermodynamic forces $\boldsymbol{\beta}$ and state variables $\boldsymbol{\varepsilon}$ and $\boldsymbol{\alpha}$ the rate $\dot{\boldsymbol{\alpha}}$ for which $\boldsymbol{\beta} = \boldsymbol{\beta}_D(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$. Once $\boldsymbol{\beta}_D^{-1}$ is known, the function $\boldsymbol{k}$ defined in (23.97) can easily be constructed.

The transformation of the evolution law formally described by (23.97) has already been used in Example 23.2. Indeed, (23.56) represents the implicit form of the evolution law while the corresponding explicit form is given by (23.57). Since the relationship between the thermodynamic flux (viscous strain rate $\dot{\varepsilon}_v$) and the conjugate thermodynamic force (viscous stress σ_v) was linear, the inversion of the function $\beta_D(\dot{\varepsilon}_v) = \eta\dot{\varepsilon}_v$ was a trivial task. In contrast to that, the transformation described by (23.97) might not be easy to perform for general nonlinear evolution laws. If the dissipative forces

$$\boldsymbol{\beta}_D = \frac{\partial\phi(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial\dot{\boldsymbol{\alpha}}} \tag{23.98}$$

are derived from a convex dissipation potential ϕ, the inverse relation

$$\dot{\boldsymbol{\alpha}} = \frac{\partial\phi^*(\boldsymbol{\beta}_D; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})}{\partial\boldsymbol{\beta}_D} \tag{23.99}$$

can be derived from the so-called *dual dissipation potential* ϕ^* obtained via the Legendre–Fenchel transform.

To explain these new concepts, we will consider a nonlinear viscous unit described by

$$\sigma_v = \beta_D(\dot{\varepsilon}_v) \tag{23.100}$$

where the thermodynamic force σ_v has the meaning of stress transmitted by the viscous unit. A possible relationship between this stress and the viscous strain rate $\dot{\varepsilon}_v$ is sketched in Figure 23.6(a). The dissipation potential $\phi(\dot{\varepsilon}_v)$ is obtained by integration of (23.98) and its value at a given $\dot{\varepsilon}_v$ is graphically represented by the area under the graph of β_D from zero to $\dot{\varepsilon}_v$; see Figure 23.6(b). On the other hand, the dual potential should be a function of the viscous stress, and its derivative at a given σ_v should be equal to the corresponding $\dot{\varepsilon}_v$ that solves (23.100). This means that the dual dissipation potential is graphically represented by the area marked in Figure 23.6(c).

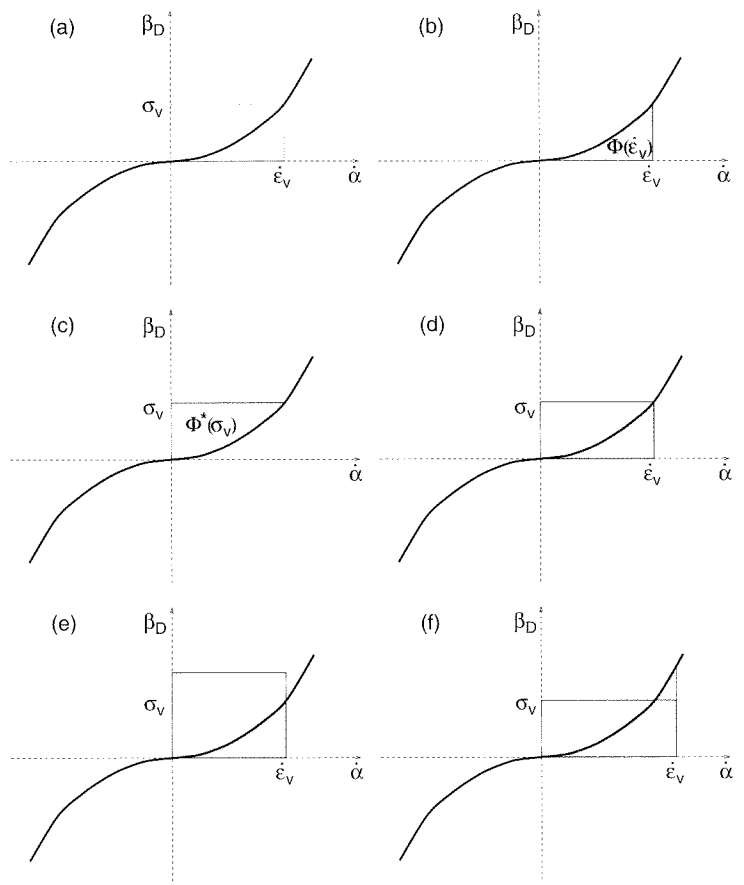


Figure 23.6 Dissipation potential ϕ and dual dissipation potential ϕ^*

It is obvious that if $\sigma_v = \beta_D(\dot{\varepsilon}_v)$, then the sum $\phi(\dot{\varepsilon}_v)+\phi^*(\sigma_v)$ is equal to the product $\sigma_v\dot{\varepsilon}_v$; see Figure 23.6(d). But we can also make another, less trivial observation: If σ_v and $\dot{\varepsilon}_v$ do **not** satisfy (23.100), then the sum $\phi(\dot{\varepsilon}_v) + \phi^*(\sigma_v)$ is larger than the product $\sigma_v\dot{\varepsilon}_v$; see Figures 23.6(e),(f) (this statement is justified if the function β_D is monotonically increasing, in which case both potentials ϕ and ϕ^* are strictly convex). Consequently, for a fixed σ_v and variable $\dot{\varepsilon}_v$, the difference $\sigma_v\dot{\varepsilon}_v - \phi(\dot{\varepsilon}_v)$ cannot exceed $\phi^*(\sigma_v)$ and is equal to $\phi^*(\sigma_v)$ only for the value of $\dot{\varepsilon}_v$ that satisfies (23.100). Based on these considerations, we can construct the dual potential from a given 'primal' potential using the formula

$$\phi^*(\sigma_v) = \max_{\dot{\varepsilon}_v\in\mathcal{R}} [\sigma_v\dot{\varepsilon}_v - \phi(\dot{\varepsilon}_v)] \tag{23.101}$$

known as the *Legendre–Fenchel transform*. A completely analogous formula,

$$\phi(\dot{\varepsilon}_v) = \max_{\sigma_v\in\mathcal{R}} [\sigma_v\dot{\varepsilon}_v - \phi^*(\sigma_v)] \tag{23.102}$$

holds for the inverse transformation of the dual potential, ϕ^*, into the primal one, ϕ. As always, duality is a two-way relationship, and the choice of the potential that is called primal is only a matter of taste. One could equally well start from the complementary law in the form (23.99) and construct $\phi(\dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ as the dual potential

of a given $\phi^*(\boldsymbol{\beta}_D; \boldsymbol{\varepsilon}, \boldsymbol{\alpha})$. In fact, some authors consider the potential that depends on thermodynamic forces as the primal one, and they call the 'dual potential' that which depends on the thermodynamic fluxes.

23.5.2 Uniaxial Perfect Plasticity

So far we have paid attention mainly to the regular case, in which the dissipation potential ϕ is continuously differentiable, strictly convex, non-negative, and vanishing for $\dot{\boldsymbol{\alpha}} = \mathbf{0}$. The dual potential then has the same properties, and the complementary law (23.82) is invertible in the standard sense. However, rate-independent models with 'switch' conditions (such as those typically encountered in classical elastoplasticity with a yield surface) lead to dissipation potentials that are not strictly convex and are not differentiable at the origin of the space of fluxes (i.e. at $\dot{\boldsymbol{\alpha}} = \mathbf{0}$). Still, after appropriate generalizations, the present theoretical framework is fully applicable.

Example 23.4: Construct the complementary law for one-dimensional perfect elastoplasticity with equal yield stress σ_0 in tension and in compression, using the Legendre–Fenchel transform of the dissipation potential.

Solution: Since the model is rate-independent, the dissipation function $\mathcal{D}(\dot{\varepsilon}_p) = \sigma_0|\dot{\varepsilon}_p|$ also plays the role of the dissipation potential $\phi(\dot{\varepsilon}_p)$. Let σ_p be the stress transmitted by the Saint–Venant unit, i.e. the dissipative force conjugate to the plastic strain ε_p. The value of the dual potential ϕ^* at a given σ_p is constructed by maximizing the expression

$$\sigma_p\dot{\varepsilon}_p - \sigma_0|\dot{\varepsilon}_p| \tag{23.103}$$

with respect to the plastic strain rate $\dot{\varepsilon}_p$. It is clear that for

$$|\sigma_p| \le \sigma_0 \tag{23.104}$$

the expression (23.103) is nonpositive for any value of $\dot{\varepsilon}_p$ and is equal to zero for $\dot{\varepsilon}_p = 0$. Consequently, $\phi^*(\sigma_p) = 0$ for all σ_p satisfying (23.104).

When σ_p is larger in magnitude than σ_0, expression (23.103) is not bounded from above and the maximum in the usual sense does not exist. To handle this situation, we allow the dual potential to attain an infinite value and we replace the symbol of maximum in (23.101) by the symbol of supremum, i.e. we write the Legendre–Fenchel transform as

$$\phi^*(\sigma_p) = \sup_{\dot{\varepsilon}_p \in \mathcal{R}} [\sigma_p\dot{\varepsilon}_p - \phi(\dot{\varepsilon}_p)] \tag{23.105}$$

For $|\sigma_p| > \sigma_0$, this formula gives $\phi^*(\sigma_p) = \infty$.

Even though the dual dissipation potential is now described by a nonstandard function

$$\phi^*(\sigma_p) = \begin{cases} 0 & \text{if } |\sigma_p| \le \sigma_0 \\ \infty & \text{if } |\sigma_p| > \sigma_0 \end{cases} \tag{23.106}$$

it still preserves the desired properties – it is convex, non-negative, and equal to zero for $\sigma_p = 0$. Next, we construct its (generalized) derivative that appears in the complementary law $\dot{\varepsilon}_p = \mathrm{d}\phi^*/\mathrm{d}\sigma_p$:

1. For $|\sigma_p| < \sigma_0$, the dual dissipation potential is differentiable in the standard sense and its derivative is equal to zero. This is consistent with the expected complementary law, since the Saint–Venant unit subjected to a stress σ_p below the yield limit remains locked and the corresponding plastic strain rate is zero.
2. Outside the interval $(-\sigma_0, \sigma_0)$, the usual derivative must be replaced by a subdifferential. Recall that the subdifferential of ϕ^* at a given point $\bar{\sigma}_p$ is the set $\partial\phi^*(\bar{\sigma}_p)$ of all the real numbers u satisfying the inequality

$$\phi^*(\sigma_p) - \phi^*(\bar{\sigma}_p) \ge u(\sigma_p - \bar{\sigma}_p) \qquad \forall \sigma_p \in \mathcal{R} \tag{23.107}$$

   (a) First consider the point $\bar{\sigma}_p = \sigma_0$. Since $\phi^*(\sigma_0) = 0$, condition (23.107) can be written as

$$\phi^*(\sigma_p) \ge u(\sigma_p - \sigma_0) \qquad \forall \sigma_p \in \mathcal{R} \tag{23.108}$$

   For $|\sigma_p| > \sigma_0$, the left-hand side is infinite and the inequality holds for any $u \in \mathcal{R}$. For $|\sigma_p| \le \sigma_0$, the left-hand side is zero and $\sigma_p - \sigma_0$ is negative (or zero for $\sigma_p = \sigma_0$); therefore, the inequality holds if u is non-negative. This means that $\partial\phi^*(\sigma_0) = \mathcal{R}_0^+ =$ set of all non-negative real numbers. The geometrical interpretation is shown in Figure 23.7.

   (b) Similarly, it can be shown that $\partial\phi^*(-\sigma_0) = \mathcal{R}_0^- =$ set of all nonpositive real numbers.

   (c) Finally, if $\bar{\sigma}_p$ is larger in magnitude than σ_0, condition (23.107) cannot be satisfied for any real number u. The reason is that if we set σ_p to zero, the inequality from (23.107) reads $-\infty \ge -u\bar{\sigma}_p$. Since the product $u\bar{\sigma}_p$ remains finite, this inequality cannot be satisfied. The subdifferential is in this case the empty set, $\emptyset$.

In summary, we have obtained the complementary law

$$\dot{\varepsilon}_p \in \partial\phi^*(\sigma_p) \equiv \begin{cases} \{0\} & \text{if } \sigma_p \in (-\sigma_0, \sigma_0) \\ \mathcal{R}_0^+ & \text{if } \sigma_p = \sigma_0 \\ \mathcal{R}_0^- & \text{if } \sigma_p = -\sigma_0 \\ \emptyset & \text{if } \sigma_p \notin [-\sigma_0, \sigma_0] \end{cases} \tag{23.109}$$

□

The complementary law constructed in the previous example seems to be quite complicated. However, upon closer examination we realize that (23.109) is equivalent to the flow rule and the loading-unloading conditions. Indeed, in terms of the yield function

$$f(\sigma_p) = |\sigma_p| - \sigma_0 \tag{23.110}$$

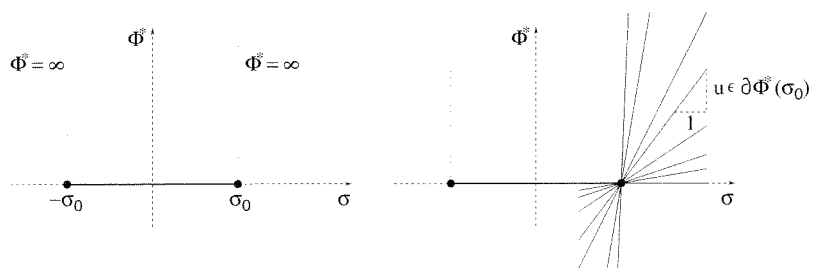


Figure 23.7 Indicator function and its subdifferential

the first line in (23.109) is characterized by $f < 0$, the second and third line by $f = 0$, and the last line by $f > 0$. The associated flow rule

$$\dot{\varepsilon}_p = \dot{\lambda}\frac{\partial f}{\partial \sigma_p} \equiv \dot{\lambda}\,\mathrm{sgn}\,\sigma_p \tag{23.111}$$

and the loading-unloading conditions

$$\dot{\lambda} \geq 0, \qquad f \leq 0, \qquad f\dot{\lambda} = 0 \tag{23.112}$$

postulate that (1) for $f < 0$ the plastic strain rate is zero, (2) for $f = 0$ the plastic strain rate is $\dot{\lambda}\,\mathrm{sgn}\,\sigma_p$, where $\dot{\lambda}$ is an arbitrary non-negative number, and (3) $f > 0$ is not allowed. Exactly the same rules are expressed by (23.109). Note that the condition $f \leq 0$ is not stated explicitly, but it is indirectly implied by the fact that for $f > 0$ the subdifferential is empty and the complementary law states $\dot{\varepsilon}_p \in \partial\phi^* = \emptyset$, which is impossible.

23.5.3 Multiaxial Perfect Plasticity

Another important observation following from Example 23.4 is that the dual potential ϕ^* defined by (23.106) is closely related to the yield function (23.110). The value of ϕ^* is zero for all plastically admissible stresses σ_p (those for which $f(\sigma_p) \leq 0$), and is infinite for all plastically inadmissible states (those for which $f(\sigma_p) > 0$). Mathematically speaking, the dual dissipation potential ϕ^* is the *indicator function* (Moreau, 1966, 1970) of the set of plastically admissible states

$$\overline{\mathcal{E}} = \{\sigma_p \in \mathcal{R} \mid f(\sigma_p) \leq 0\} \tag{23.113}$$

In general, the indicator function $\mathcal{I}_M$ of a set M is defined as

$$\mathcal{I}_M = \begin{cases} 0 & \text{if } x \in M \\ \infty & \text{if } x \notin M \end{cases} \tag{23.114}$$

The relation $\phi^* = \mathcal{I}_{\overline{\mathcal{E}}}$, which we have heuristically found in one-dimensional plasticity, has a deep significance and can be carried over to the general multidimensional case. We skip the mathematical details, which can be found in specialized literature (Moreau, 1970; Rockafellar, 1970; Ekeland and Temam, 1976; Ioffe and Tikhomirov, 1979; Maugin, 1992), and proceed directly to the main results.

Suppose that $f(\boldsymbol{\sigma})$ is a continuous convex yield function defined in the stress space, i.e. in the space $\mathcal{S}_2$ of all the symmetric second-order tensors. Having already generalized the basic concepts introduced for uniaxial plasticity in Chapter 1, we now define the *elastic domain*

$$\mathcal{E} = \{\boldsymbol{\sigma}_p \in \mathcal{S}_2 \mid f(\boldsymbol{\sigma}_p) < 0\} \tag{23.115}$$

as the set of all the elastic stress states, and we denote the closure of $\mathcal{E}$ as $\overline{\mathcal{E}}$ and its boundary as $\partial\mathcal{E}$. Note that $\partial\mathcal{E}$ is the set of stress states satisfying the yield condition, i.e. it is the yield surface, and $\overline{\mathcal{E}}$ is the set of all the plastically admissible stress states. The unstressed state should always be elastic, so we assume that $f(\mathbf{0}) < 0$ and, consequently, $\mathcal{E}$ contains the origin of the stress space. The dual dissipation potential

$$\phi^*(\boldsymbol{\sigma}_p) = \mathcal{I}_{\overline{\mathcal{E}}}(\boldsymbol{\sigma}_p) \tag{23.116}$$

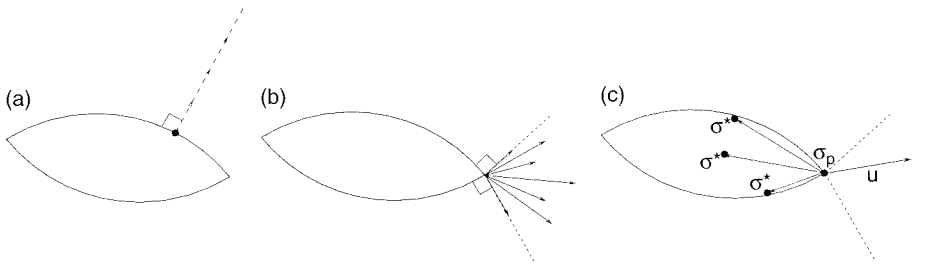


Figure 23.8 Normal cone at a) regular point, b) vertex; c) illustration of condition $\boldsymbol{u} : (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_p) \leq 0$

can be constructed as the indicator function of the set of all plastically admissible states. Arguments similar to the one-dimensional analysis in Example 23.4 lead to the subdifferential

$$\partial\phi^*(\boldsymbol{\sigma}_p) = \partial\mathcal{I}_{\overline{\mathcal{E}}}(\boldsymbol{\sigma}_p) = \begin{cases} \{\mathbf{0}\} & \text{if } \boldsymbol{\sigma}_p \in \mathcal{E} \\ \mathcal{N}(\boldsymbol{\sigma}_p) & \text{if } \boldsymbol{\sigma}_p \in \partial\mathcal{E} \\ \emptyset & \text{if } \boldsymbol{\sigma}_p \notin \overline{\mathcal{E}} \end{cases} \tag{23.117}$$

Here, $\mathcal{N}(\boldsymbol{\sigma}_p)$ denotes *the normal cone* (cone of outward normals) of the yield surface $\partial\mathcal{E}$ at its point $\boldsymbol{\sigma}_p$. At a regular point (yield function differentiable in the standard sense), this cone reduces to the semi-infinite ray described by

$$\mathcal{N}(\boldsymbol{\sigma}_p) = \left\{\boldsymbol{u} \in \mathcal{S}_2 \mid \boldsymbol{u} = \dot{\lambda}\frac{\partial f(\boldsymbol{\sigma}_p)}{\partial \boldsymbol{\sigma}}, \dot{\lambda} \geq 0\right\} \tag{23.118}$$

where $\partial f(\boldsymbol{\sigma}_p)/\partial\boldsymbol{\sigma}$ is the gradient of the yield function evaluated at $\boldsymbol{\sigma}_p$; see Figure 23.8(a). At a vertex of the yield surface (Figure 23.8(b)), the normal cone can be characterized as

$$\mathcal{N}(\boldsymbol{\sigma}_p) = \left\{\boldsymbol{u} \in \mathcal{S}_2 \mid \boldsymbol{u} : (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_p) \leq 0 \quad \forall\boldsymbol{\sigma}^* \in \overline{\mathcal{E}}\right\} \tag{23.119}$$

This means that $\mathcal{N}(\boldsymbol{\sigma}_p)$ contains all the symmetric second-order tensors having a non-negative scalar product with the difference between the given stress tensor $\boldsymbol{\sigma}_p$ on the yield surface and any other plastically admissible stress tensor $\boldsymbol{\sigma}^*$; see Figure 23.8(c). Such a definition of the normal cone is quite general and it contains (23.118) as a special case. Moreover, it is closely related to the postulate of maximum plastic dissipation, as we will soon demonstrate.

Recall that the primary motivation for introducing somewhat unusual mathematical tools such as the indicator functions or subdifferentials was a desire to extend the explicit form of the complementary law (23.99) to models for which the implicit form (23.98) is not invertible in the standard sense. It is now easy to verify that if the dual dissipation potential is given by the indicator function of the set of plastically admissible states according to (23.116), then the resulting complementary law

$$\dot{\boldsymbol{\varepsilon}}_p \in \partial\phi^*(\boldsymbol{\sigma}_p) \tag{23.120}$$

is equivalent to the associated flow rule

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda}\frac{\partial f(\boldsymbol{\sigma}_p)}{\partial \boldsymbol{\sigma}} \tag{23.121}$$

and the loading-unloading conditions (23.112). The Legendre–Fenchel transform

$$\phi(\dot{\boldsymbol{\varepsilon}}_p) = \sup_{\boldsymbol{\sigma}_p \in \mathcal{S}_2} [\boldsymbol{\sigma}_p : \dot{\boldsymbol{\varepsilon}}_p - \phi^*(\boldsymbol{\sigma}_p)] \tag{23.122}$$

is still applicable and defines the dissipation potential $\phi(\dot{\boldsymbol{\varepsilon}}_p)$. The supremum in (23.122) is taken over all the stress tensors $\boldsymbol{\sigma}_p$, but since $\phi^*(\boldsymbol{\sigma}_p) = \infty$ for $\boldsymbol{\sigma}_p \notin \overline{\mathcal{E}}$, it is sufficient to look only for the supremum over the set of all plastically admissible states, $\overline{\mathcal{E}}$.

Moreover, since the dual dissipation potential vanishes for all $\boldsymbol{\sigma}_p \in \overline{\mathcal{E}}$, equation (23.122) can be rewritten in the simpler form

$$\phi(\dot{\boldsymbol{\varepsilon}}_p) = \sup_{\boldsymbol{\sigma}_p \in \overline{\mathcal{E}}} \boldsymbol{\sigma}_p : \dot{\boldsymbol{\varepsilon}}_p \tag{23.123}$$

This formula has a straighforward graphical interpretation: For a given plastic strain rate $\dot{\boldsymbol{\varepsilon}}_p$, the value of the dissipation potential $\phi(\dot{\boldsymbol{\varepsilon}}_p)$ is obtained as the scalar product of $\dot{\boldsymbol{\varepsilon}}_p$ with the plastically admissible stress state that has the largest distance from the origin measured in the direction of $\dot{\boldsymbol{\varepsilon}}_p$; see Figure 23.9(a). If the yield function is strictly convex, such a stress state is unique and corresponds to the stress $\beta_D(\dot{\varepsilon}_p)$ actually generated by the applied plastic strain rate $\dot{\boldsymbol{\varepsilon}}_p$. If the yield function is convex but not strictly convex, then for certain rates $\dot{\boldsymbol{\varepsilon}}_p$, the scalar product $\boldsymbol{\sigma}_p : \dot{\boldsymbol{\varepsilon}}_p$ is maximized by infinitely many stress states located on a straight segment of the yield surface; see Figure 23.9(b). In either case, the dissipation potential $\phi(\dot{\boldsymbol{\varepsilon}}_p)$ is a positively homogeneous function of degree one, because

$$\phi(\lambda\dot{\boldsymbol{\varepsilon}}_p) = \sup_{\boldsymbol{\sigma}_p \in \overline{\mathcal{E}}} \boldsymbol{\sigma}_p : \lambda\dot{\boldsymbol{\varepsilon}}_p = \lambda \sup_{\boldsymbol{\sigma}_p \in \overline{\mathcal{E}}} \boldsymbol{\sigma}_p : \dot{\boldsymbol{\varepsilon}}_p = \lambda\phi(\dot{\boldsymbol{\varepsilon}}_p) \qquad \forall \lambda \geq 0 \tag{23.124}$$

The subdifferential of ϕ is the set of all those stress states that maximize the scalar product $\boldsymbol{\sigma}_p : \dot{\boldsymbol{\varepsilon}}_p$. This is consistent with the implicit form of the complementary law,

$$\boldsymbol{\sigma}_p \in \partial\phi(\dot{\boldsymbol{\varepsilon}}_p) \tag{23.125}$$

meaning that the stress state generated by the applied strain rate $\dot{\boldsymbol{\varepsilon}}_p$ can be any of those states that maximize the scalar product $\boldsymbol{\sigma}_p : \dot{\boldsymbol{\varepsilon}}_p$. As we already know, the dissipation potential $\phi(\dot{\boldsymbol{\varepsilon}}_p)$ for this class of models is identical to the dissipation function $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p)$. From that, we can directly infer the *postulate of maximum plastic dissipation*,

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p) \equiv \boldsymbol{\beta}_D(\dot{\boldsymbol{\varepsilon}}_p) : \dot{\boldsymbol{\varepsilon}}_p \geq \boldsymbol{\sigma}_p : \dot{\boldsymbol{\varepsilon}}_p \qquad \forall \boldsymbol{\sigma}_p \in \overline{\mathcal{E}} \tag{23.126}$$

The fundamental assumptions needed to derive this result were that (1) the yield function is convex, (2) the elastic domain contains the origin of the stress space (the stress-free state is elastic), and (3) the complementary law is derived from the dissipation potential (the flow rule is associated).

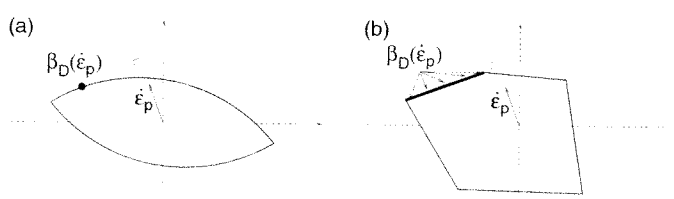


Figure 23.9 Stress state corresponding to a given plastic strain rate: a) strictly convex yield surface, b) not strictly convex yield surface

23.5.4 Hardening Plasticity

An appealing feature of the thermodynamic framework based on convex dissipation potentials is that it allows a straightforward generalization to models with additional internal variables, e.g. to hardening plasticity. In Sections 23.5.2 and 23.5.3 we worked with the internal variables represented by plastic strains $\boldsymbol{\varepsilon}_p$ and with the conjugate thermodynamic forces $\boldsymbol{\sigma}_p$ having the meaning of stress transmitted by the plastic part of the model.[5] The extension to hardening is obtained if $\boldsymbol{\varepsilon}_p$ is replaced by the more general symbol $\boldsymbol{\alpha}$, collecting the plastic strains plus additional internal variables $\boldsymbol{\kappa}$ that control hardening, and if $\boldsymbol{\sigma}_p$ is replaced by the dissipative forces $\boldsymbol{\beta}_D$ conjugate to $\boldsymbol{\alpha}$. The model is fully defined once we specify the free-energy function $\psi(\boldsymbol{\varepsilon}, \boldsymbol{\alpha})$ and the yield function $f(\boldsymbol{\beta}_D)$. The complementary law is of course tacitly assumed to be derived from a dual dissipation potential, constructed as the indicator function of the set of all plastically admissible states (i.e. of all dissipative forces $\boldsymbol{\beta}_D$ for which $f(\boldsymbol{\beta}_D) \leq 0$).

Before we proceed to examples, let us explain the specific role played by the plastic strain. In principle, all the internal variables could be formally treated in the same manner, without paying attention to their physical meaning. However, certain developments can be simplified if one admits that the free energy is the sum of two parts: (1) the elastic energy, ψ_e, which depends only on the difference between the total strain and the plastic strain, and (2) the free energy blocked in the microstructure (e.g. in dislocation pile-ups), ψ_p, which depends only on the hardening variables. The free energy is then written as

$$\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_p, \boldsymbol{\kappa}) = \psi_e(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) + \psi_p(\boldsymbol{\kappa}) \tag{23.127}$$

Let the dissipative forces conjugate to the plastic strain be denoted as $\boldsymbol{\sigma}_p$ and those conjugate to the hardening variables as $\boldsymbol{q}$. According to the evolution law (23.63), the dissipative forces must be equal to minus the quasi-conservative forces, which are given by the state law (23.28). Taking into account the specific form of the free energy (23.127), we can write

$$\boldsymbol{\sigma}_p = -\rho\frac{\partial\psi}{\partial\boldsymbol{\varepsilon}_p} = -\rho\frac{\partial\psi_e}{\partial\boldsymbol{\varepsilon}_p} \tag{23.128}$$

$$\boldsymbol{q} = -\rho\frac{\partial\psi}{\partial\boldsymbol{\kappa}} = -\rho\frac{\partial\psi_p}{\partial\boldsymbol{\kappa}} \tag{23.129}$$

In the absence of dissipative stresses, the constitutive equation (23.62) combined with the state law (23.27) yields

$$\boldsymbol{\sigma} = \rho\frac{\partial\psi}{\partial\boldsymbol{\varepsilon}} = \rho\frac{\partial\psi_e}{\partial\boldsymbol{\varepsilon}} \tag{23.130}$$

Since ψ_e is a function of the difference $\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p$, we have $\partial\psi_e/\partial\boldsymbol{\varepsilon}_p = -\partial\psi_e/\partial\boldsymbol{\varepsilon}$, which implies that

$$\boldsymbol{\sigma}_p = -\rho\frac{\partial\psi_e}{\partial\boldsymbol{\varepsilon}_p} = \rho\frac{\partial\psi_e}{\partial\boldsymbol{\varepsilon}} = \boldsymbol{\sigma} \tag{23.131}$$

---

[5] We will soon prove that if $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}_p$ enter the free energy potential only through their difference, i.e. the elastic strain $\boldsymbol{\varepsilon}_e = \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p$, and if there are no dissipative stresses $\boldsymbol{\sigma}_D$, the thermodynamic force $\boldsymbol{\sigma}_p$ is equal to the applied stress $\boldsymbol{\sigma}$.

Consequently, the dissipative forces $\boldsymbol{\sigma}_p$ conjugate to the plastic strain are equal to the stress $\boldsymbol{\sigma}$. Even though $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_p$ have a distinct physical meaning (externally applied stress and stress carried by the 'plastic part' of the model, respectively), their values, under the present assumptions, are equal at any time (which is closely related to the internal equilibrium condition in a model with serial coupling of an elastic and a plastic unit). This facilitates the formulation of the yield condition, which should be expressed in terms of the dissipative forces $\boldsymbol{\beta}_D = (\boldsymbol{\sigma}_p, \boldsymbol{q})$. For models using the free energy in the form (23.127), we do not need to distinguish between $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_p$, and we can write the yield function as $f(\boldsymbol{\sigma}, \boldsymbol{q})$. The elastic domain is now the set $\mathcal{E} = \{(\boldsymbol{\sigma}, \boldsymbol{q}) \in \mathcal{S}_2 \times \mathcal{Q} \mid f(\boldsymbol{\sigma}, \boldsymbol{q}) < 0\}$, where $\mathcal{Q}$ is the space of the dissipative forces $\boldsymbol{q}$. The indicator function $\mathcal{I}_{\overline{\mathcal{E}}}$ defines the dual dissipation potential ϕ^*, and the complementary law $(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{\kappa}}) \in \partial\phi^*(\boldsymbol{\sigma}, \boldsymbol{q})$ can equivalently be written as

$$f(\boldsymbol{\sigma}, \boldsymbol{q}) \leq 0 \tag{23.132}$$

$$(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{\kappa}}) = (\boldsymbol{0}, \boldsymbol{0}) \quad \text{if } f(\boldsymbol{\sigma}, \boldsymbol{q}) < 0 \tag{23.133}$$

$$(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{\kappa}}) \in \mathcal{N}(\boldsymbol{\sigma}, \boldsymbol{q}) \quad \text{if } f(\boldsymbol{\sigma}, \boldsymbol{q}) = 0 \tag{23.134}$$

where $\mathcal{N}(\boldsymbol{\sigma}, \boldsymbol{q})$ is the normal cone constructed at the point $(\boldsymbol{\sigma}, \boldsymbol{q})$ of the yield surface. Note that (23.132) is only an alternative way of writing $(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{\kappa}}) \in \emptyset$ if $f(\boldsymbol{\sigma}, \boldsymbol{q}) > 0$.

At a regular point of the yield surface, the relation $(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{\kappa}}) \in \mathcal{N}(\boldsymbol{\sigma}, \boldsymbol{q})$ can be replaced by

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda} \frac{\partial f(\boldsymbol{\sigma}, \boldsymbol{q})}{\partial \boldsymbol{\sigma}} \tag{23.135}$$

$$\dot{\boldsymbol{\kappa}} = \dot{\lambda} \frac{\partial f(\boldsymbol{\sigma}, \boldsymbol{q})}{\partial \boldsymbol{q}} \tag{23.136}$$

where $\dot{\lambda} \geq 0$. Equation (23.135) is the associated flow rule, and (23.136) represents the associated hardening law. These equations describe the *generalized normality rules*, because the 'vector' $(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{\kappa}})$ is orthogonal to the yield surface in the space of thermodynamic forces $(\boldsymbol{\sigma}, \boldsymbol{q})$. For this reason, the plastic multiplier rate $\dot{\lambda}$ must be the same in both equations, which implies a certain dependence between the evolutions of the plastic strain and of the hardening variable. In some cases, $\dot{\boldsymbol{\kappa}}$ can be expressed directly in terms of $\dot{\boldsymbol{\varepsilon}}_p$.

Models that use normality rules for the evolution of all internal variables are called, especially by French authors, the *generalized standard materials* (Germain, 1973; Halphen and Nguyen, 1975). If the normality rule is used only for the plastic strain (i.e. the flow rule is associated but the hardening law not necessarily), the material is called just *standard* (Radenković, 1961). Alternative names for generalized standard materials are *standard dissipative materials* or *fully associated models*. This class of material models satisfies the generalized form of the *postulate of maximum dissipation*:

$$\mathcal{D}(\dot{\boldsymbol{\alpha}}) \equiv \boldsymbol{\beta}_D(\dot{\boldsymbol{\alpha}}) \bullet \dot{\boldsymbol{\alpha}} = \sup_{\boldsymbol{\beta}^*_D \in \overline{\mathcal{E}}} (\boldsymbol{\beta}^*_D \bullet \dot{\boldsymbol{\alpha}}) \tag{23.137}$$

This form of the postulate is not limited to plasticity; the only limitation is that the admissible domain $\overline{\mathcal{E}}$ must be defined in the space of thermodynamic forces, which sometimes do not have a straightforward physical meaning. Fortunately, in plasticity the dissipative force conjugate to the plastic strain can always be identified with the

stress, and the postulate can be rewritten as

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{\kappa}}) \equiv \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p + \boldsymbol{q} \bullet \dot{\boldsymbol{\kappa}} = \max_{(\boldsymbol{\sigma}^*, \boldsymbol{q}^*) \in \overline{\mathcal{E}}} (\boldsymbol{\sigma}^* : \dot{\boldsymbol{\varepsilon}}_p + \boldsymbol{q}^* \bullet \dot{\boldsymbol{\kappa}}) \tag{23.138}$$

Example 23.5: Develop the thermodynamic formulation of associated J_2-plasticity with linear isotropic hardening.

Solution: For isotropic hardening we need one scalar variable q appearing in the yield condition along with the stress tensor $\boldsymbol{\sigma}$. This variable is a dissipative force conjugate to a certain hardening variable κ. Since q and κ are work-conjugate, we can select the physical meaning of only one of them, and the other should come out naturally from the condition that

$$\mathcal{D} = \boldsymbol{\beta}_D : \dot{\boldsymbol{\alpha}} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p + q\dot{\kappa} \tag{23.139}$$

must represent the dissipation. The first idea might be to select q as the current (uniaxial) yield stress, σ_Y, which would provide the simplest form of the yield function. However, the dissipative force should vanish in the initial unstressed state, while σ_Y is in that state equal to the initial yield stress, σ_0. Therefore, we select q as the difference $\sigma_0 - \sigma_Y$ (why not $\sigma_Y - \sigma_0$ will become clear later).

Recall that $\boldsymbol{s}$ denotes the deviatoric part of the stress tensor, and the second deviatoric invariant is given by $J_2 = \frac{1}{2}\boldsymbol{s} : \boldsymbol{s}$. Differentiation of the yield function

$$f(\boldsymbol{\sigma}, q) = \sqrt{3J_2(\boldsymbol{\sigma})} + q - \sigma_0 \tag{23.140}$$

provides the associated flow rule,

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda} \frac{\partial f(\boldsymbol{\sigma}, q)}{\partial \boldsymbol{\sigma}} = \dot{\lambda} \frac{\partial}{\partial \boldsymbol{\sigma}} \sqrt{3J_2} = \dot{\lambda} \frac{3\boldsymbol{s}}{2\sqrt{3J_2}} = \dot{\lambda} \sqrt{\frac{3}{2}} \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|} \tag{23.141}$$

and the associated hardening law,

$$\dot{\kappa} = \dot{\lambda} \frac{\partial f(\boldsymbol{\sigma}, q)}{\partial q} = \dot{\lambda} \tag{23.142}$$

By combining these equations, it is possible to identify the physical meaning of κ. Since $\boldsymbol{s}/\|\boldsymbol{s}\|$ is a unit tensor, we have

$$\|\dot{\boldsymbol{\varepsilon}}_p\| = \left| \dot{\lambda} \sqrt{\frac{3}{2}} \right| = \dot{\kappa} \sqrt{\frac{3}{2}} \tag{23.143}$$

This shows that κ is obtained by integrating the norm of the plastic strain rate scaled by the factor $\sqrt{2/3}$. It is thus equal to the accumulated plastic strain defined in (20.6).

So far, we have constructed only the complementary laws. The state law (23.129), relating the dissipative force q to the hardening variable κ, defines the shape of the hardening curve. Since $-q$ represents the increase of uniaxial yield stress (sometimes called the *drag stress*) and κ is under monotonic loading proportional to the plastic strain, the linear hardening is obtained from a quadratic expression for ψ_p. Assuming that the elastic part of the model is also linear, we can specify the free energy (23.127) as

$$\rho\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_p, \kappa) = \tfrac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) : \boldsymbol{D}_e : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) + \tfrac{1}{2}H\kappa^2 \tag{23.144}$$

where $\boldsymbol{D}_e$ is the elastic stiffness tensor and H is the hardening modulus. The elastic stress-strain law has the usual form

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) \tag{23.145}$$

and the hardening law (23.129) reads

$$q = -\rho \frac{\partial \psi}{\partial \kappa} = -H\kappa \tag{23.146}$$

In summary, all the relevant equations have been derived from the free-energy potential (23.144) and from the dual dissipation potential given by the indicator function of the set of plastically admissible states corresponding to the yield function (23.140). □

One may wonder why we have decided to work with a dissipative force q equal to $\sigma_0 - \sigma_Y$, i.e. to *minus* the drag stress. From (23.146) it transpires that, for hardening materials ($H > 0$), the conjugate variables κ and q have opposite signs. It is quite natural to work with a positive κ having the meaning of accumulated plastic strain, and accept the fact that q is negative. Of course, the opposite choice ($q = \sigma_Y - \sigma_0$) would lead to a completely equivalent model, with equation (23.142) replaced by $\dot{\kappa} = -\dot{\lambda}$ and κ having the meaning of minus the accumulated plastic strain.

It is also interesting to note that the term $q\dot{\kappa}$ that appears in expression (23.139) for the dissipation is negative. This does not violate the Second Law, which only requires that the entire dissipation be non-negative for all those processes that can actually occur. Since the evolution of internal variables cannot be controled from the outside, the inequality $\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p + q\dot{\kappa} \geq 0$ need not be satisfied for all the combinations of rates $\dot{\boldsymbol{\varepsilon}}_p$ and $\dot{\kappa}$ considered as completely independent, but only for those that can actually be generated by some imposed strain history. For any real process, the rates $\dot{\boldsymbol{\varepsilon}}_p$ and $\dot{\kappa}$ are given by (23.141) (23.142), with the same plastic multiplier rate $\dot{\lambda} \geq 0$. Consequently, the dissipation can be expressed as

$$\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p + q\dot{\kappa} = \boldsymbol{\sigma} : \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|}\sqrt{\frac{3}{2}}\dot{\lambda} + q\dot{\lambda} = \left(\sqrt{\frac{3}{2}}\|\boldsymbol{s}\| + q\right)\dot{\lambda} = \left(\sqrt{3J_2} + q\right)\dot{\lambda} \tag{23.147}$$

If $\dot{\lambda} > 0$, the yield function (23.140) must vanish, and so $\sqrt{3J_2} + q = \sigma_0 > 0$. This proves that the dissipation is non-negative for any real process.

Example 23.6: Write the general thermodynamic formulation of associated elastoplasticity with kinematic hardening and deduce its special form for J_2-plasticity with linear hardening.

Solution: A wide class of models with kinematic hardening is covered by the yield condition in the form

$$f(\boldsymbol{\sigma}, \boldsymbol{\sigma}_b) \equiv F(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b) - \sigma_0 = 0 \tag{23.148}$$

where $\boldsymbol{\sigma}_b$ is the back stress, representing the shift of the yield surface with respect to its initial position in the stress space. To comply with the thermodynamic framework, we need $\boldsymbol{\sigma}_b$ to be a dissipative force. This means that the internal variables must consist of the plastic strain $\boldsymbol{\varepsilon}_p$ and the variable $\boldsymbol{\kappa}$ conjugate to $\boldsymbol{\sigma}_b$. The free-energy potential has the general form (23.127) where $\boldsymbol{\kappa}$ is now a symmetric second-order tensor, but its physical meaning is still to be determined.

The complementary laws can be constructed from the dual dissipation potential $\phi^* = \mathcal{I}_{\overline{\mathcal{E}}}$, where $\overline{\mathcal{E}}$ is the set of plastically admissible states bounded by the yield surface (23.148). The usual procedure leads to the associated flow rule

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda}\frac{\partial f}{\partial \boldsymbol{\sigma}} = \dot{\lambda}\frac{\partial F}{\partial \boldsymbol{\sigma}} \tag{23.149}$$

and the associated hardening law

$$\dot{\boldsymbol{\kappa}} = \dot{\lambda}\frac{\partial f}{\partial \boldsymbol{\sigma}_b} = \dot{\lambda}\frac{\partial F}{\partial \boldsymbol{\sigma}_b} = -\dot{\lambda}\frac{\partial F}{\partial \boldsymbol{\sigma}} = -\dot{\boldsymbol{\varepsilon}}_p \tag{23.150}$$

with loading-unloading conditions (23.112). If the initial value of $\boldsymbol{\kappa}$ is set to a zero tensor, it follows from (23.150) that, during any real process, $\boldsymbol{\kappa}$ remains equal to minus the plastic strain. The combination of (23.63) with the state law corresponding to $\boldsymbol{\kappa}$ gives the second part of the hardening law,

$$\boldsymbol{\sigma}_b = -\rho\frac{\partial \psi_p}{\partial \boldsymbol{\kappa}} \tag{23.151}$$

The usual state law (23.130), equations (23.149)–(23.151), and the loading-unloading conditions provide a complete description of a class of elastoplastic constitutive models with kinematic hardening.

Now we can develop the particular form of the potentials and governing equations for J_2-plasticity with linear kinematic hardening. J_2-plasticity is characterized by $F(\boldsymbol{\sigma}) \equiv \sqrt{3J_2(\boldsymbol{\sigma})}$, and the yield condition (23.148) becomes

$$\sqrt{3J_2(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)} - \sigma_0 = 0 \tag{23.152}$$

The complementary laws (23.149)–(23.150) read

$$\dot{\boldsymbol{\varepsilon}}_p = \dot{\lambda}\frac{\partial F}{\partial \boldsymbol{\sigma}} = \dot{\lambda}\sqrt{\frac{3}{2}}\frac{\boldsymbol{s} - \boldsymbol{s}_b}{\|\boldsymbol{s} - \boldsymbol{s}_b\|} \tag{23.153}$$

$$\dot{\boldsymbol{\kappa}} = \dot{\lambda}\frac{\partial F}{\partial \boldsymbol{\sigma}_b} = -\dot{\lambda}\sqrt{\frac{3}{2}}\frac{\boldsymbol{s} - \boldsymbol{s}_b}{\|\boldsymbol{s} - \boldsymbol{s}_b\|} = -\dot{\boldsymbol{\varepsilon}}_p \tag{23.154}$$

where $\boldsymbol{s}_b$ is the deviatoric part of the back stress tensor (later shown to be equal to the back stress itself). Linear hardening is obtained with a quadratic potential

$$\rho\psi_p(\boldsymbol{\kappa}) = \tfrac{1}{2}\boldsymbol{\kappa} : \bar{\boldsymbol{H}} : \boldsymbol{\kappa} \tag{23.155}$$

where $\bar{\boldsymbol{H}}$ is in general an isotropic fourth-order tensor. For J_2-plasticity, it follows from (23.153) (23.154) that the plastic strain $\boldsymbol{\varepsilon}_p$ and the hardening variable $\boldsymbol{\kappa} = -\boldsymbol{\varepsilon}_p$ are purely deviatoric tensors. It is therefore sufficient to write (23.155) in the form

$$\rho\psi_p(\boldsymbol{\kappa}) = \tfrac{1}{2}\bar{H}\boldsymbol{\kappa} : \boldsymbol{\kappa} \tag{23.156}$$

where $\bar{H}$ is a scalar material parameter. Substituting into (23.151) and eliminating the hardening variable $\boldsymbol{\kappa}$, we obtain the specific form of the hardening law

$$\boldsymbol{\sigma}_b = -\rho\frac{\partial \psi_p}{\partial \boldsymbol{\kappa}} = -\bar{H}\boldsymbol{\kappa} = \bar{H}\boldsymbol{\varepsilon}_p \tag{23.157}$$

This is the total form of the Melan–Prager kinematic hardening rule (20.43). The back stress is a scalar multiple of a purely deviatoric tensor, and so it is itself a purely deviatoric tensor. □

Note that the thermodynamic formulation naturally leads to the Melan–Prager version of the kinematic hardening rule. This is true independently of the specific choice of function F appearing in the yield condition (23.148).

On the other hand, Ziegler's modification (20.50) and (20.53) cannot be cast in the format of standard dissipative materials (except for the cases in which it coincides with the Melan–Prager rule, as is the case for J_2-plasticity). This does not mean that Ziegler's law is thermodynamically inadmissible. It is only not derivable from a convex dissipation potential, and so the postulate of maximum plastic dissipation ceases to be applicable.

Also, note that essentially the same model could be constructed using a free-energy potential in the form

$$\rho\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_p) = \tfrac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) : \boldsymbol{D}_e : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p) + \tfrac{1}{2}\bar{H}\boldsymbol{\varepsilon}_p : \boldsymbol{\varepsilon}_p \tag{23.158}$$

In this case, the dissipative force $\boldsymbol{\sigma}_p$ conjugate to the plastic strain $\boldsymbol{\varepsilon}_p$ would not be equal to the stress, but to the difference between the stress and the back stress, $\boldsymbol{\sigma}_p = \boldsymbol{\sigma} - \boldsymbol{\sigma}_b$. The yield function could be defined as $f(\boldsymbol{\sigma}_p) \equiv F(\boldsymbol{\sigma}_p) - \sigma_0$, and the basic equations would be directly obtained in the form that corresponds to the model from the previous example after elimination of the internal variable $\boldsymbol{\kappa}$.

Before concluding this chapter, it may be emphasized that, as already made clear, the simplest way to ensure that all the thermodynamic restrictions are satisfied and that stability and convergence of numerical solutions guaranteed is to use normality rules for all inelastic processes. But there is a difficulty – all the potentials must be known and included in the mathematical formulation. This is often difficult or impossible to do. For example, to model the vertex effect (discussed in Section 25.4), a set of many (in practice, over 20) anisotropic potentials is needed, which is feasible in the microplane framework, but is next to impossible within the confines of the classical tensorial plasticity with isotropic invariants. As a recent plasticity theory in which a potential does not exist, one can mention the strain-gradient plasticity model based on dislocations (Gao *et al.*, 1999) and its nonlocal adaptation (Gao and Huang, 2001). As an example where insistence on a potential complicates the formulation (by leading to a double nonlocal operator), the theoretically appealing thermodynamics-based nonlocal plasticity model (Borino, Fuschi and Polizzotto, 1999) may be mentioned. So, the lack of a potential or a deviation from normality can be just a consequence of not knowing, or not using, all the potentials.

In such cases, a reduced number of potentials is used in practice. A mathematician in that case always insists on adhering to potentials with normality rule, even if it harms the capability of representing the behavior observed in experiments. From the engineering viewpoint, though, it may be preferable to violate the normality rule if it achieves a better representation of reality over a broad range of practical situations. An example is the widespread use of nonassociated plasticity for frictional soils. Of course, the tangential stiffness matrices are then nonsymmetric, and convergence of the stress return algorithm may be unachievable for some strain increments. However, an experienced engineer with a good intuitive insight can usually avoid such problems. Nevertheless, when satisfying normality is easy, its violation is a crime.

PROBLEMS

Problem 23.1: Derive the fundamental equations for the Maxwell viscoelastic model from Example 23.2, using the elastic strain instead of the viscous strain as the internal variable.

Problem 23.2: Select a suitable internal variable for the Jeffreys model (parallel coupling of the Maxwell model and a dashpot) and write the expressions for free energy and dissipation, assuming that the spring and dashpots are linear. Then set up the fundamental equations (23.62)–(23.63) and eliminate the rate of the internal variable, ending up with a differential stress-strain relation.

Problem 23.3: Find the dissipation potential $\phi(\dot{\boldsymbol{\varepsilon}}, \dot{\boldsymbol{\varepsilon}}_p, \dot{\kappa})$ corresponding to J_2-plasticity with linear isotropic hardening.

Problem 23.4: Verify the Second Law for the model developed in Example 23.6.

Problem 23.5*: Under which assumptions does a model with kinematic hardening described by Ziegler's law satisfy the Clausius–Duhem inequality?

Problem 23.6: Propose a free-energy potential that leads to an elastoplastic model with nonlinear isotropic hardening governed by the Ramberg–Osgood law $\sigma_Y = \sigma_0 + k\kappa^m$, where σ_Y is the current yield stress, σ_0 is the initial yield stress, κ is the cumulative plastic strain, and k and m are material parameters.

Problem 23.7*: It seems tempting to construct an elastoplastic model with nonlinear kinematic hardening, starting from the free-energy potential constructed in Problem 23.6 with κ replaced by $\|\boldsymbol{\kappa}\|$. What would be wrong with such a model? Hint: Plot the uniaxial stress-strain diagram corresponding to a loading cycle that leads to yielding in tension and a subsequent yielding in compression. Compare this to the model with isotropic hardening.

Problem 23.8: Write the expression for the free energy corresponding to elastoplasticity with linear mixed (kinematic and isotropic) hardening.

Problem 23.9*: Derive the general form of the hardening law for non-smooth, single-surface plasticity formulated within the framework of generalized dissipative materials.

24

Elastoplastic Constitutive Models for Large Strain

So far, it has been assumed that the material only experiences small strains and small rotations; in other words, that the components of the displacement gradient are much smaller than unity (say, smaller than 0.01). Only in this case it is possible to characterize the deformation by the symmetric part of the displacement gradient (i.e. by the usual small-strain tensor) and to neglect the difference between the initial and the deformed configuration in setting up the equilibrium equations. In ductile or slender structures, however, the displacements before failure can become quite large and the previous assumptions must be abandoned.

Three basic types of theories with an increasing level of generality and complexity can be distinguished:

1. The standard linear theory, usually called the *small-strain theory*, although a better term would be the small-displacement theory, and the most rigorous (but clumsy) term would be the small-displacement gradient theory.
2. The *large-displacement theory*, which assumes that the strains remain small but the rotations (and displacements) can become large, and may alternatively be called the large-rotation theory.
3. The most general *large-strain theory*, in which both the strains and rotations can be arbitrarily large.

Instead of the adjectives *small* and *large*, the mathematically oriented literature sometimes uses the adjectives *infinitesimal* and *finite*.

The large-displacement theory (large rotations but small strains) is appropriate for slender structures such as thin shells, plates or beams, which can change their global shape greatly even without experiencing large local deformations. The constitutive equations of standard small-strain theory can still be used, provided that (1) the strain tensor is evaluated with respect to a coordinate frame rotating with the material (this is sometimes called the *co-rotational approach*) and (2) the resulting stresses are rotated back to the global coordinate system.

This chapter is focused on the most general large-strain (or finite-strain) theory. The deformation of the material must be described by a suitably generalized strain tensor that takes into account higher-order effects and reduces to the classical small-strain tensor $\boldsymbol{\varepsilon} = (\partial \boldsymbol{u}/\partial \boldsymbol{x})_{\mathrm{sym}}$ in the limit as the displacement gradient tends to zero. It turns out that the choice of the strain measure is not unique. In the mechanics of solids, it is natural to refer the finite strain to the material coordinates, i.e. to the

coordinates indicating the position of each material point in a reference state, usually taken as the initial stress-free state. There are infinitely many definitions of material strain tensors, each of which uniquely characterizes the deformation of the material, is invariant with respect to superimposed rigid-body motions, and vanishes in the undeformed state.

Later, it will be shown that the Cauchy (true) stress is not work-conjugate to any finite-strain tensor. It turns out that each different definition of the strain measure requires a different definition of the conjugate stress tensor. This arbitrariness is projected into the formulation of the differential equations of equilibrium, whose proper form depends on the choice of the measure of stress. In general, they also contain the derivatives of the displacement field. Therefore, before we proceed to the generalization of elastoplastic constitutive equations, it is necessary to introduce certain elementary notions of finite-strain continuum mechanics. For a detailed and systematic exposition of this subject, the reader is referred to Malvern (1969), Lai, Rubin and Krempl (1999) or Holzapfel (2000).

24.1 KINEMATICS OF LARGE TRANSFORMATIONS

24.1.1 Deformation Gradient and Deformation Tensor

In the small-displacement theory, the deformation of the material is characterized by the strain tensor $\boldsymbol{\varepsilon}$ with components

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{24.1}$$

where u_i, $i = 1, 2, 3$, are the components of the displacement field $\boldsymbol{u}(\boldsymbol{x})$, and x_i, $i = 1, 2, 3$, are the Cartesian coordinates of the material point, i.e. components of the position vector $\boldsymbol{x}$. It is possible to work with other types of coordinate systems than the Cartesian one. For example, axisymmetric problems are conveniently treated in the cylindrical coordinates, and shells are naturally described in curvilinear coordinates related to their midsurface. However, this requires an appropriate generalization of the governing equations such as the kinematic relations (24.1). For the sake of simplicity, exclusively the Cartesian coordinate systems will be used.

If the displacements are assumed to be very small (theoretically infinitely small), it does not make any difference whether $\boldsymbol{x}$ is understood as the initial position of the material point, or its position after deformation. The large-strain theory admits arbitrarily large displacements, and therefore it is necessary to distinguish between the material coordinates, $\boldsymbol{X}$, and the spatial coordinates, $\boldsymbol{x}$.

The *material (Lagrangian) coordinates* are the coordinates of the material points in the initial stress-free state. Each point is permanently labeled by its initial coordinates, and so we can imagine that the material coordinates 'travel' with that point during the deformation process. On the other hand, the *spatial (Eulerian) coordinates* measure the current position of each material point in the deformed state with respect to a coordinate system 'fixed in space'. This means that if we follow the motion of a certain material particle, its material coordinates $\boldsymbol{X}$ remain constant but its spatial coordinates $\boldsymbol{x}(t)$ vary in time. The motion of the entire body is then uniquely described by function $\boldsymbol{x}(\boldsymbol{X}, t)$, which assigns to each particle $\boldsymbol{X} \in V$ and to each time instant $t \geq 0$ the spatial coordinates of that particle at time t. A mathematically rigorous notation would be $\boldsymbol{x} = \boldsymbol{\chi}(\boldsymbol{X}, t)$, where the function $\boldsymbol{\chi}$ is denoted by a different symbol

than its value $\boldsymbol{x}$. However, to simplify notation, it is a common practice to denote the function and its value by the same symbol.

For a fixed $\boldsymbol{X} = \bar{\boldsymbol{X}}$, the function $\boldsymbol{x}(\bar{\boldsymbol{X}}, t)$ describes the motion of particle $\bar{\boldsymbol{X}}$ and, for a fixed $t = \bar{t}$, the function $\boldsymbol{x}(\boldsymbol{X}, \bar{t})$ describes the *placement* of the body at time $\bar{t}$. A related concept is the *configuration*, which represents the class of placements that differ only by a rigid-body motion (translation and rotation). This terminology was proposed by Noll (1972) but is not always used consistently; in many publications the expression 'configuration' denotes what is called here 'placement'. In the present book, 'deformation' means a change of configuration while 'transformation' means a change of placement.

The entire deformation process can be thought of as a continuous sequence of placements. Usually, the *initial placement* at time $t = 0$ is taken as the reference, i.e. the material coordinates $\boldsymbol{X}$ of each point are defined as its spatial coordinates $\boldsymbol{x}$ in the initial placement. This can be mathematically written as $\boldsymbol{x}(\boldsymbol{X}, 0) = \boldsymbol{X}$. Studying the deformation process, we often focus on a certain placement corresponding to a fixed time $t = \bar{t} > 0$ and called the *current placement*. The transformation between the initial placement and the current one is defined by a function $\boldsymbol{x}(\boldsymbol{X})$. Again, in a mathematically rigorous notation we would write $\boldsymbol{x} = \bar{\boldsymbol{\chi}}(\boldsymbol{X})$, where $\bar{\boldsymbol{\chi}}(\boldsymbol{X}) \equiv \boldsymbol{\chi}(\boldsymbol{X}, \bar{t})$.

The displacement field is the difference between the current position of the particle and its initial position, i.e. $\boldsymbol{u}(\boldsymbol{X}) = \boldsymbol{x}(\boldsymbol{X}) - \boldsymbol{X}$; see Figure 24.1. Of course, the displacement field varies in time, so in general we write

$$\boldsymbol{u}(\boldsymbol{X}, t) = \boldsymbol{x}(\boldsymbol{X}, t) - \boldsymbol{X} \tag{24.2}$$

The partial derivative of the displacement with respect to time (with the material coordinate $\boldsymbol{X}$ kept fixed) is the velocity of material point

$$\boldsymbol{v}(\boldsymbol{X}, t) = \frac{\partial \boldsymbol{u}(\boldsymbol{X}, t)}{\partial t} \equiv \dot{\boldsymbol{u}}(\boldsymbol{X}, t) \tag{24.3}$$

From (24.2) it follows that $\boldsymbol{v}(\boldsymbol{X}, t) = \dot{\boldsymbol{x}}(\boldsymbol{X}, t)$.

Let us now look at the deformation of an infinitesimal material volume that has, in the initial placement, the shape of a rectangular prism with edges parallel to the coordinate axes. The deformation process transforms this prism into a parallelepiped (the straight lines are, of course, in general transformed into curved lines, but when we look at an infinitesimal volume, we can neglect the higher-order effects and replace a curved segment by a straight one). This is illustrated in a two-dimensional projection in Figure 24.2(a). In the small-displacement theory, the normal strain components characterize the relative extensions of fictitious material 'fibers' running parallel to the coordinate axes, and the shear strain components represent one half of the change of initially right angles between these fibers. However, the expressions $\varepsilon_{11} = \partial u_1 / \partial X_1$,

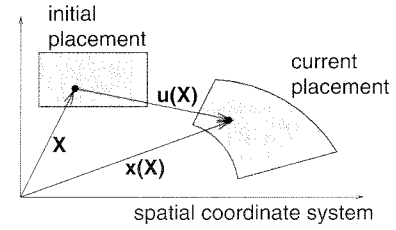


Figure 24.1 Transformation of a body from the initial into the current state

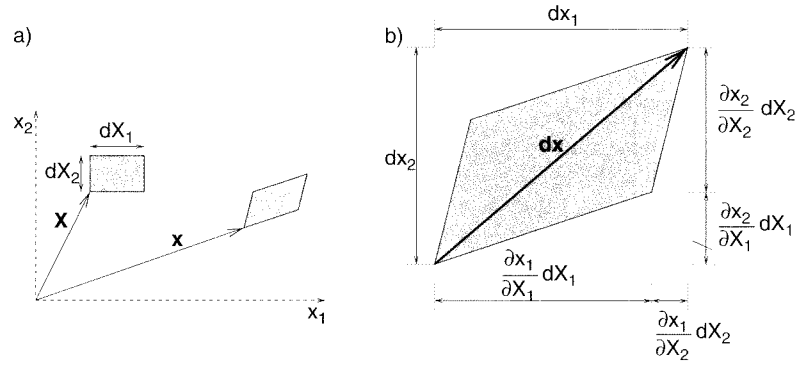


Figure 24.2 Transformation of an infinitesimal material volume

etc. are only approximations taking into account the linear terms that dominate as $\|\partial \boldsymbol{u}/\partial \boldsymbol{X}\| \to 0$.

Consider an infinitesimal segment of a fiber in an arbitrary direction, described in the initial placement by a vector $\mathrm{d}\boldsymbol{X} = \mathrm{d}X_i \boldsymbol{e}_i$, where $\boldsymbol{e}_i$ $(i = 1, 2, 3)$ are unit vectors in the global coordinate directions and the summation convention applies to the subscript i. After deformation, the vector $\mathrm{d}\boldsymbol{X}$ is mapped onto $\mathrm{d}\boldsymbol{x} = \mathrm{d}x_i \boldsymbol{e}_i$ where

$$\mathrm{d}x_i = \frac{\partial x_i}{\partial X_j} \mathrm{d}X_j, \qquad i = 1, 2, 3 \tag{24.4}$$

A graphical interpretation of this relation is sketched in Figure 24.2(b). Even though $\boldsymbol{x}$ is a function of $\boldsymbol{X}$ and t, the differential of time does not appear in (24.4) because $\mathrm{d}\boldsymbol{x}$ is understood as an infinitesimal material segment at a given time, and not as the total differential of the function $\boldsymbol{x}(\boldsymbol{X}, t)$.

Introducing the *deformation gradient*

$$\boldsymbol{F}(\boldsymbol{X}, t) = \frac{\partial \boldsymbol{x}(\boldsymbol{X}, t)}{\partial \boldsymbol{X}} \tag{24.5}$$

with components

$$F_{ij} = \frac{\partial x_i}{\partial X_j} \tag{24.6}$$

we can rewrite (24.4) as $\mathrm{d}x_i = F_{ij} \mathrm{d}X_j$ $(i = 1, 2, 3)$, and in the compact notation this gives

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{X} \tag{24.7}$$

The segment of initial length $\mathrm{d}L = \|\mathrm{d}\boldsymbol{X}\| = \sqrt{\mathrm{d}\boldsymbol{X} \cdot \mathrm{d}\boldsymbol{X}}$ is stretched by the deformation process to the current length

$$\begin{aligned} \mathrm{d}l = \|\mathrm{d}\boldsymbol{x}\| &= \sqrt{\mathrm{d}x_i \, \mathrm{d}x_i} = \sqrt{F_{ij} \mathrm{d}X_j F_{ik} \mathrm{d}X_k} \\ &= \sqrt{\mathrm{d}X_j F^T_{ji} F_{ik} \mathrm{d}X_k} = \sqrt{\mathrm{d}\boldsymbol{X} \cdot \boldsymbol{F}^T \cdot \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{X}} \end{aligned} \tag{24.8}$$

So the tensor

$$\boldsymbol{C} = \boldsymbol{F}^T \cdot \boldsymbol{F} \tag{24.9}$$

called the *Green deformation tensor*,[1] makes it possible to calculate the local extension of an arbitrarily oriented fiber passing through the given material point at which the deformation gradient $\boldsymbol{F}$ is evaluated. The ratio between the deformed and the initial length of the fiber,

$$\lambda = \frac{\mathrm{d}l}{\mathrm{d}L} = \frac{\|\mathrm{d}\boldsymbol{x}\|}{\|\mathrm{d}\boldsymbol{X}\|} = \sqrt{\frac{\mathrm{d}\boldsymbol{X} \cdot \boldsymbol{C} \cdot \mathrm{d}\boldsymbol{X}}{\mathrm{d}\boldsymbol{X} \cdot \mathrm{d}\boldsymbol{X}}} \tag{24.10}$$

is called the *stretch*. It is related to the relative extension

$$\varepsilon = \frac{\mathrm{d}l - \mathrm{d}L}{\mathrm{d}L} = \frac{\|\mathrm{d}\boldsymbol{x}\| - \|\mathrm{d}\boldsymbol{X}\|}{\|\mathrm{d}\boldsymbol{X}\|} = \frac{\|\mathrm{d}\boldsymbol{x}\|}{\|\mathrm{d}\boldsymbol{X}\|} - 1 = \lambda - 1 \tag{24.11}$$

Knowing $\boldsymbol{C}$, we can also compute the angle between two infinitesimal segments described in the initial placement by vectors $\mathrm{d}\boldsymbol{X}_A$ and $\mathrm{d}\boldsymbol{X}_B$, which are transformed into vectors $\mathrm{d}\boldsymbol{x}_A = \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{X}_A$ and $\mathrm{d}\boldsymbol{x}_B = \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{X}_B$. From the well-known geometric interpretation of the scalar product, $\mathrm{d}\boldsymbol{x}_A \cdot \mathrm{d}\boldsymbol{x}_B = \|\mathrm{d}\boldsymbol{x}_A\| \, \|\mathrm{d}\boldsymbol{x}_B\| \cos \alpha_{AB}$, and from $\mathrm{d}\boldsymbol{x}_A \cdot \mathrm{d}\boldsymbol{x}_B = \mathrm{d}\boldsymbol{X}_A \cdot \boldsymbol{F}^T \cdot \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{X}_B = \mathrm{d}\boldsymbol{X}_A \cdot \boldsymbol{C} \cdot \mathrm{d}\boldsymbol{X}_B$, it follows that the angle α_{AB} can be evaluated as

$$\alpha_{AB} = \arccos \frac{\mathrm{d}\boldsymbol{x}_A \cdot \mathrm{d}\boldsymbol{x}_B}{\|\mathrm{d}\boldsymbol{x}_A\| \, \|\mathrm{d}\boldsymbol{x}_B\|} = \arccos \frac{\mathrm{d}\boldsymbol{X}_A \cdot \boldsymbol{C} \cdot \mathrm{d}\boldsymbol{X}_B}{\sqrt{\mathrm{d}\boldsymbol{X}_A \cdot \boldsymbol{C} \cdot \mathrm{d}\boldsymbol{X}_A} \sqrt{\mathrm{d}\boldsymbol{X}_B \cdot \boldsymbol{C} \cdot \mathrm{d}\boldsymbol{X}_B}} \tag{24.12}$$

So the Green deformation tensor $\boldsymbol{C}$ completely determines the local state of deformation in an infinitesimal neighborhood of the material point at which it is evaluated.

To verify that $\boldsymbol{C}$ is symmetric, note that $\boldsymbol{C}^T = (\boldsymbol{F}^T \cdot \boldsymbol{F})^T = \boldsymbol{F}^T \cdot (\boldsymbol{F}^T)^T = \boldsymbol{F}^T \cdot \boldsymbol{F} = \boldsymbol{C}$. Consequently, $\boldsymbol{C}$ has only six independent components while the deformation gradient $\boldsymbol{F}$ is in general nonsymmetric and has nine independent components. This is because $\boldsymbol{F}$ contains the information not only on the local deformation but also on the local rotation of the material. We can also say that $\boldsymbol{F}$ characterizes the local placement while $\boldsymbol{C}$ characterizes the local configuration.

24.1.2 Polar Decomposition

Let us show how the deformation gradient $\boldsymbol{F}$ can be decomposed into the part due to pure deformation and the part due to rigid-body rotation of the local neighborhood. Since $\boldsymbol{C}$ is a symmetric, positive definite[2] tensor, all its eigenvalues are real and positive, and so we can use the *spectral decomposition*

$$\boldsymbol{C} = \sum_{I=1}^{3} \mu_I \boldsymbol{N}_I \otimes \boldsymbol{N}_I \tag{24.13}$$

---

[1] Tensor $\boldsymbol{C}$ was first used by Green (1839, 1841). Earlier, Cauchy (1828, 1829) introduced another deformation tensor, $\boldsymbol{c} = \boldsymbol{F}^{-T} \cdot \boldsymbol{F}^{-1}$, which is in fact $\boldsymbol{C}$ computed for the inverse transformation (from the current to the initial placement). Therefore, $\boldsymbol{C}$ is often referred to as the *right Cauchy–Green deformation tensor* (and $\boldsymbol{c}^{-1} = \boldsymbol{F} \cdot \boldsymbol{F}^T$ is then the *left Cauchy–Green deformation tensor*). We opt here for the shorter name, Green deformation tensor.

[2] The positive definiteness of $\boldsymbol{C}$ can be proven as follows: $\boldsymbol{X} \cdot \boldsymbol{C} \cdot \boldsymbol{X} = \boldsymbol{X} \cdot \boldsymbol{F}^T \cdot \boldsymbol{F} \cdot \boldsymbol{X} = (\boldsymbol{F} \cdot \boldsymbol{X})^T \cdot (\boldsymbol{F} \cdot \boldsymbol{X}) > 0$ for any nonzero vector $\boldsymbol{X}$ provided that $\boldsymbol{F}$ is regular. However, regularity of $\boldsymbol{F}$ is guaranteed because of the physical meaning of its determinant, $J = \det \boldsymbol{F}$, which characterizes the relative change of volume between the initial and currect configuration; see (24.53). A vanishing J would mean that the material is compressed into a zero volume, which is physically impossible.

where μ_I $(I = 1,2,3)$ are the eigenvalues of $\boldsymbol{C}$, and $\boldsymbol{N}_I$ $(I = 1,2,3)$ are the corresponding unit eigenvectors that satisfy the orthonormality conditions $\boldsymbol{N}_I \cdot \boldsymbol{N}_J = \delta_{IJ}$ = Kronecker delta. For a fiber aligned with the principal direction $\boldsymbol{N}_J$, we obtain from (24.10) the stretch

$$\lambda_J = \sqrt{\frac{\boldsymbol{N}_J \cdot \boldsymbol{C} \cdot \boldsymbol{N}_J}{\boldsymbol{N}_J \cdot \boldsymbol{N}_J}} = \sqrt{\sum_{I=1}^{3} \mu_I \boldsymbol{N}_J \cdot \boldsymbol{N}_I \otimes \boldsymbol{N}_I \cdot \boldsymbol{N}_J}$$
$$= \sqrt{\sum_{I=1}^{3} \mu_I \delta_{JI} \delta_{IJ}} = \sqrt{\mu_J}, \qquad J = 1,2,3 \tag{24.14}$$

This means that μ_I, the eigenvalues of $\boldsymbol{C}$, are the squares of the principal stretches λ_I associated with the principal directions $\boldsymbol{N}_I$. The tensor

$$\boldsymbol{U} = \sum_{I=1}^{3} \lambda_I \boldsymbol{N}_I \otimes \boldsymbol{N}_I = \sum_{I=1}^{3} \sqrt{\mu_I} \boldsymbol{N}_I \otimes \boldsymbol{N}_I \tag{24.15}$$

has the same principal directions as $\boldsymbol{C}$ and its eigenvalues are the square roots of the eigenvalues of $\boldsymbol{C}$. $\boldsymbol{U}$ represents what is known as the *tensorial square root* of $\boldsymbol{C}$, denoted as $\boldsymbol{C}^{1/2}$, because

$$\boldsymbol{U}^2 = \boldsymbol{U} \cdot \boldsymbol{U} = \sum_{I,J=1}^{3} \lambda_I \boldsymbol{N}_I \otimes \boldsymbol{N}_I \cdot \boldsymbol{N}_J \otimes \boldsymbol{N}_J \lambda_J$$
$$= \sum_{I,J=1}^{3} \lambda_I \lambda_J \boldsymbol{N}_I \otimes \delta_{IJ} \boldsymbol{N}_J = \sum_{I=1}^{3} \lambda_I^2 \boldsymbol{N}_I \otimes \boldsymbol{N}_I = \boldsymbol{C} \tag{24.16}$$

Since $\boldsymbol{U}$ is symmetric, we have $\boldsymbol{U}^T = \boldsymbol{U}$ and we can also write $\boldsymbol{C} = \boldsymbol{U}^T \cdot \boldsymbol{U}$. So, the tensor $\boldsymbol{U}$ is one of those (infinitely many) local deformation gradients that lead to the same Green deformation tensor $\boldsymbol{C}$ (note that the field $\boldsymbol{U}(\boldsymbol{X})$ is not necessarily the deformation gradient of any function $\boldsymbol{x}(\boldsymbol{X})$, but this is irrelevant since we analyze the local state of each material point separately). At the same time, tensor $\boldsymbol{U}$ is a very special local deformation gradient, because it is symmetric. Since $\boldsymbol{N}_I$ are the principal directions of $\boldsymbol{U}$, and λ_I are the corresponding eigenvalues, we have $\boldsymbol{U} \cdot \boldsymbol{N}_I = \lambda_I \boldsymbol{N}_I$ $(I = 1,2,3)$. This means that, in a deformation process for which $\boldsymbol{U}$ is the deformation gradient, the fibers in the three mutually orthogonal principal directions are only stretched but do not rotate and remain mutually orthogonal. Therefore, $\boldsymbol{U}$ corresponds to a pure stretching of the material. Any other deformation gradient $\boldsymbol{F}$ leading to the same Green deformation tensor $\boldsymbol{C}$ as $\boldsymbol{U}$ (i.e. to $\boldsymbol{U}^2$) can only differ from $\boldsymbol{U}$ by a rigid-body rotation of the local neighborhood.

So, we can decompose the total deformation gradient $\boldsymbol{F}$ into pure stretching, described by the tensor $\boldsymbol{U} = \boldsymbol{C}^{1/2} = (\boldsymbol{F}^T \cdot \boldsymbol{F})^{1/2}$, followed by a rigid-body rotation, described by a tensor $\boldsymbol{R}$. The stretching transforms an infinitesimal vector $\mathrm{d}\boldsymbol{X}$ into $\boldsymbol{U} \cdot \mathrm{d}\boldsymbol{X}$ and the subsequent rotation transforms $\boldsymbol{U} \cdot \mathrm{d}\boldsymbol{X}$ into $\boldsymbol{R} \cdot (\boldsymbol{U} \cdot \mathrm{d}\boldsymbol{X}) = (\boldsymbol{R} \cdot \boldsymbol{U}) \cdot \mathrm{d}\boldsymbol{X}$, which means that the resulting deformation gradient is

$$\boldsymbol{F} = \boldsymbol{R} \cdot \boldsymbol{U} \tag{24.17}$$

Equation (24.17) defines the *polar decomposition* of the deformation gradient $\boldsymbol{F}$ into the *right stretch tensor* $\boldsymbol{U}$ and the *rotation tensor* $\boldsymbol{R}$. The right stretch tensor is symmetric, because it is defined as the tensorial square root of the symmetric, positive definite tensor $\boldsymbol{C} = \boldsymbol{F}^T \cdot \boldsymbol{F}$. The rotation tensor $\boldsymbol{R}$ represents a (possibly large) rotation of the local neighborhood as a rigid body, and so it can be expected to be orthogonal in the sense that $\boldsymbol{R}^T \cdot \boldsymbol{R} = \boldsymbol{R} \cdot \boldsymbol{R}^T = \boldsymbol{I}$. This can be easily verified if we express, from (24.17), $\boldsymbol{R} = \boldsymbol{F} \cdot \boldsymbol{U}^{-1}$ and then compute

$$\boldsymbol{R}^T \cdot \boldsymbol{R} = (\boldsymbol{F} \cdot \boldsymbol{U}^{-1})^T \cdot (\boldsymbol{F} \cdot \boldsymbol{U}^{-1}) = \boldsymbol{U}^{-T} \cdot \boldsymbol{F}^T \cdot \boldsymbol{F} \cdot \boldsymbol{U}^{-1}$$
$$= \boldsymbol{U}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{U}^{-1} = \boldsymbol{U}^{-1} \cdot \boldsymbol{U}^2 \cdot \boldsymbol{U}^{-1} = \boldsymbol{I} \tag{24.18}$$

In this chapter, $\boldsymbol{I}$ denotes the second-order unit tensor (Kronecker delta), which was denoted as $\boldsymbol{\delta}$ in the preceding chapters. Here we adhere to the traditional notation of continuum mechanics.

The tensor $\boldsymbol{U}$ is called the right stretch tensor because there is an alternative polar decomposition (in which pure stretching follows rotation), $\boldsymbol{F} = \boldsymbol{V} \cdot \boldsymbol{R}$, with the same rotation tensor $\boldsymbol{R}$ but a different symmetric tensor $\boldsymbol{V}$, called the *left stretch tensor*. $\boldsymbol{V}$ has the same eigenvalues λ_I as $\boldsymbol{U}$, and its eigenvectors, $\boldsymbol{R} \cdot \boldsymbol{N}_I$, are obtained by rotating the eigenvectors of $\boldsymbol{U}$.

24.1.3 Strain Tensors

According to (24.17), the deformation gradient can be decomposed into the part due to pure stretching and the part due to a rigid-body rotation. Note that this decomposition is multiplicative and is applied to the entire deformation gradient

$$\boldsymbol{F} = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}} = \frac{\partial}{\partial \boldsymbol{X}} (\boldsymbol{X} + \boldsymbol{u}) = \boldsymbol{I} + \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \tag{24.19}$$

This seems to be quite different from the decomposition used in the small-displacement theory, which is additive and operates on the displacement gradient $\partial \boldsymbol{u}/\partial \boldsymbol{X}$. The symmetric part $(\partial \boldsymbol{u}/\partial \boldsymbol{X})_{\mathrm{sym}} = \boldsymbol{\varepsilon}$ is the small-strain tensor, and the skew-symmetric part $(\partial \boldsymbol{u}/\partial \boldsymbol{X})_{\mathrm{skew}} = \boldsymbol{\omega}$ represents a small rotation (in (24.1) we defined the components of the small-strain tensor using derivatives with respect to $\boldsymbol{x}$, but that was before introducing the distinction between the material coordinates $\boldsymbol{X}$ and the spatial coordinates $\boldsymbol{x}$). The multiplicative polar decomposition (24.17) can be reduced to the additive decomposition

$$\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} = \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)_{\mathrm{sym}} + \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)_{\mathrm{skew}} = \boldsymbol{\varepsilon} + \boldsymbol{\omega} \tag{24.20}$$

under the assumption that the components of the displacement gradient are small with respect to unity. Neglecting terms of order $O(\|\partial \boldsymbol{u}/\partial \boldsymbol{X}\|^2)$, we can write

$$\boldsymbol{C} = \boldsymbol{F}^T \cdot \boldsymbol{F} = \left(\boldsymbol{I} + \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)^T \cdot \left(\boldsymbol{I} + \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)$$
$$= \boldsymbol{I} + \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} + \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)^T + \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)^T \cdot \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}}$$
$$\approx \boldsymbol{I} + 2 \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)_{\mathrm{sym}} = \boldsymbol{I} + 2\boldsymbol{\varepsilon} \tag{24.21}$$

$$\boldsymbol{U} = \boldsymbol{C}^{1/2} \approx (\boldsymbol{I} + 2\boldsymbol{\varepsilon})^{1/2} \approx \boldsymbol{I} + \boldsymbol{\varepsilon} \tag{24.22}$$

$$\boldsymbol{U}^{-1} = \boldsymbol{C}^{-1/2} \approx (\boldsymbol{I} + 2\boldsymbol{\varepsilon})^{-1/2} \approx \boldsymbol{I} - \boldsymbol{\varepsilon} \tag{24.23}$$

$$\boldsymbol{R} = \boldsymbol{F} \cdot \boldsymbol{U}^{-1} \approx \left(\boldsymbol{I} + \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}}\right) \cdot (\boldsymbol{I} - \boldsymbol{\varepsilon}) \approx \boldsymbol{I} + \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} - \boldsymbol{\varepsilon} = \boldsymbol{I} + \boldsymbol{\omega} \tag{24.24}$$

From (24.22) and (24.21) we obtain alternative approximations

$$\boldsymbol{\varepsilon} \approx \boldsymbol{U} - \boldsymbol{I} \tag{24.25}$$

and

$$\boldsymbol{\varepsilon} \approx \tfrac{1}{2}(\boldsymbol{C} - \boldsymbol{I}) \tag{24.26}$$

which differ only by terms of order $O(\|\partial \boldsymbol{u}/\partial \boldsymbol{X}\|^2)$. This provides motivation for the definition of strain measures

$$\boldsymbol{E}^{(1)} \equiv \boldsymbol{U} - \boldsymbol{I} \tag{24.27}$$

and

$$\boldsymbol{E}^{(2)} \equiv \tfrac{1}{2}(\boldsymbol{C} - \boldsymbol{I}) = \tfrac{1}{2}(\boldsymbol{U}^2 - \boldsymbol{I}) \tag{24.28}$$

respectively called the *Biot strain tensor* (Biot, 1934) and the *Green's Lagrangian strain tensor* (introduced by Green (1839), and called Lagrangian because it is referred to the Lagrangian coordinates, i.e. to the coordinates of material points in the initial state, as opposed to the spatial, Eulerian coordinates). Both $\boldsymbol{E}^{(1)}$ and $\boldsymbol{E}^{(2)}$ are appropriate measures of strain in the sense that in the initial undeformed configuration they are zero, for (infinitely) small displacement gradients they coincide with the usual small-strain tensor, and for arbitrarily large displacements they depend only on the stretching part of the deformation gradient (i.e. only on $\boldsymbol{U}$ but not on $\boldsymbol{R}$).

Equations (24.27) and (24.28) are two special cases of a more general formula

$$\boldsymbol{E}^{(m)} \equiv \tfrac{1}{m}(\boldsymbol{U}^m - \boldsymbol{I}) \tag{24.29}$$

which defines an entire class of tensors, called the *Doyle–Ericksen strain tensors* (Doyle and Ericksen, 1956) parameterized by a real number $m \neq 0$. For a discussion of many aspects of the Doyle–Ericksen strain tensors, see Bažant and Cedolin (1991, Chapter 11). The spectral decomposition of (24.29) is

$$\boldsymbol{E}^{(m)} = \sum_{I=1}^{3} \frac{1}{m}(\lambda_I^m - 1)\, \boldsymbol{N}_I \otimes \boldsymbol{N}_I \tag{24.30}$$

This formula can even be extended to the limit case $m = 0$ by introducing the *logarithmic strain tensor*, also called the *Hencky strain tensor* (Biezeno and Hencky, 1928, 1929),

$$\boldsymbol{E}^{(0)} \equiv \ln \boldsymbol{U} = \sum_{I=1}^{3} (\ln \lambda_I) \boldsymbol{N}_I \otimes \boldsymbol{N}_I \tag{24.31}$$

This strain measure is often favored in materials science. It has several convenient properties: (1) the logarithmic strain tensor corresponding to $\boldsymbol{F}$ is minus the logarithmic strain tensor corresponding to $\boldsymbol{F}^{-1}$ (compression-tension symmetry); (2) the volumetric-deviatoric decomposition of the logarithmic strain tensor is additive: and (3) the logarithmic strains corresponding to two subsequent transformations

with the same principal stretch directions are additive. Computationally, however, the Hencky strain is less efficient than Green's Lagrangian strain because the spectral decomposition must be calculated. However, very accurate yet efficient approximations of $\ln \boldsymbol{U}$ exist; $\ln \boldsymbol{U} \approx \frac{1}{2}(\boldsymbol{U} + \boldsymbol{U}^{-1})$ and $\ln \boldsymbol{U} \approx \boldsymbol{U}^{1/2} - \boldsymbol{U}^{-1/2}$ (whose increments are easy to calculate). More generally (Bažant, 1998),

$$\ln \boldsymbol{U} \approx (\boldsymbol{U}^m - \boldsymbol{U}^{-m})/2m = (\boldsymbol{E}^{(m)} + \boldsymbol{E}^{(-m)})/2 \tag{24.32}$$

For principal stretches between 2/3 and 3/2, the error is under 2.8% for $m = 1$, and under 0.7% for $m = 1/2$.

24.2 EQUILIBRIUM AND CONSTITUTIVE EQUATIONS

24.2.1 Equilibrium Equations

After explaining the elementary kinematic aspects of the large-strain theory, we are ready to look at the tensors that characterize the stress state, and at the corresponding equilibrium equations. Recall that, in the small-displacement theory, the Cauchy equations of equilibrium (D.74) have the well-known form

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \bar{b}_i = 0, \qquad i = 1, 2, 3, \tag{24.33}$$

where $\bar{b}_i$, $i = 1, 2, 3$, are the components of the body force vector. These equations can be derived from the equilibrium conditions of an infinitesimal material prism and they remain valid even in the general case of arbitrarily large displacements, provided that all the symbols are interpreted properly. Since the state of equilibrium is reached in the current placement, the elementary prism must be cut out from the deformed body, and the tractions and volume forces acting on this prism must be considered as force intensities per unit area or volume in the current (deformed) configuration. This means that in (24.33)

1. x_j must be considered as the *spatial* coordinates (not as the material coordinates),
2. $\bar{b}_i$ are components of the body force per unit *deformed* volume (not per unit initial volume),
3. each stress component $\sigma_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{t}_j$ is the projection of traction $\boldsymbol{t}_j$ on the direction parallel to the ith coordinate axis, where $\boldsymbol{t}_j$ is the traction acting on an elementary surface with a normal parallel to the jth coordinate axis and taken as a force per unit *deformed* area.

The stress tensor $\boldsymbol{\sigma}$ defined according to point 3 is called the *Cauchy stress* or the *true stress*. From the definition of Cauchy stress it follows that an elementary surface that has in the current placement area $\mathrm{d}a$ and unit normal $\boldsymbol{n}$ transmits an elementary force

$$\mathrm{d}\boldsymbol{f} = \boldsymbol{\sigma} \cdot \boldsymbol{n}\, \mathrm{d}a \tag{24.34}$$

It is easy to see that the traction boundary conditions must have exactly the same form (D.81) as in the small-displacement case, i.e.

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \bar{\boldsymbol{t}} \tag{24.35}$$

where $\bar{\boldsymbol{t}}$ is the prescribed surface traction per unit deformed area. The body forces (such as the gravity forces or inertial forces in the sense of d'Alembert's principle) are

often proportional to the mass, and so it is convenient to write the body forces per unit deformed volume as

$$\bar{\boldsymbol{b}} = \rho \bar{\boldsymbol{g}} \tag{24.36}$$

where ρ is the density (mass per unit volume) in the deformed configuration and $\bar{\boldsymbol{g}}$ is the body force per unit mass. As an example, for gravity loading, $\bar{\boldsymbol{g}}$ is the gravity acceleration vector, and for inertial forces it is the acceleration $\ddot{\boldsymbol{u}} = \partial^2 \boldsymbol{u}/\partial t^2$.

24.2.2 Internal Power and Dissipation

We are now able to describe the kinematic aspects of a deformation process in terms of the polar decomposition of the deformation gradient and one of the strain measures related to the stretching part of that decomposition. We can also set up the equilibrium equations in terms of Cauchy stress derivatives with respect to the spatial coordinates. It remains to link the stress tensor to the strain tensor by the constitutive equations. Since the same state of deformation can be characterized by many different strain tensors, the precise form of the constitutive equations depends not only on the material behavior but also on the choice of the strain measure.

To construct the constitutive equations in a rigorous and systematic manner, it is useful to exploit a thermodynamic framework. For that we need to generalize the basic expressions derived for the small-displacement case in Section 23.1. We start from the external power

$$\mathcal{P}_{\text{ext}} = \int_v \rho \bar{\boldsymbol{g}} \cdot \boldsymbol{v} \, \mathrm{d}v + \int_s \bar{\boldsymbol{t}} \cdot \boldsymbol{v} \, \mathrm{d}s \tag{24.37}$$

representing the rate of work done by body forces $\rho \bar{\boldsymbol{g}}$ (per unit deformed volume) and surface tractions $\bar{\boldsymbol{t}}$ (per unit deformed area) on the displacement rate $\boldsymbol{v} = \dot{\boldsymbol{u}} = \dot{\boldsymbol{x}}$, i.e. on the time derivative of the position vector of each material particle. The integrals are taken over the spatial domain v occupied by the deformed body and over its surface s. For large displacements, the difference between the current domain v and the initial domain V cannot be neglected.

Equation (24.37) is a generalized form of (23.2). The corresponding generalization of expression (23.3) for the kinetic energy is

$$\mathcal{K} = \frac{1}{2} \int_v \rho \, \boldsymbol{v} \cdot \boldsymbol{v} \, \mathrm{d}v \tag{24.38}$$

An application of the Green–Gauss theorem analogous to (23.6)–(23.7) leads to $\mathcal{P}_{\text{ext}} = \dot{\mathcal{K}} + \mathcal{P}_{\text{int}}$, with the internal power given by

$$\mathcal{P}_{\text{int}} = \int_v \boldsymbol{\sigma} : \frac{\partial \boldsymbol{v}(\boldsymbol{x}, t)}{\partial \boldsymbol{x}} \, \mathrm{d}v \tag{24.39}$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor and $\partial \boldsymbol{v}/\partial \boldsymbol{x}$ is the spatial gradient of the velocity field, henceforth denoted as $\boldsymbol{L}$ and called simply the *velocity gradient*.

The same expression as (24.39) (only with integration over V instead of v) was obtained in Section 23.1. It was then argued that, due to the symmetry of $\boldsymbol{\sigma}$, $\partial \boldsymbol{v}/\partial \boldsymbol{x}$ can be replaced by its symmetric part, which is equal to the rate of the small-strain tensor, $\dot{\boldsymbol{\varepsilon}}$.

In the general large-strain theory, the Cauchy stress $\boldsymbol{\sigma}$ remains symmetric, and so $\partial \boldsymbol{v}/\partial \boldsymbol{x}$ can again be replaced by its symmetric part. This motivates the decomposition of the velocity gradient

$$\boldsymbol{L} = \frac{\partial \boldsymbol{v}(\boldsymbol{x}, t)}{\partial \boldsymbol{x}} = \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \right)_{\text{sym}} + \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \right)_{\text{skew}} = \boldsymbol{D} + \boldsymbol{W} \tag{24.40}$$

into the symmetric part

$$\boldsymbol{D} = \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \right)_{\text{sym}} = \frac{1}{2} \left[\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} + \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \right)^T \right] = \frac{1}{2} \left(\boldsymbol{L} + \boldsymbol{L}^T \right) \tag{24.41}$$

called the *deformation-rate tensor* (or the *stretching tensor*, or the *velocity strain*), and the skew-symmetric part

$$\boldsymbol{W} = \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \right)_{\text{skew}} = \frac{1}{2} \left[\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} - \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \right)^T \right] = \frac{1}{2} \left(\boldsymbol{L} - \boldsymbol{L}^T \right) \tag{24.42}$$

called the *spin tensor* (or the *vorticity tensor*). Since the spatial coordinates $\boldsymbol{x}$ are different from the material coordinates $\boldsymbol{X}$, the deformation-rate tensor $\boldsymbol{D} = (\partial \boldsymbol{v}/\partial \boldsymbol{x})_{\text{sym}}$ is different from the rate of the small-strain tensor, $\dot{\boldsymbol{\varepsilon}} = \partial(\partial \boldsymbol{u}/\partial \boldsymbol{X})_{\text{sym}}/\partial t = (\partial \boldsymbol{v}/\partial \boldsymbol{X})_{\text{sym}}$. Note that the correct definition of $\boldsymbol{\varepsilon}$ is $(\partial \boldsymbol{u}/\partial \boldsymbol{X})_{\text{sym}}$, which is equal to $(\partial \boldsymbol{u}/\partial \boldsymbol{x})_{\text{sym}}$ only if the displacements are negligible.

Repeating the procedure explained for the small-strain theory in Section 23.1 and restricting attention to isothermal processes, we can express the mechanical dissipation per unit volume in the deformed configuration as

$$\mathcal{D} = \boldsymbol{\sigma} : \boldsymbol{D} - \rho \dot{\psi} \tag{24.43}$$

where ρ is the mass density in the deformed configuration (variable in time).

24.2.3 Hyperelasticity

In the simplest case of an elastic material, the dissipation is zero and no internal variables are needed. The material state is uniquely characterized by one selected strain measure, henceforth denoted as $\boldsymbol{E}$. The stress-strain law can be derived from the free-energy potential $\psi(\boldsymbol{E})$. According to (24.43), for a nondissipative material under isothermal conditions we must have

$$\boldsymbol{\sigma} : \boldsymbol{D} - \rho \frac{\partial \psi}{\partial \boldsymbol{E}} : \dot{\boldsymbol{E}} = 0 \tag{24.44}$$

If we could find a strain measure $\boldsymbol{E}$ such that $\dot{\boldsymbol{E}} = \boldsymbol{D}$, we could deduce from (24.44) the elastic stress-strain law in the form $\boldsymbol{\sigma} = \rho \, \partial \psi / \partial \boldsymbol{E}$, analogous to the small-displacement case. However, no holonomic (i.e. path-independent) strain measure with the desired property exists. Fortunately, we can still exploit equation (24.44) with any strain measure $\boldsymbol{E}$, but the relationship between its rate $\dot{\boldsymbol{E}}$ and the deformation-rate tensor $\boldsymbol{D}$ must be found and substituted into (24.44) in order to identify the relationship between the Cauchy stress tensor $\boldsymbol{\sigma}$ and the gradient of the free-energy potential, $\partial \psi / \partial \boldsymbol{E}$.

The elastic constitutive laws derived from an energy potential (Green, 1839, 1841; Kirchhoff, 1852) are called *hyperelastic*, as opposed to *hypoelastic* laws, postulated directly in the rate form (Truesdell, 1955a). Any hyperelastic law can be transformed into the rate form, but the converse is not true.

Example 24.1: Derive the hyperelastic stress-strain law ensuing from a free-energy potential that is defined as a function of the Green's Lagrangian strain tensor $\boldsymbol{E}^{(2)}$.

Solution: It is useful to first establish an auxiliary relation between the time derivative of the deformation gradient and the spatial velocity gradient:

$$\dot{\boldsymbol{F}} = \frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}}\right) = \frac{\partial^2 \boldsymbol{x}}{\partial \boldsymbol{X}\partial t} = \frac{\partial}{\partial \boldsymbol{X}}\left(\frac{\partial \boldsymbol{x}}{\partial t}\right) = \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{X}} = \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \cdot \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}} = \boldsymbol{L} \cdot \boldsymbol{F} \tag{24.45}$$

According to (24.28) and (24.9), the Green's Lagrangian strain $\boldsymbol{E} \equiv \boldsymbol{E}^{(2)}$ is defined as

$$\boldsymbol{E} = \tfrac{1}{2}(\boldsymbol{C} - \boldsymbol{I}) = \tfrac{1}{2}\left(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I}\right) \tag{24.46}$$

and so its time derivative is

$$\begin{aligned}\dot{\boldsymbol{E}} &= \tfrac{1}{2}\left(\dot{\boldsymbol{F}}^T \cdot \boldsymbol{F} + \boldsymbol{F}^T \cdot \dot{\boldsymbol{F}}\right) = \tfrac{1}{2}\left(\boldsymbol{F}^T \cdot \boldsymbol{L}^T \cdot \boldsymbol{F} + \boldsymbol{F}^T \cdot \boldsymbol{L} \cdot \boldsymbol{F}\right) \\ &= \tfrac{1}{2}\boldsymbol{F}^T \cdot \left(\boldsymbol{L} + \boldsymbol{L}^T\right) \cdot \boldsymbol{F} = \boldsymbol{F}^T \cdot \boldsymbol{D} \cdot \boldsymbol{F}\end{aligned} \tag{24.47}$$

The scalar product $(\partial\psi/\partial\boldsymbol{E}) : \dot{\boldsymbol{E}}$ can now be expressed as

$$\begin{aligned}\frac{\partial \psi}{\partial \boldsymbol{E}} : \dot{\boldsymbol{E}} &= \frac{\partial \psi}{\partial \boldsymbol{E}} : (\boldsymbol{F}^T \cdot \boldsymbol{D} \cdot \boldsymbol{F}) = \frac{\partial \psi}{\partial E_{ij}} F_{ki} D_{kl} F_{lj} \\ &= F_{ki}\frac{\partial \psi}{\partial E_{ij}} F^T_{jl} D_{kl} = \left(\boldsymbol{F} \cdot \frac{\partial \psi}{\partial \boldsymbol{E}} \cdot \boldsymbol{F}^T\right) : \boldsymbol{D}\end{aligned} \tag{24.48}$$

Substituting this into (24.44) we obtain

$$\left(\boldsymbol{\sigma} - \rho \boldsymbol{F} \cdot \frac{\partial \psi}{\partial \boldsymbol{E}} \cdot \boldsymbol{F}^T\right) : \boldsymbol{D} = 0 \tag{24.49}$$

Since this equality must be satisfied for any $\boldsymbol{D}$, the expression in the parentheses must vanish; the resulting stress-strain law

$$\boldsymbol{\sigma} = \rho \boldsymbol{F} \cdot \frac{\partial \psi}{\partial \boldsymbol{E}} \cdot \boldsymbol{F}^T \tag{24.50}$$

links the Cauchy stress to the gradient of the free energy with respect to the Green's Lagrangian strain tensor.

Strictly speaking, $\boldsymbol{D}$ is not completely arbitrary because it must always be symmetric. Therefore, from (24.49) we can conclude only that the symmetric part of $\boldsymbol{\sigma} - \rho\boldsymbol{F} \cdot (\partial\psi/\partial\boldsymbol{E}) \cdot \boldsymbol{F}^T$ must be zero. However, since both $\boldsymbol{\sigma}$ and $\boldsymbol{E}$ are symmetric, this is equivalent to (24.50). □

Note that the stress-strain law (24.50) also contains the deformation gradient, and so it is somewhat more complicated than the analogous law valid in the small-displacement theory, $\boldsymbol{\sigma} = \rho_0\, \partial\psi/\partial\boldsymbol{\varepsilon}$ (the mass density in the initial configuration

is now denoted ρ_0 while ρ denotes the current mass density). The reason is that the Cauchy stress is not work-conjugate to the Green's Lagrangian strain. The quasiconservative thermodynamic force $\boldsymbol{S}$ conjugate to the Green's Lagrangian strain (which will from now on be denoted simply $\boldsymbol{E}$) is defined by the state law

$$\boldsymbol{S} = \rho_0 \frac{\partial \psi}{\partial \boldsymbol{E}} \tag{24.51}$$

and, as it follows from (24.50), the Cauchy stress is obtained by the transformation

$$\boldsymbol{\sigma} = \frac{\rho}{\rho_0}\boldsymbol{F} \cdot \boldsymbol{S} \cdot \boldsymbol{F}^T \tag{24.52}$$

The density ratio ρ/ρ_0 can be expressed in terms of the deformation gradient if we realize that, due to mass conservation, we must have $\rho\, \mathrm{d}v = \rho_0\, \mathrm{d}V$ where $\mathrm{d}V$ is the volume occupied by a certain elementary mass in the initial configuration and $\mathrm{d}v$ is the volume occupied by the same mass in the current configuration. It is known from integral calculus that $\mathrm{d}v/\,\mathrm{d}V = \det(\partial\boldsymbol{x}/\partial\boldsymbol{X}) = \det\boldsymbol{F}$. The determinant of the deformation gradient,

$$J = \det \boldsymbol{F} = \frac{\mathrm{d}v}{\mathrm{d}V} = \frac{\rho_0}{\rho} \tag{24.53}$$

is called the *Jacobian* of the transformation between the initial configuration and the current one. Using the Jacobian, equation (24.52) can be rewritten as

$$\boldsymbol{\sigma} = \frac{1}{J}\boldsymbol{F} \cdot \boldsymbol{S} \cdot \boldsymbol{F}^T \tag{24.54}$$

Since the components of the deformation gradient are dimensionless, the thermodynamic force $\boldsymbol{S}$ has the dimension of stress. It is called the *second Piola–Kirchhoff stress tensor* and its physical meaning is not as straightforward as the meaning of Cauchy stress. Nevertheless, it is perfectly legitimate to use $\boldsymbol{S}$ as a stress measure and substitute (24.54) into the equilibrium equations (24.33). After some algebraic manipulations, the resulting equations can be presented in the elegant form

$$\frac{\partial(F_{ik}S_{kj})}{\partial X_j} + \rho_0 \bar{g}_i = 0 \tag{24.55}$$

This motivates the definition of the *first Piola–Kirchhoff stress tensor*,

$$\boldsymbol{P} = \boldsymbol{F} \cdot \boldsymbol{S} = J\boldsymbol{\sigma} \cdot \boldsymbol{F}^{-T} \tag{24.56}$$

which is nonsymmetric but leads to the equilibrium equations in the initial placement with exactly the same formal structure as the Cauchy equations, namely

$$\frac{\partial P_{ij}}{\partial X_j} + \rho_0 \bar{g}_i = 0 \tag{24.57}$$

So, depending on the particular choice of the stress and strain measures, we obtain different forms of the kinematic, constitutive and equilibrium equations, even though the material itself remains the same. In fact, each of the strain tensors $\boldsymbol{E}^{(m)}$ defined in (24.29) is conjugate to a certain stress tensor $\boldsymbol{S}^{(m)}$ that could be transformed into the Cauchy stress tensor $\boldsymbol{\sigma}$. The transformation laws for stress measures other than

the second Piola–Kirchhoff stress, i.e. for $m \neq 2$, are more involved than (24.52); see Bažant and Cedolin (1991, Chapter 11). The scalar product $\boldsymbol{S}^{(m)} : \dot{\boldsymbol{E}}^{(m)}$ has the meaning of stress power per unit initial volume, while $\boldsymbol{\sigma} : \boldsymbol{D}$ is the stress power per unit deformed volume. In terms of work-conjugate quantities, the hyperelastic stress-strain laws generally have the form

$$\boldsymbol{S}^{(m)} = \rho_0 \frac{\partial \psi}{\partial \boldsymbol{E}^{(m)}} \tag{24.58}$$

It is important to realize that the *expression* for free energy depends on the selected strain measure, while the *value* of free energy for a given material state is of course independent of the choice of the strain measure. If the free energy is described by a quadratic function (with constant coefficients) in terms of one specific strain measure, the corresponding law (24.58) is linear; however, after transformation to any other strain measure the free energy is not quadratic any more and the corresponding stress-strain law becomes nonlinear. There is no reason to expect that there exists a preferred strain measure for which the free energy remains quadratic for arbitrarily large strains. This means that linear small-strain elasticity does not have a unique extension to the large-strain range. The meaning of 'linear' depends on the choice of the strain measure, which is not unique.

A number of hyperelastic stress-strain laws have been proposed in the literature, and their suitability depends on the specific type of material. As an illustrative example, we will present a particular case of the compressible extension of the *neo-Hookean material* (Rivlin, 1948). To provide some motivation for the specific choice of the free-energy potential, let us start from the potential used in linear small-strain elasticity,

$$\rho_0 \psi = \frac{1}{2}\left(K - \frac{2}{3}G\right)(\boldsymbol{I} : \varepsilon)^2 + G\varepsilon : \varepsilon = \frac{1}{2}\left(K - \frac{2}{3}G\right)\left(\sum_{I=1}^{3} \varepsilon_I\right)^2 + G\sum_{I=1}^{3} \varepsilon_I^2 \tag{24.59}$$

where $K - 2G/3$ and G are Lamé parameters (normally denoted λ and μ, but we want to avoid confusion with stretches λ), $\boldsymbol{I}$ is the second-order unit tensor (Kronecker delta), ε is the small-strain tensor, and ε_I $(I = 1, 2, 3)$ are its eigenvalues (principal strains).

The simplest extension to the large-strain range would consist in the replacement of the small-strain tensor ε by the Green's Lagrangian strain tensor $\boldsymbol{E}$. In that case, the resulting relations between $\boldsymbol{E}$ and the second Piola–Kirchhoff stress $\boldsymbol{S}$ would be exactly the same as the standard stress-strain equations of linear isotropic small-strain elasticity (D.24). This is acceptable if we have in mind applications in which the strains actually remain small and only large rotations need to be taken into account. However, if the strains can become large, it is reasonable to modify the model such that it would satisfy at least some basic physical requirements for the extreme cases of very large tensile and compressive deformations.

Extreme tension is characterized by one or more principal stretches λ_I growing without any bounds, and extreme compression is characterized by one or more principal stretches λ_I approaching zero. In the former case, the corresponding principal Green's Lagrangian strains $E_I = (\lambda_I^2 - 1)/2$ tend to infinity and the free energy (24.59) with ε_I replaced by E_I also tends to infinity. In the latter case, the principal strains $E_I = (\lambda_I^2 - 1)/2$ tend to -0.5, and the free energy evaluated from the quadratic expression (24.59) remains finite. The gradient of the free energy remains

finite, too, which means that the material can be squeezed into a zero volume by applying a finite pressure. To remedy this obvious paradox, it is necessary to use some non-polynomial terms in the free-energy expression (because all polynomials, not only quadratic ones, would remain bounded as E_I approaches -0.5). A good candidate is a logarithmic function of the stretch λ_I, which is unbounded in both extreme cases, $\lambda_I \to \infty$ and $\lambda_I \to 0$. Indeed, one suitable formulation, popular in applications to large-strain elasto plasticity (Weber and Anand, 1990; Eterovic and Bathe, 1990; Perić, Owen and Honnor, 1992; Simo, 1992), is based on the replacement of the small-strain tensor ε by the logarithmic strain (24.31). Derivation of the corresponding stress-strain law is left to the reader as an exercise (Problem 24.2). Here we consider another model, which is easier to handle, because the logarithm appears only at the Jacobian $J = \lambda_1\lambda_2\lambda_3$ and not at each principal stretch separately.

Example 24.2: Derive the specific form of hyperelastic constitutive equations corresponding to the free energy (Simo and Pister, 1984)

$$\rho_0 \psi = \tfrac{1}{4}\left(K - \tfrac{2}{3}G\right)(J^2 - 2\ln J - 1) + G\left(\boldsymbol{E} : \boldsymbol{I} - \ln J\right) \tag{24.60}$$

Solution: For simplicity, the free energy has been written in terms of the Green's Lagrangian strain tensor $\boldsymbol{E}$ and the Jacobian J, but it ought to be understood as a function of $\boldsymbol{E}$, with $J = \sqrt{\det \boldsymbol{C}} = \sqrt{\det(\boldsymbol{I} + 2\boldsymbol{E})}$. Since the expression for free energy depends only on the invariants of $\boldsymbol{E}$, the present material model is isotropic. To differentiate (24.60) with respect to $\boldsymbol{E}$, we will need the gradient $\partial J/\partial \boldsymbol{E}$, which can be evaluated as follows:

$$\frac{\partial J}{\partial \boldsymbol{E}} = 2\frac{\partial J}{\partial \boldsymbol{C}} = 2\frac{\partial \sqrt{\det \boldsymbol{C}}}{\partial \boldsymbol{C}} = 2\frac{1}{2\sqrt{\det \boldsymbol{C}}}\frac{\partial(\det \boldsymbol{C})}{\partial \boldsymbol{C}} = \frac{1}{J}(\det \boldsymbol{C})\, \boldsymbol{C}^{-T} = J\boldsymbol{C}^{-1} \tag{24.61}$$

Here we have exploited the relation

$$\frac{\partial(\det \boldsymbol{C})}{\partial \boldsymbol{C}} = (\det \boldsymbol{C})\boldsymbol{C}^{-T} \tag{24.62}$$

which is valid for general second-order tensors but is especially easy to verify for symmetric tensors, using the spectral decomposition and formula (D.61) for the derivative of an eigenvalue.

With (24.61) at hand, it is easy to derive the constitutive equation for the second Piola–Kirchhoff stress tensor,

$$\begin{aligned} \boldsymbol{S} = \rho_0 \frac{\partial \psi}{\partial \boldsymbol{E}} &= \frac{1}{4}\left(K - \frac{2}{3}G\right)\left(2J\frac{\partial J}{\partial \boldsymbol{E}} - \frac{2}{J}\frac{\partial J}{\partial \boldsymbol{E}}\right) + G\left(\boldsymbol{I} - \frac{1}{J}\frac{\partial J}{\partial \boldsymbol{E}}\right) \\ &= \tfrac{1}{2}\left(K - \tfrac{2}{3}G\right)(J^2 - 1)\boldsymbol{C}^{-1} + G\left(\boldsymbol{I} - \boldsymbol{C}^{-1}\right) \end{aligned} \tag{24.63}$$

Since $\boldsymbol{I}$ is a spherical tensor, the principal directions of the second Piola–Kirchhoff stress tensor $\boldsymbol{S}$ coincide with the principal directions of $\boldsymbol{C}^{-1}$, denoted as $\boldsymbol{N}_I$, $I = 1, 2, 3$, which are the same as the principal directions of $\boldsymbol{C}$ and of the right stretch tensor $\boldsymbol{U}$. This is a general property of isotropic hyperelastic materials. Using (24.54), the foregoing formula can be converted into the expression for Cauchy stress,

$$\boldsymbol{\sigma} = \frac{1}{J}\boldsymbol{F} \cdot \boldsymbol{S} \cdot \boldsymbol{F}^T = \frac{1}{2}\left(K - \frac{2}{3}G\right)\left(J - \frac{1}{J}\right)\boldsymbol{I} + \frac{G}{J}\left(\boldsymbol{F} \cdot \boldsymbol{F}^T - \boldsymbol{I}\right) \tag{24.64}$$

The principal directions of $\boldsymbol{\sigma}$ coincide with the principal directions of $\boldsymbol{F}\cdot\boldsymbol{F}^T = \boldsymbol{R}\cdot\boldsymbol{U}^2\cdot\boldsymbol{R}^T = \boldsymbol{R}\cdot\boldsymbol{C}\cdot\boldsymbol{R}^T$, which are equal to $\boldsymbol{R}\cdot\boldsymbol{N}_I$, i.e. to the principal directions of $\boldsymbol{U}$ rotated by $\boldsymbol{R}$. The principal Cauchy stresses

$$\begin{aligned}\sigma_I &= \frac{1}{2}\left(K-\frac{2}{3}G\right)\left(J-\frac{1}{J}\right)+\frac{G}{J}\left(\lambda_I^2-1\right) \\ &= \frac{1}{2}\left(K-\frac{2}{3}G\right)\left(\lambda_1\lambda_2\lambda_3-\frac{1}{\lambda_1\lambda_2\lambda_3}\right)+G\frac{\lambda_I^2-1}{\lambda_1\lambda_2\lambda_3}, \qquad I=1,2,3\end{aligned} \tag{24.65}$$

can easily be calculated from the principal stretches λ_I. If any of the principal stretches tends to zero (while the other principal stretches remain finite), the corresponding principal Cauchy stress tends to minus infinity. This means that the material keeps a positive volume at arbitrarily large pressures. □

24.2.4 Elasticity in Rate Form (Hypoelasticity)

Upon closer examination of (24.63) and (24.64), we realize that the second Piola–Kirchhoff stress $\boldsymbol{S}$ depends only on the right stretch tensor $\boldsymbol{U}$ (since $\boldsymbol{C}^{-1}=\boldsymbol{U}^{-2}$), while the Cauchy stress $\boldsymbol{\sigma}$ also depends on the rotation tensor $\boldsymbol{R}$. If, after some previous deformation, the material rotates for a while as a rigid body, the second Piola–Kirchhoff stress remains constant but the Cauchy stress does not. This is not surprising, because the Cauchy stress is the actual stress acting in the current placement and referred to the spatial axes. If the material rotates, the traction on a physical plane rotating with the material (when referred to that plane) remains constant, but its projection onto the coordinate axes 'fixed' in space and also the projection of the normal to the physical plane are variable, and so the Cauchy stress components must be variable as well. The principal Cauchy stresses given by (24.65) nevertheless only depend on the principal stretches, and thus remain constant during a rigid-body rotation, although the corresponding principal directions $\boldsymbol{R}\cdot\boldsymbol{N}_I$ rotate due to the change of $\boldsymbol{R}$.

These facts must be taken into account when writing the constitutive equations for Cauchy stress $\boldsymbol{\sigma}$ in the rate form. In analogy to the small-displacement theory, one could be tempted to write such equations in the form $\dot{\boldsymbol{\sigma}}=\boldsymbol{\mathcal{E}}:\boldsymbol{D}$, where $\dot{\boldsymbol{\sigma}}$ is the time derivative of Cauchy stress, $\boldsymbol{\mathcal{E}}$ is the elastic stiffness tensor[3] and $\boldsymbol{D}$ is the deformation-rate tensor, which is analogous to the small-strain rate $\dot{\boldsymbol{\varepsilon}}$ because it is computed as the symmetric part of the spatial velocity gradient.

During rigid-body rotations, however, $\boldsymbol{D}$ is zero but $\dot{\boldsymbol{\sigma}}$ is not. To be able to write the rate of stress as a function of the deformation-rate tensor, we need to define a so-called *objective stress rate*, which may be obtained by subtracting the effect of rigid-body rotation from the time derivative of stress.

The effect of rigid-body rotation can be figured out from the time derivative of the stress transformation formula (24.54),

$$\dot{\boldsymbol{\sigma}}=-\frac{\dot{J}}{J^2}\boldsymbol{F}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T+\frac{1}{J}\left(\dot{\boldsymbol{F}}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\dot{\boldsymbol{S}}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\boldsymbol{S}\cdot\dot{\boldsymbol{F}}^T\right) \tag{24.66}$$

[3] In other chapters of this book we denote the material stiffness tensor as $\boldsymbol{D}$, but in the large-strain theory this symbol is usually reserved for the symmetric part of the spatial velocity gradient, and so an alternative notation for stiffness is needed.

During a rigid-body rotation, we have $\dot{\boldsymbol{U}}=\boldsymbol{O}$, and so $\dot{J}=\mathrm{d}(\det\boldsymbol{U})/\mathrm{d}t=0$. Since the second Piola-Kirchhoff stress $\boldsymbol{S}$ only depends on $\boldsymbol{U}$, its time derivative must also vanish. So the corresponding rate of change of Cauchy stress that is entirely due to rigid-body rotation is

$$\begin{aligned}\dot{\boldsymbol{\sigma}}_R &= \frac{1}{J}\left(\dot{\boldsymbol{F}}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\boldsymbol{S}\cdot\dot{\boldsymbol{F}}^T\right)=\frac{1}{J}\left(\boldsymbol{L}\cdot\boldsymbol{F}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T\cdot\boldsymbol{L}^T\right) \\ &= \boldsymbol{L}\cdot\boldsymbol{\sigma}+\boldsymbol{\sigma}\cdot\boldsymbol{L}^T\end{aligned} \tag{24.67}$$

This motivates the definition of the *Oldroyd rate* of Cauchy stress (Oldroyd, 1950),

$$\mathring{\boldsymbol{\sigma}}\equiv\dot{\boldsymbol{\sigma}}-\boldsymbol{L}\cdot\boldsymbol{\sigma}-\boldsymbol{\sigma}\cdot\boldsymbol{L}^T \tag{24.68}$$

which is identically zero if the (previously deformed) material experiences only a rigid-body rotation. Such an objective stress rate could be used in the rate form of constitutive equations. However, for reasons that will become clear later, it is often more convenient to work with an objective rate of another stress tensor, τ, called the *Kirchhoff stress* and defined as

$$\tau=J\sigma=\boldsymbol{F}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T \tag{24.69}$$

Its Oldroyd rate is defined by a formula analogous to (24.68).

$$\mathring{\tau}\equiv\dot{\tau}-\boldsymbol{L}\cdot\tau-\tau\cdot\boldsymbol{L}^T \tag{24.70}$$

Substituting (24.69) and its time derivative into (24.70) and recalling that $\dot{\boldsymbol{F}}=\boldsymbol{L}\cdot\boldsymbol{F}$, we obtain a remarkable relation that sheds more light on the physical meaning of the Oldroyd rate:

$$\begin{aligned}\mathring{\tau} &= \dot{\boldsymbol{F}}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\dot{\boldsymbol{S}}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\boldsymbol{S}\cdot\dot{\boldsymbol{F}}^T-\boldsymbol{L}\cdot\boldsymbol{F}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T-\boldsymbol{F}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T\cdot\boldsymbol{L}^T \\ &= \boldsymbol{F}\cdot\dot{\boldsymbol{S}}\cdot\boldsymbol{F}^T=\boldsymbol{F}\cdot\frac{\partial}{\partial t}\left(\boldsymbol{F}^{-1}\cdot\tau\cdot\boldsymbol{F}^{-1}\right)\cdot\boldsymbol{F}^T\end{aligned} \tag{24.71}$$

The last expression means that, in order to evaluate the Oldroyd rate $\mathring{\tau}$, the Kirchhoff stress τ is first 'pulled back' to the initial placement, then differentiated with respect to time, and finally 'pushed forward' to the current placement. In mathematical language, this operation is called the *Lie derivative* (Marsden and Hughes, 1983).

Example 24.3: Derive the rate form of the constitutive equations from Example 24.2 in terms of the Oldroyd rate of Kirchhoff stress and the deformation-rate tensor.

Solution: Equation (24.64) from Example 24.2 can be rewritten in terms of Kirchhoff stress,

$$\boldsymbol{\tau}=J\boldsymbol{\sigma}=\tfrac{1}{2}\left(K-\tfrac{2}{3}G\right)\left(J^2-1\right)\boldsymbol{I}+G(\boldsymbol{F}\cdot\boldsymbol{F}^T-\boldsymbol{I}) \tag{24.72}$$

Differentiating with respect to time we obtain

$$\begin{aligned}\dot{\boldsymbol{\tau}} &= \tfrac{1}{2}\left(K-\tfrac{2}{3}G\right)2J\dot{J}\boldsymbol{I}+G(\dot{\boldsymbol{F}}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\dot{\boldsymbol{F}}^T) \\ &= \left(K-\tfrac{2}{3}G\right)J^2(\boldsymbol{D}:\boldsymbol{I})\boldsymbol{I}+G(\boldsymbol{L}\cdot\boldsymbol{F}\cdot\boldsymbol{F}^T+\boldsymbol{F}\cdot\boldsymbol{F}^T\cdot\boldsymbol{L}^T)\end{aligned} \tag{24.73}$$

Here we have exploited relation (24.45) and the formula for the rate of Jacobian,

$$\dot{J} = \frac{\mathrm{d}(\det \boldsymbol{F})}{\mathrm{d}t} = \frac{\partial(\det \boldsymbol{F})}{\partial \boldsymbol{F}} : \dot{\boldsymbol{F}} = (\det \boldsymbol{F})\, \boldsymbol{F}^{-T} : (\boldsymbol{L} \cdot \boldsymbol{F}) = J F^{-1}_{ji} L_{ik} F_{kj}$$
$$= J L_{ik} \delta_{ki} = J L_{ii} = J \boldsymbol{L} : \boldsymbol{I} = J \boldsymbol{D} : \boldsymbol{I} \tag{24.74}$$

Substituting (24.72) and (24.73) into the definition (24.69) of the Oldroyd rate of Kirchhoff stress, we get

$$\overset{\circ}{\boldsymbol{\tau}} = \left(K - \tfrac{2}{3}G\right) J^2 (\boldsymbol{D} : \boldsymbol{I}) \boldsymbol{I} + G(\boldsymbol{L} \cdot \boldsymbol{F} \cdot \boldsymbol{F}^T + \boldsymbol{F} \cdot \boldsymbol{F}^T \cdot \boldsymbol{L}^T)$$
$$- \tfrac{1}{2}\left(K - \tfrac{2}{3}G\right)\left(J^2 - 1\right)(\boldsymbol{L} + \boldsymbol{L}^T) - G(\boldsymbol{L} \cdot \boldsymbol{F} \cdot \boldsymbol{F}^T + \boldsymbol{F} \cdot \boldsymbol{F}^T \cdot \boldsymbol{L}^T - \boldsymbol{L} - \boldsymbol{L}^T)$$
$$= \left(K - \tfrac{2}{3}G\right)\left[J^2 (\boldsymbol{D} : \boldsymbol{I}) \boldsymbol{I} - (J^2 - 1)\boldsymbol{D}\right] + 2G\boldsymbol{D} \tag{24.75}$$

Clearly, the stress rate $\overset{\circ}{\boldsymbol{\tau}}$ uniquely depends on the deformation-rate tensor $\boldsymbol{D}$ (and on the current state, entering through the Jacobian J). It is independent of the spin tensor $\boldsymbol{W}$, i.e. of the skew-symmetric part of the spatial velocity gradient $\boldsymbol{L}$. We can even write (24.75) in the form

$$\overset{\circ}{\boldsymbol{\tau}} = \boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\tau}}} : \boldsymbol{D} \tag{24.76}$$

where

$$\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\tau}}} = \left(K - \tfrac{2}{3}G\right) J^2 (\boldsymbol{I} \otimes \boldsymbol{I}) + \left[2G - (J^2 - 1)\left(K - \tfrac{2}{3}G\right)\right] \boldsymbol{I}_s \tag{24.77}$$

is the *Eulerian tangent material stiffness tensor* (the adjective Eulerian means that $\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\tau}}}$ relates two rates used in the spatial description – the deformation rate $\boldsymbol{D}$ and the objective stress rate $\overset{\circ}{\boldsymbol{\tau}}$). Recall that $\boldsymbol{I}_s$ is the symmetric fourth-order unit tensor. The stiffness tensor $\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\tau}}}$ depends on the current state (through J), and in the initial configuration it reduces to the stiffness tensor (D.25) known from small-strain elasticity,

$$\boldsymbol{\mathcal{E}}_0 = \left(K - \tfrac{2}{3}G\right)(\boldsymbol{I} \otimes \boldsymbol{I}) + 2G\boldsymbol{I}_s \tag{24.78}$$

□

It is obvious that the stiffness tensor (24.77) is only valid in conjunction with the Oldroyd rate of Kirchhoff stress, $\overset{\circ}{\boldsymbol{\tau}}$. Let us explore how the result changes if $\overset{\circ}{\boldsymbol{\tau}}$ is replaced by the Oldroyd rate of Cauchy stress, $\overset{\circ}{\boldsymbol{\sigma}}$. Instead of differentiating (24.64), we will derive the general relation between the rates $\overset{\circ}{\boldsymbol{\tau}}$ and $\overset{\circ}{\boldsymbol{\sigma}}$. Substituting $\boldsymbol{\tau} = J\boldsymbol{\sigma}$ into (24.70) and using (24.74) and (24.68), we obtain

$$\overset{\circ}{\boldsymbol{\tau}} = \dot{\boldsymbol{\tau}} - \boldsymbol{L} \cdot \boldsymbol{\tau} - \boldsymbol{\tau} \cdot \boldsymbol{L}^T = \dot{J}\boldsymbol{\sigma} + J\dot{\boldsymbol{\sigma}} - J\boldsymbol{L} \cdot \boldsymbol{\sigma} - J\boldsymbol{\sigma} \cdot \boldsymbol{L}^T = J(\boldsymbol{D} : \boldsymbol{I})\boldsymbol{\sigma} + J\overset{\circ}{\boldsymbol{\sigma}} \tag{24.79}$$

Consequently,

$$\overset{\circ}{\boldsymbol{\sigma}} = J^{-1}\overset{\circ}{\boldsymbol{\tau}} - \boldsymbol{\sigma}(\boldsymbol{I} : \boldsymbol{D}) \tag{24.80}$$

and the stiffness tensor that appears in the rate equation $\overset{\circ}{\boldsymbol{\sigma}} = \boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\sigma}}} : \boldsymbol{D}$ is in general given by

$$\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\sigma}}} = J^{-1}\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\tau}}} - \boldsymbol{\sigma} \otimes \boldsymbol{I} \tag{24.81}$$

The surprising property is that, due to the presence of the term $\boldsymbol{\sigma} \otimes \boldsymbol{I}$, $\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\sigma}}}$ is in general nonsymmetric. Note that the stiffness tensor $\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\tau}}}$ given by (24.77) is symmetric. This is not by chance.

The lack of major symmetry of the stiffness tensor $\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\sigma}}}$ can be proven if the hypoelastic formulation is related to the hyperelastic law (24.51), rewritten in the rate form

$$\dot{\boldsymbol{S}} = \rho_0 \frac{\partial^2 \psi}{\partial \boldsymbol{E}^2} : \dot{\boldsymbol{E}} = \boldsymbol{\mathcal{E}} : \dot{\boldsymbol{E}} \tag{24.82}$$

The tangential stiffness tensor $\boldsymbol{\mathcal{E}}$, linking the rate of the second Piola–Kirchhoff stress to the rate of Green's Lagrangian strain, is always symmetric because of the symmetry of the second derivatives $\partial^2\psi/\partial E_{ij}E_{kl} = \partial^2\psi/\partial E_{kl}E_{ij}$. To transform (24.82) into the spatial description that exploits the deformation-rate tensor $\boldsymbol{D}$ and some objective rate of Cauchy or Kirchhoff stress, we substitute the relation $\dot{\boldsymbol{E}} = \boldsymbol{F}^T \cdot \boldsymbol{D} \cdot \boldsymbol{F}$, which was derived in Example 24.1. In the subscript notation we obtain

$$\dot{S}_{ij} = \mathcal{E}_{ijkl}\dot{E}_{kl} = \mathcal{E}_{ijkl}F_{rk}D_{rs}F_{sl} = (\mathcal{E}_{ijkl}F_{rk}F_{sl})D_{rs} \tag{24.83}$$

To preserve symmetry of the stiffness tensor, it is now natural to multiply (24.83) by $F_{pi}F_{qj}$. The resulting relation, $F_{pi}F_{qj}\dot{S}_{ij} = (F_{pi}F_{qj}\mathcal{E}_{ijkl}F_{rk}F_{sl})D_{rs}$, can be written in the compact notation as

$$\boldsymbol{F} \cdot \dot{\boldsymbol{S}} \cdot \boldsymbol{F}^T = \tilde{\boldsymbol{\mathcal{E}}} : \boldsymbol{D} \tag{24.84}$$

where $\tilde{\boldsymbol{\mathcal{E}}}$ is a symmetric Eulerian stiffness tensor with components

$$\tilde{\mathcal{E}}_{pqrs} = F_{pi}F_{qj}F_{rk}F_{sl}\mathcal{E}_{ijkl} \tag{24.85}$$

The expression on the left-hand side of (24.84) is recognized as the Oldroyd rate of Kirchhoff stress, see (24.71), from which $\overset{\circ}{\boldsymbol{\tau}} = \dot{\boldsymbol{\tau}} - \boldsymbol{L} \cdot \boldsymbol{\tau} - \boldsymbol{\tau} \cdot \boldsymbol{L}^T = \boldsymbol{F} \cdot \dot{\boldsymbol{S}} \cdot \boldsymbol{F}^T$. So the stiffness tensor $\tilde{\boldsymbol{\mathcal{E}}}$ from (24.84) must be identical to the tensor $\boldsymbol{\mathcal{E}}_{\overset{\circ}{\boldsymbol{\tau}}}$ from (24.76), and equation (24.85) relates its components to the stiffness tensor $\boldsymbol{\mathcal{E}} = \rho_0\partial^2\psi/\partial\boldsymbol{E}^2$, derived directly from the elastic potential. The transformation (24.83) preserves the major symmetry.

The Oldroyd rate of Kirchhoff stress divided by the Jacobian is the *Truesdell stress rate* (Truesdell, 1955b). It is usually understood as a rate of Cauchy stress, because it can be expressed as

$$\overset{\circ}{\boldsymbol{\sigma}} = J^{-1}\boldsymbol{F} \cdot \dot{\boldsymbol{S}} \cdot \boldsymbol{F}^T = \dot{\boldsymbol{\sigma}}^T - \boldsymbol{L} \cdot \boldsymbol{\sigma} - \boldsymbol{\sigma} \cdot \boldsymbol{L}^T + \boldsymbol{\sigma}(\boldsymbol{L} : \boldsymbol{I}) \tag{24.86}$$

The definition of the Oldroyd rate was motivated by the fact that during a rigid-body motion the time derivative of Cauchy stress is $\boldsymbol{L}\cdot\boldsymbol{\sigma}+\boldsymbol{\sigma}\cdot\boldsymbol{L}^T$; see (24.67). During a rigid-body motion, the deformation-rate tensor $\boldsymbol{D}$ vanishes and the velocity gradient $\boldsymbol{L}$ is equal to the spin tensor $\boldsymbol{W}$. So the rate of Cauchy stress produced by the rigid-body motion can equivalently be written as $\boldsymbol{W} \cdot \boldsymbol{\sigma} + \boldsymbol{\sigma} \cdot \boldsymbol{W}^T$ or as $\boldsymbol{W} \cdot \boldsymbol{\sigma} - \boldsymbol{\sigma} \cdot \boldsymbol{W}$ (because $\boldsymbol{W}$ is skew-symmetric). This leads to the *Jaumann–Zaremba stress rate* (Jaumann, 1911), $\dot{\boldsymbol{\sigma}} - \boldsymbol{W} \cdot \boldsymbol{\sigma} + \boldsymbol{\sigma} \cdot \boldsymbol{W}$, which is objective, but in the general case (when $\boldsymbol{D} \neq \boldsymbol{O}$) differs from the Oldroyd rate. In fact, it is possible to define a number of other objective stress rates that all vanish during rigid-body rotations but differ during a general deformation process.

A summary of the most frequently used rates is given in Table 24.1. They are written in terms of Kirchhoff stress, but can be applied also to the Cauchy stress. However, the rates of Cauchy stress are inappropriate for stability and bifurcation investigations because they lead to nonsymmetric stiffness tensors even if the rate equations are derived from an elastic potential, which is also related to the fact that

Table 24.1 Objective stress rates and parameters m of the Doyle–Ericksen strain tensor associated by incremental work (Bažant, 1971)

Rate of Kirchhoff stress	Formula	m
Oldroyd (Lie derivative)	$\dot{\boldsymbol{\tau}} - \boldsymbol{L}\cdot\boldsymbol{\tau} - \boldsymbol{\tau}\cdot\boldsymbol{L}^T$	2
Biot	$\dot{\boldsymbol{\tau}} - \boldsymbol{L}\cdot\boldsymbol{\tau} - \boldsymbol{\tau}\cdot\boldsymbol{L}^T + (\boldsymbol{\tau}\cdot\boldsymbol{D})_{\text{sym}}$	1
Jaumann–Zaremba (corotational)	$\dot{\boldsymbol{\tau}} - \boldsymbol{W}\cdot\boldsymbol{\tau} - \boldsymbol{\tau}\cdot\boldsymbol{W}^T$	0
Cotter–Rivlin (convected)	$\dot{\boldsymbol{\tau}} + \boldsymbol{L}^T\cdot\boldsymbol{\tau} + \boldsymbol{\tau}\cdot\boldsymbol{L}$	−2
Green–Naghdi–McInnis	$\dot{\boldsymbol{\tau}} - \dot{\boldsymbol{R}}\cdot\boldsymbol{R}^T\cdot\boldsymbol{\tau} + \boldsymbol{\tau}\cdot\dot{\boldsymbol{R}}\cdot\boldsymbol{R}^T$	–

the product $\mathring{\boldsymbol{\sigma}} : \boldsymbol{D}$ is not the correct expression for the second-order work. Recognizing this fact, Bažant (1971a) proposed a class of rates

$$\hat{\boldsymbol{\sigma}}^{(m)} = \dot{\boldsymbol{\sigma}} - \boldsymbol{L}\cdot\boldsymbol{\sigma} - \boldsymbol{\sigma}\cdot\boldsymbol{L}^T + \boldsymbol{\sigma}(\boldsymbol{L}:\boldsymbol{I}) + \left(1 - \frac{m}{2}\right)(\boldsymbol{\sigma}\cdot\boldsymbol{D} + \boldsymbol{D}\cdot\boldsymbol{\sigma}) \tag{24.87}$$

where m is any real constant. He showed that $\hat{\boldsymbol{\sigma}}^{(m)}$ is associated with an approximation of the Doyle–Ericksen finite-strain tensor $\boldsymbol{E}^{(m)}$.

Note that $\hat{\boldsymbol{\sigma}}^{(2)}$ is in fact the Truesdell stress rate. For special values of m, the rates $\hat{\boldsymbol{\sigma}}^{(m)}$ multiplied by the Jacobian correspond to the objective rates of *Kirchhoff* stress defined in Table 24.1: The Oldroyd rate is equal to $J\hat{\boldsymbol{\sigma}}^{(2)}$, the Jaumann rate to $J\hat{\boldsymbol{\sigma}}^{(0)}$, and the Cotter–Rivlin rate (Cotter and Rivlin, 1955) to $J\hat{\boldsymbol{\sigma}}^{(-2)}$. The rate $J\hat{\boldsymbol{\sigma}}^{(1)}$ is the Biot rate (Biot, 1965).

The arbitrariness of the form of the constitutive relation due to the choice of the objective stress rate does not imply a physical arbitrariness. It has been shown (Bažant, 1971a; Bažant and Cedolin, 1991, Section 11.4) that if the tangential stiffness tensor associated with Truesdell's stress rate $\hat{\boldsymbol{\sigma}}^{(2)}$ is $\mathcal{E}^{(2)}_{ijkl}$, then the tangential stiffness tensor for any other rate associated with the strain tensor $\boldsymbol{E}^{(m)}$ is

$$\mathcal{E}^{(m)}_{ijkl} = \mathcal{E}_{ijkl} + \frac{2-m}{4}(\sigma_{ik}\delta_{jl} + \sigma_{jk}\delta_{il} + \sigma_{il}\delta_{jk} + \sigma_{jl}\delta_{ik}) \tag{24.88}$$

Obviously, there is no reason to expect $\mathcal{E}_{ijkl}$, or any other $\mathcal{E}^{(m)}_{ijkl}$, to be independent of $\boldsymbol{\sigma}$. The computational literature is replete with studies ignoring the consequences of the arbitrariness in the choice of m; the tangential moduli corresponding to some chosen objective stress rate are often assumed constant without any experimental justification.

24.3 LARGE-STRAIN PLASTICITY

The theory of plasticity at finite strains is a wide and complex subject. There are a number of competing theoretical formulations and numerical algorithms, and it seems that no universally accepted general theory has emerged so far. Nevertheless, one can distinguish two main classes of elastoplastic models for materials subjected to large strains. One of them postulates the basic equations directly in the rate form while the other starts from thermodynamic potentials.

24.3.1 Rate Formulation

The first approach is conceived as a direct generalization of the incremental equations of phenomenological small-strain elastoplasticity. Such formulations have been developed by, e.g. McMeeking and Rice (1975), Nagtegaal and Jong (1981), Key and Krieg (1982), Nagtegaal (1982) and Pinsky, Ortiz and Pister (1983). To ensure objectivity in the sense of invariance with respect to rigid-body rotations of the coordinate frame, the stress rate entering the constitutive equations cannot be interpreted as the usual time derivative, but as one of the objective stress rates defined in Section 24.2.4. The role of the strain rate is played by the deformation-rate tensor $\boldsymbol{D}$, which is additively split into the elastic part, $\boldsymbol{D}_e$, and the plastic part, $\boldsymbol{D}_p$. The elastic part of the constitutive law is written in the rate form

$$\mathring{\boldsymbol{\tau}} = \boldsymbol{\mathcal{E}} : (\boldsymbol{D} - \boldsymbol{D}_p) \tag{24.89}$$

where $\mathring{\boldsymbol{\tau}}$ is an objective rate of the Kirchhoff stress tensor and $\boldsymbol{\mathcal{E}}$ is a suitable Eulerian elastic stiffness tensor. Such an elastic law written in terms of spatial rates can describe only materials with initial elastic isotropy. The advantage of using the Kirchhoff stress rather than the Cauchy stress is that, upon discretization, such a formulation results into a symmetric stiffness matrix. As we have seen in Section 24.2.4, the rates of Cauchy stress require a nonsymmetric tensor $\boldsymbol{\mathcal{E}}$ whenever a hyperelastic potential exists.

In the simplest case of perfect elastoplasticity, the elastic domain is described by the yield function $f(\boldsymbol{\tau})$, where $\boldsymbol{\tau}$ is the Kirchhoff stress, and the evolution of plastic strain can be prescribed e.g. by the associated flow rule

$$\boldsymbol{D}_p = \dot{\lambda}\frac{\partial f}{\partial \boldsymbol{\tau}} \tag{24.90}$$

combined with the usual loading-unloading conditions,

$$f \le 0, \qquad \dot{\lambda} \ge 0, \qquad \dot{\lambda} f = 0 \tag{24.91}$$

In the present section, λ denotes again the plastic multiplier and not the stretch.

Extensions of the rate formulation to isotropic and kinematic hardening are also possible. Due to its formal similarity with the small-strain theory, this framework is widely popular and has been implemented in a number of commercial finite element packages. However, it has certain drawbacks stemming from the incremental form of the stress-strain law (24.89).

In contrast to the small-strain limit, the choice of the objective rate $\mathring{\boldsymbol{\tau}}$ is not unique. In many practical applications, $\boldsymbol{\mathcal{E}}$ is taken as a constant isotropic fourth-order tensor. However, if $\boldsymbol{\mathcal{E}}$ is taken as a constant for one specific choice of the stress rate, after transformation into another stress rate the corresponding stiffness becomes variable; see equation (24.88). There is no reason to assume that there exists one special rate for which $\boldsymbol{\mathcal{E}}$ is constant (independent of the current state of the material). Moreover, it can be shown that hypoelastic formulations with a constant Eulerian stiffness tensor are not derivable from a hyperelastic potential (Simo and Pister, 1984). Such formulations do not satisfy thermodynamic restrictions, the stress response is path-dependent even inside the elastic domain, and spurious energy dissipation or production ('negative' dissipation) may occur.

Some of the hypoelastic formulations lead to a pathological behavior, e.g. to an oscillating stress response under monotonic simple shear for a model with kinematic

hardening that uses the Jaumann rate (Lehmann, 1972; Nagtegaal and Jong, 1981; Nagtegaal, 1982). This oscillatory behavior disappears when the Jaumann rate is replaced by the Green–Naghdi rate (Dienes, 1979; Dafalias, 1983) or by the Truesdell rate (Haupt and Tsamakis, 1986).

Numerical implementation of this class of models requires special incrementally objective algorithms (Hughes and Winget, 1980; Rubinstein and Atluri, 1983; Pinsky, Ortiz and Pister, 1983; Simo and Hughes, 1998).

24.3.2 Formulation Based on Thermodynamic Potentials

In the present text, meant as a simple introduction, we will pay more attention to models developed within a thermodynamic framework, which naturally leads to the hyperelastic form of the elastic part of the stress-strain law. The starting point is the choice of a suitable free-energy potential ψ that depends on a strain measure and on a set of internal variables representing the plastic deformation and hardening effects.

Most models of this type introduce the notion of an *unstressed intermediate placement*, which is a fictitious state of the material after elastic unloading of each infinitesimal material volume, without taking into account the compatibility conditions. The intermediate placement is physically attainable only if the current state of the body is homogeneous (uniform in space) and if the zero stress state still belongs to the elastic domain (which might not be true after strong kinematic hardening). In a general case of nonhomogeneous deformation we only pay attention to a local neighborhood of each material point, in which the current deformation is characterized by the deformation gradient $\boldsymbol{F}$. After local elastic unloading, the permanent deformation at that point can be characterized by a tensor $\boldsymbol{F}_p$, but the entire field $\boldsymbol{F}_p(\boldsymbol{X})$ is in general not the gradient of any globally defined displacement field. The transformation between the intermediate placement and the current one is locally defined by a tensor $\boldsymbol{F}_e$, and the superposition of the 'plastic' transformation and the 'elastic' one must give the total transformation from the initial to the current placement; see Figure 24.3. Thus the idea of an intermediate placement directly leads to a *multiplicative decomposition of the deformation gradient*,

$$\boldsymbol{F} = \boldsymbol{F}_e \cdot \boldsymbol{F}_p \tag{24.92}$$

proposed by Flory (1961). Mandel (1964b, 1972), Lee and Liu (1967) and Lee (1969).

The physical motivation for the above decomposition is provided by the idealized description of elastoplastic deformation in crystals. The total deformation is decomposed into (1) *plastic slip* along one or more densely packed crystallographic planes, and (2) *elastic distortion* of the crystal lattice. It is assumed that, after removal

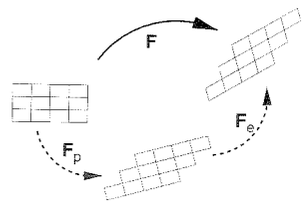


Figure 24.3 Multiplicative decomposition of the deformation gradient

of the applied loads, the geometric structure of the lattice is the same as in the initial configuration, but some layers of atoms have been shifted by an integer multiple of the interatomic distance which, from the macroscopic point of view, is perceived as permanent deformation (in reality, the relative slip between atomic layers is not a spatially uniform process – it results from the motion of *dislocations*, i.e. line defects in the perfect geometric structure of the lattice).

Accepting this schematic idea, we can expect the elastic properties of the material to be unaffected by the plastic flow. This assumption of *elastic isomorphy* (Noll, 1967, 1972; Bertram, 1998) is widely used, even though, for polycrystals, it does not have an unlimited validity. Polycrystalline materials are assemblies of grains, each of which has a regular lattice structure, and the orientation of the crystallographic axes varies from grain to grain. Depending on the fabrication process, the random orientation of individual grains may or may not be uniformly distributed in space. For example, rolled metal sheets usually have a clearly pronounced anisotropic structure. For the sake of simplicity, we restrict ourselves to the case of a uniform probability density of the grain orientation, which results into macroscopically isotropic elastic properties. The plastic deformation process may lead to texture development and induce elastic anisotropy, but the assumption of isotropic elastic behavior seems to be reasonable at early stages. Of course, this assumption does not preclude modeling of anisotropic plastic effects, such as kinematic hardening.

For a material with a given internal structure (even an anisotropic one), the specification of the total deformation gradient $\boldsymbol{F}$ and the plastic 'deformation gradient' $\boldsymbol{F}_p$ fully defines the current and the intermediate (local) placements. This means that the elastic energy stored in the material must be a unique function of these tensors. Moreover, rigid-body rotations of the current or intermediate placements cannot change the elastic energy. In terms of the polar decompositions $\boldsymbol{F} = \boldsymbol{R} \cdot \boldsymbol{U}$ and $\boldsymbol{F}_p = \boldsymbol{R}_p \cdot \boldsymbol{U}_p$ (where $\boldsymbol{R}$ and $\boldsymbol{R}_p$ are orthogonal tensors and $\boldsymbol{U}$ and $\boldsymbol{U}_p$ are symmetric tensors), this means that the elastic energy must depend only on $\boldsymbol{U}$ and $\boldsymbol{U}_p$ or, equivalently, that it must be a unique function of $\boldsymbol{C} = \boldsymbol{F}^T \cdot \boldsymbol{F} = \boldsymbol{U}^2$ and $\boldsymbol{C}_p = \boldsymbol{F}_p^T \cdot \boldsymbol{F}_p = \boldsymbol{U}_p^2$.

Limiting ourselves to isotropic hardening, we may write the free energy (per unit mass) as

$$\bar{\psi}(\boldsymbol{C}, \boldsymbol{C}_p, \kappa) = \bar{\psi}_e(\boldsymbol{C}, \boldsymbol{C}_p) + \psi_p(\kappa) \tag{24.93}$$

where κ is a scalar hardening variable, $\bar{\psi}_e$ is the elastically stored energy, and ψ_p is the plastic part of the free energy, i.e. the energy blocked in dislocations (the bars in $\bar{\psi}$ and $\bar{\psi}_e$ are used only because we want to reserve the symbols ψ and ψ_e for functions of other arguments that will be introduced later).

More specifically, for *materials that are and remain elastically isotropic*, the elastic energy can be expressed as an isotropic function of the elastic stretch tensor $\boldsymbol{U}_e$ defined by the polar decomposition $\boldsymbol{F}_e = \boldsymbol{R}_e \cdot \boldsymbol{U}_e$ or, equivalently, as an isotropic function ψ_e of $\boldsymbol{C}_e = \boldsymbol{F}_e^T \cdot \boldsymbol{F}_e = \boldsymbol{U}_e^2$. Since $\boldsymbol{F}_e = \boldsymbol{F} \cdot \boldsymbol{F}_p^{-1}$, we have

$$\boldsymbol{C}_e = \boldsymbol{F}_e^T \cdot \boldsymbol{F}_e = \boldsymbol{F}_p^{-T} \cdot \boldsymbol{F}^T \cdot \boldsymbol{F} \cdot \boldsymbol{F}_p^{-1} = \boldsymbol{F}_p^{-T} \cdot \boldsymbol{C} \cdot \boldsymbol{F}_p^{-1} \tag{24.94}$$

and we can write the free energy as

$$\psi(\boldsymbol{C}, \boldsymbol{F}_p, \kappa) = \psi_e(\boldsymbol{F}_p^{-T} \cdot \boldsymbol{C} \cdot \boldsymbol{F}_p^{-1}) + \psi_p(\kappa) \tag{24.95}$$

Equation (24.95) means that we have chosen $\boldsymbol{F}_p$ as the internal variable characterizing the plastic deformation. In the present context (elastic isotropy) we could replace it by

$\boldsymbol{U}_p$, because $\boldsymbol{U}_p^{-T}\cdot\boldsymbol{C}\cdot\boldsymbol{U}_p^{-1} = \boldsymbol{R}_p^T\cdot\boldsymbol{C}_e\cdot\boldsymbol{R}_p$ is a tensor with the same principal invariants as $\boldsymbol{C}_e$ ($\boldsymbol{R}_p$ is a pure rotation matrix) and the isotropic function ψ_e must have the same value for both tensors. Nevertheless, the choice of $\boldsymbol{F}_p$ instead of $\boldsymbol{U}_p$ simplifies certain steps and is potentially useful for generalizations to elastically anisotropic materials.

Recall that the stress power per unit initial volume of the material is given by

$$\boldsymbol{S}:\dot{\boldsymbol{E}} = \tfrac{1}{2}\boldsymbol{S}:\dot{\boldsymbol{C}} \tag{24.96}$$

where $\boldsymbol{S}$ is the second Piola–Kirchhoff stress, $\boldsymbol{E}$ is the Green's Lagrangian strain, and $\boldsymbol{C}$ is the Green deformation tensor. For inelastic materials, the dissipation, i.e. the difference between the stress power and the rate of change of free energy, does not vanish. Refering all densities to a unit initial volume, we can write

$$\mathcal{D} = \boldsymbol{S}:\dot{\boldsymbol{E}} - \rho_0\dot{\psi} = \frac{1}{2}\boldsymbol{S}:\dot{\boldsymbol{C}} - \rho_0\frac{\partial\psi}{\partial\boldsymbol{C}}:\dot{\boldsymbol{C}} - \rho_0\frac{\partial\psi}{\partial\boldsymbol{F}_p}:\dot{\boldsymbol{F}}_p - \rho_0\frac{\partial\psi}{\partial\kappa}\dot{\kappa} \tag{24.97}$$

In the absence of dissipative stress, the standard procedure (analogous to the small-strain case, treated in Chapter 23) leads to the constitutive law

$$\boldsymbol{S} = 2\rho_0\frac{\partial\psi}{\partial\boldsymbol{C}} = 2\rho_0\frac{\partial\psi_e}{\partial\boldsymbol{C}_e}:\frac{\partial\boldsymbol{C}_e}{\partial\boldsymbol{C}} \tag{24.98}$$

Denoting

$$\boldsymbol{S}_e = 2\rho_0\frac{\partial\psi_e}{\partial\boldsymbol{C}_e} \tag{24.99}$$

and evaluating

$$\frac{\partial C_{e,kl}}{\partial C_{ij}} = \frac{\partial}{\partial C_{ij}}\left(F^{-1}_{p,mk}C_{mn}F^{-1}_{p,nl}\right) = F^{-1}_{p,mk}\delta_{im}\delta_{jn}F^{-1}_{p,nl} = F^{-1}_{p,ik}F^{-1}_{p,jl} \tag{24.100}$$

we can rewrite (24.98) as

$$S_{ij} = 2\rho_0\frac{\partial\psi}{\partial C_{ij}} \qquad S_{e,kl}\frac{\partial C_{e,kl}}{\partial C_{ij}} = F^{-1}_{p,ik}S_{e,kl}F^{-1}_{p,jl} \tag{24.101}$$

or, in the compact notation, as

$$\boldsymbol{S} = \boldsymbol{F}_p^{-1}\cdot\boldsymbol{S}_e\cdot\boldsymbol{F}_p^{-T} \tag{24.102}$$

Substituting this into the expression for Cauchy stress (24.54),

$$\boldsymbol{\sigma} = \frac{1}{J}\boldsymbol{F}\cdot\boldsymbol{S}\cdot\boldsymbol{F}^T = \frac{1}{J}\boldsymbol{F}_e\cdot\boldsymbol{F}_p\cdot\boldsymbol{F}_p^{-1}\cdot\boldsymbol{S}_e\cdot\boldsymbol{F}_p^{-T}\cdot\boldsymbol{F}_p^T\cdot\boldsymbol{F}_e^T = \frac{1}{J}\boldsymbol{F}_e\cdot\boldsymbol{S}_e\cdot\boldsymbol{F}_e^T \tag{24.103}$$

we realize that the symmetric tensor $\boldsymbol{S}_e$ obtained by differentiation of the elastic potential ψ_e with respect to $\boldsymbol{C}_e$ has the meaning of the second Piola–Kirchhoff stress with the intermediate placement taken as the reference. Equation (24.99) is the hyperelastic stress-strain law linking the stress tensor $\boldsymbol{S}_e$ to the Green tensor of the elastic deformation, $\boldsymbol{C}_e$. Combining that with the stress transformation law (24.102) and the 'strain transformation law' (24.94), we can link the second Piola–Kirchhoff tensor $\boldsymbol{S}$ to the total strain represented by $\boldsymbol{C}$ and the plastic deformation represented by $\boldsymbol{F}_p$.

In addition to the elastic law we need to postulate the evolution equations for the internal variables $\boldsymbol{F}_p$ and κ. Adopting the framework of standard dissipative materials (which is generally acceptable for metals but may be inappropriate for other materials, especially those with internal friction), we derive the evolution laws from the dual dissipation potential $\phi^*(\boldsymbol{\beta}_D)$, defined as the indicator function of the set of plastically admissible states $\bar{\mathcal{E}} = \{\boldsymbol{\beta}_D \mid f(\boldsymbol{\beta}_D) \le 0\}$. Here, $\boldsymbol{\beta}_D = (\boldsymbol{T}_D, q_D)$ are the dissipative thermodynamic forces conjugate to the internal variables $\boldsymbol{\alpha} = (\boldsymbol{F}_p, \kappa)$. In small-strain plasticity, the dissipative force conjugate to the plastic strain turns out to be equal to the stress, and the formulation of the yield conditions in terms of stress is very natural. Here we can expect the dissipative force $\boldsymbol{T}_D$ conjugate to $\boldsymbol{F}_p$ to be some measure of stress, but most likely not the Cauchy stress $\boldsymbol{\sigma}$. We could formally write the yield condition as

$$f(\boldsymbol{T}_D, q_D) = 0 \tag{24.104}$$

but it appears more natural to postulate a 'physical' yield condition in terms of the Cauchy stress, for instance in the form

$$\bar{f}(\boldsymbol{\sigma}, q_D) \equiv \bar{F}(\boldsymbol{\sigma}) - \sigma_0 + q_D = 0 \tag{24.105}$$

and then transform it into (24.104).

The physical meaning of $\boldsymbol{T}_D$ can be identified if we look at its quasiconservative counterpart,

$$\boldsymbol{T}_Q = \rho_0\frac{\partial\psi}{\partial\boldsymbol{F}_p} = \rho_0\frac{\partial\psi_e}{\partial\boldsymbol{C}_e}:\frac{\partial(\boldsymbol{F}_p^{-T}\cdot\boldsymbol{C}\cdot\boldsymbol{F}_p^{-1})}{\partial\boldsymbol{F}_p} = -\boldsymbol{C}_e\cdot\boldsymbol{S}_e\cdot\boldsymbol{F}_p^{-T} \tag{24.106}$$

A detailed derivation of this identity is left to the reader as an exercise (see Problem 24.6). Since the dissipation (24.97) must reduce to

$$\mathcal{D} = \boldsymbol{T}_D:\dot{\boldsymbol{F}}_p + q_D\dot{\kappa} \tag{24.107}$$

for any $\dot{\boldsymbol{F}}_p$ and $\dot{\kappa}$, we have, in analogy to (23.128)–(23.129),

$$\boldsymbol{T}_D = -\boldsymbol{T}_Q = -\rho_0\frac{\partial\psi}{\partial\boldsymbol{F}_p} = \boldsymbol{C}_e\cdot\boldsymbol{S}_e\cdot\boldsymbol{F}_p^{-T} \tag{24.108}$$

$$q_D = -q_Q = -\rho_0\frac{\partial\psi}{\partial\kappa} = -\rho_0\psi'_p \tag{24.109}$$

The tensor

$$\boldsymbol{M} = \boldsymbol{C}_e\cdot\boldsymbol{S}_e \tag{24.110}$$

is called the *Mandel stress* (Mandel, 1972, 1974). It is closely related to the Kirchhoff stress tensor $\boldsymbol{\tau} = J\boldsymbol{\sigma}$. Recall that $\boldsymbol{S}_e = 2\rho_0\,\partial\psi_e/\partial\boldsymbol{C}_e$. Since ψ_e is an isotropic function, the symmetric tensors $\boldsymbol{C}_e = \boldsymbol{U}_e^2$ and $\partial\psi_e/\partial\boldsymbol{C}_e$ must have the same principal directions $\boldsymbol{N}_{eI}$ ($I = 1, 2, 3$). Consequently, $\boldsymbol{C}_e\cdot\boldsymbol{S}_e = \boldsymbol{U}_e^2\cdot\boldsymbol{S}_e = \boldsymbol{U}_e\cdot\boldsymbol{S}_e\cdot\boldsymbol{U}_e$, and substituting the inverse form of (24.103) we obtain

$$\begin{aligned}\boldsymbol{M} &= \boldsymbol{C}_e\cdot\boldsymbol{S}_e = \boldsymbol{U}_e\cdot\boldsymbol{S}_e\cdot\boldsymbol{U}_e = \boldsymbol{U}_e\cdot(\boldsymbol{F}_e^{-1}\cdot J\boldsymbol{\sigma}\cdot\boldsymbol{F}_e^{-T})\cdot\boldsymbol{U}_e\\ &= \boldsymbol{U}_e\cdot\boldsymbol{U}_e^{-1}\cdot\boldsymbol{R}_e^{-1}\cdot\boldsymbol{\tau}\cdot\boldsymbol{R}_e^{-T}\cdot\boldsymbol{U}_e^{-1}\cdot\boldsymbol{U}_e = \boldsymbol{R}_e^T\cdot\boldsymbol{\tau}\cdot\boldsymbol{R}_e\end{aligned} \tag{24.111}$$

So, in this particular case, the Mandel stress tensor $\boldsymbol{M}$ is equal to the Kirchhoff stress tensor $\boldsymbol{\tau}$ rotated back to the intermediate placement.

A rotation of the coordinate system does not change the invariants, and so any isotropic yield condition formulated in terms of the invariants of Kirchhoff stress can be directly rewritten in terms of the invariants of Mandel stress without modifying the elastic domain. Our goal was to formulate a physically reasonable yield condition in terms of the dissipative thermodynamic force $\boldsymbol{T}_D$. According to (24.108) and (24.110), we have $\boldsymbol{M} = \boldsymbol{T}_D \cdot \boldsymbol{F}_p^T$, and since $\boldsymbol{F}_p$ is an internal variable, it can enter the yield condition as a parameter. Recall that the dual dissipation potential ϕ^* is primarily a function of $\boldsymbol{\beta}_D$, but it can also depend on the state variables.

In metals, the volumetric changes usually remain small, because the elastic volume change can only become large under extremely large pressures and the plastic volume change vanishes due to the isochoric character of the plastic flow. The Jacobian $J = \det \boldsymbol{F} = \det(\boldsymbol{F}_e \cdot \boldsymbol{F}_p) = (\det \boldsymbol{F}_e)(\det \boldsymbol{F}_p) = J_e J_p$ is then very close to 1, and the difference between the Cauchy stress $\boldsymbol{\sigma}$ and the Kirchhoff stress $\boldsymbol{\tau} = J\boldsymbol{\sigma}$ is negligible. In this case, it does not matter whether the 'physical' yield condition is formulated in terms of Cauchy stress or Kirchhoff stress. If the volumetric changes are large and the 'physical' yield condition is formulated in terms of Cauchy stress, it is still possible to rewrite it in terms of Mandel stress with the Jacobian J taken as another parameter. Note that $J = \sqrt{\det \boldsymbol{C}}$ and that $\boldsymbol{C}$ is a state variable.

A given, physically motivated, isotropic yield condition (24.105) can be rewritten in terms of the dissipative thermodynamic forces as

$$f(\boldsymbol{T}_D, q_D; \boldsymbol{F}_p) = F(\boldsymbol{T}_D \cdot \boldsymbol{F}_p^T) - \sigma_0 + q_D = 0 \tag{24.112}$$

and the associated complementary laws read

$$\dot{\boldsymbol{F}}_p \quad \dot{\lambda}\frac{\partial f}{\partial \boldsymbol{T}_D} = \dot{\lambda}\frac{\partial \bar{F}}{\partial \boldsymbol{M}} \cdot \boldsymbol{F}_p \tag{24.113}$$

$$\dot{\kappa} = \dot{\lambda}\frac{\partial f}{\partial q_D} = \dot{\lambda} \tag{24.114}$$

along with the usual loading-unloading conditions (24.91).

Equation (24.113) specifies the direction of evolution of all the components of the nonsymmetric tensor $\boldsymbol{F}_p$. Recall that the spatial velocity gradient $\boldsymbol{L} = \partial \boldsymbol{v}/\partial \boldsymbol{x}$ is related to the deformation gradient and its rate through (24.45), from which $\boldsymbol{L} = \dot{\boldsymbol{F}} \cdot \boldsymbol{F}^{-1}$. In analogy to that, one can define the 'velocity gradient' associated with the intermediate placement, $\boldsymbol{L}_p = \dot{\boldsymbol{F}}_p \cdot \boldsymbol{F}_p^{-1}$, which is not the gradient of any velocity field, but locally characterizes the rate of plastic deformation. From (24.113), it directly follows that

$$\boldsymbol{L}_p = \dot{\boldsymbol{F}}_p \cdot \boldsymbol{F}_p^{-1} = \dot{\lambda}\frac{\partial \bar{F}}{\partial \boldsymbol{M}} \tag{24.115}$$

For elastically isotropic materials, $\partial \bar{F}/\partial \boldsymbol{M}$ is always a symmetric tensor, and so $\boldsymbol{L}_p$ evaluated from (24.115) turns out to be also symmetric.

A critical reader may have noticed a logical flaw in the arguments leading to (24.113), which was pointed out by Lubliner (1986, 1990). To keep the derivation simple, we used the entire tensor $\boldsymbol{F}_p$ as an internal variable. However, the free energy must be invariant with respect to rigid-body rotations of the intermediate placement, i.e. it depends only on the stretching part of $\boldsymbol{F}_p$. Since a part of $\boldsymbol{F}_p$ has no influence on the free energy, the corresponding part of the evolution law cannot be obtained by thermodynamic arguments. To explain that, consider a perfectly plastic material

(κ and q_D omitted from the formulation). In view of (24.107), (24.108) and (24.110), the dissipation rate is

$$\mathcal{D} = \boldsymbol{T}_D : \dot{\boldsymbol{F}}_p = (\boldsymbol{M} \cdot \boldsymbol{F}_p^{-T}) : (\boldsymbol{L}_p \cdot \boldsymbol{F}_p) = \boldsymbol{M} : \boldsymbol{L}_p \tag{24.116}$$

The postulate of maximum plastic dissipation,

$$\boldsymbol{T}_D : \dot{\boldsymbol{F}}_p = \max_{F(T_D^*) \le \sigma_0} \boldsymbol{T}_D^* : \dot{\boldsymbol{F}}_p \tag{24.117}$$

can be written in the equivalent form,

$$\boldsymbol{M} : \boldsymbol{L}_p = \max_{\bar{F}(M^*) \le \sigma_0} \boldsymbol{M}^* : \boldsymbol{L}_p \tag{24.118}$$

The maximum should be taken over those tensors $\boldsymbol{M}^*$ that are plastically admissible and symmetric. Otherwise we would have to assume the yield condition to remain valid in its original form if the 'plastic part of the material model' (in the sense of the plastic unit in a rheologic model) is subjected to nonsymmetric stresses, which is highly questionable. However, if only symmetric tensors $\boldsymbol{M}^*$ are considered, (24.118) implies only that the symmetric part of $\boldsymbol{L}_p$ must be a non-negative multiple of the gradient of $\bar{F}$, while the skew-symmetric part can be completely arbitrary (since a symmetric tensor $\boldsymbol{M}^*$ does no work on it). If we denote the symmetric part of $\boldsymbol{L}_p$ as $\boldsymbol{D}_p$ (in analogy to the symmetric part $\boldsymbol{D}$ of the spatial velocity gradient $\boldsymbol{L}$), the associated flow rule that follows from (24.118) should be written as

$$\boldsymbol{D}_p = \dot{\lambda}\frac{\partial \bar{F}}{\partial \boldsymbol{M}} \tag{24.119}$$

In (24.117), the tensor $\boldsymbol{T}_D^*$ need not be symmetric but its product with the given tensor $\boldsymbol{F}_p^T$ must be symmetric. So $\boldsymbol{T}_D^*$ is still subject to three constraints, even though they are not as easy to see as the symmetry constraints on $\boldsymbol{M}^*$. Again, this has implications for the corresponding evolution law, which should be written in the form

$$\dot{\boldsymbol{F}}_p = \dot{\lambda}\frac{\partial f}{\partial \boldsymbol{T}_D} + \boldsymbol{W}_p \cdot \boldsymbol{F}_p \tag{24.120}$$

where $\boldsymbol{W}_p$ is an arbitrary skew-symmetric tensor.

Even though the derivation of (24.113) was not completely rigorous, this equation can still be used as the flow rule. The reason is that the choice of the skew-symmetric tensor $\boldsymbol{W}_p$ in (24.120) has no influence on the behavior of the isotropic model, and the simplest choice obviously is $\boldsymbol{W}_p = \boldsymbol{O}$. This additional assumption removes the ambiguity in the definition of the intermediate placement. As already mentioned, all tensors $\boldsymbol{F}_p$ that differ only by a rigid-body rotation represent the same intermediate configuration, and the elastic part of the free energy depends only on the intermediate and current configurations, but not on the specific placements.

Alternatively, one could make the assumption that $\boldsymbol{F}_p$ is a pure stretching tensor, i.e. that the rotation tensor $\boldsymbol{R}_p$ in the polar decomposition $\boldsymbol{F}_p = \boldsymbol{R}_p \cdot \boldsymbol{U}_p$ remains a unit tensor and $\boldsymbol{F}_p = \boldsymbol{U}_p$ is symmetric. This is the case if $\dot{\boldsymbol{F}}_p = \boldsymbol{L}_p \cdot \boldsymbol{F}_p = (\boldsymbol{D}_p + \boldsymbol{W}_p) \cdot \boldsymbol{U}_p$ is always symmetric, which gives three independent conditions for the three independent components of the skew-symmetric tensor $\boldsymbol{W}_p$, provided that $\boldsymbol{D}_p$ is determined from the flow rule.

To summarize, a large-strain elastoplastic model of an elastically isotropic material with isotropic hardening and associated flow is fully specified by the choice of the elastic potential, $\psi_e(\boldsymbol{C}_e)$, the plastic part of the free-energy potential, $\psi_p(\kappa)$, and the expression for the 'equivalent stress', $\bar{F}(\boldsymbol{\tau})$, which enters the yield function in the form of (24.105). These functions determine all the governing equations of the material model formulated within the present constitutive framework.

Example 24.4: Derive the specific form of equations corresponding to a material characterized by (1) the elastic potential (24.60), (2) the plastic part of free-energy potential

$$\rho_0\psi_p(\kappa) = \frac{\sigma_0\kappa_0}{n+1}\left(\frac{\kappa}{\kappa_0}\right)^{n+1} \tag{24.121}$$

with n and κ_0 = material parameters, and (3) the von Mises yield function in terms of the Kirchhoff stress,

$$\bar{f}(\boldsymbol{\tau}, q_D) = \sqrt{3J_2(\boldsymbol{\tau})} - \sigma_0 + q_D \tag{24.122}$$

Solution: The potential (24.60), introduced in Example 24.2, was written in terms of the Jacobian J and the Green strain tensor $\boldsymbol{E}$ of the *total* deformation. These strain measures now have to be replaced by the measures characterizing the *elastic* part of deformation, i.e. by $\boldsymbol{E}_e = (\boldsymbol{C}_e - \boldsymbol{I})/2$ and $J_e = \det \boldsymbol{F}_e$. Since we will differentiate the potential with respect to the elastic Green deformation tensor $\boldsymbol{C}_e$, we rewrite it as

$$\rho_0\psi_e(\boldsymbol{C}_e) = \tfrac{1}{4}(K - \tfrac{2}{3}G)(J_e^2 - 2\ln J_e - 1) + \tfrac{1}{2}G(\boldsymbol{C}_e : \boldsymbol{I} - 2\ln J_e - 3) \tag{24.123}$$

where $J_e = \sqrt{\det \boldsymbol{C}_e}$. Clearly, the resulting hyperelastic law will be completely analogous to (24.63). So we can skip the derivation and write directly

$$\boldsymbol{S}_e = 2\rho_0\frac{\partial\psi_e}{\partial\boldsymbol{C}_e} = \tfrac{1}{2}\left(K - \tfrac{2}{3}G\right)(J_e^2 - 1)\boldsymbol{C}_e^{-1} + G\left(\boldsymbol{I} - \boldsymbol{C}_e^{-1}\right) \tag{24.124}$$

After transformation to the Mandel stress, we get

$$\boldsymbol{M} = \boldsymbol{C}_e \cdot \boldsymbol{S}_e = \tfrac{1}{2}\left(K - \tfrac{2}{3}G\right)(J_e^2 - 1)\boldsymbol{I} + G\left(\boldsymbol{C}_e - \boldsymbol{I}\right) \tag{24.125}$$

As could have been expected, for elastically isotropic materials the Mandel stress $\boldsymbol{M}$ has the same principal directions as the right elastic stretch tensor $\boldsymbol{U}_e$, and its principal values can be easily calculated from the principal elastic stretches.

From the plastic part of the free-energy potential it is possible to establish the relation between the hardening variable κ and its thermodynamically conjugate variable,

$$q_D = -q_Q = -\rho_0\frac{\partial\psi}{\partial\kappa} = -\rho_0\frac{\mathrm{d}\psi_p}{\mathrm{d}\kappa} = -\frac{\mathrm{d}}{\mathrm{d}\kappa}\left[\frac{\sigma_0\kappa_0}{n+1}\left(\frac{\kappa}{\kappa_0}\right)^{n+1}\right] = -\sigma_0\left(\frac{\kappa}{\kappa_0}\right)^{n} \tag{24.126}$$

The result corresponds to a power-type hardening law.

Finally, the associated flow rule (24.115) is easy to construct because it is formally the same as for small-strain J_2-plasticity. The gradient of $\bar{F}(\boldsymbol{M}) = \sqrt{3J_2(\boldsymbol{M})}$ is

$$\frac{\partial\bar{F}}{\partial\boldsymbol{M}} = \frac{\partial}{\partial\boldsymbol{M}}\left[\sqrt{3J_2(\boldsymbol{M})}\right] = \frac{1}{2\sqrt{3J_2(\boldsymbol{M})}}3\frac{\partial J_2(\boldsymbol{M})}{\partial\boldsymbol{M}}$$

$$= \frac{3\boldsymbol{M}_{\mathrm{dev}}}{2\sqrt{3J_2(\boldsymbol{M})}} = \sqrt{\frac{3}{2}}\frac{\boldsymbol{M}_{\mathrm{dev}}}{\|\boldsymbol{M}_{\mathrm{dev}}\|} \tag{24.127}$$

and the flow rule reads

$$\boldsymbol{L}_p = \dot{\lambda}\sqrt{\frac{3}{2}}\frac{\boldsymbol{M}_{\mathrm{dev}}}{\|\boldsymbol{M}_{\mathrm{dev}}\|} \tag{24.128}$$

Here, $\boldsymbol{M}_{\mathrm{dev}}$ is the deviatoric part of the Mandel stress. In view of (24.125), it can be evaluated as

$$\boldsymbol{M}_{\mathrm{dev}} = G\boldsymbol{C}_{e,\mathrm{dev}} = G[\boldsymbol{C}_e - \tfrac{1}{3}(\boldsymbol{C}_e : \boldsymbol{I})\boldsymbol{I}] \tag{24.129}$$

The complete model is described by the elastic law (24.125), flow rule (24.128), hardening law (24.126) with hardening variable κ defined by the evolution law $\dot{\kappa} = \dot{\lambda}$, and loading-unloading conditions (24.91) with the yield function given by $f(\boldsymbol{M}, q_D) = \sqrt{3J_2(\boldsymbol{M})} - \sigma_0 + q_D$. □

For the sake of simplicity, we have discussed in this section only models with isotropic hardening. Extensions to kinematic hardening have been proposed e.g. by Dogui and Sidoroff (1985), Tsamakis (1996) and Svendsen *et al.* (1998). For numerical implementation of large-strain plasticity, see Simo and Hughes (1998).

24.3.3 Comment on Plastic Buckling and Bifurcation

Ceneralization of the theory of plasticity to finite strain automatically provides the capability of tackling the problems of plastic buckling and bifurcation, and finite element programs for large strain plasticity can in principle handle such problems. However, the understanding of plastic buckling can be achieved only through specialized analytical solutions for plastic columns, plates, shells and three-dimensional bodies. Even the most elementary problem of plastic buckling of columns was marred by controversies for over half a century, until finally clarified by Shanley (1947). Likewise, the surprisingly pronounced role of residual stresses in steel columns took several decades to understand, which happened at about the same time.

The important and exciting field of plastic buckling and equilibrium path bifurcation cannot be covered in this book, in view of space limitations. However, a detailed exposition is found in Chapters 8 and 10 of the book by Bažant and Cedolin (1991), which may be used to supplement the present text and incorporated into an advanced course on plasticity.

PROBLEMS

Problem 24.1: Is

$$\boldsymbol{E}^{(f)} = \sum_{I=1}^{3} f(\lambda_I)\boldsymbol{N}\otimes\boldsymbol{N} \tag{24.130}$$

with $f(\lambda) = \sinh(\lambda - 1)$ an admissible finite-strain tensor? What about $f(\lambda) = \sin(\lambda - 1)$? What about $f(\lambda) = \lambda - 1$ for $\lambda \le 2$ and $\lambda^2/4$ for $\lambda > 2$?

Problem 24.2: Derive the specific form of hyperelastic constitutive equations corresponding to the free energy

$$\rho_0 \psi(\boldsymbol{E}^{(0)}) = \left(\frac{K}{2} - \frac{G}{3}\right)\left(\boldsymbol{I} : \boldsymbol{E}^{(0)}\right)^2 + G\boldsymbol{E}^{(0)} : \boldsymbol{E}^{(0)} \tag{24.131}$$

where $\boldsymbol{E}^{(0)} \equiv \ln \boldsymbol{U}$ is the Hencky strain tensor.

Problem 24.3: Derive the rate form of the constitutive equations from Example 24.2 in terms of the Jaumann rate of Cauchy stress. Check whether the resulting Eulerian tangent stiffness tensor is symmetric.

Problem 24.4: How would you have to change the tangent stiffness tensor from Problem 24.3 if you were to use the Jaumann rate of Kirchhoff stress?

Problem 24.5: Show that the Green–Naghdi–McInnis rate and the Cotter–Rivlin rate are objective (see Table 24.1 for their definition).

Problem 24.6: Derive equation (24.106).
Hint: First derive the formula $\partial A^{-1}_{ij}/\partial A_{kl} = -A^{-1}_{ik} A^{-1}_{lj}$, valid for any regular second-order tensor $\boldsymbol{A}$.

Problem 24.7: Using the indical notation, verify the derivation of (24.113) and (24.116).

25

Crystal Plasticity and Microplane Constitutive Models

All the plastic constitutive models presented so far have been formulated directly in terms of the stress and strain tensors and their invariants. This classical approach, however, has probably entered a period of diminishing returns, in which a great further effort can only provide minor improvements of the constitutive model. There is another modeling approach, which is in principle more powerful, since it makes it easier to model various physical phenomena occurring in the microstructure.

25.1 MICROMECHANICALLY BASED MODELS FOR CRYSTALLINE MATERIALS

The simplest and most commonly used phenomenological theory for infinitesimal elastic-plastic deformation of polycrystalline metals is the classical, rate-independent, Prandl–Reuß flow theory with isotropic hardening based on the von Mises yield criterion. Such a theory is completely unrelated to the microscopic mechanisms of plastic deformation. More sophisticated approaches are based on micromechanical considerations, and try to reproduce the physical phenomena whose macroscopic manifestation is plastic deformation. A detailed exposition of such advanced theories is beyond the scope of the present book. However, to stimulate the reader's interest and to provide a starting point for further reading, the micromechanically based models for crystalline materials will now be briefly reviewed.

25.1.1 Crystallographic Models

The models reviewed in this section assume that the microscopic source of plastic deformation is *crystallographic slip*. Such an assumption is based on physical considerations. These models produce constitutive laws that can be used to interpret experiments on large-strain plastic behavior of metals. They also provide a basis for formulating phenomenological constitutive laws. For instance Christoffersen and Hutchinson (1979) used the results from Hutchinson's (1970) small-strain, rate-independent polycrystal calculations to specify corner characteristics in their J_2-corner theory.

The idea for the crystallographic approach to material modeling was inspired by experimental observations of crystallographic behavior of metals. *Single crystal* tests by Taylor and Elam (1923, 1925, 1926) showed that, under high stress, slip occurs on

certain crystallographic planes along certain directions. They observed that slip on a given plane depends on the resolved shear stress on that plane, and is independent of the normal stress on the plane. The stress at which slip occurs is called the *critical stress*. The increase of the critical stress with the magnitude of slip is known as *self-hardening*, while *latent hardening* is the increase in critical stress in unslipped systems.

Taylor (1934) showed in his *dislocation theory* that sliding occurs in such a manner that perfect crystal structure is reformed after each atomic jump. The lattice structure of the bulk of the material remains essentially the same after the occurrence of slip, hence the elastic modulus remains the same. The elastic strain is caused by the elastic deformation of the lattice structure, and the plastic strain is caused by the movement of dislocations. A discussion of the dislocation mechanism background of this continuum slip description of plastic flow is given by Asaro (1983a). A survey of numerical algorithms for single-crystal plasticity has recently been published by Miehe and Schröder (2001).

The so-called *Taylor-type models* and *self-consistent schemes* explicitly represent the deformation by crystallographic slip in *polycrystals*. Polycrystals are continuous three-dimensional collections of grains (i.e. crystallites), each of which can deform by the mechanism of crystallographic slip. Taylor (1938) proposed a model that strictly enforces compatibility by imposing the same set of strains (aggregate strains) on each grain. This idea comes from the experimental observation that most grains of a polycrystal undergo about the same strain.

Taylor (1938) used his model to analyze simple tension or compression of single phase FCC (Face-Centered Cubic) polycrystals. In Taylor's numerical calculations, the polycrystal is made initially isotropic by choosing the grains to have a uniform coverage of all crystallographic orientations. The only source of inelastic deformation is assumed to be the crystallographic slip. The calculation is then based on determining the combination of slip systems and corresponding shear strains, and the stress state in each grain required to produce the specified strain. The selection of slip systems required to produce an arbitrary strain is not necessarily unique. Taylor made the physically intuitive assumption that among all the possible choices for combinations of active slip systems, the appropriate choice was that for which the cumulative shears are minimized (actually, that for which the net internal work is minimized, which reduces to minimization of cumulative shears if all the shear systems have the same shear strength and hardening modulus). Taylor could predict textures in FCC crystals in agreement with experiments of axisymmetric tension and compression.

Bishop and Hill (1951) proposed a polycrystal theory based on the postulate of maximum work. They used inequalities between external work, computed as the product of macroscopic stress and strain increments, and internal work, computed as the products of the crystallographic shear strength and assumed slip increments integrated over the volumes of grains, to set bounds on the critical stress state required to induce yield. Their primary interest was to determine yield surfaces.

The Taylor-type models have been extensively used in the past 50 years to predict texture development and stress-strain response. The development in this field is in the direction of increasing complexity, but most of the models retain the basic assumptions and structure of the Taylor model. We will now outline the common assumptions and the basic formulation of the Taylor model, as developed and used by, among others, Rice (1971), Hill and Rice (1972), Asaro and Rice (1977), Asaro (1983b), Asaro and Needleman (1985), Harren *et al.* (1989), Bronkhorst, Kalindindi and Anand (1992), and Butler and McDowell (1998).

The stress response at each macroscopic continuum material point is given by the volume-averaged response of the multitude of microscopic single crystalline grains comprising the material point. It is assumed that all the grains have equal volume, and that the local deformation gradient in each grain is homogeneous and identical to the macroscopic deformation gradient $\boldsymbol{F}$ at the continuum material point. The mechanics of crystal deformation by slip is represented in two parts: the plastic deformation is given by a set of plastic simple shears from the reference configuration on the slip system of the crystallite; and the lattice with its embedded material elastically deforms and rigidly rotates from this plastically sheared state to reach the current configuration (Rice, 1971; Asaro and Rice, 1971). This motivated the multiplicative decomposition of the deformation gradient,

$$\boldsymbol{F} = \boldsymbol{F}_e \cdot \boldsymbol{F}_p \tag{25.1}$$

already discussed in Section 24.3.2. Here, tensor $\boldsymbol{F}_p$ describes the material shear flow along the various slip systems of the crystallite, and tensor $\boldsymbol{F}_e$ describes the elastic distortion of the lattice along with the rigid rotation of the crystallite. The plastic shear flow from the reference configuration is written as

$$\boldsymbol{L}_p = \dot{\boldsymbol{F}}_p \cdot \boldsymbol{F}_p^{-1} = \sum_\alpha \dot{\gamma}_\alpha \boldsymbol{s}_\alpha \otimes \boldsymbol{n}_\alpha \tag{25.2}$$

where $\dot{\gamma}_\alpha$ is the shear rate on slip system α. The slip system α is defined by the unit crystallographic vectors $\boldsymbol{s}_\alpha$ and $\boldsymbol{n}_\alpha$, where $\boldsymbol{s}_\alpha$ is the direction of slip in the actual configuration and $\boldsymbol{n}_\alpha$ is normal to the slip plane in the actual configuration. The plastic response of a crystallite is defined in terms of the resolved shear stress on each slip system. Each of the slip systems is active as long as the resolved shear stress on that system does not vanish. The plastic shear rate on the slip system α is taken to be governed by a constitutive function of the form

$$\dot{\gamma}_\alpha = g(\tau_\alpha, q_\alpha) \tag{25.3}$$

where τ_α is the resolved shear stress for the slip system α, and q_α is the shear resistance of the slip system α. This equation can be given in the form of a power law (Hutchinson, 1965; Pan and Rice, 1983; Pierce, Asaro and Needleman, 1983).

The evolution of q_α is assumed to be governed by a hardening law in the form

$$\dot{q}_\alpha = \sum_\beta h_{\alpha\beta} |\dot{\gamma}_\beta| \tag{25.4}$$

where $h_{\alpha\beta}$ is the rate of strain hardening on slip system α due to shearing on the slip system β (latent hardening is thus taken into account).

The elastic response of the crystallite is governed by the elastic moduli set up to take into account the anisotropy due to crystallite orientation. In some cases, researchers consider the elastic response of polycrystalline aggregate, but neglect the elastic anisotropy of the FCC single crystal, and the elasticity tensor is the usual elasticity tensor.

Once the behavior of a crystallite is defined, the constitutive response of a polycrystalline aggregate is obtained using an averaging scheme, described in Hill (1972) and Asaro and Needleman (1985).

The previous constitutive model can be used in two types of finite element calculations:

1. Those where the integration point represents a material point in a polycrystalline sample and the constitutive response is given through a Taylor-type polycrystal model; and
2. those where the integration point represents a material point in a single grain and the constitutive response is given through a single crystal model without invoking the Taylor assumptions.

In this second kind of computation, both equilibrium (in the weak sense) and compatibility are satisfied. In Taylor-type computations, the compatibility is satisfied, but not equilibrium between grains.

Bronkhorst, Kalindindi and Anand (1992) presented an evaluation of the Taylor model, done by comparing finite-element results to experimental results. They concluded that the Taylor-type model is in a reasonable first-order agreement with the observation of the texture formation, and also with the overall stress-strain response of single-phase copper. Some deficiencies in the prediction from the Taylor model (finite element computations of the first kind) may be a consequence of the strong kinematic constraint on the deformation of the individual grains of the polycrystal in this model. In the finite element computations of the second kind, the constraint on the deformation of the individual grains is relaxed, and the computation is run using initially random grain orientations. In this second case, each element represents a single grain. When such an approach is adopted, a typical 'orange peel' effect can be observed on the unconstrained surfaces of the FE mesh. This is caused by differences in the orientations of the grains that intersect the free surfaces, and is observed also in physical experiments.

To improve the performance of the model, the strong kinematic constraint characteristic of the Taylor model can be relaxed by introducing *intergranular constraint relations* of various sorts. Butler and McDowell (1998) presented a reformulation of the kinematics of the plastic velocity gradient, introducing a new concept for taking into account the grain subdivision using additional plastic rotation associated with generation of the geometrically necessary dislocations. The modeling of grain subdivision is motivated by the recent experimental studies of microtexture formation (Hughes and Hansen, 1993; Hughes, 1995). The subdivision process facilitates adherence to the Taylor assumption of uniform deformation at the scale of individual grains. Butler and McDowell (1998) introduce a generalization of the decomposition of deformation gradient,

$$\boldsymbol{F} = \boldsymbol{F}_e \cdot \boldsymbol{F}_{\mu s} \cdot \boldsymbol{F}_p \tag{25.5}$$

which includes a part, $\boldsymbol{F}_{\mu s}$, that is associated with the subdivision, distinct from the contribution of dislocation glide, $\boldsymbol{F}_p$. Here, $\boldsymbol{F}_e = \boldsymbol{R} \cdot \boldsymbol{U}_e$ includes the elastic stretch and, for small elastic strains, essentially rigid rotation relative to the lattice.

Another way to tackle the problem of excessive kinematic constraint in Taylor assumption is by enforcing the so-called *relaxed constraints* to reflect the inhomogeneity of strain from grain to grain that is expected to develop with increasing deformation. When grains become too distorted, enforcement of stress equilibrium rather than full compatibility may perhaps be more reasonable; this is taken as the

basis of the relaxed constraint model (Honneff and Mecking, 1978; Canova, Kocks and Jonas, 1984; Rollet, Stout and Kocks, 1989).

Another approach to polycrystal modeling is the *self-consistent scheme* suggested for this purpose by Kröner (1961), Budianski and Wu (1962) and Hill (1965). This approach is used to produce rate-independent models that attempt to account for grain interaction by considering each grain to be an inclusion embedded in an infinite homogeneous matrix whose moduli are the overall moduli of the crystal, to be determined as an average over all grains. Hutchinson (1970) used this approach to determine the yield surfaces and stress-strain response for uniaxial tension and uniaxial tension followed by shear.

Brown (1970) and Berveiller and Zaoui (1979) extended the self-consistent models to account for rate dependence. In self-consistent models the assumption of identical grain deformation is relaxed and these models also account, approximately, for intergranular equilibrium. They do so by enforcing equilibrium between the individual grains and the aggregate average. However, in self-consistent models as well as in Taylor-type models, the deformation within each grain is presumed homogeneous.

25.1.2 Slip Theory of Plasticity

Another approach that draws inspiration from crystallographic experimental observation is the *slip theory of plasticity*, proposed by Batdorf and Budianski (1949) as an adaptation of the idea of Taylor (1938). This approach differs from the strictly crystallographic ones, because the material is modeled as a continuum, neglecting the crystalline structure and the anisotropy of the lattice. The plastic slip is assumed to be the only source of plastic deformation. The slip in any direction along parallel planes of any particular orientation in the material gives rise to a plastic shear strain, which depends only on the history of the corresponding component of shear stress. The plastic strain due to any system of applied stress is found by:

1. considering the history of the shear stress component in each direction on each plane of the material,
2. finding the corresponding plastic shear strain,
3. transforming this plastic shear strain into plastic strains in some fixed system of coordinates, and
4. summing over all the slip directions and slip plane orientations. In their original paper, Batdorf and Budianski proposed a semigraphical method for the computations.

The shear stress required to produce slip is assumed to be independent of the normal stress and the amount of slip. The plastic shear deformation resulting from slip on a plane of a given orientation depends only upon the history of the component of the shear stress in the direction of slip on that plane.

Batdorf and Budianski's approach differs substantially from the Taylor-type models and self-consistent models in that it neglects the actual crystallographic structure of the material and the anisotropy that it implies. The metal is modeled by considering it to be a macroscopically isotropic continuum, and slip is possible on a plane of any direction at any given point, and not only on the planes whose orientation is determined by crystallographic considerations. The theory contemplates an infinitesimal plastic shear strain associated with each infinitesimal fraction of the continuum comprising all possible planes. Given a plastic shear strain on a given plane,

the contribution of this infinitesimal shear strain to the strains in the macroscopic reference system is expressed by projecting the shear strain along the global coordinate directions.

The initial critical shear stress for slip on a plane is given by one half of the elastic limit in pure tension or compression since, in tension test, the maximum shear stress is one half of the applied uniaxial stress. Beyond this initial value, the critical stress increases according to a characteristic shear function to be determined experimentally, from tests on the polycrystalline continuum, not on the single crystal. The slip theory of plasticity can therefore be considered to be 'semi-phenomenological', since the assumed mechanism of plastic deformation is suggested by crystallographic observation. However, the mathematical description of slip is derived in a phenomenological way.

Batdorf and Budianski (1949) analyzed tests on thin aluminum alloy cylinders, compressed into the plastic range and then twisted while the compressive strain is held constant. They showed that their slip theory yields more accurate predictions than the deformation theory of plasticity or the J_2-flow theory in the cases characterized by an abrupt change in the loading path direction in the stress space.

25.2 MICROPLANE MODELS FOR CONCRETE

25.2.1 History of Microplane Models

The essential idea of the microplane modeling approach is to characterize the constitutive law in terms of *stress and strain vectors* rather than second-order tensors. This idea can be traced back to the aforementioned pioneering work of Taylor (1938), and was also implied in the slip theory of Batdorf and Budianski (1949), discussed in the preceding section. Taylor proposed that the constitutive law of polycrystalline metals be defined by relations between the stress and strain vectors acting on planes of various orientations within the material, considered originally as the slip planes in polycrystals and later named more generally the *microplanes* (Bažant, 1984b). The macroscopic strain and stress tensors are determined as a summation of all these vectors under the assumption of a static or kinematic constraint. Tensorial invariance is automatically satisfied by considering all possible spatial orientations (infinitely many in the case of randomly inhomogeneous isotropic materials but only the crystallographic planes in the case of crystals). After Batdorf and Budianski (1949), who were the first to extend Taylor's idea and develop a realistic statically constrained model for plasticity of polycrystalline metals, many others subsequently refined or modified this approach to metals (Kröner, 1961; Budianski and Wu, 1962; Lin and Ito, 1965, 1966; Hill, 1966a, 1966b; Rice, 1970). Extensions for the hardening inelastic response of soils and rocks were also made (Zienkiewicz and Pande, 1977; Pande and Sharma, 1981, 1982; Pande and Xiong, 1982). All these models used the so-called *static constraint* – the assumption that the stress vector acting on a given plane in the material, later called the microplane, is the projection of the macroscopic stress tensor. As seen in Section 25.1, applications to plasticity continue to be refined until today.

Bažant (1984b) and Bažant and Oh (1985) found that, for stability reasons, a static constraint prevents the model from being generalized for postpeak strain-softening damage typical of quasi-brittle materials. The extension to softening damage requires replacing the static constraint by a *kinematic constraint*. In this constraint, which was also suggested by Taylor (1938) but remained unexplored until 1983, the strain vector on any inclined plane in the material is the projection of the macroscopic strain tensor. Bažant and Oh (1985) considered not only inelastic shear strains, but also inelastic normal strains on the microplanes and evaluated the macroscopic stress tensor from the microplane stresses based on the principle of virtual work.

The formulations for metals emanating from Batdorf and Budianski's work were called the *slip theory of plasticity*. This term, however, is unsuitable for general material models, for example those for inelastic strains due to cracking, where the inelastic behavior on the microscale does not physically represent plastic slip, and inelastic volume changes are important. For this reason, the neutral term *microplane model* (Bažant, 1984b) was coined.

After generalizing the microplane model for both tensile and compressive damage (Bažant and Prat, 1988), the microplane model and the corresponding numerical algorithm reached in Caner and Bažant (2000a) and Bažant *et al.* (2000a, 2000b) its latest effective formulation for concrete, named model M4. Microplane formulations have also been developed for anisotropic clays (Bažant and Prat, 1987) and for soils (Prat and Bažant, 1991). For a detailed review and examination of the basic principles, see Carol and Bažant (1997), Bažant *et al.*, (2000b), Brocca and Bažant (2000), and Carol, Jirásek and Bažant (2001).

For both the formulations with kinematic and static constraints, the material properties are characterized by relations between the stress and strain components on the microplanes. The tensorial invariance restrictions need not be enforced directly in the constitutive relations, which is an advantageous feature of the microplane formulation, as well as the formulations in Section 25.1. They are automatically satisfied by superimposing the responses from the microplanes of all orientations in a suitable manner. This is done by means of the principle of virtual work (Bažant, 1984b).

The characterization of the constitutive law in terms of vectors rather than tensors endows the model with conceptual clarity and allows more realistic modeling of oriented processes such as plastic or frictional slip, or cracking. For example, the classical modeling of internal friction as a relation between the mean stress and the deviatoric invariant J_2 is not very realistic; the friction is a relation between the normal and tangential components of stress acting on a given plane, and different frictional responses occur on different planes. This can be directly represented by the microplane model. An important benefit is an automatic modeling of the vertex effect; see Section 25.4.

25.2.2 Kinematic Constraint

The classical constitutive models directly relate the components σ_{ij} and ε_{ij} of the stress and strain tensors $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$. The microplane constitutive model is defined by a relation between the stresses and strains acting on a generic plane called the microplane, having an arbitrary orientation characterized by its unit normal $\boldsymbol{n}$. The basic hypothesis, which ensures stability of postpeak strain softening (Bažant 1984c), is that the strain vector $\boldsymbol{\varepsilon}^n$ on the microplane (Figure 25.1(a)) is the projection of $\boldsymbol{\varepsilon}$, i.e. $\varepsilon_i^n = \varepsilon_{ij} n_j$. The normal strain on the microplane is

$$\varepsilon_N = n_i \varepsilon_i^n = n_i \varepsilon_{ij} n_j = N_{ij} \varepsilon_{ij} \tag{25.6}$$

where $N_{ij} = n_i n_j$.

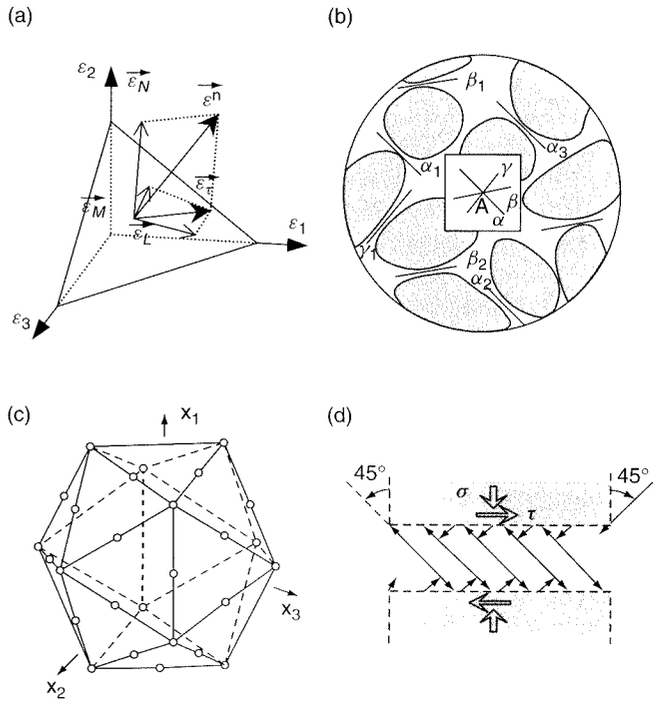


Figure 25.1 (a) Strain components on the microplane; (b) microplane model ensuing by separate homogenization of slips and openings on weak planes $\alpha_1, \alpha_2, \ldots$ of various orientations $\alpha, \beta, \gamma, \ldots$ within the representative volume of material; (c) example of 21-point optimal Gaussian integration formula (the circled points at the vertices and mid-edges of icosahedron represent the directions of the microplane normals); (d) explanation of shear-dilatant cross coupling engendered by interaction of microplanes (after Bažant and Gambarova, 1984)

The shear strains on each microplane are characterized by their components in chosen directions M and L given by orthogonal unit coordinate vectors $\boldsymbol{m}$ and $\boldsymbol{l}$, of components m_i and l_i, lying within the microplane. The shear strain components in the directions of $\boldsymbol{m}$ and $\boldsymbol{l}$ are $\varepsilon_M = m_i \varepsilon_i^n = m_i \varepsilon_{ij} n_j$ and $\varepsilon_L = l_i \varepsilon_i^n = l_i \varepsilon_{ij} n_j$. By virtue of the symmetry of tensor ε_{ij},

$$\varepsilon_M = M_{ij} \varepsilon_{ij}, \qquad \varepsilon_L = L_{ij} \varepsilon_{ij} \tag{25.7}$$

where $M_{ij} = (m_i n_j + m_j n_i)/2$ and $L_{ij} = (l_i n_j + l_j n_i)/2$ (Bažant and Prat, 1988). In fact, ε_N, ε_M and ε_L are the tensorial strain components ε_{nn}, ε_{nm} and ε_{nl} with respect to the local coordinate system $(\boldsymbol{n}, \boldsymbol{m}, \boldsymbol{l})$ aligned with the microplane. The remaining tensorial strain components ε_{mm}, ε_{ll} and ε_{ml} are assumed to have no influence on the stress vector on the microplane with normal $\boldsymbol{n}$ (this is certainly a simplification), but they are taken into account on microplanes with other normals.

Due to the previous kinematic constraint relating the strains on the micro-level (microplane) and macro-level (continuum), the static equivalence (or equilibrium) of stresses between the macro- and micro-levels can only be enforced approximately, in a weak form. This is done by the principle of virtual work (Bažant, 1984c), written as

$$\frac{2\pi}{3} \sigma_{ij} \delta\varepsilon_{ij} = \int_\Omega \left(\sigma_N \delta\varepsilon_N + \sigma_M \delta\varepsilon_M + \sigma_L \delta\varepsilon_L \right) \, \mathrm{d}\Omega \tag{25.8}$$

where the integral is taken over all the possible orientations of the microplane, i.e. over the set of all unit vectors $\boldsymbol{n}$. Geometrically, this set can be represented by the surface of a unit sphere. Since microplanes with normals $\boldsymbol{n}$ and $-\boldsymbol{n}$ are identical, it is sufficient to consider only one half of the sphere. Therefore, the domain of integration Ω is taken as the surface of a unit hemisphere.

Equation (25.8) means that the virtual work of macrostresses (continuum stresses) within a unit sphere must be equal to the virtual work of microstresses (microplane stress components) regarded as the tractions on the surface of the sphere. The factor $2\pi/3$ on the left-hand side is one half of the volume of a unit sphere. The integral on the right-hand side physically represents a homogenization of different contributions coming from planes of various orientations within the material as depicted in Figure 25.1(b). For a detailed physical justification, see Bažant, Xiang and Prat (1996b).

Since the virtual macroscopic strains $\delta\varepsilon_{ij}$ and virtual microplane strains $\delta\varepsilon_N$, $\delta\varepsilon_M$ and $\delta\varepsilon_L$ are considered as variations of the actual strains, they are linked by the kinematic constraint. Substituting $\delta\varepsilon_N = N_{ij}\delta\varepsilon_{ij}$, $\delta\varepsilon_L = L_{ij}\delta\varepsilon_{ij}$ and $\delta\varepsilon_M = M_{ij}\delta\varepsilon_{ij}$, and noting that identity (25.8) must hold for any virtual strain tensor $\delta\varepsilon_{ij}$, one gets the integral stress-evaluation formula (Bažant, 1984c),

$$\sigma_{ij} \;=\; \frac{3}{2\pi} \int_\Omega \left(\sigma_N N_{ij} + \sigma_M M_{ij} + \sigma_L L_{ij} \right) \, \mathrm{d}\Omega \tag{25.9}$$

In numerical calculations, the integral is approximated by an optimal Gaussian integration formula for a spherical surface (Stroud, 1971; Bažant and Oh, 1986),

$$\sigma_{ij} \;\approx\; 6 \sum_{\mu=1}^{N_m} w_\mu \left(\sigma_N^{(\mu)} N_{ij}^{(\mu)} + \sigma_M^{(\mu)} M_{ij}^{(\mu)} + \sigma_L^{(\mu)} L_{ij}^{(\mu)} \right) \tag{25.10}$$

representing a weighted sum over the microplanes of orientations $\boldsymbol{n}_\mu$, with weights w_μ normalized so that $\sum_\mu w_\mu = 1/2$ (Bažant and Oh, 1985, 1986). For a given number of integration points on the spherical surface, the optimum Gaussian formula maximizes the degree of the polynomial that can be integrated exactly and minimizes the coefficient of the next term of the polynomial. For sufficient accuracy, the differences in response curves obtained after various rigid-body rotations of the set of microplanes (integration points) must be small enough. Based on such studies (Bažant and Oh, 1986), the most efficient formula that still yields acceptable accuracy (differences in early postpeak response within a few percent of the mean) involves $N_m = 21$ microplanes (Figure 25.1(c)). The values of $N_{ij}^{(\mu)}$, $M_{ij}^{(\mu)}$ and $L_{ij}^{(\mu)}$ for all the microplanes $\mu = 1, \ldots N_m$ are common to all integration points of all finite elements (see Section 22.1), and are calculated and stored in advance.

The following closed-form expressions are useful for theoretical derivations (Lubarda and Krajcinovic, 1993):

$$\int_\Omega \mathrm{d}\Omega = 2\pi, \qquad \int_\Omega n_i \, \mathrm{d}\Omega = 0 \tag{25.11}$$

$$\int_\Omega N_{ij} \, \mathrm{d}\Omega = \int_\Omega n_i n_j \, \mathrm{d}\Omega = \frac{2\pi}{3} \delta_{ij} \tag{25.12}$$

$$\int_\Omega N_{ij} N_{kl} \, \mathrm{d}\Omega = \int_\Omega n_i n_j n_k n_l \, \mathrm{d}\Omega = \frac{2\pi}{15} \left(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} \right) \tag{25.13}$$

The numerical integration formula with 21 microplanes is exact for all polynomials up to the 9th degree, and so it evaluates the integrals in (25.11)–(25.13) exactly. An important consequence is that the linear elastic behavior of isotropic materials is reproduced exactly, without any influence of the orientation of the microplane system with respect to the principal axes of stress and strain.

25.2.3 Constitutive Relations for a Microplane

For computational efficiency, it is important to avoid iteration during the evaluation of the stress history corresponding to a given strain history. This can be achieved with explicit versions of the kinematically constrained microplane model, which postulate the constitutive laws on the microplane level as explicit formulae for the evaluation of microplane stresses σ_N, σ_M and σ_L from microplane strains ε_N, ε_M and ε_L. Despite our avoidance of tensors, there is an enormous number of possibilities. Lacking a complete micromechanical model, one must partly rely on intuitive understanding of the physical mechanism involved. To a large extent, this is an art.

For concrete, the response to uniaxial compressive stress exhibits postpeak strain softening while the response to hydrostatic compression does not. To capture this difference, it is helpful to decompose the normal strain on each microplane into a volumetric part and a deviatoric part:

$$\varepsilon_N = N_{ij}\varepsilon_{ij} = N_{ij}(\varepsilon_V\delta_{ij} + e_{ij}) = \varepsilon_V + N_{ij}e_{ij} = \varepsilon_V + \varepsilon_D \tag{25.14}$$

Here, e_{ij} are the deviatoric strain components, and $\varepsilon_D = N_{ij}e_{ij}$ is called the (normal) *deviatoric microplane strain*. Since $N_{ij}\delta_{ij} = n_i n_i = 1$, the normal projection of the volumetric part of the strain tensor is the same for all the microplanes, and is equal to the volumetric strain, $\varepsilon_V = \delta_{ij}\varepsilon_{ij}/3 = \varepsilon_{ii}/3$. Note that ε_V is the mean value of ε_N over the unit hemisphere, and

$$\varepsilon_D = \varepsilon_N - \varepsilon_V = N_{ij}\varepsilon_{ij} - \frac{\delta_{ij}}{3}\varepsilon_{ij} = \left(N_{ij} - \frac{\delta_{ij}}{3}\right)\varepsilon_{ij} \tag{25.15}$$

is the fluctuating part of ε_N with zero mean.

If σ_V and σ_D are the volumetric and deviatoric microplane stresses work-conjugate with microplane strains ε_V and ε_D, the principle of virtual work (25.8) must be replaced by

$$\frac{2\pi}{3}\sigma_{ij}\delta\varepsilon_{ij} = \int_\Omega (\sigma_V\delta\varepsilon_V + \sigma_D\delta\varepsilon_D + \sigma_M\delta\varepsilon_M + \sigma_L\delta\varepsilon_L)\ \mathrm{d}\Omega \tag{25.16}$$

The kinematic constraint links the microplane strains to the macroscopic strain tensor, and the same constraint holds for the virtual strains. So, we have $\delta\varepsilon_V = \delta\varepsilon_{kk}/3 = (\delta_{ij}/3)\delta\varepsilon_{ij}$, $\delta\varepsilon_D = \delta\varepsilon_N - \delta\varepsilon_V = (N_{ij} - \delta_{ij}/3)\delta\varepsilon_{ij}$, $\delta\varepsilon_L = L_{ij}\delta\varepsilon_{ij}$, and $\delta\varepsilon_M = M_{ij}\delta\varepsilon_{ij}$. Substitution of these relations into (25.16) yields a modified integral formula for the macroscopic stress (Carol, Jirásek and Bažant, 2001),

$$\begin{aligned}\sigma_{ij} &= \frac{3}{2\pi}\int_\Omega \left[\sigma_V\frac{\delta_{ij}}{3} + \sigma_D\left(N_{ij} - \frac{\delta_{ij}}{3}\right) + \sigma_M M_{ij} + \sigma_L L_{ij}\right]\ \mathrm{d}\Omega \\ &= (\sigma_V - \bar\sigma_D)\delta_{ij} + \frac{3}{2\pi}\int_\Omega (\sigma_D N_{ij} + \sigma_M M_{ij} + \sigma_L L_{ij})\ \mathrm{d}\Omega\end{aligned} \tag{25.17}$$

where

$$\bar\sigma_D = \frac{1}{2\pi}\int_\Omega \sigma_D\ \mathrm{d}\Omega \tag{25.18}$$

is the mean value of the deviatoric stress σ_D over all the microplanes.

Note that formula (25.9) is valid for models without the volumetric-deviatoric split, while (25.17) is valid for models with the split. If $\sigma_N = \sigma_V + \sigma_D$ were substituted into (25.9), the term $-\bar\sigma_D\delta_{ij}$ that appears in (25.17) would be missing. The need for this term was justified by Carol, Jirásek and Bažant (2001); without it, the model would not be thermodynamically consistent.

For the model with the volumetric-deviatoric split, the linear elastic stress-strain relations on the microplanes may be written as

$$\sigma_V = E_V\varepsilon_V, \qquad \sigma_D = E_D\varepsilon_D, \qquad \sigma_M = E_T\varepsilon_M, \qquad \sigma_L = E_T\varepsilon_L \tag{25.19}$$

where E_V, E_D, E_T = microplane elastic moduli, whose relationship to the macroscopic Young's modulus E and Poisson ratio ν is $E_V = E/(1-2\nu)$, $E_D = 5\ E/[(2+3\mu)(1+\nu)]$, $E_T = \mu E_D$ (Bažant and Prat, 1988); here μ = parameter which may be chosen, preferably as $\mu = 1$ (Carol, Bažant and Prat, 1991; Bažant, Xiang and Prat, 1996).

For purely elastic deformations,

$$-\bar\sigma_D = \int_\Omega \sigma_D\ \mathrm{d}\Omega = \int_\Omega E_D\varepsilon_D\ \mathrm{d}\Omega = E_D\varepsilon_{ij}\int_\Omega \left(N_{ij} - \frac{\delta_{ij}}{3}\right)\ \mathrm{d}\Omega = 0 \tag{25.20}$$

and so the presence or absence of the term $-\bar\sigma_D\delta_{ij}$ in (25.17) does not matter. For inelastic deformations, though, $\int_\Omega \sigma_D\ \mathrm{d}\Omega \neq 0$, and so this term does have an effect. In the special case that the tensile and compressive stress-strain relations for the deviatoric strain are symmetric or nearly symmetric, this effect is very small (but generally nonzero). The stronger the asymmetry of the tensile and compressive deviatoric boundaries, the greater is the effect of the term $-\bar\sigma_D\delta_{ij}$ in (25.17) on σ_{ij}. This makes it possible to exploit the asymmetry of the tensile and compressive deviatoric boundaries for controlling dilatancy and pressure sensitivity of deviatoric deformations (Figure 25.1(d)).

25.2.4 Microplane Model M4 for Concrete

In the most recent version of the microplane model for concrete developed by Bažant and coworkers at Northwestern University, called M4 (Bažant *et al.*, 2000b; Caner and Bažant, 2000a), all the inelastic behavior is characterized on the microplane level by the so-called *stress-strain boundaries*. These boundaries may be regarded as strain-dependent yield limits, generally exhibiting strain softening (Figure 25.2). Within the boundaries, the response is incrementally elastic, although the elastic moduli may undergo progressive degradation as a result of damage. Exceeding the boundary stress is never allowed. Travel along the boundary is permitted only if the strain increment is of the same sign as the stress, otherwise elastic unloading occurs. These simple rules for the boundaries suffice to obtain, on the macro-level, the Bauschinger effect as well as realistic hysteretic loops during cyclic loading. Besides, these rules lead to the accumulation of residual stress in the microplane system, which is the cause of fatigue. Despite the abrupt drop of slope when the stress reaches the boundary, the macroscopic response is typically quite smooth, thanks to the fact that different microplanes enter the unloading or reloading regime at different times.

Experience with data fitting has shown that the stress at each of the normal, volumetric and deviatoric boundaries can be assumed to depend only on its conjugate strain, and that a simple friction condition can be used;

$$\sigma_N^b = F_N(\varepsilon_N), \quad \sigma_D^b = F_D(\varepsilon_D), \quad \sigma_V^b = F_V(\varepsilon_V), \quad \sigma_T^b = F_T(\sigma_N^b) \tag{25.21}$$

where superscript ^{b} labels the boundary values and F_N, F_D, F_V and F_T are functions of one variable defining the constitutive behavior (Figure 25.2). For algebraic expressions for the curves, see Bažant *et al.* (2000b).

Note that even though the boundaries on the microplane level are path-independent, the macroscopic response is strongly path-dependent. The reasons is that a vast number of combinations of loading and unloading on various microplanes can occur. Even for proportional loading, some of the microplanes often unload.

The inelastic cross effects between nonconjugate strain and stress components on the macroscale (e.g. between σ_{11} and ε_{22} or ε_{12}) have been shown to be adequately captured by interactions among various microplanes engendered by the kinematic constraint. The frictional boundary (Figure 25.2(A).(a)) is the only exception with a cross effect between nonconjugate variables on the microplane level but, on the other hand, it is the only strain-independent, non-softening boundary (except for a minor strain-dependent adjustment during the initial damage), and thus represents a yield surface.

The roundness or sharpness of the peaks of the response stress-strain curves is controled by the length of the horizontal segments of the boundaries (Figures 25.2(B), (C), (b),(c)). These segments are strain-independent and represent microplane yield limits. In contrast to the softening segments of the boundaries, the horizontal boundaries (yield limits) lead to a very different distribution of the inelastic strains over the microplanes. When the softening segment of a boundary becomes active, all the inelastic strains tend to localize into very few microplanes while most microplanes unload. On the other hand, for the horizontal segments, such localization among spatial orientations does not take place; rather, the yielding gradually spreads over many microplanes and no microplanes unload under monotonic loading. The capacity of simultaneous yielding (but not simultaneous softening) of many microplanes causes the peak of the stress-strain curve to be smoothly curved even though the transition from elasticity to yielding on the microplane is sudden. By extending the range of the horizontal boundaries it is possible to generate a plateau of any length on the macroscopic stress-strain curves.

The *compressive deviatoric boundary* controls the axial crushing strain of concrete in compression when the lateral confinement is too weak to prevent crushing. The *tensile deviatoric boundary* simulates the transverse crack opening of axial distributed cracks in compression and controls the volumetric expansion and lateral strains in unconfined compression tests.

The *frictional boundary* approaches a horizontal asymptote as the compressive stress magnitude increases. The existence of a horizontal asymptote means that, at very high confining pressures (larger than the uniaxial compression strength), concrete becomes a plastic but frictionless material. That this is indeed so has recently been confirmed by a new type of experiment – the so-called 'tube-squash' test (Bažant, Kim and Brocca, 1999; Caner and Bažant, 2000b; Brocca and Bažant, 2001), in which a thick highly ductile steel alloy tube is filled with concrete and squashed to about a half of its initial length. Very high pressures develop and shear angles above 70° are

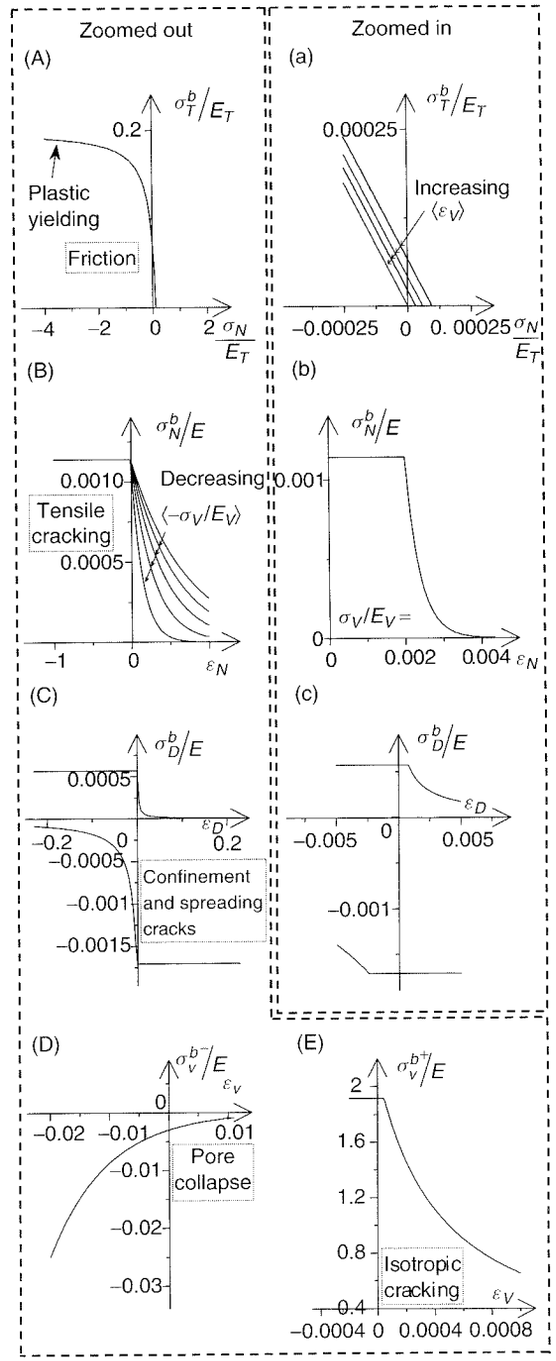


Figure 25.2 (A,B,C,D,E) Stress-strain boundaries plotted at large scales; and (a,b,c) boundaries from (A,B,C) replotted at small scales (for $E = 25$ GPa and reference values used for other parameters)

achieved in this test. Yet no cracks nor voids are visible to the naked eye on a cut through the deformed material, and cohesion is not lost (the uniaxial compression strength of the cores drilled out from the deformed material is still about 30% of their virgin strength). Such plastic behavior occurs in concrete near the nose of a penetrating missile or explosively driven anchor.

The inelastic behavior under hydrostatic pressure (as well as uniaxial compressive strain with no lateral expansion) exhibits no strain-softening but progressively stronger hardening caused primarily by the collapse and closure of pores. It is simulated by a progressively steepening *compressive volumetric boundary* (Figure 25.2(D)). Imposition of *tensile volumetric boundary* helps to control isotropic cracking damage.

The property that one microplane strain component may be unloading while another is loading is an advantage compared to the classical theory of plasticity. Such behavior is admissible because the microplane stress and strain components are not tensors and thus do not have to satisfy any tensorial invariance conditions. When some microplane strains on some microplane are unloading while other microplane strains on the same microplane or on another microplane are simultaneously loading (or reloading), residual stresses balancing each other (with a zero macroscopic stress resultant) are being accumulated and locked up in the kinematically constrained microplane system. This simulates accumulation of strain energy in the microstructure of material, which is known to be the physical source of fatigue.

Concrete is an inelastic material exhibiting a broader and more complex range of inelastic phenomena than metals. At the same time, compared to other materials, there is now an enormous scope of test data and a larger number of test types that need to be modeled. With the latest microplane model, M4 (Bažant *et al.*, 2000b), it has been possible to simulate essentially all the known basic types of inelastic response of concrete, and the simulation has been closer than achieved before. Model M4 has, admittedly, many empirical parameters, however, it has been set up in such a way that there are only four free parameters that need to be adjusted by the users to simulate their type of concrete (the number of parameters in sophisticated models has often been deplored but given the richness of different types of response, a significant reduction in the number of parameters is hardly possible). These four parameters can be identified explicitly from the given uniaxial compression strength and the corresponding strain, and from the shape of the stress-strain curve for hydrostatic compression.

25.2.5 Advantages of the Microplane Model

It might be useful to give the following summary of the advantages of microplane model compared to the classical tensorial models.

1. The constitutive law is written in terms of vectors rather than tensors. The modeler need not worry about tensorial invariance because it is automatically satisfied by combining the responses from microplanes of all possible orientations.
2. The inelastic physical phenomena associated with surfaces, such as the slip, friction, tensile cracking, lateral confinement on a given plane or its spreading, can be characterized directly in terms of the stress and strain on the surface on which they take place. This contrasts with tensorial models in which, for instance, a relation between the invariants I_1 and J_2 is regarded as friction although it is only a poor overall measure of friction, unable to capture frictional slip on any particular plane.
3. In computational practice, macroscopic tensorial plastic-damage models with only one or two loading surfaces are generally used. For such models, most materials appear to exhibit large apparent deviations from the normality rule for the plastic strain increments. Aside from friction, the reason is that in reality many simultaneous yield (or loading) surfaces intersect at every point of a plastic loading path in the stress space (Bažant, 1978, 1980). However, with the classical multi-surface plasticity proposed by Koiter (1953b) and described in Section 20.4, it seems next to impossible to fit more extensive test data. The microplane model is equivalent to infinitely many simultaneous (active or inactive) loading surfaces of all possible orientations, at least one for each microplane. This enables the model to automatically capture these apparent deviations from normality, including those caused by dilatancy in frictional slip (Bažant and Gambarova, 1984).
4. Unlike the classical plasticity models used in computational practice, the microplane model captures the so-called vertex effect, i.e. the fact that an apparent vertex (or corner) exists at every point of the yield surface detected by radial loading; see Section 25.4 for more details for both metals and concrete.
5. The interaction of microplanes due to the kinematic constraint suffices to provide all the main cross effects on the macro-level, such as the pressure sensitivity of inelastic shear strain and the dilatancy. It is therefore possible on the microplane level to use simple one-to-one stress-strain laws relating a stress component to its associated strain component. Microplane friction, though, is an exception – a yield surface in terms of two (but not more than two) microplane stress components is needed.
6. A vast number of combinations of loading or unloading on various microplanes is possible, and some microplanes even unload for monotonic loading on the macroscale. This is an important source of path dependence in the microplane model, and allows a simple yet physically realistic representation of the Bauschinger effect at reverse loading and of hysteresis under cyclic loading.
7. During unload-reload cycles, residual stresses on microplanes (i.e. self-equilibrated stresses in balance with a zero stress tensor of the continuum) are being automatically locked up within the microplane system, and their strain energy gradually accumulated, thus weakening the response. This naturally models fatigue, as well as hysteresis during cyclic loading.
8. With the microplane approach, anisotropic materials are not appreciably more difficult to model than isotropic materials. It suffices to use either orientation-dependent model parameters (Brocca, Bažant and Daniel, 2001) or an orientation-dependent weight function in the integral expression for the virtual work on the micro-level.
9. When the stress-strain boundaries on the microplanes include a horizontal segment (non-softening yield surface), the yielding spreads during loading over many microplanes. This gradual spread of yielding automatically produces on the macro-level a rising stress-strain curve with a gradually diminishing slope, even if the transition from elastic response to yielding on the microplanes is abrupt. So one need not bother with the modeling of strain-hardening on the microplane level (unless some precise shape of hardening stress-strain curve is to be matched).

10. The philosophies of the microplane approach and finite elements blend well. While the latter represents a spatial discretization (with respect to distance), the former can be regarded as an angular discretization (with respect to orientations). In both, the principle of virtual work is used in analogous ways, suitable for explicit temporal integration. The microplane concept also blends well with the nonlocal concept to be explained in Section 26.3.1. While the latter captures interactions of damage at distance, the former captures interactions among angular orientations.
11. The inherent conceptual simplicity of the model, gained by dealing with stress and strain components rather than tensors and their invariants, facilitates understanding and intuition. This is important since the modeling of complex materials is as much an art as a science.

The penalty to pay for these advantages is an increase in computational work and storage requirements, which is about 4 to 10-fold within the constitutive subroutine compared to the classical plasticity models. This is due to the need to deal with stress components on all the microplanes, whose number must be at least 21 for acceptable accuracy (Bažant and Oh, 1986). This penalty, however, becomes relatively less burdensome for very large structural systems in which the computational work is dominated by solving the system rather than the material subroutine. Thanks to the rapid rise of computer power, the microplane model is no longer computationally forbidding. Concrete structures discretized with hundreds of thousands of finite elements are being analyzed today at Waterways Experiment Station (WES) with the wave code (hydrocode) EPIC containing an explicit microplane constitutive subroutine developed at Northwestern (Bažant *et al.*, 2000a).

It may be noted that the microplane model is a necessary consequence of mathematical homogenization (smearing) applied separately for various orientations. In concrete, the major contribution to inelastic strain comes from the slip and opening on weak planes within the microstructure, typically lying in the thin contact layers between aggregate pieces (Figure 25.1(b)). This slip and opening may be considered to conform kinematically to the strain of the macroscopic continuum at point A lying at the center of the representative volume of material. One may now isolate within the representative volume of the material all the weak planes $\alpha_1, \alpha_2, \alpha_3$ whose normals have orientations with spherical coordinate angles between (ϕ, θ) and $(\phi + \mathrm{d}\phi, \theta + \mathrm{d}\phi)$. Macroscopic homogenization (smearing) of the slips and openings on the weak planes $\alpha_1, \alpha_2, \alpha_3$ yields a continuum that exhibits at point A the corresponding homogenized inelastic strains on plane α of the same orientation (ϕ, θ). Considering all the possible orientations, one thus obtains by such separate homogenization processes a number of planes of many orientations (Figure 25.1(c)), all of them located at A and all with strains constrained kinematically to the same continuum strain tensor.

How does the microplane model automatically generate dilatancy in shear? Figure 25.1(d) (after Bažant and Gambarova (1984)) explains that. The shear strain in a band prevented from expanding causes normal tensile and compressive strains at microplanes inclined by $\pm 45^\circ$. Because the microplane strength of concrete is much smaller in tension than in compression, the compressive normal stress vectors on microplanes inclined by 45° are, at large shear strain, much larger than the tensile normal stress vectors on microplanes inclined by -45°, causing a large compressive normal stress resultant σ across the band. If this resultant is not opposed by a restraint against lateral expansion of the band, the band will expand. This represents dilatancy.

Thus the constitutive law for a microplane need not involve any cross coupling between the normal stress and shear strain on a microplane.

25.3 MICROPLANE MODELS FOR METAL PLASTICITY

We will now briefly present three microplane models for metal plasticity, which differ in various respect from the Taylor models described in Section 25.1: (1) a kinematically constrained microplane model version of the J_2-flow theory; (2) a microplane version of the slip theory of plasticity, which uses a static constraint; and (3) a kinematically constrained microplane model for plasticity based on independent yield conditions for shear and deviatoric components of strain.

25.3.1 Microplane Model Version of J_2-plasticity

The von Mises yield condition (15.10) is formulated in terms of the deviatoric invariant $J_2 = \frac{1}{2}\boldsymbol{s} : \boldsymbol{s} = \frac{1}{2}\|\boldsymbol{s}\|^2$, where $\boldsymbol{s}$ is the deviatoric part of the stress tensor. On the microplane level, the counterpart of tensor $\boldsymbol{s}$ is the vector $\boldsymbol{s}^n = \sigma_D \boldsymbol{n} + \sigma_M \boldsymbol{m} + \sigma_L \boldsymbol{l}$, and so a yield condition analogous to the von Mises condition may be proposed in the form

$$f(\sigma_D, \sigma_M, \sigma_L) \equiv \sigma_D^2 + \sigma_L^2 + \sigma_M^2 - k^2 = 0 \tag{25.22}$$

where k is a parameter related to the macroscopic yield stress.

When condition (25.22) is met, the microplane starts yielding and the deviatoric stress vector $\boldsymbol{s}^n$ must stay on the yield surface. In analogy to the macroscopic (tensorial) elastoplasticity, one can split the microplane strain vector $\boldsymbol{\varepsilon}^n$ into an elastic part and a plastic part, link the elastic part to the microplane stress vector $\boldsymbol{\sigma}^n$ by (25.19), and postulate the associated flow rule governing the evolution of the plastic part. The complete set of constitutive relations on the microplane level consists of the equations

$$\sigma_V = E_V \varepsilon_V \tag{25.23}$$

$$\sigma_D = E_D(\varepsilon_D - \varepsilon_D^p), \qquad \sigma_M = E_T(\varepsilon_M - \varepsilon_M^p), \qquad \sigma_L = E_T(\varepsilon_L - \varepsilon_L^p) \tag{25.24}$$

$$\dot{\varepsilon}_D^p = \dot{\lambda}\frac{\partial f}{\partial \sigma_D} = 2\dot{\lambda}\sigma_D, \qquad \dot{\varepsilon}_M^p = \dot{\lambda}\frac{\partial f}{\partial \sigma_M} = 2\dot{\lambda}\sigma_M, \qquad \dot{\varepsilon}_L^p = \dot{\lambda}\frac{\partial f}{\partial \sigma_L} = 2\dot{\lambda}\sigma_L \tag{25.25}$$

and the usual loading-unloading conditions (15.36).

Note that the model, called MP1, remains kinematically constrained. For a given increment of the macroscopic strain tensor, the microplane strains are evaluated by projection on the individual planes, and the microplane stresses and plastic strains are solved from the previous equations on each microplane separately. Finally, the macroscopic stress tensor is constructed by substituting the microplane stresses into (25.17).

The onset of yielding does not take place on all microplanes simultaneously. During first loading, different microplanes reach yielding at different times. This progressive approach to yielding saturation gives a smooth transition from the elastic response to the plastic response. Figure 25.3 shows the uniaxial stress-strain curves obtained with J_2-plasticity and with model MP1. With a J_2-flow plastic law, the material response changes from plastic to elastic suddenly. With the microplane model this transition occurs gradually, as the yield condition is met successively in different planes.

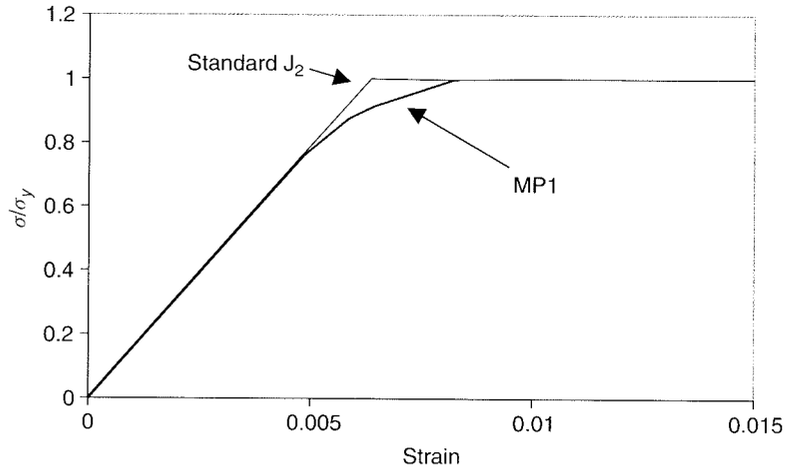


Figure 25.3 Uniaxial stress-strain curves for an elastic-perfectly plastic material as reproduced by model MP1 and a standard J_2-flow model

Under proportional loading, the response of model MP1 approaches the *saturated state of yielding*, in which all the microplanes undergo yielding at the same time (Carol and Bažant, 1997). This is similar to the limit state of an elastoplastic structure. The stresses and elastic strains in the saturated state remain constant and only the plastic strains grow proportionally. Combining this fact with the kinematic constraints (25.7) and (25.15), the stress evaluation formula (25.17) and the flow rule (25.25), it is possible to show that in the saturated state model MP1 satisfies the static constraint, i.e. that the microplane stresses are projections of the macroscopic stress tensor (see Section 25.3.2). This permits us to relate the parameter k that appears in the microplane yield condition (25.22) to the macroscopic yield stress. Under the static constraint, we have

$$\|\boldsymbol{s}^n\|^2 = s_k^n s_k^n = s_{ki} n_i s_{kj} n_j = N_{ij} s_{ki} s_{kj} \tag{25.26}$$

and

$$\int_\Omega \|\boldsymbol{s}^n\|^2 \, \mathrm{d}\Omega = \int_\Omega N_{ij} \, \mathrm{d}\Omega \; s_{ki} s_{kj} = \frac{2\pi}{3} \delta_{ij} s_{ki} s_{kj} = \frac{2\pi}{3} s_{ki} s_{ki} = \frac{2\pi}{3} \|\boldsymbol{s}\|^2 = \frac{4\pi}{3} J_2 \tag{25.27}$$

In the saturated state of yielding, $\|\boldsymbol{s}^n\|^2 = k^2$ on all the microplanes, and so

$$J_2 = \frac{3}{4\pi} \int_\Omega \|\boldsymbol{s}^n\|^2 \, \mathrm{d}\Omega = \frac{3k^2}{4\pi} \int_\Omega \mathrm{d}\Omega = \frac{3k^2}{2} \tag{25.28}$$

So, the macroscopic shear yield stress τ_0 corresponding to the saturated state of model MP1 is $\tau_0 = \sqrt{J_2} = k\sqrt{3/2}$.

Model MP1 cannot reproduce the so-called vertex effect, which will be described in Section 25.4, nor the accompanying apparent deviations from normality if the principal strain axes rotate. Because of the way in which the yield surface is defined on each microplane (indeed equivalent to a macroscopic J_2-plasticity model), the increment of the stress component on the microplane level will be tangential to the yield surface when the increment in the macroscopic deviatoric stress is normal to the actual deviatoric surface. Therefore, the initial shear stiffness predicted by this model

in the case of shear following uniaxial pre-loading is the elastic shear stiffness, as it is when a J_2-plasticity model is used.

A version of MP1 was originally introduced by Carol and Bažant (1997) who proposed an algorithm employing a yield condition rewritten in terms of microplane strains instead of stresses.

25.3.2 Microplane Version of Slip Theory of Plasticity

A formulation with static constraint, which is suitable for hardening plastic metals, equates the stress components on each microplane to the projections of the macroscopic stress tensor. The microplane stress components are

$$\sigma_N = N_{ij}\sigma_{ij}, \quad \sigma_M = M_{ij}\sigma_{ij}, \quad \sigma_L = L_{ij}\sigma_{ij} \tag{25.29}$$

As usual in plasticity, the strain tensor is additively decomposed into an elastic part and a plastic part. The elastic part of the strain tensor is related to the stress tensor by the usual linear elastic relations. The plastic part of the strain tensor is obtained by summing the contributions of individual microplanes, calculated from the microplane stresses. The complementary virtual work equation provides, in analogy to (25.9),

$$\varepsilon_{ij}^p = \frac{3}{2\pi} \int_\Omega (\varepsilon_N^p N_{ij} + \varepsilon_M^p M_{ij} + \varepsilon_L^p L_{ij}) \, d\Omega \tag{25.30}$$

Microplane model MP2 (Brocca and Bažant, 2000) is very similar to the slip theory of plasticity (Batdorf and Budianski, 1949), which has been described in Section 25.1.2. This approach requires adopting the static constraint. The shear slip on the microplane level is assumed to be the only source of plastic deformation, and so the normal component of the microplane plastic strain, ε_N^p, vanishes. Yielding occurs when the shear component of stress on the microplane level exceeds a limit value, and hardening is driven by the effective plastic strain κ, defined by its rate $\dot{\kappa} = \sqrt{(\dot{\varepsilon}_M^p)^2 + (\dot{\varepsilon}_L^p)^2}$. The yield condition, therefore, is

$$f(\sigma_M, \sigma_L, \kappa) \equiv \sqrt{\sigma_M^2 + \sigma_L^2} - h(\kappa) = 0 \tag{25.31}$$

where $h(\kappa)$ is a hardening function (on the microplane level). When yielding occurs, slip produces plastic shear strains on the microplane level. ε_M^p and ε_L^p, which can be determined from the associated flow rule

$$\dot{\varepsilon}_M^p = \dot{\lambda} \frac{\partial f}{\partial \sigma_M} = \dot{\lambda} \frac{\sigma_M}{\sqrt{\sigma_M^2 + \sigma_L^2}} \tag{25.32}$$

$$\dot{\varepsilon}_L^p = \dot{\lambda} \frac{\partial f}{\partial \sigma_L} = \dot{\lambda} \frac{\sigma_L}{\sqrt{\sigma_M^2 + \sigma_L^2}} \tag{25.33}$$

and from the consistency condition ($\dot{f} = 0$ during yielding). The components of the plastic strain tensor ε_{ij}^p are then obtained by superposition of the effect of slip on all the planes according to formula (25.30) with $\varepsilon_N^p = 0$.

When slip is the only source of plastic deformation, the distribution of plastic strain among the various planes is very non-uniform. The kinematic constraint used by model MP1 gives a more 'uniform' distribution of plastic slip strains

among the microplanes, and thus the MP1 response is inevitably too stiff. The difference is in the integration scheme and the way the overall plastic strain tensor is computed.

Model MP2 can represent a vertex on the yield surface. The consequence of such a vertex is, for example, a reduction of shear stiffness in the case of a loading path in which shear follows uniaxial pre-loading into the yielding region. In this regard, see the numerical results on the vertex effect given in Section 25.4.

A displacement-based finite element computation requires a subroutine that updates the stress for a given strain. Models with the static constraint are computationally less efficient than explicit models with the kinematic constraint, because the stress history cannot be obtained from a given strain history by simple substitution into explicit formulae. The stress evaluation algorithm requires iteration. Obviously, this increases the computational time and makes MP2 less convenient than MP3, to be described next.

A statically constrained model is inherently less stable than a kinematically constrained model, because in a statically constrained model the strain distribution over the microplanes can localize into one or several microplanes in a spurious way. This problem arises in the case of microplane constitutive laws which imply strain softening or elastic-perfectly plastic behavior. However, spurious localization can sometimes also occur with mild hardening, when geometric nonlinearities are taken into account (in a manner similar to that demonstrated by Rudnicky and Rice, 1975). For this reason it would, for example, be very difficult to extend MP2 to the finite strain.

25.3.3 Microplane Model for Plasticity with Kinematic Constraint

A model with kinematic constraint, which will be called MP3, is necessary for developing a microplane model for plasticity that can represent a vertex in the yield surface and is at the same time efficient and stable for finite-strain explicit finite element analysis. As mentioned, a kinematic constraint is not suitable for dealing with a model in which plastic slip is the only source of plastic deformation. To achieve a satisfactory performance, it is possible to assume that the contribution to the overall plastic deformation is not only given by the shear components of strain on each microplane, but also by the deviatoric part of the normal component of strain on each microplane. This is a convenience assumption whose physical justification is still unclear. However, it is acceptable in a model that does not aim at a crystallographic description of material response. Independent yield conditions can be assumed for the shear stresses and the deviatoric part of the normal stresses on the microplane level. The model is constructed as a multisurface extension of MP1, with two independent yield conditions,

$$f_1(\sigma_M, \sigma_L, \kappa_1) \equiv \sigma_M^2 + \sigma_L^2 - h_1(\kappa_1)^2 = 0, \qquad f_2(\sigma_D, \kappa_2) \equiv \sigma_D^2 - h_2(\kappa_2)^2 = 0 \quad (25.34)$$

Plastic flow is governed by Koiter's rule (20.122) and by the associated flow rule with the corresponding loading-unloading conditions. The hardening variables κ_1 and κ_2 are defined by the rate equations $\dot{\kappa}_1 = \sqrt{(\dot{\varepsilon}_M^p)^2 + (\dot{\varepsilon}_L^p)^2}$ and $\dot{\kappa}_2 = |\dot{\varepsilon}_D^p|$. The hardening functions $h_1(\kappa_1)$ and $h_2(\kappa_2)$ are determined experimentally.

25.4 VERTEX EFFECT

The main problem with the classical tensorial models, and the main reason for using the microplane model, is the vertex effect, which will now be briefly discussed.

25.4.1 Physical Source and Modeling of Vertex Effect

As a convenient way of avoiding material instability and satisfying thermodynamic restrictions with simple plasticity models, the inelastic strain rate is taken to be normal to the yield surface. Adherence to the normality rule has been the basic feature of all the practical constitutive relations for plasticity of metals. A similar rule has also been adopted for the fracturing (Dougill, 1976) and plastic-fracturing constitutive models for concrete with loading surfaces in the strain space (Bažant and Kim, 1979b; Lin *et al.*, 1987; Červenka, Červenka and Eligehausen, 1998). The normality rule implies that, for load increments that are parallel to the current yield surface (or to the loading surface), called 'loading to the side', the response is purely elastic. Deviations from such behavior were hinted at by tests of plastic buckling of plates by Bleich (1952) and confirmed by Gerard and Becker (1951, 1957), who tested axially compressed thin-walled cruciform steel columns that buckle in the plastic range by torsion. The critical load of such columns is proportional to the tangent inelastic stiffness for loading to the side. As it was found, this stiffness can be much smaller than the critical load for the elastic stiffness, in fact even less than one half of the elastic critical load; see Figure 25.5. A reduction of the torsional stiffness of concrete cylinders after preloading in compression was demonstrated by Caner, Bažant and Cervenka, (2001); see Figure 25.6.

The existence of inelastic strain increments for loading to the side implies that there must be a corner, or vertex, on the yield surface (or loading surface) at the current state point, traveling during loading with the state point as the material hardens or softens. Therefore, the phenomenon is called the *vertex effect*. This effect is the strongest for abrupt changes of loading direction in the stress space or strain space. However, the vertex effect is important for all nonproportional loading paths (even smooth paths) when the principal stress (or principal strain) directions rotate significantly.

In general, the vertex effect develops for any significant deviations from proportional loading involving significant rotation of the principal axes. It obviously creates difficulties for the classical tensorial models of plasticity on which this book is focussed. The vertex effect has been intensely debated in the literature up to mid 1980's, and then the interest waned, apparently out of frustration. No tensorial models exhibiting the vertex effect are available for large-scale finite element analysis, in which, for metals, von Mises J_2 plasticity reigns.

One way to cope with the vertex effect is to assume a so-called 'traveling vertex', i.e. a vertex on the current yield surface that travels with the current state point along the loading path. Such *vertex enhancements of classical plasticity* have been formulated for characterizing the first infinitesimal deviation from the proportional loading path and were used to study the bifurcation instabilities at the onset of localization of yielding into shear bands (Rice, 1975, 1976; Rudnicky and Rice, 1975; Christoffersen and Hutchinson, 1979; Hughes and Shakib, 1986). These approaches, however, are

not effective for predicting what happens afterwards, after large deviations from the previous loading direction in the stress or strain space take place, after large rotations of the principal axes.

Hencky's *deformation theory of plasticity* (Hencky, 1924) was found to predict an approximately correct vertex behavior for initial infinitesimal deviations from a proportional loading path; e.g. it gives approximately correct predictions for the tests of Gerard and Becker; see Budianski (1959). However, since it represents a path-independent model, it cannot be used for highly nonproportional loading paths. The *endochronic theory*, which represents an incrementally nonlinear modification of plasticity, provides vertex behavior for arbitrary nonproportional paths (Bažant, 1978), but has no physical basis and turned out to be less powerful than the microplane model.

The physical source of the vertex effect was clarified long ago by the models described in Section 25.1, which include various sophisticated models for polycrystalline metals, especially the slip theory of plasticity of Batdorf and Budianski (1949), and further refinements of Taylor's (1938) model (Phillips and Gray, 1961; Naghdi, Essenburg and Koff, 1958; Ivey, 1961; Bertsch and Findley, 1962; Budianski *et al.*, 1951); in detail see Brocca and Bažant (2000). The vertex effect is caused by the development of multiple yield surfaces. Many surfaces are active simultaneously. After a sudden change of loading path direction in the stress or strain space, accompanied by a sudden rotation of the principal strain directions, some yield surfaces suddenly go into unloading and deactivate, while other yield surfaces suddenly become active. While, for proportional loading, the vertex effect plays no role, for highly nonproportional loading paths with significantly rotating principal axes, it is very important. Such loading typically arises in dynamics, for example in problems of impact or penetration of missiles, shock and blast effects on hardened structures, and earthquake effects on buildings or bridges. The vertex effect is crucial for plastic bifurcation and buckling problems; see e.g. Hutchinson (1974) or Bažant and Cedolin (1991). The vertex effect can be observed in nonproportional axial-torsional tests of cylinders, but is missed in the so-called 'true' triaxial tests of cube specimens, in which the principal stress directions do not rotate against the material.

Formulations of multisurface plasticity within the context of the classical tensorial models with invariants are available (Section 20.4) but difficult to implement and identify from test data. No models with enough surfaces exist. For soils and concrete, models involving two yield surfaces (one for the deviatoric space and one for the coupling with mean stress) are typically used, but for capturing the vertex effect, two yield surfaces are insufficient by far. To model the vertex effect on the torsional stiffness, one would have to use at least ten simultaneously active yield surfaces in the deviatoric plane alone.

25.4.2 Experimental Observations and Numerical Results for Metals

For metals, the vertex effect was experimentally discovered by loading not too slender but sufficiently thin cruciform metalic columns (Figure 25.4) by axial compression into the plastic range and observing their torsional buckling (Gerard and Becker, 1957). These experiments exploited the fact that the torsional buckling of thin cruciform columns depends only on the tangential plastic shear modulus G_t. A salient feature of this experiment is that, at the beginning of torsion, the principal stress directions

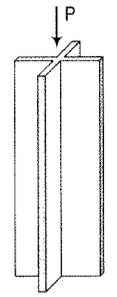


Figure 25.4 Cruciform column subjected to axial load

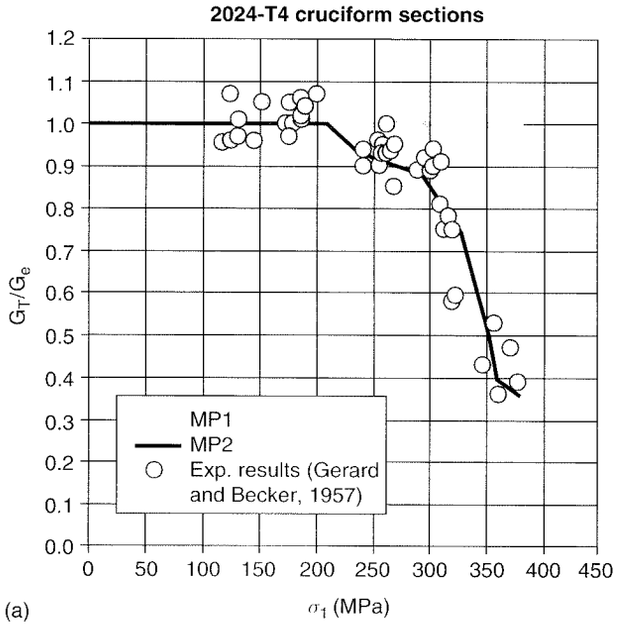
rotate in proportion to the rotation angle while the values of principal stresses remain constant (to the first order in the rotation angle, as can be readily noted from the Mohr circle). Such loading paths are the stumbling block for most plasticity models, (for concrete and soils see Bažant, 1993) for which the endochronic model (Bažant, 1978) was the only exception before the microplane model.

The critical compressive stress of a cruciform column buckling in torsion is

$$\sigma_{cr} = -\frac{G_t J}{2I} = -G_t \left(\frac{h}{b}\right)^2 \tag{25.35}$$

where G_t is the tangent plastic shear modulus, h is the wall thickness, b is the width of the flange plates, $J = 4bh^3/3$ is the stiffness of the cross section in simple torsion, and $I = 2hb^3/3$ (e.g. Bažant and Cedolin, 1991, Section 8.1). Standard J_2-plasticity yields for this loading direction $G_t = G =$ elastic shear modulus, because the incremental deviatoric stress is tangent to the yield surface, and thus the response is predicted as elastic (Bažant and Cedolin 1991, Section 8.1). This means that the bifurcation load should be unaffected by plasticity if the J_2-flow theory is valid.

The results by Gerard and Becker clearly show that the axial compressive stress σ_1 does affect the tangential shear modulus, provided that the column is loaded into the plastic region. The effect is large. Brocca and Bažant (2000) assumed (for aluminum 2024-T4) the following material properties: Young's modulus E = 10.6×10^6 psi, Poisson's ratio $\nu = 1/3$, and yield stress (at 0.2% strain) 40 ksi (1 ksi = 6.895 MPa), and used them to fit the torsional test data of Gerard and Becker. Figure 25.5(a) shows a comparison of the variation of G_t predicted by the microplane models MP1 and MP2 with the experimental results. As expected, model MP1 (equivalent to J_2-plasticity) predicts no reduction in the tangential shear modulus with increasing uniaxial stress. On the other hand, model MP2 agrees with the experimental results up to the maximum applied compressive stress very accurately. Figure 25.5(b) shows the results obtained with model MP3. An accurate fit of the test data is achieved again, but only up to $|\sigma_1| = 325$ MPa. After that, for increasing values of pre-applied uniaxial stress, the tangential shear modulus tends asymptotically toward a limit minimum value. The reason is the kinematic constraint. In a statically constrained model (such as MP2), a very large plastic strain can localize into a limited number of microplanes making the overall response very compliant. With the kinematic constraint, the plastic strain on the microplanes at which the yield condition is met is constrained by the elastic strain of other microplanes, which are mainly those that undergo unloading after the start of torsion. If all the microplanes are assumed to yield before the start of torsion, the maximum reduction of G_t can be shown to be $0.5G$ for MP3.



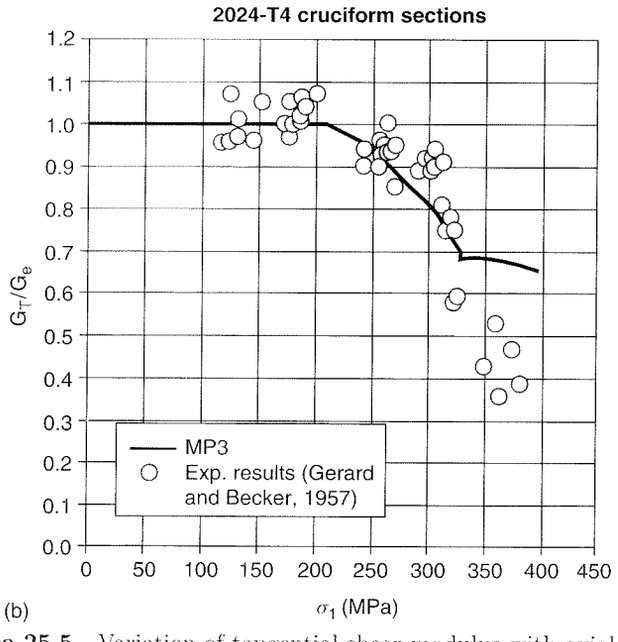
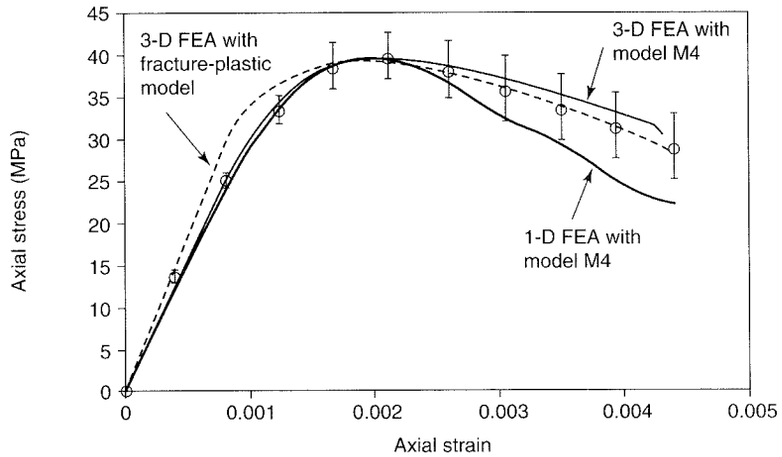


Figure 25.5 Variation of tangential shear modulus with axial stress



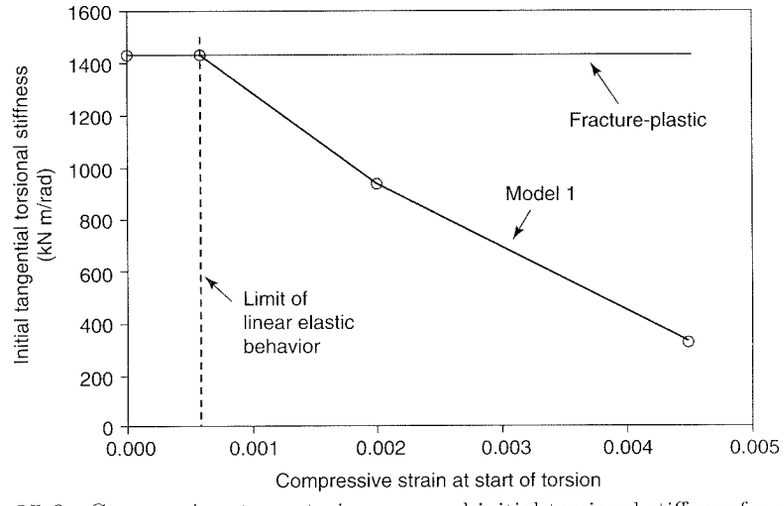


Figure 25.6 Compressive stress-strain curve and initial torsional stiffness for concrete

25.4.3 Experimental Observations and Numerical Results for Concrete

For concrete, cruciform columns buckling in torsion cannot be produced. A stiff testing machine that can suddenly superpose torsion on high compressive deformation is needed. Exploiting the development of such a machine, Caner, Bažant and Červenka (2001) recently tested concrete cylinders which were first subjected to uniaxial compressive stress reaching to the peak compressive load and well into the postpeak softening range. Solid cylinders were used because of the difficulty of producing thin walls with concrete and keeping such specimens stable in the postpeak range. Consequently, the overall response in these tests was simulated with a three-dimensional finite element code. However, the initial tangent plastic shear modulus G_t can be easily and unambiguously deduced from the measured initial torsional stiffness (assuming the initial distribution of shear strains along the radius to be linear).

The measured values of initial torsional stiffness are plotted in Figure 25.6, where they are compared with the predictions of two models for concrete: (1) microplane model M4, and (2) the plastic fracturing model of Červenka, Červenka and

Eligehausen (1998), which is one of the constitutive models of the classical type, i.e. formulated in terms of tensors and their invariants and utilizing only a few loading surfaces, of which there is only one in the deviatoric space. Models of the latter type are inherently incapable of simulating the vertex effect, and this is confirmed by the computed values in Figure 25.6. On the other hand, the results for the microplane model accurately predict the measured initial torsional stiffness. The results show that the vertex effect in strain softening concrete is very large.

PROBLEMS

Problem 25.1: Prove in detail that microplane model MP1 in the state of saturated yielding satisfies the static constraint.

Problem 25.2*: Set up a simple thermodynamic framework for a class of kinematically constrained microplane models that includes model MP1.
Hint: Start from the assumption that the free energy density and the dual dissipation potential can be expressed as integrals over all microplanes of certain microplane-related potentials.

Part V

Time-Dependent Inelastic Behavior of Metals and Concrete

26

Models for Localization of Softening and Size Effect

Classical plasticity, as well as any theory in which the material behavior is fully characterized in terms of stresses and strains (without reference to any characteristic length), exhibits no size effect. This means that the nominal strength (defined as the maximum load divided by the characteristic cross section area of the structure) is independent of structure size D. This chapter is focused on problems in which the nominal strength depends on D, i.e. the size effect is present. Such problems are of two kinds – those where the size effect is caused by material softening (damage), and those where it is not. The former is addressed first, and the latter is postponed to Section 26.4.

26.1 LACK OF OBJECTIVITY OF CONSTITUTIVE MODELS WITH SOFTENING

In classical plasticity, hardening has been considered as a natural material behavior, and softening has not received much attention. However, many non-metallic engineering as well as natural materials exhibit under certain load scenarios a strain-softening behavior. Typical examples are concrete and rock under tension or unconfined compression, or heavily consolidated soils under shear. The idea of modeling such materials as strain-softening continua, which emerged in the 1960s, was not immediately accepted by the entire scientific community. It was necessary to overcome a number of objections regarding the ill-posedness and lack of objectivity of such formulations. Indeed, most of the early analyses were not truly objective and, upon mesh refinement, their results would not have converged to a meaningful solution. Instead of going into theoretical details, let us explain the nature of the problem by a simple one-dimensional example.

Consider a straight bar of a constant cross section A and of total length L, subjected to uniaxial tension; see Figure 26.1. The material is assumed to be elastoplastic with isotropic softening. The yield condition can be written in the form

$$f(\sigma, \kappa) \equiv |\sigma| - h(\kappa) = 0 \tag{26.1}$$

where κ is the cumulative plastic strain (defined by the rate equation $\dot{\kappa} = |\dot{\varepsilon}_p|$), and $h(\kappa)$ is the function describing the evolution of the current yield stress. Under monotonic tensile loading, κ is equal to the actual plastic strain, ε_p. For the simplest case of linear softening, the stress-strain diagram is shown in Figure 26.2(a). Yielding

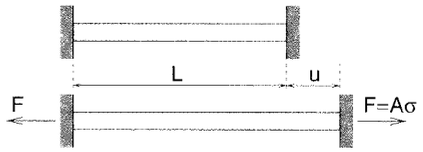


Figure 26.1 Bar under uniaxial tension

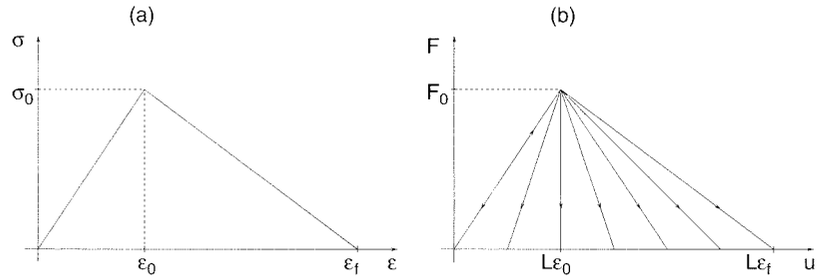


Figure 26.2 a) Stress-strain diagram with linear softening, b) fan of possible post-peak branches of the load-displacement diagram

starts at strain $\varepsilon_0 = \sigma_0/E$, where σ_0 is the initial yield stress and E is Young's modulus of elasticity. The yield stress is then gradually decreasing until it vanishes at $\varepsilon = \varepsilon_f$. For still larger strains, the yield stress remains equal to zero, which means that the material has completely lost cohesion and cannot carry any stresses. So, the softening law reads

$$h(\kappa) = \langle \sigma_0 + H\kappa \rangle \tag{26.2}$$

where $\langle \ldots \rangle$ is the positive-part operator, and $H = -\sigma_0/\varepsilon_f$ is the (negative) plastic modulus.

If the bar is loaded in tension by a displacement u applied at one of the supports, the response remains linearly elastic up to $u = u_0 \equiv L\varepsilon_0$. At this state, the force F transmitted by the bar (reaction at the support) attains its maximum value, $F_0 = A\sigma_0$. After that, the resistance of the bar is degrading. At each cross section, the stress can decrease either at increasing strain (by softening) or at decreasing strain (by elastic unloading). The equation of equilibrium implies that the stress distribution must remain uniform along the bar. However, any stress $\bar{\sigma}$ between zero and σ_0 can be generated by infinitely many strain paths that are bounded by two extreme cases – elastic unloading from the peak to strain ε_u, and sustained softening to strain ε_s; see Figure 26.3(a). This means that the strain distribution need not be uniform. For example, any piecewise constant strain distribution that jumps between the two extreme strain values ε_u and ε_s represents a valid solution; see Figure 26.3(b). Let us denote by L_s the cumulative length of the softening regions and by $L_u = L - L_s$ the cumulative length of the unloading regions. When the stress vanishes, the strain in the softening region is $\varepsilon_s = \varepsilon_f$ and the strain in the unloading region is $\varepsilon_u = 0$; therefore, the total elongation of the bar is $u_f = L_s\varepsilon_s + L_u\varepsilon_u = L_s\varepsilon_f$. The length L_s is undetermined – it can have any value between zero and L. This means that the problem has infinitely many solutions, and the corresponding post-peak branches of the load-displacement diagram fill the fan shown in Figure 26.2(b). This fan is bounded on one side by the solution with uniform softening ($u_f = L\varepsilon_f$) and on the other side

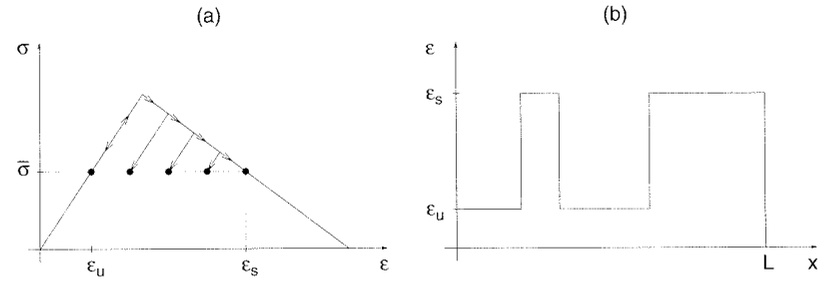


Figure 26.3 a) Multiple strain values corresponding to the same stress level, b) piecewise constant strain distribution

by the solution with uniform unloading ($u_f = 0$). The latter limit represents the case when the bar is unloaded just before any damage takes place. Any other solution represents a possible process in which a part of the bar completely loses its residual strength and the bar fails. It is not immediately clear which of these solutions is the 'correct' one, i.e. which of them reflects the actual failure process.

The ambiguity is removed if imperfections are taken into account. The real material properties and cross sectional dimensions cannot be perfectly uniform. Suppose that the strength in a small region is slightly lower than in the remaining portion of the bar. When the applied stress reaches the reduced strength, softening starts and the stress decreases. Consequently, the material outside the weaker region must unload elastically, because its strength has not been exhausted. This leads to the conclusion that the size of the softening region is dictated by the size of the region with minimum strength. Such a region can be arbitrarily small, and so the corresponding softening branch is arbitrarily close to the elastic branch of the load-displacement diagram. Thus, the standard strain-softening continuum formulation leads to a solution that has several pathological features (Bažant, 1976):

1. The softening region is infinitely small.
2. The load-displacement diagram always exhibits snapback, independently of the size of the structure and of the ductility of the material.
3. The total amount of energy dissipated during the failure process is zero.

From the mathematical point of view, these annoying features are related to the so-called *loss of ellipticity* of the governing differential equation. In the present one-dimensional context, loss of ellipticity occurs when the tangent modulus ceases to be positive. The boundary value problem becomes ill-posed, i.e. it does not have a unique solution with continuous dependence on the given data (an infinitesimal change in the data can cause a finite change in the solution). From the numerical point of view, ill-posedness is manifested by pathological sensitivity of the results to the size of finite elements. For example, consider a bar discretized by N_e two-node elements with a linear displacement interpolation and, consequently, with a constant strain interpolation. Suppose that the weakest section is located at the bar center. If the numerical algorithm properly captures the most localized solution, the softening region extends over one element, and we have $L_s = L/N_e$. Therefore, the post-peak branch strongly depends on the number of elements, and it approaches the initial elastic branch as the number of elements tends to infinity; see Figure 26.4(a).

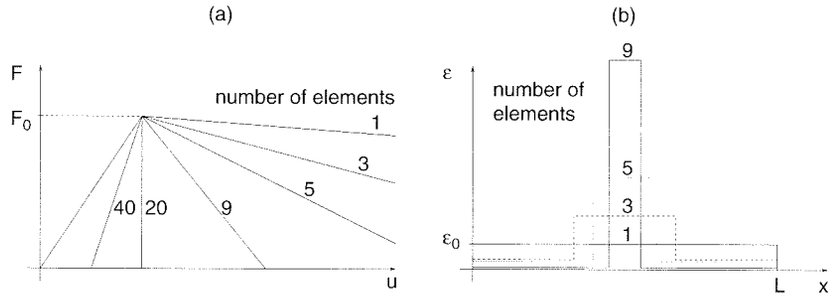


Figure 26.4 Effect of mesh refinement on the numerical results: a) load-displacement diagrams, b) strain distributions

constructed for a stress-strain law with $\varepsilon_f/\varepsilon_0 = 20$. The strain distributions at $u = 2u_0$ for various mesh refinements are plotted in Figure 26.4(b), under the assumption that the imperfection is located at the center of the bar. In the limit for an infinitely fine mesh, they tend to $2u_0\delta(x - L/2)$, where $\delta(x - L/2)$ denotes the Dirac distribution centered at $x = L/2$. The limit solution represents a displacement jump at the center, with zero strain everywhere else.

26.2 ATTAINING OBJECTIVITY WITH MESH-ADJUSTED SOFTENING MODULUS

Pathological sensitivity of the finite element results to the element size is not acceptable and must be avoided. The simplest remedy, frequently used in engineering applications, is based on an adjustment of the stress-strain diagram depending on the size of the element. Such techniques were proposed for shear softening in plasticity by Pietruszczak and Mróz (1981), and for tensile softening due to smeared cracking by Bažant and Oh (1983); the latter is known as the *crack band model*. It is best explained using the concept of a cohesive zone.

For many materials (with the exception of perfectly brittle ones), the formation of a macroscopic stress-free crack is preceded by the development of a fracture process zone, i.e. a region characterized by a highly localized strain and by the development and growth of microcracks or other defects (voids or slips), which reduce the cohesion of the material and lead to softening. Cohesive models lump the inelastic effects and replace the process zone by a discontinuity surface, across which the displacement field has a jump. The displacement jump is described by the *separation vector*, which represents the sum of contributions of all the microdefects over the thickness of the process zone. If the effect of strains in the plane tangent to the discontinuity surface is neglected, the traction transmitted by the discontinuity can be considered as a function of the separation vector. The resulting traction-separation law provides an objective (albeit somewhat simplified) description of the decohesion process. Models of this type originated in the pioneering work of Dugdale (1960) and Barenblatt (1962). They have been developed for metals and composites under the name of *cohesive zone models* (Needleman, 1987a; Tvergaard and Hutchinson, 1992) and for concrete under the name of *fictitious crack models* (Hillerborg, Modéer and Peterson, 1976).

The cohesive traction and the separation vector can be decomposed into the components normal and tangential with respect to the discontinuity surface. If the tangential components are negligible, it is sufficient to specify the relation between the normal traction, t_n, and the normal component of the displacement jump (opening of the cohesive crack), $[[u_n]]$. The area under the traction-separation curve represents the energy required for a complete decohesion of the material per unit area of the discontinuity surface, and it is often referred to as the Mode-I fracture energy, G_F.

The numerical solution of the decohesion process can be based on an explicit representation of the displacement discontinuity, either by separation of neighboring finite elements along their interfaces, or by special discontinuous enhancements of the finite element interpolation. Alternatively, it is possible to smear the displacement discontinuity over a certain finite distance, L_s, and transform it into inelastic strain. This directly leads to a stress-strain law with softening. However, it is important to bear in mind that the objective description of the decohesion process is given by the traction-separation law, and the corresponding stress-strain law depends on the smearing distance, L_s. To obtain the desired correspondence between the discrete model (with a displacement jump) and the smeared model (with a finite band of localized strain), the smearing distance must be equal to (or at least a good approximation of) the width of the numerically resolved localization band, which depends on the discretization scheme, e.g. on the finite element mesh. Therefore, the parameters defining the softening part of the stress-strain law (e.g. the softening modulus) cannot be considered as objective material properties.

The technique schematically illustrated in Figure 26.5 adjusts the softening law depending on the local mesh characteristics, i.e. it works with a *mesh-adjusted softening modulus*. For the sake of simplicity, we consider a uniaxial situation. The traction normal to the discontinuity, t_n, corresponds to the applied uniaxial stress σ, and the separation is characterized by the displacement jump, $[[u]]$. Suppose that the traction-separation law reads

$$\sigma = f_u([[u]]) \tag{26.3}$$

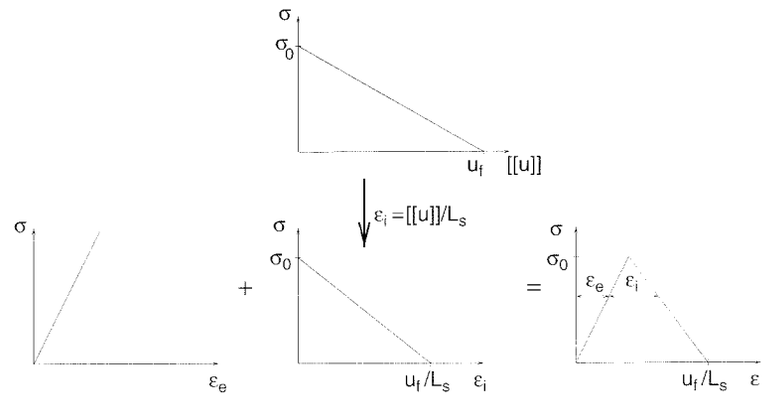


Figure 26.5 Stress-strain diagram derived from a cohesive traction-separation law

If the separation is smeared over a distance L_s, the corresponding inelastic strain is

$$\varepsilon_i = \frac{[[u]]}{L_s} \tag{26.4}$$

Therefore, the softening law for the smeared (continuum) model must be written as

$$\sigma_Y = f_u(L_s \varepsilon_i) \tag{26.5}$$

The softening modulus

$$H = \frac{\mathrm{d}\sigma_Y}{\mathrm{d}\varepsilon_i} = L_s f_u'(L_s \varepsilon_i) \tag{26.6}$$

depends on the slope of the traction-separation diagram, f_u', which is an objective material property, and on the smearing distance, L_s, which depends on the chosen finite element mesh.

The appropriate value of L_s must be deduced from the mesh characteristics. As shown in Section 26.1, the inelastic strain in the uniaxial problem localizes into a single element, and so L_s is equal to the length of that element. In multiple dimensions, the inelastic strain in general localizes into a band of elements running across the mesh (for concrete often called the *crack band*). Usually, the band is the smallest possible pattern that still allows separation of nodes on its opposite sides. The correct value of L_s is affected not only by the mesh size, but also by the inclination of the crack band with respect to the mesh lines. This is illustrated in Figure 26.6(b), which explains why the correct value of L_s for a band propagating along the diagonals of a regular square mesh is $\sqrt{2}$ times larger than for a band aligned with the element sides (Figure 26.6(a)).

Based on similar considerations, Bažant (1985) proposed the basic rules for the choice of the equivalent element size, L_s, and later Rots (1988) refined them for a number of special situations on the basis of extensive numerical studies. A more rigorous approach was developed by Oliver (1989). In practical simulations, it seems to be reasonable to compute L_s as the size of the element projected onto the crack normal; see Figure 26.6(c). When concrete with a dense reinforcing mesh is modeled as a continuum, the L_s value (as well as the fracture energy) is greatly affected by the properties of steel bars (Bažant and Cedolin, 1983).

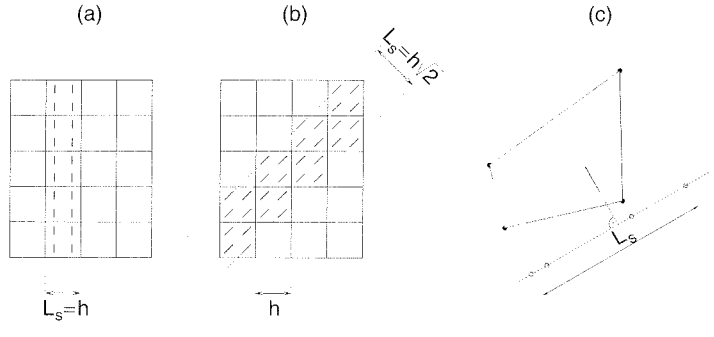


Figure 26.6 a) Localization band parallel to element sides, b) localization band across element diagonals, c) equivalent element size

Special care must be taken of higher-order elements, for which the effective width of the crack band also depends on the integration scheme. For example, in a 3-node uniaxial element with a quadratic displacement interpolation and a 3-point Gauss integration, only two integration points are softening, while the third is unloading, and the effective width of the softening region is 13/18 times the length of the element.

The above considerations apply to uniaxial stress states (not necessarily in one dimension) and to softening that starts immediately at the onset of nonlinearity. In general, the onset of localization can be detected by eigenvalue analysis of the so-called *acoustic tensor* (Rudnicky and Rice, 1975). Before localization, the constitutive law formulated as a stress-strain relation is considered as objective, mesh-independent, and the adjustment according to the element size is done only in the localized regime. In a general multiaxial case, it is not easy to detect the onset of localization and to define which geometric characteristic of the element should play the role of the size L_s. These problems can be bypassed if the constitutive model is enriched by a localization limiter that imposes a certain size of the localization band considered as an additional material parameter.

The concept of mesh-adjusted softening modulus has been explained for crack band models with smeared cracking, but very similar considerations can be done for other types of constitutive models. In softening plasticity, the inelastic strain is interpreted as the plastic strain rather than the cracking strain, and the direction normal to the crack is replaced by the direction normal to the band of localized strain, determined by localization analysis.

26.3 LOCALIZATION LIMITERS

In the preceding section we have described a simple technique reducing the sensitivity of numerical results to the finite element mesh (or any other type of computational grid). The approach based on a mesh-dependent softening modulus has gained a wide popularity in engineering applications, because it is easy to understand and implement. The load-displacement diagram and the energy dissipated by fracture are no longer pathologically sensitive to the mesh size, but fracture still localizes into a single layer of elements. Consequently, the results may suffer by sensitivity to the element shape and orientation. The formulas for the effective size of a general element are only approximative and partly empirical. These deficiencies are removed or alleviated by more sophisticated techniques that enrich the standard continuum and supply additional information on the internal structure of the material. Such techniques can enforce a realistic and mesh-independent size of the region of localized strain; therefore they are called *localization limiters*.

A wide class of localization limiters is based on the concept of a *nonlocal continuum*, which was introduced for elastic materials in the 1960s (Krumhansl, 1965; Rogula, 1965; Eringen, 1966a) and exploited as a localization limiter two decades later (Bažant, 1984b; Pijaudier-Cabot and Bažant, 1987). Nonlocal models abandon the assumption that the stress at a given point is uniquely determined by the history of strain and temperature at this point only. They take into account possible interactions with other material points. Theoretically, the stress at a point can depend on the strain history in the entire body, but the long-range interactions certainly diminish with increasing distance, and can be neglected when the distance exceeds a certain limit called the *interaction radius*.

Another class of localization limiters relies on enrichments by terms that contain gradients of the state variables. Some of these models use *gradients of internal variables*, e.g. of the cumulative plastic strain (Aifantis, 1984; Bažant, 1984b; Schreyer and Chen, 1986; Zbib and Aifantis, 1988; Belytschko and Lasry, 1989; Mühlhaus and Aifantis, 1991; Vardoulakis and Aifantis, 1991; de Borst and Mühlhaus, 1991, 1992; Pamin, 1994; Lorentz, 1999), while other models deal with *strain gradients*, i.e. higher-order gradients of the displacement field (Triantafyllidis and Aifantis, 1986; Fleck and Hutchinson, 1993; Chambon, Caillerie and Hassan, 1998). This latter group of models is in fact a particular case of generalized continua, such as *micromorphic continua* (Eringen, 1966b), *continua with microstructure* (Mindlin, 1964), or the *multipolar theory* (Green and Rivlin, 1964), which were all inspired by the pioneering work of the Cosserat brothers (Cosserat and Cosserat, 1909). The *Cosserat continuum* (or *micropolar continuum*) enhances the kinematic description of deformation by an additional field of local rotations, which can (but need not) depend on the rotations corresponding to the displacement field, i.e. on the skew-symmetric part of the displacement gradient for the small-displacement theory, or on the rotational part of the polar decomposition in the large-displacement theory. Micropolar extensions of softening continuum models can also serve as localization limiters (Mühlhaus and Vardoulakis, 1987; de Borst, 1991; Steinmann and Willam, 1992), but only for localization modes that activate the rotational degrees of freedom, i.e. not for pure splitting failure.

All the theories mentioned so far preserve the rate-independent nature of the material model. Incorporation of *rate-dependent* viscous terms (Needleman, 1988; Loret and Prevost, 1990; Sluys, 1992; Dubé, Pijaudier-Cabot and La Borderie, 1996; Wang, Sluys and de Borst, 1997) or delayed damage (Ladevèze, 1992) can limit localization in dynamic problems and delay localization in quasi-static problems. However, as the rate of loading decreases, the regularizing effect fades away.

26.3.1 Nonlocal Continuum

A computationally efficient and theoretically sound localization limiter is provided by the concept of nonlocal averaging, which is in principle applicable to any type of constitutive model. The idea of a nonlocal continuum originally appeared in elasticity (Krumhansl, 1965; Rogula, 1965; Eringen, 1966a; Kröner, 1967) and in phenomenological hardening plasticity (Eringen, 1981, 1983). Early studies on nonlocal elasticity, motivated by homogenization of the atomic theory of Bravais lattices, aimed at a better description of phenomena taking place in crystals on a scale comparable to the range of interatomic forces. They showed that nonlocal continuum models can approximate the dispersion of short elastic waves and improve the description of interactions between crystal defects such as vacancies, interstitial atoms and dislocations.

During the last quarter of a century, it became clear that neither distributed damage in materials nor transitions to discrete microstructural models can be adequately characterized by local constitutive relations between stress and strain tensors. Early extensions of the nonlocal concept to strain-softening materials, leading to the so-called imbricate continuum (Bažant, 1984b), were later improved by the nonlocal damage theory (Pijaudier-Cabot and Bažant, 1987; Bažant and Pijaudier-Cabot, 1988) and adapted for concrete (Saouridis, 1988; Saouridis and Mazars, 1992).

Nonlocal formulations were elaborated for a wide spectrum of models, including softening plasticity (Bažant and Lin, 1988b; Vermeer and Brinkgreve, 1994; Nilsson, 1994; Planas, Guinea and Elices, 1996; Strömberg and Ristinmaa, 1996; Nilsson, 1997; Borino, Fuschi and Polizzotto, 1999), hardening crystal plasticity (Kratochvil, 1988; Gao and Huang, 2001), progressively cavitating porous plastic solids (Leblond, Perrin and Devaux, 1994; Tvergaard and Needleman, 1995; Needleman and Tvergaard, 1998), smeared crack models (Bažant and Lin, 1988a; Jirásek and Zimmermann, 1998), and microplane models (Bažant and Ožbolt, 1990; Ožbolt and Bažant, 1996). Physical interpretation of nonlocal averaging was studied among others by Bažant (1987b, 1994), Drugan and Willis (1996) and Lacy, McDowell and Talreja (1999).

Generally speaking, integral-type nonlocal models replace one or more variables (typically, state variables) by their nonlocal counterparts obtained by weighted averaging over a spatial neighborhood of each point under consideration. If $f(\boldsymbol{x})$ is some 'local' field in a domain V, the corresponding nonlocal field is defined as

$$\overline{f}(\boldsymbol{x}) = \int_V \alpha(\boldsymbol{x},\boldsymbol{\xi}) f(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \tag{26.7}$$

where $\alpha(\boldsymbol{x},\boldsymbol{\xi})$ is a given *nonlocal weight function*. In an infinite isotropic body, the weight function depends only on the distance between the 'source', $\boldsymbol{\xi}$, and the 'receiver', $\boldsymbol{x}$. In the vicinity of a boundary, the weight function is usually scaled such that the nonlocal operator does not alter a uniform field. This can be achieved by setting

$$\alpha(\boldsymbol{x},\boldsymbol{\xi}) = \frac{\alpha_0(\|\boldsymbol{x}-\boldsymbol{\xi}\|)}{\int_V \alpha_0(\|\boldsymbol{x}-\boldsymbol{\zeta}\|)\,\mathrm{d}\boldsymbol{\zeta}} \tag{26.8}$$

where $\alpha_0(r)$ is a monotonically decreasing non-negative function of the distance $r = \|\boldsymbol{x}-\boldsymbol{\xi}\|$.

Based on the analysis of microcrack interactions, Bažant (1994) suggested replacing the traditional averaging operator (26.7) by a more complicated implicit scheme, which takes into account not only the distance between $\boldsymbol{x}$ and $\boldsymbol{\xi}$, but also the relative orientation of principal strain axes at these points. Ožbolt and Bažant (1996) reported that this modification leads to an improved performance, especially when tensile and shear fractures are to be simulated with the same material model. However, the numerical implementation seems to be quite complicated. In the present section, we restrict our attention to standard nonlocal averaging (26.7).

The weight function is often taken as the Gauss distribution function

$$\alpha_0(r) = \exp\left(-\frac{n_{dim} r^2}{2\ell^2}\right) \tag{26.9}$$

where ℓ is called the *internal length* of the nonlocal continuum, and n_{dim} is the spatial dimension of the problem (1, 2, or 3). Another possible choice is the bell-shaped function

$$\alpha_0(r) = \begin{cases} \left(1-\dfrac{r^2}{R^2}\right)^2 & \text{if } 0 \le r \le R \\ 0 & \text{if } R \le r \end{cases} \tag{26.10}$$

where R is a parameter proportional to the internal length. Since R corresponds to the largest distance of point $\boldsymbol{\xi}$ that affects the nonlocal average at point $\boldsymbol{x}$, it is called

the *interaction radius*. Theoretically, the Gauss function (26.9) has an unbounded support, i.e. its interaction radius is $R = \infty$. In practical calculations, the weight function is truncated at the distance where its value becomes negligible.

The length scales imposed by different types of weight functions can be compared if the internal length ℓ for an arbitrary function $\alpha_0(r)$ is defined as the central polar radius of inertia of an infinite body of mass density $\alpha_0(r)$, where $r = \|\boldsymbol{x}\|$ is the radial coordinate (distance from the center of mass). In an n_{dim}-dimensional space we have

$$\ell = \sqrt{\frac{\displaystyle\int_0^\infty r^{n_{dim}+1}\,\alpha_0(r)\,dr}{\displaystyle\int_0^\infty r^{n_{dim}-1}\,\alpha_0(r)\,dr}} \tag{26.11}$$

Parameter ℓ of the Gauss function (26.9) satisfies relation (26.11) for any spatial dimension. For the bell-shaped function (26.10), we obtain in one dimension $\ell = R/\sqrt{7}$, in two dimensions $\ell = R/2$, and in three dimensions $\ell = R/\sqrt{3}$. A normalized plot of both weight functions in one dimension is shown in Figure 26.7.

From the purely phenomenological point of view, the choice of the variable to be averaged remains to some extent arbitrary. For example, there exist nonlocal formulations that average the total strain, the plastic strain, the damage parameter, the compliance parameter, the inelastic stress, the secant stiffness tensor, the damage energy release rate, or the equivalent strain. One must be careful when selecting a certain formulation from the literature, because almost all of them capture properly the onset of localization (bifurcation from a uniform state), but some fail to give physically reasonable results at later stages of the localization process (Jirásek, 1998).

A simple nonlocal formulation of plasticity with isotropic strain softening is obtained if the current yield stress is computed from the nonlocal average of the softening variable κ (cumulative plastic strain). The yield function (20.12) is replaced by

$$f(\boldsymbol{\sigma}, \bar{\kappa}) \equiv F(\boldsymbol{\sigma}) - h(\bar{\kappa}) \tag{26.12}$$

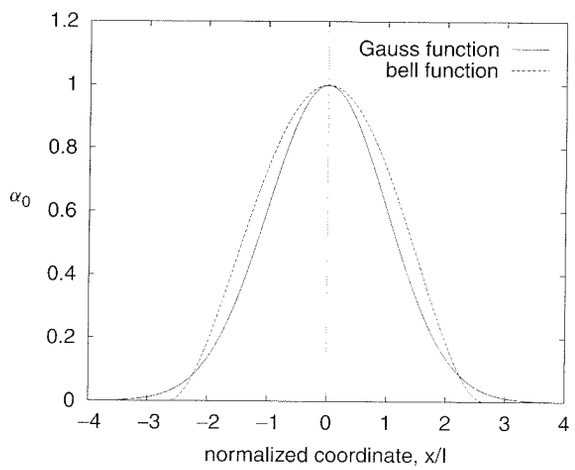


Figure 26.7 Nonlocal weight functions α_0

where

$$\bar{\kappa}(\boldsymbol{x}) = \int_V \alpha(\boldsymbol{x}, \boldsymbol{\xi})\kappa(\boldsymbol{\xi})\,d\boldsymbol{\xi} \tag{26.13}$$

is the nonlocal softening variable. The flow rule and the evolution law for the local softening variable κ are kept unchanged.

As shown by Planas, Elices and Guinea (1993), this simple nonlocal plasticity model is in fact equivalent to a cohesive zone model. The plastic strain is still localized into one single cross section but, in contrast to the local model, the energy dissipation is finite, and its amount is controled by the initial yield stress, softening modulus, and value of the scaled weight function (26.8) at $\boldsymbol{x} = \boldsymbol{\xi}$.

The singular character of the strain distribution generated by this particular nonlocal model is an exception. As a rule, nonlocal models lead to solutions characterized by a high degree of regularity. This is the case, for instance, for nonlocal damage and smeared crack models. For nonlocal softening plasticity, regularity can be achieved by a modification of the variable driving the yield stress degradation. Vermeer and Brinkgreve (1994) proposed defining the variable $\bar{\kappa}$ that enters the yield condition (26.12) as a linear combination of the local cumulative plastic strain and its nonlocal average:

$$\bar{\kappa}(\boldsymbol{x}) = (1-m)\kappa(\boldsymbol{x}) + m\int_V \alpha(\boldsymbol{x}, \boldsymbol{\xi})\kappa(\boldsymbol{\xi})\,d\boldsymbol{\xi} \tag{26.14}$$

The same definition of softening variable was also considered by Strömberg and Ristinmaa (1996) and by Planas, Elices and Guinea (1996).

Note that equation (26.13) is obtained as a special case with $m = 1$. Another special case with $m = 0$ corresponds to the local softening plasticity model. One might expect the typical values of m used by the generalized model to lie between these two cases but, interestingly, the plastic region has a nonzero width only for $m > 1$.

Figure 26.8(a) shows the load-displacement diagram for the bar under uniaxial tension (Figure 26.1) analyzed with the nonlocal plasticity model using $m = 2$. Strain distributions generated by an applied displacement $u = 3u_0$ are plotted in

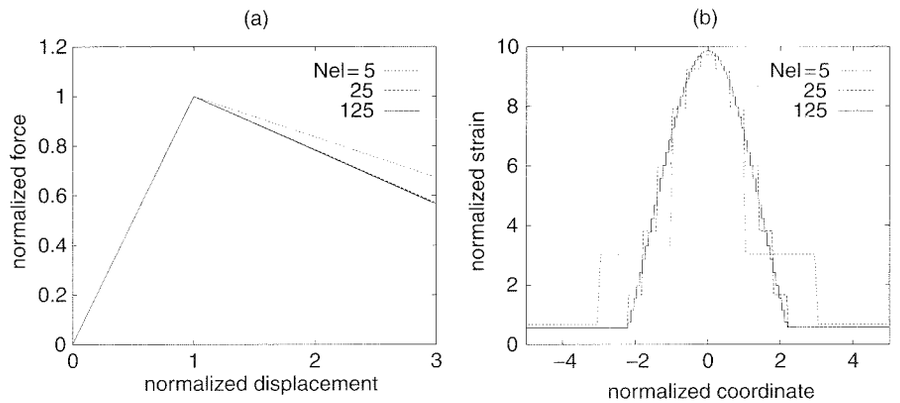


Figure 26.8 Convergence of a) load-displacement diagram, b) strain distribution; Nel = number of elements

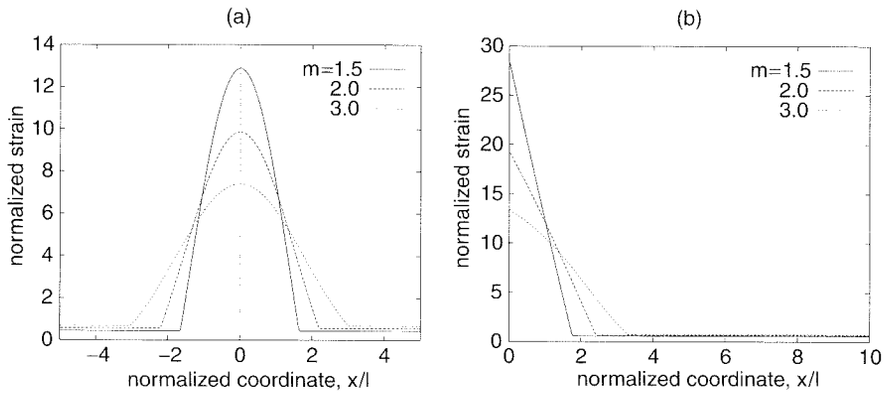


Figure 26.9 Strain distribution according to nonlocal plasticity for different values of parameter m: a) plastic region inside the bar, b) plastic region at the boundary

Figure 26.8(b). With an increasing number of finite elements, the solution rapidly converges to the exact one.

The strain distributions generated by the applied displacement $u = 3u_0$ for various values of m are plotted in Figure 26.9. Far from the boundary, the distribution of plastic strain inside the plastic region is symmetric (Figure 26.9(a)). If the plastic region is adjacent to the physical boundary of the bar, the largest plastic strain is attained right on the boundary (Figure 26.9(b)). The width of the plastic region is an increasing function of the parameter m and tends to zero for $m \to 1^+$. For a given m, the width of the plastic region is proportional to the internal length imposed by the nonlocal averaging function. It is worth noting that the shape of the plastic strain profile does not depend on the softening modulus, and that the width of the plastic region does not change during the softening process (in our simple one-dimensional example).

Since the yield function (26.12) depends on the nonlocal softening variable, the loading-unloading conditions have a nonlocal character, and the stress return algorithm cannot be performed at each Gauss integration point separately. This is the price to pay for the enhancement of the standard continuum description, which restores the well-posedness of the boundary value problem and leads to numerical solutions that do not exhibit the pathological sensitivity to the finite element discretization.

26.3.2 Models with Gradients of Internal Variables

The gradient models can be considered as the differential counterpart of integral nonlocal formulations. Instead of introducing nonlocality through spatial interaction integrals, we can account for it by incorporating higher-order gradients into the constitutive model. It is important to distinguish between two categories of gradient models:

1. *Explicit gradient models* directly enhance the constitutive equations or evolution laws by certain gradient terms.
2. *Implicit gradient models* enhance the constitutive equations or evolution laws by a nonlocal field, which is defined implicitly as the solution of a differential equation with the local field on the right-hand side.

In mathematical language, explicit gradient models belong to the class of *weakly nonlocal theories*, in the sense that the constitutive response of each material point depends only on an infinitesimal neighborhood of that point (because the gradients can be calculated from the distribution of the function in an arbitrarily small neighborhood). In contrast to that, implicit gradient models along with integral-type nonlocal models belong to the class of *strongly nonlocal theories*, in the sense that the constitutive response of each material point depends on the entire body, or at least on a finite neighborhood of that point (for integral-type nonlocal models using a weight function with a bounded support). For precise mathematical definitions of weak and strong nonlocality, see Rogula (1982).

Explicit Gradient Models

As a prototype of an explicit formulation with second-order gradients of internal variables, consider the gradient-dependent softening plasticity theory developed by Zbib and Aifantis (1988), Mühlhaus and Aifantis (1991), de Borst and Mühlhaus (1991, 1992), Pamin (1994), and others. The yield stress is assumed to depend not only on the value of the softening variable κ (cumulative plastic strain), but also on its Laplacean. The yield function can again be written in the form (26.12), but the 'nonlocal' softening variable is now defined by the differential expression

$$\bar{\kappa} = \kappa + c\,\nabla^2\kappa \tag{26.15}$$

where c is a material parameter proportional to the square of the internal length, and

$$\nabla^2 = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla} = \sum_{i=1}^{n_{dim}} \frac{\partial^2}{\partial x_i^2} \tag{26.16}$$

is the Laplacean operator in n_{dim}-dimensional space. If the solution remains homogeneous, the plastic strain is constant in space, its second derivative vanishes, and (26.12) reduces to the original 'local' yield function. After the onset of localization, the second-gradient term is activated and prevents localization of the plastic strain into a set of zero measure (i.e. into a surface in three dimensions, a curve in two dimensions, or a point in one dimension).

Let us illustrate the effect of the second-gradient term by analyzing the one-dimensional localization problem. We consider again the bar under uniaxial tension (Figure 26.1) and a material with linear softening (Figure 26.2(a)). Before the applied positive stress σ reaches the initial yield stress, the response is linearly elastic along the entire bar. At peak stress, some cross sections start yielding while others unload elastically. Let us assume that yielding takes place in a contiguous zone of length L_s. To satisfy equilibrium, the stress along the bar must be uniform, and so the stress rate is also uniform (i.e. constant in space). Inside the plastic zone, the stress rate must be equal to the rate at which the current yield stress degrades. Substituting the one-dimensional rate form of (26.15) into the consistency condition

$$\dot{f} = \dot{\sigma} - H\dot{\bar{\kappa}} = 0 \tag{26.17}$$

we obtain a second-order differential equation

$$H(\dot{\kappa} + c\,\dot{\kappa}'') = \dot{\sigma} \tag{26.18}$$

which is valid as long as $\kappa + c\,\kappa''$ remains smaller than $\varepsilon_f = \sigma_0/|H|$. Equation (26.18) is a linear differential equation with constant coefficients and, since the stress rate $\dot{\sigma}$ is independent of the spatial variable x, also with a constant right-hand side. Its general solution (for $H < 0$ and $c > 0$) is

$$\dot{\kappa}(x) = \frac{\dot{\sigma}}{H} + C\cos\frac{x - x_s}{l} \tag{26.19}$$

where $l = \sqrt{c}$ is a material constant with the dimension of length, and C and x_s are integration constants, to be determined from appropriate boundary conditions.

The cumulative plastic strain $\kappa(x)$ must be continuously differentiable with respect to the spatial coordinate x, otherwise its second derivative that appears in the yield function would not be bounded and the solution would not be plastically admissible. Consequently, the rate $\dot{\kappa}(x)$ must be continuously differentiable, too. The consistency condition (26.17) applies only to the plastic region, a region in which $\dot{\kappa}(x)$ given by (26.19) is positive. Inside the elastic region(s), the plastic strain rate vanishes. Continuity is preserved only if expression (26.19) vanishes at the internal boundary separating the elastic and plastic regions. To preserve continuous differentiability, the spatial derivative of the expression in (26.19) at the elasto-plastic interface must also vanish. This provides interface conditions for the determination of the integration constants. The peculiar feature of the explicit gradient model is that these conditions are not prescribed on the physical boundary of the body, but on the (a priori unknown) elasto-plastic interface. In a general multi-dimensional case, the condition imposed at the elasto-plastic interface is taken as

$$\frac{\partial\dot{\kappa}}{\partial n} \equiv \boldsymbol{n}\cdot\boldsymbol{\nabla}\dot{\kappa} = 0 \tag{26.20}$$

where $\boldsymbol{n}$ is the unit normal to the interface. This condition ensures C^1-continuity of the softening variable and simplifies the variational formulation of the problem.

Going back to the one-dimensional example, we look for the most localized solution with one plastic region surrounded by elastically unloading material. In general, the integration constants C and x_s should be solved from interface conditions $\dot{\kappa}'(x_l) = 0$ and $\dot{\kappa}'(x_r) = 0$ where x_l and x_r are the coordinates of the 'left' and 'right' boundary of the plastic region. These coordinates are not prescribed, but they must satisfy the conditions $\dot{\kappa}(x_l) = 0$ and $\dot{\kappa}(x_r) = 0$, where $\dot{\kappa}(x)$ is again the general solution (26.19) that contains the integration constants. In total, we have four coupled equations for four unknowns – two integration constants, C and x_s, and two coordinates of the elasto-plastic interface, x_l and x_r.

Instead of working out the algebraic details, we can construct the solution graphically, which will also increase our insight into the problem. Geometrically, the interface conditions mean that the graph of the harmonic function $C\cos((x - x_s)/l)$ must have a common tangent with the graph of the function $-\dot{\sigma}/H$, which is a straight horizontal line. Inside the plastic region, the graph of $C\cos((x - x_s)/l)$ must be above that line. This means that the constant $-\dot{\sigma}/H$ must be negative (i.e. the stress rate $\dot{\sigma}$ must be negative), and that $C = \dot{\sigma}/H$ while x_s is arbitrary; see Figure 26.10(a). At point x_s, the plastic strain rate attains its maximum, and so this point is at the center of the plastic region. The elasto-plastic interfaces are located at $x_l = x_s - \pi l$ and $x_r = x_s + \pi l$. Since parameter l controls the size of the localization zone, $L_s = x_r - x_s = 2\pi l$, it plays the role of a characteristic material

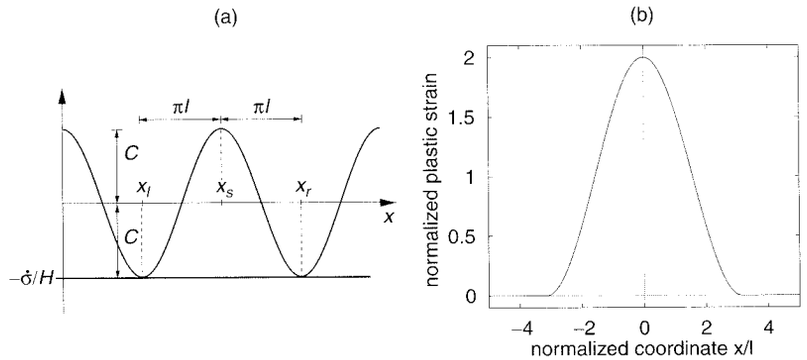


Figure 26.10 a) Graphical interpretation of the interface conditions, b) distribution of plastic strain according to the explicit gradient model

length. It is interesting to note that the size of the plastic region is independent of the softening modulus H, and the spatial distribution of the plastic strain rate remains proportional to $\cos((x - x_s)/l)$. The resulting plastic strain distribution, described by

$$\varepsilon_p(x) = \kappa(x) = \begin{cases} \dfrac{\sigma - \sigma_0}{H}\left(1 + \cos\dfrac{x - x_s}{l}\right) & \text{if } |x - x_s| \le \pi l \\ 0 & \text{if } |x - x_s| \ge \pi l \end{cases} \tag{26.21}$$

is plotted in Figure 26.10(b) for $x = x_s$. The total strain distribution is obtained by adding the uniform elastic strain, σ/E.

The graphical interpretation of the interface conditions reveals that the solution of the present localization problem is not unique. A valid solution exists for an arbitrary value of integration constant x_s, i.e. for an arbitrary position of the center of the plastic region (of course, at a sufficient distance from the physical boundary). All these solutions differ only by a shift along the bar axis. The source of nonuniqueness is in the highly idealized formulation of the problem, with the material and geometric properties assumed to be perfectly uniform along the bar. The actual position of the localized plastic region would be determined by random imperfections. If these imperfections are taken into account by the solution procedure, the harmonic curve and straight line that must have common tangents at the elastoplastic interfaces are slightly perturbed and the number of admissible solutions becomes finite (in a generic case).

It is instructive to discuss how the gradient term limits localization. Around the point that experiences the largest plastic strain, the second derivative of the softening variable is negative, and the Laplacean term in (26.15) makes $\bar{\kappa}$ smaller than κ. Consequently, the residual strength is higher than it would be for a standard local model. If the softening zone were too narrow, the negative curvature of the strain distribution around its peak would have a large magnitude, and the yield stress would be higher than at other points. Since the applied stress is constant along the bar, the sections around the center of the localization zone could not yield, and the zone would grow laterally, in order to reduce the magnitude of the negative curvature and thus restore a constant value of the current yield stress.

Implicit Gradient Models

Due to the presence of second derivatives of internal variables, the numerical implementation of explicit gradient models is not easy. It is often necessary to use C^1-

continuous finite elements, or mixed elements with an independent interpolation of one or more primary unknown fields in addition to the usual displacement interpolation (Pamin, 1994; Comi and Perego, 1996). There are also problems related to the enforcement of the interface conditions at the evolving elasto-plastic boundary. New developments in this research area indicate that a more robust implementation can be achieved with implicit gradient models. Such formulations, first proposed for gradient damage (Peerlings *et al.*, 1996), have recently been adapted for gradient plasticity (Engelen, Geers and Baaijens, 2001) and even extended to large strain (Geers, Engelen and Ubachs, 2001).

The implicit gradient model for softening plasticity defines the nonlocal softening variable $\bar{\kappa}$ as the solution of the *Helmholtz differential equation*

$$\bar{\kappa} - c\,\boldsymbol{\nabla}^2\bar{\kappa} = \kappa \tag{26.22}$$

with the Neumann boundary condition

$$\frac{\partial\bar{\kappa}}{\partial n} \equiv \boldsymbol{n}\cdot\boldsymbol{\nabla}\bar{\kappa} = 0 \tag{26.23}$$

In contrast to the explicit gradient model, the boundary conditions are imposed on the actual physical boundary of the body, and $\boldsymbol{n}$ is thus the unit normal to the physical boundary.

As shown by Peerlings *et al.* (2001), the definition of $\bar{\kappa}$ as the solution of the boundary value problem (26.22)–(26.23) is equivalent to nonlocal averaging with the Green function of the boundary value problem playing the role of the weight function. For instance, in one dimension and on an infinite domain, the Green function of the Helmholtz equation (26.22) is given by

$$G(x,\xi) = \frac{1}{2l}\exp\left(-\frac{|x-\xi|}{l}\right) \tag{26.24}$$

where $l = \sqrt{c}$ is again a characteristic material length (different from, but proportional to the internal length of the nonlocal continuum defined in (26.11)). The graph of $G(x,\xi)$ for fixed $\xi = 0$ is shown in Figure 26.11. This function has similar properties

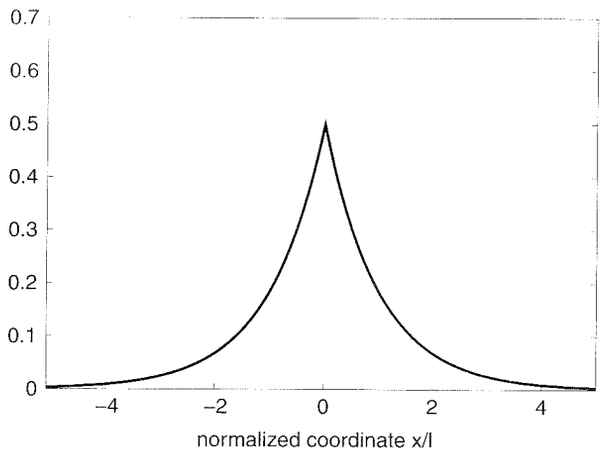


Figure 26.11 Green function of Helmholtz equation in one dimension

to the usual nonlocal weight functions – it depends only on the distance $|x-\xi|$, is non-negative and monotonically decreases with increasing distance. The only unusual feature of the Green function is that it is not continuously differentiable at $x = \xi$. In three dimensions, the Green function even becomes unbounded at $x = \xi$ (with a singularity of the $1/r$ type), but it remains square-integrable.

The nonlocal softening variable $\bar{\kappa}$, obtained as the solution of the boundary value problem (26.22)–(26.23), could directly be used in the yield function (26.12). However, the implicit gradient model would then exhibit the same behavior as the simple integral-type nonlocal plasticity model – plastic strain would localize into a single cross section. Engelen, Geers and Baaijens (2001) adopted a more sophisticated hardening-softening law, motivated by the concept of ductile damage. They postulated the yield condition in the form

$$f(\boldsymbol{\sigma},\kappa,\bar{\kappa}) \equiv F(\boldsymbol{\sigma}) - h(\kappa,\bar{\kappa}) = 0 \tag{26.25}$$

where

$$h(\kappa,\bar{\kappa}) = (\sigma_0 + H\kappa)[1-\omega_p(\bar{\kappa})] \tag{26.26}$$

Parameter H is here taken as positive, and the term $\sigma_0 + H\kappa$ represents linear plastic hardening. Due to the growth and coalescence of voids, the 'effective' yield stress is reduced by a scalar factor $1-\omega_p$, similar to the reduction of the effective stress in continuum damage mechanics. While the hardening process is driven by the local value of the cumulative plastic strain, the degradation of yield stress due to ductile damage is driven by the nonlocal average of the cumulative plastic strain. In contrast to damage mechanics, only the current yield stress is reduced, but the elastic stiffness remains constant. The damage parameter ω_p grows from 0 to 1, but the onset of ductile damage does not need to coincide with the onset of yielding. In the simplest case, the dependence of ω_p on $\bar{\kappa}$ is described by a piecewise linear law,

$$\omega_p(\bar{\kappa}) = \begin{cases} 0 & \text{for } \bar{\kappa} \le \kappa_i \\ \dfrac{\bar{\kappa}-\kappa_i}{\kappa_c-\kappa_i} & \text{for } \kappa_i \le \bar{\kappa} \le \kappa_c \\ 1 & \text{for } \bar{\kappa} \ge \kappa_c \end{cases} \tag{26.27}$$

in which κ_i and κ_c are material parameters that specify the (nonlocal) plastic strain at damage initiation and at complete damage (characterized by zero residual yield stress).

To illustrate the localization properties of the Geers–Engelen model, let us look again at the simple uniaxial tensile test. Initially, the material hardens and the strain distribution remains uniform. Bifurcation to a localized solution can occur only when the tangent stiffness ceases to be positive. This cannot happen before the damage threshold, $\kappa = \kappa_i$. Since the plastic strain distribution up to $\kappa = \kappa_i$ remains uniform, it is possible, for the purpose of bifurcation analysis, to set $\kappa_i = 0$ without any loss of generality. The general case could be obtained simply by replacing κ, κ_c and σ_0 by $\kappa - \kappa_i$, $\kappa_c - \kappa_i$ and $\sigma_0 + H\kappa_i$, respectively.

With parameter κ_i set to zero, the dependence of the yield stress on the local and nonlocal cumulative plastic strain is described by

$$h(\kappa,\bar{\kappa}) = (\sigma_0 + H\kappa)\left(1 - \frac{\bar{\kappa}}{\kappa_c}\right) \tag{26.28}$$

Differentiating this with respect to time and substituting the result into the consistency condition $\dot{f} = \dot{\sigma} - (\partial h/\partial\kappa)\dot{\kappa} - (\partial h/\partial\bar{\kappa})\dot{\bar{\kappa}} = 0$, we obtain the equation

$$H\left(1 - \frac{\bar{\kappa}}{\kappa_c}\right)\dot{\kappa} - (\sigma_0 + H\kappa)\frac{\dot{\bar{\kappa}}}{\kappa_c} = \dot{\sigma} \tag{26.29}$$

that must be valid inside the plastic region. For convenience, we introduce a parameter $\kappa_0 = \sigma_0/H$ and rewrite (26.29) as

$$(\kappa_0 + \kappa)\dot{\bar{\kappa}} - (\kappa_c - \bar{\kappa})\dot{\kappa} = -\frac{\kappa_c\dot{\sigma}}{H} \tag{26.30}$$

For the implicit gradient model, the local and nonlocal fields are linked by the one-dimensional version of the Helmholtz equation (26.22), and the same equation must hold for their rates. Substituting

$$\dot{\kappa} = \dot{\bar{\kappa}} - l^2\dot{\bar{\kappa}}'' \tag{26.31}$$

into (26.30), we can eliminate the local rate, $\dot{\kappa}$. The resulting equation,

$$(\kappa_0 + \kappa - \kappa_c + \bar{\kappa})\dot{\bar{\kappa}} + l^2(\kappa_c - \bar{\kappa})\dot{\bar{\kappa}}'' = -\frac{\kappa_c\dot{\sigma}}{H} \tag{26.32}$$

makes it possible to compute the rate $\dot{\bar{\kappa}}$ if the current distribution of κ and $\bar{\kappa}$ is given.

At damage initiation, we have $\kappa = \bar{\kappa} = 0$, and equation (26.32) reduces to

$$(\kappa_0 - \kappa_c)\dot{\bar{\kappa}} + l^2\kappa_c\dot{\bar{\kappa}}'' = -\frac{\kappa_c\dot{\sigma}}{H} \tag{26.33}$$

This is a second-order differential equation with constant coefficients, and the form of its general solution depends on the sign of $\kappa_0 - \kappa_c$. For $\kappa_c < \kappa_0$ we obtain

$$\dot{\bar{\kappa}}(x) = C_1\cos\left(\sqrt{\frac{\kappa_0 - \kappa_c}{\kappa_c}}\frac{x - x_s}{l}\right) - \frac{\kappa_c\dot{\sigma}}{(\kappa_0 - \kappa_c)H} \tag{26.34}$$

where C_1 and x_s are integration constants. Substituting (26.34) into (26.31), we get the distribution of the local rate of the softening variable,

$$\dot{\kappa}(x) = \dot{\bar{\kappa}}(x) - l^2\dot{\kappa}''(x) = \frac{\kappa_0}{\kappa_c}C_1\cos\left(\sqrt{\frac{\kappa_0 - \kappa_c}{\kappa_c}}\frac{x - x_s}{l}\right) - \frac{\kappa_c\dot{\sigma}}{(\kappa_0 - \kappa_c)H} \tag{26.35}$$

This solution is valid only inside the plastic region, characterized by $\dot{\kappa}(x) > 0$. In the elastic region, the rate of the yield function does not necessarily vanish, and so equation (26.29) is not applicable. However, equation (26.31) remains valid and $\dot{\kappa}(x)$ must vanish, which makes it possible to express

$$\dot{\bar{\kappa}}(x) = C_2\cosh\frac{x}{l} + C_3\sinh\frac{x}{l} \tag{26.36}$$

with two additional integration constants, C_2 and C_3 (which are of course different in each contiguous part of the elastic region). The nonlocal softening variable $\bar{\kappa}(x)$ must be continuously differentiable, which provides two interface conditions at each elasto-plastic boundary. On the physical boundary, condition (26.23) in one dimension reduces to $\bar{\kappa}' = 0$.

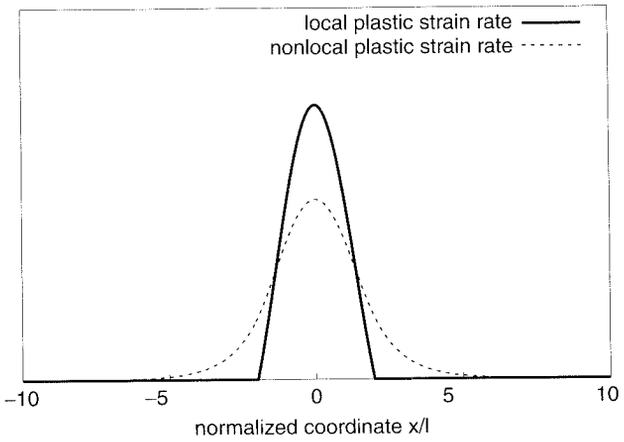


Figure 26.12 Distribution of the local and nonlocal plastic strain rates at bifurcation from a homogeneous state for the implicit gradient model

The boundary and interface conditions make it possible to determine the integration constants and construct the distribution of the plastic strain rate. The graphical interpretation of these conditions and the algebraic details are left to the reader as an excercise (Problem 26.2). Here we only present the final result for an infinite domain; see Figure 26.12. Let us emphasize that this solution describes only the distribution of the plastic strain rate at bifurcation from a homogeneous state. At later stages of softening, the size of the active plastic domain decreases.

26.4 PLASTICITY AT CONTINUUM LIMIT ON APPROACH TO NANO-SCALE

The nonlocal and gradient models discussed in Section 26.3 are motivated by localization of softening, and their principal aim is to achieve objectivity of continuum modeling and numerical simulation. There exist other enhanced continuum models motivated by the need to improve the classical continuum description in situations that require the incorporation of an internal length parameter even for some problems with no strain softening at all. Such models, for example, are needed to capture the dispersion of elastic waves for wave lengths approaching the scale of the atomic lattice (Eringen, 1972), to set the length scale of surface instabilities that develop in elastic bodies under large strains (Exadaktylos and Vardoulakis, 1998), or to reproduce various types of size effects without softening (Fleck and Hutchinson, 1993; Aifantis, 1999).

Since the main subject of this book is plasticity, it is appropriate to focus on the recently appeared strain-gradient models for metal plasticity, pioneered by Fleck and Hutchinson (1993, 1997) and refined by Gao *et al.* (1999), Huang *et al.* (2000) and Gao and Huang (2001). This family of models is motivated by very strong size effects observed on the micron scale in torsional tests of wires, Rockwell micro-hardness tests and micro-bending tests. The ordinary plasticity with which all the preceding chapters have been concerned cannot capture such size effects, but the gradient plasticity is a continuum model that can. Naturally, these size effects are important for miniature electronic devices and MEMS (micro-electro-mechanical systems).

The phenomenological strain-gradient plasticity theory of Fleck and Hutchinson (1997) is a nonlinear extension of the elastic theory of Toupin (1962) and Mindlin (1964, 1965). These early investigators developed a generalization of linear elasticity that uses the strain gradient as an additional observable state variable. The *strain gradient* $\boldsymbol{\eta}$ is a third-order tensor with components $\eta_{ijk} = \partial^2 u_k / \partial x_i \partial x_j$, and its conjugate force $\boldsymbol{\tau}$ is called the *higher-order stress*. The internal power density is enriched by the product $\boldsymbol{\tau} \bullet \dot{\boldsymbol{\eta}} = \tau_{ijk}\dot{\eta}_{ijk}$. The external power is enriched by the product of the so-called *surface double tractions* with the normal derivative of the displacement rate. Since both the strain $\boldsymbol{\varepsilon}$ and the strain gradient $\boldsymbol{\eta}$ are linked to the displacement field, the principle of virtual work (after integration by parts, which causes the spatial derivatives to move from the kinematic variables to the internal forces) furnishes only one set of momentum balance equations,

$$\frac{\partial \sigma_{i,j}}{\partial x_j} - \frac{\partial^2 \tau_{kji}}{\partial x_k \partial x_j} + \bar{b}_i = 0 \tag{26.37}$$

In addition to the standard terms, these equations contain the second derivatives of the higher-order stress $\boldsymbol{\tau}$. There are also two natural (Neumann-type) boundary conditions relating projections of the internal forces ($\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$) to the surface tractions and double tractions.

In the Toupin–Mindlin theory, the constitutive equations represent linear relations between $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$, and between $\boldsymbol{\tau}$ and $\boldsymbol{\eta}$, and thus are restricted to linear elastic behavior. For isotropic materials, there are, in addition to the two standard elastic moduli, five elastic constants that link $\boldsymbol{\tau}$ to $\boldsymbol{\eta}$ and have the dimension of stress times length squared. If they are independent, the square roots of their ratios to Young's modulus imply five material length parameters.

Fleck, Hutchinson and their coworkers developed a family of models extending the Toupin–Mindlin theory into the plastic range. They considered both the deformation theory of plasticity (Fleck *et al.*, 1994) and the flow theory of plasticity (Fleck and Hutchinson, 1993). In these initial publications, only the part of the strain gradient related to rotations was taken into account. The effect of the stretching gradients was added later (Fleck and Hutchinson, 1997).

In the deformation version of the theory (Fleck *et al.*, 1994), the free energy density is defined as a power function of the so-called *combined strain*, which is a scalar measure of the strain and strain-gradient tensors. Once this measure is defined, the stress-strain relations are obtained in the standard way as the state laws. No internal variables are needed and the dissipation need not be considered.

The deformation theory of plasticity of course cannot realistically describe unloading, nor path dependence. However, it is useful for monotonic loading paths that are almost proportional, and for the initial sharp deviations from such paths. Some special modifications are needed for incompressible materials – only the deviatoric parts of $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ are computed from the state laws, while the hydrostatic parts become Lagrange multipliers associated with the incompressibility constraint.

In parallel, the strain-gradient formulation was also developed for the flow theory of plasticity with kinematic hardening (Fleck and Hutchinson, 1993, 1997; Hutchinson, 1997). The structure of the constitutive equations is similar to the standard flow theory. The equivalent stress measure (generalization of von Mises stress) is also affected by the higher-order stress (more precisely, by its deviatoric part), and this dependence is then reflected in the associated flow rule for the plastic strain gradient.

This evolution law governs the dissipative force conjugate to the hardening scalar variable. The elastic part of the free energy potential is given by a quadratic expression taken from the Toupin–Mindlin theory, with five length parameters, and the definition of the equivalent stress contains additional material length parameters because it combines the stresses and higher-order stresses.

According to Fleck and Hutchinson, the five elastic material length parameters appearing in the elastic energy density expression may be considered as very small compared to the plastic length parameters appearing in the expression for the equivalent stress (and, consequently, in the yield condition). This means that the gradient terms affect the elastic solution only on an extremely small scale. On the other hand, once yielding starts, the gradient effects of plasticity become activated on a larger scale. Here one could see an analogy with the usual nonlocal models for quasibrittle materials, for which the elastic behavior ought to be modeled as local (Pijaudier-Cabot and Bažant, 1987).

Physical explanation of the strain-gradient effects in crystalline materials is based on the existence of two kinds of dislocations (line defects in the crystal lattice):

- The *statistically stored dislocations*, whose cumulative contributions to a homogeneous strain field cancel out; see Figure 26.13(b). Their effect is to reduce the tangent modulus for strain hardening.
- The *geometrically necessary dislocations*, which are required to make the plastic strain field nonuniform and, in particular, to produce a gradient of plastic strain (which is equivalent to curvature of the displacement field); see Figure 26.13(e).

The latter is a well-known phenomenon in materials science (Nye, 1953; Weertman and Weertman, 1964; Cottrell, 1964; Ashby, 1970). The point is that although dislocations are always likely to be present (Figure 26.13(b)), they are not necessary to produce the distorsion of a lattice by a uniform strain (Figure 26.13(c)). The geometrically

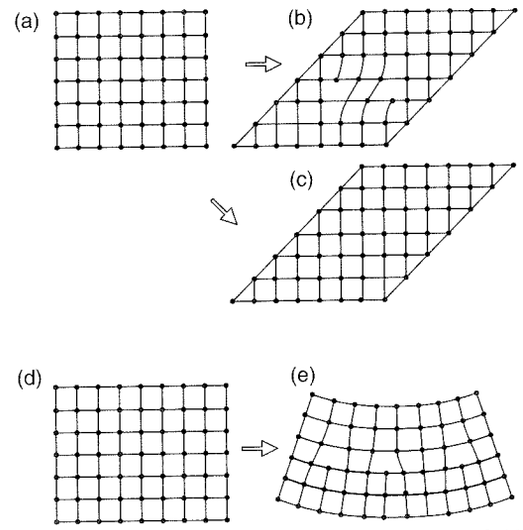


Figure 26.13 (a,d) Initial states of square lattices with 56 and 63 atoms, respectively; (b,c,e) plastically deformed lattices: (b,c) homogeneous strain; (e) strain gradient (lattice curvature); (b) 'statistically stored' edge dislocations, whose contributions to overall deformation cancel out; (e) geometrically necessary edge dislocations

necessary dislocations are the cause of the strain gradient effect occurring when the nanoscale is approached.

Recently, a number of authors have tried to derive gradient plasticity models from dislocation kinematics (Estrin *et al.*, 1998; Gao *et al.*, 1999; Huang *et al.*, 2000; Acharya and Bassani, 2000; Menzel and Steinmann, 2000). The main objective of the strain-gradient enrichment of plasticity is to capture the size effect observed in experiments, particularly (1) the micro-bending tests, (2) the micro-torsion tests of thin wires, and (3) the Rockwell micro-hardness of copper polycrystals and single crystals. Huang *et al.* (2000) demonstrated that the mechanism-based strain-gradient plasticity of Gao *et al.* (1999) meets this objective very well.

It has been shown (Bažant, 2001) that the asymptotic size effect for $D \to 0$ is of the type $\sigma_N \propto \lambda_\varepsilon^2 (b/D)^{5/2}$, which is a much stronger size effect than that in linear elastic fracture mechanisms. For the transition from the asymptotic small-size behavior to the classical plasticity, the following approximate size effect law with correct small-size and large-size asymptotes has been proposed (Bažant, 2001):

$$\sigma_N = C_0 \sigma_Y \left[1 + \left(\frac{D_0}{D}\right)^{5/r}\right]^{r/2} \tag{26.38}$$

in which D_0, C_0 and r are constants, C_0 and r being dimensionless; C_0 can be determined from the classical plasticity solution for large sizes, and D_0 and r can then be identified by matching either experiment a numerical solutions for very small sizes. It has also been shown that in the limit $D \to 0$ the dominant asymptotic term corresponds to a load-deflection curve of the form $P = ku^{3/2}$, where P is the load, u is the deflection under the load, and k is a constant. This curve begins with a zero tangent stiffness, and then the tangent stiffness gradually increases. Such asymptotic behavior does not seem reasonable, and so it must be observed that Gao *et al.*'s theory can only be realistic for not too small sizes.

PROBLEMS

Problem 26.1*: How does the solution of the one-dimensional localization problem for the explicit gradient plasticity model change when $\kappa + c\,\kappa''$ exceeds ε_f? How does the solution change if the linear softening law is replaced by an exponential law, in the local case written as $h(\kappa) = \sigma_0[1 - \exp(-\kappa/\varepsilon_f)]$?

Problem 26.2*: Find the distribution of the plastic strain rate at bifurcation from a homogeneous state for the Geers–Engelen implicit gradient plasticity model described in Section 26.3.2. Parameters of the ductile damage law (26.27) are $\kappa_i = 0$ and $\kappa_c = \kappa_0/2$. Solve the problem in one dimension on a) an infinite domain, b) a finite interval of size $L = 8l$. Is the position of the plastic region arbitrary?

Problem 26.3*: Solve the previous problem with parameters of the ductile damage law (26.27) give as $\kappa_i = 0$ and $\kappa_c = 2\kappa_0$.
Hint: The bifurcation does not occur at damage initiation.

Problem 26.4*: Compare the Vermeer–Brinkgreve integral-type nonlocal plasticity model and the Geers–Engelen implicit gradient plasticity model. Suppose that the Vermeer–Brinkgreve model is used with the Green function of the Helmholtz equation as the nonlocal weight function in (26.14), but the softening law is kept as linear. Show that, with an appropriate choice of the model parameters, the distribution of the plastic strain rate at bifurcation from a homogeneous state in one dimension would be exactly the same. In terms of the parameters of the Geers-Engelen model, what would be the appropriate choice of parameters m and H of the Vermeer-Brinkgreve model?

27

Viscoplasticity

All of the exposition of inelastic behavior has so far dealt with rate-independent material models, i.e. with models whose response to a given loading path does not depend on the time scale of the loading process. In reality, the material behavior is affected by the rate at which the loading is applied and by its duration. This is manifested for example by

- creep (increase of strain at constant applied stress),
- relaxation (decrease of stress at constant strain),
- rate-dependent strength (increase of resistance under fast loading).

A natural framework for the modeling of rate-dependent plastic deformation is provided by the theory of viscoplasticity (sometimes called elastoviscoplasticity, to emphasize that the response at low stress levels is not rigid but elastic). Later we will show that classical elastoplasticity can be regarded as a singular limit case of viscoplasticity, with a sudden jump from a zero plastic strain rate inside the elastic domain to an arbitrarily fast plastic flow on the yield surface. The effect of viscosity is to regularize the model in the sense that the transition from the purely elastic response to yielding becomes smooth. Plastic flow cannot become infinitely fast because microstructural rearrangements causing flow take place on a certain time scale that is characteristic of the material.

Viscoplastic models have traditionally been used for metals, especially at high stresses and temperatures (above one-third of the melting temperature on the absolute scale). Concrete exhibits a time-dependent response (creep) even at low stress levels and normal temperatures, and this is usually described by linear viscoelastic models, to be discussed in Chapters 28 and 29. Deviations from linearity are mainly due to cracking, which itself is also rate-dependent. Even though the truly plastic deformation of concrete occurs only under highly confined compression, the rate-dependent inelastic deformation can formally be modeled as viscoplastic strain, provided that the degradation of strength is taken into account by an appropriate softening law.

The rate-independent flow theory of plasticity is built around the concept of yield function, f, which depends on the stress $\boldsymbol{\sigma}$ and, if hardening or softening is considered, on the thermodynamic forces $\boldsymbol{q}$ conjugate to the hardening-softening variables $\boldsymbol{\kappa}$. The yield function $f(\boldsymbol{\sigma}, \boldsymbol{q})$ defines the elastic domain, which is bounded by the yield surface. In the usual rate-independent theory, the states for which $f < 0$ are elastic, the states for which $f = 0$ are plastic, and the states for which $f > 0$ are inadmissible. Plastic flow can take place only if the current state satisfies the yield condition,

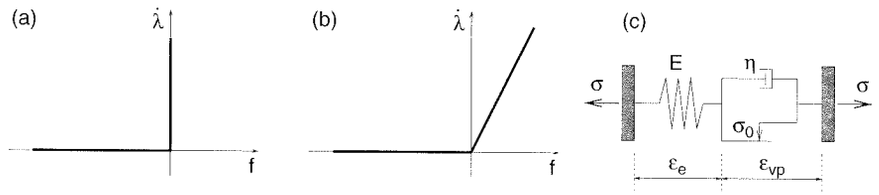


Figure 27.1 a) Classical plasticity, b) viscoplasticity, c) Bingham model

$f(\boldsymbol{\sigma}, \boldsymbol{q}) = 0$. All this is described by the loading-unloading conditions

$$f \leq 0, \qquad \dot{\lambda} \geq 0, \qquad \dot{\lambda} f = 0 \tag{27.1}$$

The theory of viscoplasticity regularizes the model behavior by allowing stress states that lie outside the yield surface. The sudden jump from a purely elastic behavior to an arbitrarily fast plastic flow (Figure 27.1(a)) is replaced by a continuous increase of the plastic strain rate as a function of the stress level (Figure 27.1(b)). The yield surface $f(\boldsymbol{\sigma}, \boldsymbol{q}) = 0$ represents the static limit, i.e. the set of states for which the plastic flow evolves at an infinitely slow rate. The difference between the actual stress and the static yield stress is called the *overstress*. It is caused by viscous effects, and it increases with an increasing rate of the viscoplastic strain, $\dot{\boldsymbol{\varepsilon}}_{vp}$.

27.1 BINGHAM MODEL

To introduce the basic ideas, let us analyze a simple rheological model of perfect viscoplasticity in the uniaxial setting.

Example 27.1: Set up the basic equations describing the Bingham model (Bingham, 1919) in Figure 27.1(c).

Solution: The spring of stiffness E corresponds to the elastic part of the response, and the friction unit starts sliding when the magnitude of the applied stress reaches the yield limit, σ_0. During plastic flow, the linear dashpot of viscosity η produces an overstress, $\sigma_v = \eta \dot{\varepsilon}_{vp}$. For a while, we restrict our attention to yielding under tension. In this regime, the total stress is $\sigma = \sigma_0 + \sigma_v$, and so the viscoplastic strain rate can be expressed as

$$\dot{\varepsilon}_{vp} = \frac{\sigma_v}{\eta} = \frac{\sigma - \sigma_0}{\eta} \tag{27.2}$$

The total stress must also be equal to the stress transmitted by the elastic unit,

$$\sigma = E\varepsilon_e = E(\varepsilon - \varepsilon_{vp}) \tag{27.3}$$

Differentiating this with respect to time and substituting $\dot{\varepsilon}_{vp}$ from (27.2), we can eliminate the viscoplastic strain and arrive at the differential equation

$$\eta \dot{\sigma} + E\sigma = E\left(\eta \dot{\varepsilon} + \sigma_0\right) \tag{27.4}$$

which characterizes the constitutive behavior in the regime of viscoplastic yielding under tension. In the elastic range, we must set $\dot{\varepsilon}_{vp} = 0$ and replace (27.4) by $\dot{\sigma} = E\dot{\varepsilon}$, and for viscoplastic yielding under compression we must replace σ_0 in (27.4) by $-\sigma_0$. □

Example 27.2: Develop the basic equations of the Bingham model using the thermodynamic approach.

Solution: It is natural to select the viscoplastic strain ε_{vp} as the internal variable, and express the elastic strain as $\varepsilon_e = \varepsilon - \varepsilon_{vp}$. The free energy stored in the elastic spring is $\frac{1}{2}E\varepsilon_e^2$. In terms of the state variables, it can be written as

$$\rho\psi(\varepsilon, \varepsilon_{vp}) = \tfrac{1}{2}E(\varepsilon - \varepsilon_{vp})^2 \tag{27.5}$$

Differentiation of (27.5) provides the state laws

$$\sigma_Q = \rho\frac{\partial \psi}{\partial \varepsilon} = E(\varepsilon - \varepsilon_{vp}) \tag{27.6}$$

$$\beta_Q = \rho\frac{\partial \psi}{\partial \varepsilon_{vp}} = -E(\varepsilon - \varepsilon_{vp}) \tag{27.7}$$

i.e., the laws that define the quasi-conservative stress σ_Q and the quasi-conservative thermodynamic force β_Q.

The viscous dashpot and the friction unit both contribute to the dissipation, which can be expressed as

$$\mathcal{D}(\dot{\varepsilon}_{vp}) = \eta\dot{\varepsilon}_{vp}^2 + \sigma_0|\dot{\varepsilon}_{vp}| = \beta_D(\dot{\varepsilon}_{vp})\,\dot{\varepsilon}_{vp} \tag{27.8}$$

From this we can identify the dissipative thermodynamic force conjugate to the viscoplastic strain,

$$\beta_D(\dot{\varepsilon}_{vp}) = \eta\dot{\varepsilon}_{vp} + \sigma_0 \operatorname{sgn}\dot{\varepsilon}_{vp} \tag{27.9}$$

However, equation (27.9) is a case valid only if $\dot{\varepsilon}_{vp} \neq 0$. For a vanishing viscoplastic strain rate, the dissipation also vanishes and the identity $\mathcal{D} = \beta_D\dot{\varepsilon}_{vp}$ is satisfied with an arbitrary value of β_D. To also cover the latter case, it is convenient to introduce the dissipation potential

$$\phi(\dot{\varepsilon}_{vp}) = \tfrac{1}{2}\eta\dot{\varepsilon}_{vp}^2 + \sigma_0|\dot{\varepsilon}_{vp}| \tag{27.10}$$

and write the complementary law in the form

$$\beta_D(\dot{\varepsilon}_{vp}) \in \partial\phi(\dot{\varepsilon}_{vp}) \tag{27.11}$$

The structure of the model equations is very similar to standard elastoplasticity. As we will later see, the main difference is that the dissipation potential (27.10) is strictly convex, which implies the dual dissipation potential to be differentiable in the standard sense. The constitutive law (23.62) and the evolution law (23.63) take the simple form

$$\sigma = \sigma_Q \tag{27.12}$$

$$0 = \beta_Q + \beta_D \tag{27.13}$$

and, when combined with the state laws (27.6)–(27.7), they indicate that the dissipative force conjugate to the viscoplastic strain is equal to the stress. We will take this into account and replace the symbol β_D by σ, which facilitates the physical interpretation of the resulting equations.

The dissipation potential (27.10), graphically presented in Figure 27.2(a), is differentiable in the standard sense everywhere except the origin, where its

subdifferential is equal to the interval $[-\sigma_0, \sigma_0]$. The complementary law (27.11) can, therefore, be written as

$$\sigma = \frac{\mathrm{d}\phi(\dot{\varepsilon}_{vp})}{\mathrm{d}\dot{\varepsilon}_{vp}} = \eta\dot{\varepsilon}_{vp} + \sigma_0 \operatorname{sgn}\dot{\varepsilon}_{vp} \quad \text{if } \dot{\varepsilon}_{vp} \neq 0 \tag{27.14}$$

$$\sigma \in \partial\phi(0) = [-\sigma_0, \sigma_0] \qquad \text{if } \dot{\varepsilon}_{vp} = 0 \tag{27.15}$$

Note that the first equation corresponds to (27.9) and the second provides a rigorous description of the elastic regime.

To obtain an explicit expression for the viscoplastic strain rate, we must invert relations (27.14)–(27.15). In the present simple case, this could be done directly, especially since a graphical representation of the complementary law is available; see Figure 27.2(b). Nevertheless, it is instructive to demonstrate the systematic procedure based on the construction of the dual dissipation potential by the Legendre–Fenchel transform. According to (23.101), we have

$$\phi^*(\sigma) = \max_{\dot{\varepsilon}_{vp}\in\mathcal{R}} (\sigma\dot{\varepsilon}_{vp} - \phi(\dot{\varepsilon}_{vp})) = \max_{\dot{\varepsilon}_{vp}\in\mathcal{R}} (\sigma\dot{\varepsilon}_{vp} - \sigma_0|\dot{\varepsilon}_{vp}| - \tfrac{1}{2}\eta\dot{\varepsilon}_{vp}^2) \tag{27.16}$$

From Figure 27.3(a), it is clear that, if $|\sigma| \leq \sigma_0$, the expression $\sigma\dot{\varepsilon}_{vp} - \phi(\dot{\varepsilon}_{vp})$ to be maximized is negative for all $\dot{\varepsilon}_{vp} \neq 0$ and is zero for $\dot{\varepsilon}_{vp} = 0$. In this case, $\phi^*(\sigma) = 0$. On the other hand, if $|\sigma| > \sigma_0$, the difference $\sigma\dot{\varepsilon}_{vp} - \phi(\dot{\varepsilon}_{vp})$ is positive for some $\dot{\varepsilon}_{vp}$; see Figures 27.3(b),(c). The maximum is obtained at that $\dot{\varepsilon}_{vp}$ for which

$$\frac{\mathrm{d}}{\mathrm{d}\dot{\varepsilon}_{vp}} (\sigma\dot{\varepsilon}_{vp} - \phi(\dot{\varepsilon}_{vp})) = \sigma - \frac{\mathrm{d}\phi(\dot{\varepsilon}_{vp})}{\mathrm{d}\dot{\varepsilon}_{vp}} = \sigma - \sigma_0 \operatorname{sgn}\dot{\varepsilon}_{vp} - \eta\dot{\varepsilon}_{vp} = 0 \tag{27.17}$$

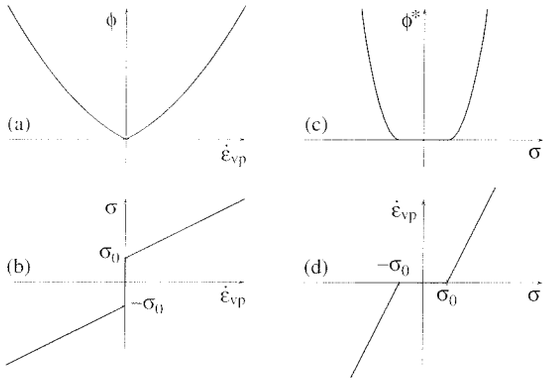


Figure 27.2 a) Dissipation potential, b) complementary law in the implicit form, c) dual dissipation potential, d) complementary law in the explicit form

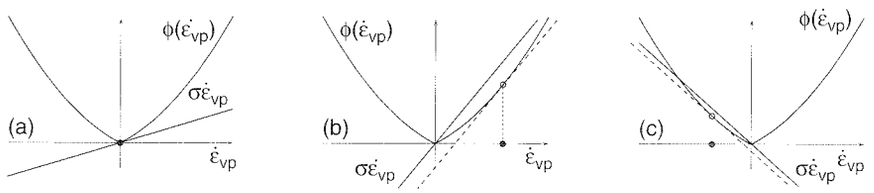


Figure 27.3 Illustration of the Legendre–Fenchel transform of the dissipation potential for three different values of stress σ; the filled circle marks the viscoplastic strain rate that maximizes the difference $\sigma\dot{\varepsilon}_{vp} - \phi(\dot{\varepsilon}_{vp})$

From Figures 27.3(b),(c), it is clear that the value of $\dot{\varepsilon}_{vp}$ that solves (27.17) has the same sign as σ. Replacing $\operatorname{sgn}\dot{\varepsilon}_{vp}$ by $\operatorname{sgn}\sigma$, we can write the solution as

$$\dot{\varepsilon}_{vp} = \frac{\sigma - \sigma_0 \operatorname{sgn}\sigma}{\eta} \tag{27.18}$$

Substituting this into (27.16) we obtain

$$\begin{aligned}\phi^*(\sigma) &= \frac{\sigma(\sigma - \sigma_0 \operatorname{sgn}\sigma)}{\eta} - \frac{\sigma_0 \operatorname{sgn}\sigma(\sigma - \sigma_0 \operatorname{sgn}\sigma)}{\eta} - \frac{\eta}{2}\left(\frac{\sigma - \sigma_0 \operatorname{sgn}\sigma}{\eta}\right)^2 \\ &= \frac{(\sigma - \sigma_0 \operatorname{sgn}\sigma)^2}{2\eta} = \frac{(|\sigma| - \sigma_0)^2}{2\eta}\end{aligned} \tag{27.19}$$

Both cases, $|\sigma| \leq \sigma_0$ and $|\sigma| > \sigma_0$, can be covered by one single formula,

$$\phi^*(\sigma) = \frac{1}{2\eta}\langle|\sigma| - \sigma_0\rangle^2 \tag{27.20}$$

The dual dissipation potential, plotted in Figure 27.2(c), is continuously differentiable (in the standard sense). The dual (explicit) form of the complementary law can be conveniently written as

$$\dot{\varepsilon}_{vp} = \frac{\partial\phi^*(\sigma)}{\partial\sigma} = \frac{\operatorname{sgn}\sigma}{\eta}\langle|\sigma| - \sigma_0\rangle \tag{27.21}$$

and it indeed corresponds to the inverse of relations (27.14)–(27.15), as shown in Figure 27.2(d). □

27.2 PERZYNA MODEL

Inspecting the results of Example 27.2, we realize that the viscoplastic flow rule (27.21) can be interpreted as a regularized version of the flow rule of uniaxial perfect elastoplasticity. Indeed, for a decreasing value of the viscosity η the stress transmitted by the dashpot of the Bingham model (at a given strain rate) gets reduced, and in the limit $\eta = 0$ it completely vanishes. In terms of the diagram in Figure 27.2(d) this means that the inclined lines get progressively steeper and in the limit become vertical; see Figure 27.4. The elastic domain is not affected by the viscosity and is always equal to the interval $\mathcal{E} = (-\sigma_0, \sigma_0)$. The corresponding yield function can be defined as

$$f(\sigma) = |\sigma| - \sigma_0 \tag{27.22}$$

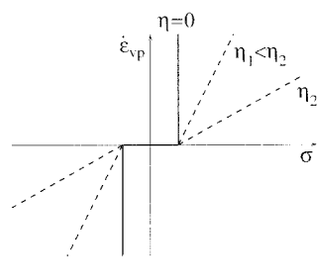


Figure 27.4 Rate-independent elastoplasticity as a limit of viscoplasticity

Since, in the rate-independent limit, the dual dissipation potential $\phi^* = \mathcal{I}_{\mathcal{E}}$ is no longer differentiable in the standard sense, the corresponding flow rule cannot be obtained by substituting $\eta = 0$ into (27.21). Nevertheless, we know from Chapter 23 that ϕ^* remains subdifferentiable and the complementary law $\dot{\varepsilon}_{vp} \in \partial\phi^*(\sigma)$ is equivalent to the flow rule

$$\dot{\varepsilon}_{vp} = \dot{\lambda}\frac{\mathrm{d}f(\sigma)}{\mathrm{d}\sigma} \tag{27.23}$$

and the loading-unloading conditions (27.1).

The yield function (27.22) and its gradient $\mathrm{d}f/\mathrm{d}\sigma = \operatorname{sgn}\sigma$ also appear in the viscoplastic flow rule (27.21), which can be rewritten as

$$\dot{\varepsilon}_{vp} = \frac{1}{\eta}\langle f(\sigma)\rangle\frac{\mathrm{d}f(\sigma)}{\mathrm{d}\sigma} \tag{27.24}$$

This equation is valid for an arbitrary σ, i.e. it covers not only the case of viscoplastic flow, but also the purely elastic case, in which $f(\sigma) < 0$, $\langle f(\sigma)\rangle = 0$ and $\dot{\varepsilon}_{vp} = 0$. This means that (27.24) is a regularized complementary law that replaces the flow rule (27.23) as well as the loading-unloading conditions (27.1). In a slightly different interpretation, we could say that the rate of the plastic multiplier $\dot{\lambda}$ from (27.23), in the rate-independent limit given by $\dot{\lambda} = 0$ for $f < 0$ and $\dot{\lambda} \geq 0$ for $f = 0$, is replaced by a continuous function $\dot{\lambda}(\sigma) = \langle f(\sigma)\rangle/\eta$, which remains equal to zero for $f < 0$ and increases with increasing positive f.

The above analysis of the uniaxial model permits a straightforward generalization to multiple dimensions. The dual dissipation potential, in the rate-independent case defined as the indicator function of the set of states for which $f(\boldsymbol{\sigma}) \leq 0$, is replaced by a continuously differentiable function

$$\phi^*(\boldsymbol{\sigma}) = \frac{1}{2\eta}\langle f(\boldsymbol{\sigma})\rangle^2 \tag{27.25}$$

and the corresponding viscoplastic flow rule reads

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \frac{\partial\phi^*(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}} = \frac{1}{\eta}\langle f(\boldsymbol{\sigma})\rangle\frac{\partial f(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}} \tag{27.26}$$

The quadratic form of the dual dissipation potential (27.25) was motivated by the assumed linearity of the viscous unit. Experiments indicate that the real behavior is highly nonlinear. This is reflected, for example, by the classical *Norton law* (formulated by Norton, 1929) for uniaxial stress and generalized by Odqvist, (1933b) to multiaxial stress states. Norton's law neglects the elastic domain and describes the strain rate during the secondary creep stage in metals under uniaxial tension as a power function (with a high exponent) of the applied stress. In a general case, a nonlinear dependence between the overstress and the viscoplastic strain rate can be incorporated into the model by a suitable modification of the dual dissipation potential,

$$\phi^*(\boldsymbol{\sigma}) = \Phi(\langle f(\boldsymbol{\sigma})\rangle) \tag{27.27}$$

where $\Phi(f)$ is a continuously differentiable and strictly convex function with a vanishing derivative at the origin (due to the presence of the positive-part operator, it is sufficient to define Φ only for non-negative arguments). The flow rule (27.26) is then generalized to

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \frac{\partial\phi^*(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}} = \frac{\partial}{\partial\boldsymbol{\sigma}}\Phi(\langle f(\boldsymbol{\sigma})\rangle) = \Phi'(\langle f(\boldsymbol{\sigma})\rangle)\frac{\partial f(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}} \tag{27.28}$$

Owing to the assumed properties of Φ, the derivative $\Phi' \equiv \mathrm{d}\Phi/\mathrm{d}f$ is an increasing function that vanishes at the origin. This ensures that $\dot{\boldsymbol{\varepsilon}}_{vp} = \mathbf{0}$ for $f(\boldsymbol{\sigma}) \leq 0$ and that $\dot{\boldsymbol{\varepsilon}}_{vp}$ monotonically increases with an increasing positive value of the yield function (which is a scalar measure of the overstress).

Example 27.3: Develop the basic equations of J_2-viscoplasticity with the rate dependence described by a power law.

Solution: We start from the standard yield function of J_2-elastoplasticity,

$$f(\boldsymbol{\sigma}) = \sqrt{3J_2(\boldsymbol{\sigma})} - \sigma_0 \tag{27.29}$$

where $J_2(\boldsymbol{\sigma}) = \frac{1}{2}\|\boldsymbol{s}\|^2 = \frac{1}{2}\boldsymbol{s} : \boldsymbol{s}$ is the second invariant of the deviatoric stress, $\boldsymbol{s}$. Under uniaxial stress, this yield function reduces to (27.22), i.e. (27.29) is indeed a scalar measure of the overstress. To obtain the dependence of the viscoplastic strain rate on the overstress described by a power law with exponent n, we must select the potential function Φ as a power function with exponent $n + 1$. If the yield function has the dimension of stress, Φ must have the dimension of stress divided by time. For the sake of dimensionality, we normalize f by the uniaxial yield stress σ_0 and write

$$\Phi(f) = \frac{\sigma_0^2}{(n+1)\eta}\left(\frac{f}{\sigma_0}\right)^{n+1} \tag{27.30}$$

The derivative of this potential function is

$$\Phi'(f) = \frac{\sigma_0}{\eta}\left(\frac{f}{\sigma_0}\right)^n \tag{27.31}$$

Note that, since the dimension of the viscosity parameter η is stress times time, $\sigma_0\eta$ has the dimension of one over time. Taking into account the fact that the gradient of the yield function (27.29) is $\partial f(\boldsymbol{\sigma})/\partial\boldsymbol{\sigma} = \sqrt{3/2}\,\boldsymbol{s}/\|\boldsymbol{s}\|$ and substituting this and (27.31) into (27.28) leads to

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \Phi'(\langle f(\boldsymbol{\sigma})\rangle)\frac{\partial f(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}} = \frac{\sigma_0}{\eta}\left\langle\frac{f}{\sigma_0}\right\rangle^n\sqrt{\frac{3}{2}}\frac{\boldsymbol{s}}{\|\boldsymbol{s}\|} = \frac{\sigma_0}{\eta}\left\langle\sqrt{\frac{3}{2}}\frac{\|\boldsymbol{s}\|}{\sigma_0} - 1\right\rangle^n\sqrt{\frac{3}{2}}\frac{\boldsymbol{s}}{\|\boldsymbol{s}\|} \tag{27.32}$$

This flow rule must be combined with the elastic stress-strain law, which has the usual form

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_{vp}) \tag{27.33}$$

Under uniaxial stress, (27.32) reduces to

$$\dot{\varepsilon}_{vp} = \frac{\sigma_0}{\eta}\left\langle\frac{|\sigma|}{\sigma_0} - 1\right\rangle^n\operatorname{sgn}\sigma \tag{27.34}$$

which contains (27.21) as a special case with $n = 1$. □

The model developed in the previous example belongs to the class of viscoplastic models originating in the work of Perzyna (1963, 1966, 1971). The general framework can easily be extended to hardening materials. The yield function $f(\boldsymbol{\sigma}, \boldsymbol{q})$ becomes

dependent on the dissipative forces conjugate to the hardening variables, $\boldsymbol{\kappa}$, and the associated complementary law

$$\dot{\boldsymbol{\kappa}} = \frac{\partial \phi^*(\boldsymbol{\sigma}, \boldsymbol{q})}{\partial \boldsymbol{q}} = \frac{\partial}{\partial \boldsymbol{q}} \Phi(\langle f(\boldsymbol{\sigma}, \boldsymbol{q}) \rangle) = \Phi'(\langle f(\boldsymbol{\sigma}, \boldsymbol{q}) \rangle) \frac{\partial f(\boldsymbol{\sigma}, \boldsymbol{q})}{\partial \boldsymbol{q}} \tag{27.35}$$

represents one part of the hardening law (the other part being the relation $\boldsymbol{q} = -\rho\, \partial \psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_{vp}, \boldsymbol{\kappa}) / \partial \boldsymbol{\kappa}$, having its origin in the state law). In the flow rule (27.28), the yield function now depends also on $\boldsymbol{q}$, but otherwise the equation remains unchanged.

Equations (27.28) and (27.35) have been derived from a dual dissipation potential. If the yield function $f(\boldsymbol{\sigma}, \boldsymbol{q})$ is convex and the function $\Phi(f)$ is strictly convex and has its minimum at $f = 0$, then the dual dissipation potential $\phi^*(\boldsymbol{\sigma}, \boldsymbol{q}) = \Phi(\langle f(\boldsymbol{\sigma}, \boldsymbol{q}) \rangle)$ has the properties that were shown in Section 23.4 to guarantee a non-negative dissipation.

It is also possible to consider a more general, nonassociated version of the Perzyna model, in which (27.28) and (27.35) are replaced by

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \Phi'(\langle f(\boldsymbol{\sigma}, \boldsymbol{q}) \rangle) \frac{\partial g(\boldsymbol{\sigma}, \boldsymbol{q})}{\partial \boldsymbol{\sigma}} \tag{27.36}$$

$$\dot{\boldsymbol{\kappa}} = \Phi'(\langle f(\boldsymbol{\sigma}, \boldsymbol{q}) \rangle) \frac{\partial g(\boldsymbol{\sigma}, \boldsymbol{q})}{\partial \boldsymbol{q}} \tag{27.37}$$

where $g(\boldsymbol{\sigma}, \boldsymbol{q})$ is a viscoplastic potential. In this general case, the dissipation is given by

$$\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{vp} + \boldsymbol{q} \bullet \dot{\boldsymbol{\kappa}} = \left(\boldsymbol{\sigma} : \frac{\partial g}{\partial \boldsymbol{\sigma}} + \boldsymbol{q} \bullet \frac{\partial g}{\partial \boldsymbol{q}} \right) \Phi'(\langle f(\boldsymbol{\sigma}, \boldsymbol{q}) \rangle) \tag{27.38}$$

Clearly, if $\Phi' \geq 0$ for all $f \geq 0$, the sign of the dissipation depends on the expression in the parentheses. The dissipation inequality is verified if this expression is non-negative for all $\boldsymbol{\sigma}$ and $\boldsymbol{q}$ satisfying $f(\boldsymbol{\sigma}, \boldsymbol{q}) > 0$. Whether this is true depends on the specific choice of functions f and g.

The Perzyna formulation of viscoplasticity was used, for instance, by Chaboche and coworkers. They developed a sophisticated but rather complex model for metals, which became popular in applications. The basic version of the Chaboche model (Chaboche, 1978; Chaboche and Rousselier 1983) is in fact associated J_2-viscoplasticity with nonlinear isotropic hardening, nonlinear kinematic hardening of the Armstrong-Frederick type (described in Section 20.1.4), and a power-law dependence between the viscoplastic strain rate and the overstress. This 'simple' version can describe the basic features of rate dependence and the phenomenon of ratchetting, observed in cyclic experiments with a nonzero mean value of the applied stress. Chaboche and Nouailhas (1989) presented a hierarchical model with a number of additional features. The complete model takes into account the plastic strain memorization and the time recovery for isotropic and kinematic hardening.

27.3 MODELS OF DUVAUT–LIONS TYPE

There is another important class of viscoplastic models, motivated by an alternative generalization of the Bingham model. The three-dimensional version of equation (27.2) can be written in the form

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \frac{1}{\eta} (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}) \tag{27.39}$$

where $\boldsymbol{\sigma}$ is the actual stress and $\bar{\boldsymbol{\sigma}}$ is its projection on the static yield surface. If $\bar{\boldsymbol{\sigma}}$ is understood as the point that minimizes the distance from $\boldsymbol{\sigma}$ among all points $\bar{\boldsymbol{\sigma}}^*$ for which $f(\bar{\boldsymbol{\sigma}}^*) \leq 0$, equation (27.39) remains valid even in the elastic domain, in which $\bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma}$ and $\dot{\boldsymbol{\varepsilon}}_{vp} = \mathbf{0}$.

In the original formulation, due to Duvaut and Lions (1972), $\bar{\boldsymbol{\sigma}}$ was obtained by minimizing the distance in the Euclidean metric. The evolution of stress during relaxation at a fixed value of the total strain is then described by the differential equation

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\boldsymbol{\varepsilon}}_{vp}) = -\boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}_{vp} = -\frac{1}{\eta} \boldsymbol{D}_e : (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}) \tag{27.40}$$

If $\dot{\boldsymbol{\sigma}}$ is colinear with the overstress $\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}$, the stress point travels in the stress space along a straight line towards the point representing $\bar{\boldsymbol{\sigma}}$, which remains stationary during the entire relaxation process. The stress state $\bar{\boldsymbol{\sigma}}$ can then be interpreted as the relaxation limit. However, $\dot{\boldsymbol{\sigma}}$ computed from (27.40) is colinear with $\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}$ only if the multiplication by the elastic stiffness $\boldsymbol{D}_e$ does not change the direction of the overstress. This is most notably the case if the yield condition is pressure-independent, because then the overstress $\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}$ is a purely deviatoric tensor and $\boldsymbol{D}_e : (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}) = 2G(\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}})$.

The difference between the flow rules used by the Perzyna model and by the Duvaut–Lions model is illustrated in Figure 27.5. For the associated Perzyna model, the direction of viscoplastic flow is normal to the equipotential surface ($f = \text{const.}$) passing through the current stress point. For the Duvaut–Lions model, the direction of viscoplastic flow is normal to the static yield surface $f = 0$ at the point $\bar{\boldsymbol{\sigma}}$ closest to the current stress point. Of course, for the von Mises yield surface both formulations coincide, but in general they are different.

For general, pressure-dependent yield conditions, equation (27.40) would lead to a relaxation rate $\dot{\boldsymbol{\sigma}}$ not colinear with the overstress. The projected stress state $\bar{\boldsymbol{\sigma}}$ would then vary during the relaxation process and could not be physically interpreted as the relaxation limit. Simo, Kennedy and Govindjee (1988) proposed a modification of the flow rule (27.39) in the form

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \frac{1}{\tau} \boldsymbol{C}_e : (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}) \tag{27.41}$$

where τ is a material parameter with the dimension of time, and $\bar{\boldsymbol{\sigma}}$ is defined as the closest point in the metric generated by the elastic compliance tensor $\boldsymbol{C}_e = \boldsymbol{D}_e^{-1}$. In other words, $\bar{\boldsymbol{\sigma}}$ minimizes the expression $(\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}^*) : \boldsymbol{C}_e : (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}^*)$ among all stress states $\bar{\boldsymbol{\sigma}}^*$ satisfying $f(\bar{\boldsymbol{\sigma}}^*) \leq 0$.

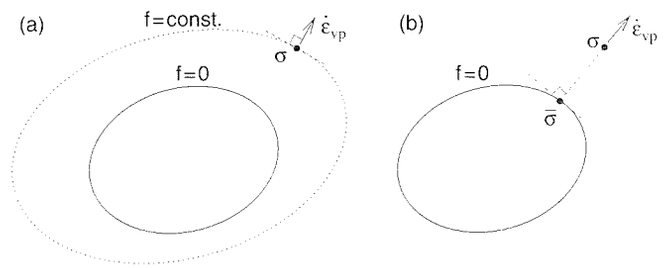


Figure 27.5 Viscoplastic flow according to a) Perzyna, b) Duvaut and Lions

For the flow rule (27.41), the stress rate during relaxation ($\dot{\boldsymbol{\varepsilon}} = 0$) is given by

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_e : (\dot{\boldsymbol{\varepsilon}} - \dot{\boldsymbol{\varepsilon}}_{vp}) = -\boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}}_{vp} = -\frac{1}{\tau}(\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}) \tag{27.42}$$

and the stress during relaxation always tends to $\bar{\boldsymbol{\sigma}}$. The solution of the linear differential equation (27.42) satisfying the initial condition $\boldsymbol{\sigma}(0) = \bar{\boldsymbol{\sigma}} + \Delta\boldsymbol{\sigma}$ is

$$\boldsymbol{\sigma}(t) = \bar{\boldsymbol{\sigma}} + \Delta\boldsymbol{\sigma} \exp(-t/\tau) \tag{27.43}$$

The parameter τ sets the time scale of the relaxation process and therefore can be called the *relaxation time*. For pressure-independent yield conditions, the flow rules (27.39) and (27.41) are equivalent if the relaxation time is set equal to $\tau = \eta/2G$.

Equation (27.41) with a constant relaxation time describes a linear dependence between the viscoplastic strain rate and the overstress. Nonlinearity of the viscous part of the model can be incorporated by treating the relaxation time as a variable depending on the Euclidean norm of the overstress, $\|\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}\| = \sqrt{(\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}) : (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}})}$, or on the energy norm of the overstress, $\sqrt{(\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}) : \boldsymbol{C}_e : (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}})}$. For example, a power law similar to (27.32) can be obtained with

$$\tau = \frac{\eta}{3G} \left\langle \sqrt{\frac{3}{2}} \frac{\|\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}\|}{\sigma_0} \right\rangle^{1-n} \tag{27.44}$$

If hardening is considered, it is natural to expect the evolution of the dissipative forces $\boldsymbol{q}$ conjugate to the hardening variables $\boldsymbol{\kappa}$ to be governed by an equation analogous to (27.42),

$$\dot{\boldsymbol{q}} = -\frac{1}{\tau}(\boldsymbol{q} - \bar{\boldsymbol{q}}) \tag{27.45}$$

where $\bar{\boldsymbol{q}}$ is given by the projection of the current state on the yield surface in the space $(\boldsymbol{\sigma}, \boldsymbol{q})$. However, it is not immediately clear how to define an appropriate metric in that space.

The systematic development of hardening viscoplastic models of the Duvaut–Lions type can be based on the thermodynamic approach. The formulation again takes inspiration from the simple one-dimensional Bingham model, which dissipates energy by two distinct mechanisms – in the rate-independent friction unit and in the dashpot. The total dissipation is the sum of the plastic dissipation and the viscous dissipation. The dissipative force associated with the friction unit is $\bar{\sigma} = \sigma_0 \operatorname{sgn} \dot{\varepsilon}_{vp}$ and the dissipative force associated with the linear dashpot is $\sigma_v = \eta\dot{\varepsilon}_{vp}$. Admitting such an additive split for the general case (Ristinmaa and Ottosen, 1998), we can write

$$\mathcal{D} = \boldsymbol{\beta}_D \bullet \dot{\boldsymbol{\alpha}} = \bar{\boldsymbol{\beta}} \bullet \dot{\boldsymbol{\alpha}} + \boldsymbol{\beta}_v \bullet \dot{\boldsymbol{\alpha}} \tag{27.46}$$

where $\bar{\boldsymbol{\beta}}$ is the part of the dissipative forces associated with the rate-independent mechanism and $\boldsymbol{\beta}_v$ is the additional part due to viscous effects. These forces are related to the flux $\dot{\boldsymbol{\alpha}}$ by complementary laws, written in terms of the corresponding dissipation potentials as

$$\bar{\boldsymbol{\beta}} \in \partial\bar{\phi}(\dot{\boldsymbol{\alpha}}), \qquad \boldsymbol{\beta}_v \in \partial\phi_v(\dot{\boldsymbol{\alpha}}) \tag{27.47}$$

This is, of course, equivalent to postulating that the dissipation potential is additively decomposed as $\phi(\dot{\boldsymbol{\alpha}}) = \bar{\phi}(\dot{\boldsymbol{\alpha}}) + \phi_v(\dot{\boldsymbol{\alpha}})$, but it was more instructive to start from the dissipation, which has a direct physical meaning.

If the dissipation potentials $\bar{\phi}$ and ϕ_v have the usual properties (convexity, positivity, zero value for $\dot{\boldsymbol{\alpha}} = \boldsymbol{0}$), there exist dual dissipation potentials $\bar{\phi}^*$ and ϕ_v^*, and the complementary laws (27.47) can be written in the inverse form

$$\dot{\boldsymbol{\alpha}} \in \partial\bar{\phi}^*(\bar{\boldsymbol{\beta}}), \qquad \dot{\boldsymbol{\alpha}} \in \partial\phi_v^*(\boldsymbol{\beta}_v) \tag{27.48}$$

Equations (27.46)–(27.48) correspond to a very wide class of models with an additive split of the dissipative forces. Now we will specify the nature of the internal variables and define a particular form of the dual dissipation potentials leading to the generalized Duvaut–Lions models.

As usual in viscoplasticity, the internal variables $\boldsymbol{\alpha}$ are the viscoplastic strain $\boldsymbol{\varepsilon}_{vp}$ and the hardening variables $\boldsymbol{\kappa}$. The dissipative force conjugate to $\boldsymbol{\varepsilon}_{vp}$ could be denoted by a special symbol, say $\boldsymbol{\sigma}_{vp}$, but as we have repeatedly mentioned, this force remains equal to the stress and therefore can be right away denoted as $\boldsymbol{\sigma}$. Finally, $\boldsymbol{q}$ denotes the dissipative force conjugate to $\boldsymbol{\kappa}$. The additive split of the dissipative forces, $\boldsymbol{\beta}_D = \bar{\boldsymbol{\beta}} + \boldsymbol{\beta}_v$, is now applied to the individual components:

$$\boldsymbol{\sigma} = \bar{\boldsymbol{\sigma}} + \boldsymbol{\sigma}_v \tag{27.49}$$

$$\boldsymbol{q} = \bar{\boldsymbol{q}} + \boldsymbol{q}_v \tag{27.50}$$

Our intention is, of course, to construct the theory such that $\bar{\boldsymbol{\sigma}}$ is the stress projection on the static yield surface and $\boldsymbol{\sigma}_v = \boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}$ is the overstress. The dual dissipation potential $\bar{\phi}^*(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})$ should describe the rate-independent plastic effects, and so it is set equal to the indicator function $\mathcal{I}_{\bar{\mathcal{E}}}$ of the closure of the elastic domain. This part of the model, corresponding to standard associated elastoplasticity, is fully defined by the yield function $f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})$. The dual dissipation potential $\phi_v^*(\boldsymbol{\varepsilon}_{vp}, \boldsymbol{\kappa})$ describes viscous effects, and is assumed to be continuously differentiable. With the present choice of variables and potentials, the complementary laws (27.48) have the form

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \dot{\lambda}\frac{\partial f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})}{\partial\bar{\boldsymbol{\sigma}}}, \qquad \dot{\boldsymbol{\kappa}} = \dot{\lambda}\frac{\partial f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})}{\partial\bar{\boldsymbol{q}}} \tag{27.51}$$

$$\dot{\lambda} \geq 0, \qquad f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}}) \leq 0, \qquad \dot{\lambda} f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}}) = 0 \tag{27.52}$$

$$\dot{\boldsymbol{\varepsilon}}_{vp} = \frac{\partial\phi_v^*(\boldsymbol{\sigma}_v, \boldsymbol{q}_v)}{\partial\boldsymbol{\sigma}_v}, \qquad \dot{\boldsymbol{\kappa}} = \frac{\partial\phi_v^*(\boldsymbol{\sigma}_v, \boldsymbol{q}_v)}{\partial\boldsymbol{q}_v} \tag{27.53}$$

These equations correspond to the Duvaut–Lions viscoplastic models in the generalized sense. In particular, equation (27.39) can be obtained with

$$\phi_v^*(\boldsymbol{\sigma}_v, \boldsymbol{q}_v) = \frac{1}{2\eta}\boldsymbol{\sigma}_v : \boldsymbol{\sigma}_v + \phi_{vq}^*(\boldsymbol{q}_v) \tag{27.54}$$

and equation (27.41) can be obtained with

$$\phi_v^*(\boldsymbol{\sigma}_v, \boldsymbol{q}_v) = \frac{1}{2\tau}\boldsymbol{\sigma}_v : \boldsymbol{C}_e : \boldsymbol{\sigma}_v + \phi_{vq}^*(\boldsymbol{q}_v) \tag{27.55}$$

The part of the potential that depends on $\boldsymbol{q}_v$ can be defined, for instance, by a quadratic expression

$$\phi_{vq}^*(\boldsymbol{q}_v) = \tfrac{1}{2}\boldsymbol{q}_v \bullet \boldsymbol{Q} \bullet \boldsymbol{q}_v \tag{27.56}$$

where $\boldsymbol{Q}$ is a constant tensor, assumed to be positive definite (to make sure that ϕ^*_{vq} is strictly convex). This choice leads to linear complementary equations

$$\dot{\boldsymbol{\kappa}} = \boldsymbol{Q} \bullet \boldsymbol{q}_v \tag{27.57}$$

To relate this law to (27.45), we must find the link between the rates $\dot{\boldsymbol{\kappa}}$ and $\dot{\boldsymbol{q}}$. Since the free-energy potential has the decoupled form $\psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_{vp}, \boldsymbol{\kappa}) = \psi_e(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_{vp}) + \psi_p(\boldsymbol{\kappa})$, the quasi-conservative forces conjugate to $\boldsymbol{\kappa}$ are given by $\boldsymbol{q}_Q = \rho\, \partial\psi_p / \partial\boldsymbol{\kappa}$. The dissipative forces are $\boldsymbol{q} = -\boldsymbol{q}_Q$, and differentiation with respect to time provides the rate equation

$$\dot{\boldsymbol{q}} = -\boldsymbol{H}(\boldsymbol{\kappa}) \bullet \dot{\boldsymbol{\kappa}} \tag{27.58}$$

where

$$\boldsymbol{H}(\boldsymbol{\kappa}) = \rho \frac{\partial^2 \psi_p(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}^2} \tag{27.59}$$

is the tensor of tangent plastic moduli. Substituting (27.57) into (27.58) leads to the equation

$$\dot{\boldsymbol{q}} = -\boldsymbol{H}(\boldsymbol{\kappa}) \bullet \dot{\boldsymbol{\kappa}} = -\boldsymbol{H}(\boldsymbol{\kappa}) \bullet \boldsymbol{Q} \bullet \boldsymbol{q}_v = -\boldsymbol{H}(\boldsymbol{\kappa}) \bullet \boldsymbol{Q} \bullet (\boldsymbol{q} - \bar{\boldsymbol{q}}) \tag{27.60}$$

which coincides with (27.45) if $\boldsymbol{H}(\boldsymbol{\kappa}) \bullet \boldsymbol{Q}$ is a multiple of the unit tensor. For linear hardening, the tensor $\boldsymbol{H}$ is constant (independent of $\boldsymbol{\kappa}$), and it is sufficient to set $\boldsymbol{Q} = \tau^{-1}\boldsymbol{H}^{-1}$. This is the most usual form of the hardening Duvaut–Lions model, considered by Simo, Kennedy and Govindjee (1988). However, even if the hardening is nonlinear, we can achieve a perfect agreement with (27.45) by generalizing (27.56) to

$$\phi^*_{vq}(\boldsymbol{q}_v; \boldsymbol{\kappa}) = \tfrac{1}{2} \boldsymbol{q}_v \bullet \boldsymbol{H}^{-1}(\boldsymbol{\kappa}) \bullet \boldsymbol{q}_v \tag{27.61}$$

Recall that the dual dissipation potential can depend on the state variables taken as parameters, and so its dependence on $\boldsymbol{\kappa}$ is acceptable. Nevertheless, problems occur if the tensor of plastic moduli ceases to be positive definite, which happens in the softening regime. If $\boldsymbol{H}$ has some negative eigenvalues but their number does not change (e.g. this is the case for linear softening), it is still invertible and the only problem is that the potential (27.56) is not convex. Under certain conditions it is possible to show that the model remains thermodynamically admissible; see Problem 27.4. However, if some eigenvalues of $\boldsymbol{H}$ smoothly change from positive to negative as the hardening variables $\boldsymbol{\kappa}$ evolve (which corresponds to a smooth transition from hardening to softening), at some state the tensor $\boldsymbol{H}$ must become singular and thus not invertible, and close to this state certain components of its inverse blow up. In this case, it is better to give up the physical meaning of $\bar{\boldsymbol{q}}$ as the relaxation limit, and to keep $\boldsymbol{Q}$ in (27.56) as a constant, positive definite tensor, different from $\boldsymbol{H}^{-1}(\boldsymbol{\kappa})$.

One advantage of the present thermodynamic formulation is that the dissipation inequality is automatically satisfied, provided that the (static) yield function $f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})$ and the potential $\phi^*_v(\boldsymbol{\sigma}_v, \boldsymbol{q}_v)$ are convex. For the potential in the form (27.54) or (27.55) it is sufficient to verify that $\phi^*_{vq}(\boldsymbol{q}_v)$ is convex, which is the case for example for the quadratic potential (27.56) with a positive definite $\boldsymbol{Q}$.

Another advantage is that the clear decomposition into the rate-independent plastic part and the viscous part allows an easy extension to nonassociated flow. To this end, we replace the yield function $f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})$ in (27.51) by a suitable plastic potential $g(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})$. The line connecting the current state $(\boldsymbol{\sigma}, \boldsymbol{q})$ with the relaxation limit $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})$ is now

orthogonal (in the energy metric) to the equipotential surface, $g = \text{const.}$, instead of the yield surface, $f = 0$, and $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})$ loses the property of being the closest point on the yield surface. The dissipation

$$\begin{aligned} \mathcal{D} &= \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_{vp} + \boldsymbol{q} \bullet \dot{\boldsymbol{\kappa}} = \bar{\boldsymbol{\sigma}} : \dot{\boldsymbol{\varepsilon}}_{vp} + \bar{\boldsymbol{q}} \bullet \dot{\boldsymbol{\kappa}} + \boldsymbol{\sigma}_v : \dot{\boldsymbol{\varepsilon}}_{vp} + \boldsymbol{q}_v \bullet \dot{\boldsymbol{\kappa}} \\ &= \dot{\lambda} \left(\bar{\boldsymbol{\sigma}} : \frac{\partial g(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})}{\partial \bar{\boldsymbol{\sigma}}} + \bar{\boldsymbol{q}} : \frac{\partial g(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}})}{\partial \bar{\boldsymbol{q}}} \right) + \boldsymbol{\sigma}_v : \frac{\partial \phi^*_v(\boldsymbol{\sigma}_v, \boldsymbol{q}_v)}{\partial \boldsymbol{\sigma}_v} + \boldsymbol{q}_v : \frac{\partial \phi^*_v(\boldsymbol{\sigma}_v, \boldsymbol{q}_v)}{\partial \boldsymbol{q}_v} \end{aligned} \tag{27.62}$$

is non-negative if ϕ^*_v is convex and if, at the same time, the expression in parentheses is nonnegative for all $\bar{\boldsymbol{\sigma}}$ and $\bar{\boldsymbol{q}}$ satisfying $f(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{q}}) = 0$. Note that this condition is somewhat less restrictive than for the nonassociated Perzyna model, where a similar expression had to be nonnegative for all $\boldsymbol{\sigma}$ and $\boldsymbol{q}$ satisfying $f(\boldsymbol{\sigma}, \boldsymbol{q}) > 0$.

The present exposition of the thermodynamically-based Duvaut–Lions model exploited certain ideas from Ristinmaa and Ottosen (1998) and Johansson (1996). Algorithmic aspects of Perzyna and Duvaut–Lions viscoplasticity have been addressed, e.g. by Simo and Hughes (1998), Runesson, Ristinmaa and Mähler (1999), Simo and Govindjee (1991), Perić (1993) and Hartmann, Lührs and Haupt (1997). Finally, let us mention that there exist other types of viscoplastic formulations, e.g. the so-called consistency model (Wang, Sluys and de Borst, 1997), or the generalized viscoplasticity (Auricchio and Taylor, 1994), which is the rate-dependent extension of generalized plasticity (Lubliner, 1991).

PROBLEMS

Problem 27.1: Using the thermodynamic approach, derive the basic equations governing the response of the uniaxial Schwedoff model in Figure 27.6.

Problem 27.2: Show that for perfect J_2-viscoplasticity the Duvaut Lions formulation with a variable relaxation time (27.44) is exactly equivalent to the Perzyna formulation with the power-type potential function (27.30).

Problem 27.3*: Write the free energy and the dual dissipation potential from which it is possible to derive the basic version of Chaboche model.
Hint: Combine the information on Perzyna viscoplasticity and Chaboche model (Section 27.2), Armstrong–Frederick hardening rule (Section 20.1.4), and thermodynamic formulation of hardening plasticity (Section 23.5.4).

Problem 27.4: Derive the conditions under which the associated Duvaut–Lions model with linear softening ($\boldsymbol{H}^{-1}$ in (27.61) constant but not positive definite) is thermodynamically admissible.

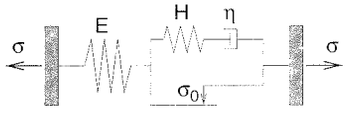


Figure 27.6 Schwedoff model

28

Material Models for Concrete Creep and Shrinkage

In contrast to structural metals, which exhibit time-dependent behavior only under elevated temperatures or under very high stresses, strains in concrete grow in time even if the applied stresses are much smaller than the material strength. Although, with a few exceptions, the time-dependent deformations of concrete normally have little effect on the safety against collapse, they play an important role in serviceability and durability of structures, a problem of enormous economic consequences.

The graduate curricula usually do not provide enough room for a full exposition of this vast and complex subject. However, no graduate student of structural engineering should leave a university without at least hearing about it. Therefore, a short, elementary and practically oriented exposition of the fundamentals is included in this book, even though this inelastic phenomenon is qualitatively very different and its modeling disconnected from plasticity.

28.1 CONCRETE CREEP AND SHRINKAGE AND ITS PROPERTIES

The time-dependent inelastic strains of concrete consist of (1) *shrinkage strain*, ε_{sh} (negative), which is independent of stress σ, and (2) additional mechanical strain ε_v produced by stress, called in general *creep*. The existence of shrinkage has been known since the invention of modern concrete and the existence of creep was discovered by Hatt in 1907. For the service stress levels of concrete structures, it is generally assumed that creep depends on stress σ linearly, and the modeling of this phenomenon is based on linear viscoelasticity (which explains the subscript v in ε_v). During drying or at high stress levels, concrete creep is nonlinear, but the term viscoplasticity, used for nonlinear creep of metals at high temperature, is unsuitable because the deformations are not of plastic nature (except under extremely high confining pressure). Nonlinear effects are introduced into the formulation by the influence of humidity on creep parameters and by the superimposed cracking strain, ε_c.

When a specimen of concrete is exposed to a dry atmosphere, it gradually shrinks (Figure 28.1(a)), the shrinkage strain ε_{sh} being an increasing function of time t which approaches a finite bound at a gradually decreasing rate (Figures 28.2(a) and 28.8). Most of the shrinkage of normal concretes is the *drying shrinkage*, which is caused mainly by increase of capillary tension of pore water and solid surface tension of pore walls, engendered by diffusion of pore water out of the specimen. When a specimen

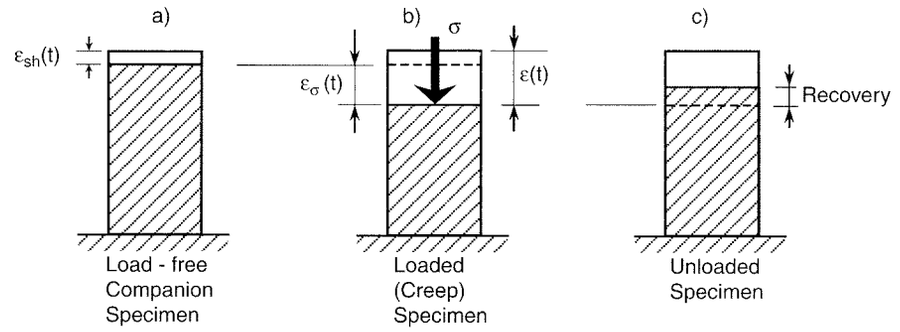


Figure 28.1 a) Shrinkage deformation, b) mechanical deformation (elastic plus creep) produced by stress, and c) recovery after unloading

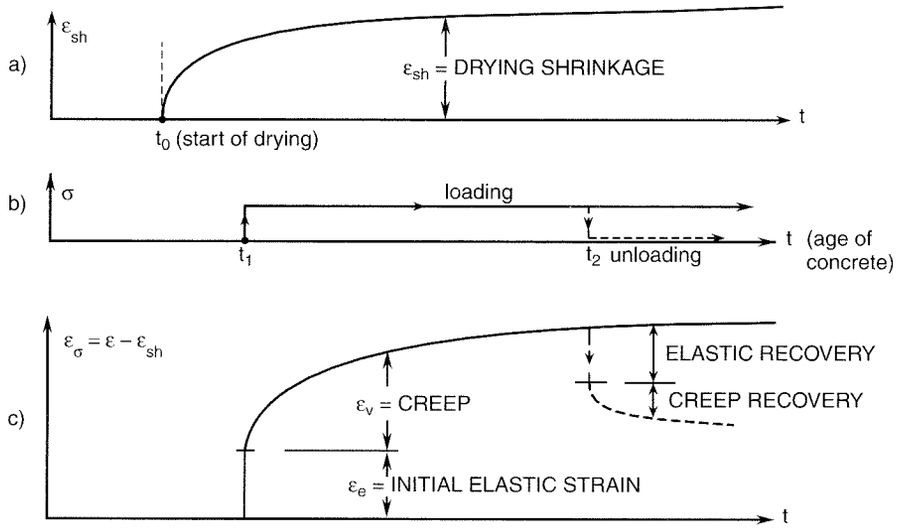


Figure 28.2 Curves of shrinkage, creep and recovery after unloading

is immersed in water, it exhibits swelling (positive ε_{sh}), which is normally an order of magnitude less than drying shrinkage, and thus often negligible. The swelling is caused by an increase of the water content of hindered adsorbed layers only a few molecules thick, acting as part of the solid microstructure.

A part of shrinkage, called the *autogeneous shrinkage* (or chemical shrinkage or endogeneous shrinkage) is caused by volume changes associated with the chemical reactions of cement hydration. The autogeneous shrinkage is the only shrinkage that occurs in sealed specimens. It represents only about 5% of the drying shrinkage of normal concretes, and is usually neglected. However, for the modern high strength concretes cast with a very low water-cement ratio (0.35 or less), the autogeneous shrinkage is of a comparable magnitude as the drying shrinkage and must be taken into account. After concrete dries to a relative humidity 80% or less in the pores, the chemical reactions of hydration stop, and so does the autogeneous shrinkage. Since much of drying shrinkage occurs when the pore humidity drops well below 80%, little autogeneous shrinkage is superposed on drying shrinkage. The cores of cross sections of normal concrete structures take years to dry out. The cores of thick cross sections (thicker than about 40 cm) never dry out during the lifetime, and thus the autogeneous

shrinkage may be happening in the core even if the surface is exposed to a drying atmosphere.

When a concrete specimen is loaded (say, at age t_1), it deforms more than the load-free companion specimen (Figure 28.1(b)). The difference is the *mechanical strain* $\varepsilon_\sigma(t)$, consisting of (1) the *elastic strain* ($\varepsilon_e = \sigma/E(t_1)$ in the case of uniaxial stress σ, $E(t_1)$ = Young's modulus at age t_1), also called the instantaneous strain or short-time strain, and (2) *creep (or viscoelastic) strain* $\varepsilon_v(t)$. The time-dependent shear strain is purely due to creep because shrinkage is volumetric and does not produce shear strains. At constant stress, creep increases at a gradually decreasing rate and, as far as known, without any bound (Figures 28.2(b),(c)). If concrete is unloaded at some age t_2 (Figure 28.1(c)), there is an instantaneous strain recovery corresponding to Young's modulus $E(t_2)$ at age t_2, followed by a long-time monotonic partial recovery of creep strain at a gradually decreasing rate (Figure 28.2(c), dashed curve).

The creep occurring at constant moisture content (i.e. in the absence of moisture exchange with the environment) is called the *basic creep*. It is caused by breakage and reformation of bonds at various highly stressed sites within the colloidal microstructure of the calcium silicate hydrate gels in the hardened cement paste.

Simultaneous drying causes additional creep, called the *drying creep* or Pickett effect (Pickett, 1956). It has a complex physical origin – the drying elevates the local stress peaks within the microstructure of calcium silicate hydrates, and thus increases the rate of bond breakages (Bažant *et al.*, 1997). A large part of the observed drying creep in compression is an apparent creep (Wittmann, 1974; Bažant and Wu, 1974), which in reality represents stress-induced shrinkage, treated as creep for convenience. Explanation: the drying of concrete specimens is nonuniform and the differences in local shrinkage produce self-equilibrated stresses large enough to cause extensive microcracking. This tends to reduce the observed drying shrinkage of the companion specimen, in comparison to what it would be if there were no microcracking. Therefore, the shrinkage customarily measured on load-free specimens is in fact an apparent shrinkage which is significantly less than the true shrinkage in absence of microcracking that takes place under sufficient compressive stress, simultaneously with compressive creep. Drying creep evolves in time similarly to shrinkage and exhibits a similar dependence on thickness D (Figure 28.9), while the basic creep is independent of D.

An important property of concrete is its *aging*. What is meant by aging is very different from the aging of other materials, e.g. the gradual degradation of strength of fiber composites. In concrete, the aging causes the strength as well as the elastic modulus E to increase with age t at a gradually decreasing rate (note that the elastic recovery in Figure 28.2(c) is smaller than the initial elastic strain). It causes a large reduction of creep, i.e. concrete specimens loaded at a higher age creep less. An important cause of the aging is the chemical process of *hydration*. The hydration products, chiefly the tri-calcium silicate hydrate gel (tobermorite gel), gradually fill the pores of hardened cement paste; hence, the mean size of capillary pores decreases, which gradually stiffens and strengthens the microstructure. However, the aging process continues for many years, unless the concrete has dried below about 80% relative humidity in the capillary pores. The source of such long-time aging is gradual relaxation of stress peaks on the nanoscale in the microstructure.

Since aging plays such an important role, *the time is always measured from the set of concrete*, i.e. from the instant when concrete first becomes a solid (typically

several hours after mixing). This means that, in creep calculations, the time variable t always corresponds to the age of concrete. During a certain initial period, concrete remains in the formwork and cannot dry, or is artificially prevented from drying (by spraying, covers, membranes) to achieve proper curing. For shrinkage calculations it is important to know the approximate age t_0 at the start of drying, which roughly corresponds to the age at the removal of formwork or the end of other drying protection in the process of concrete curing.

28.2 LINEARITY OF CREEP AND PRINCIPLE OF SUPERPOSITION IN TIME

If only stresses less than about 50% of the strength limit (for the same ratio of principal stresses) are considered (which includes the service stress range of structures), the creep law is, in the current design practice, generally assumed to be linear with respect to the stress history, $\sigma(t)$. The linearity hypothesis agrees very well with test results in the case of basic creep (creep at no moisture content changes), but has a considerable error in the case of drying, which is neglected.

The linearity, first of all, means that, at constant uniaxial stress σ applied at concrete age t', the corresponding strain $\varepsilon(t)$ at any time $t \geq t'$ may be written as

$$\varepsilon(t) \;=\; \sigma\, J(t,t') \tag{28.1}$$

where $J(t,t')$ is the uniaxial *compliance function*, a material property characterizing creep. The typical curves of compliance J versus current age t at various constant values of the age at load application, t', are shown in Figures 28.3 and 28.4(a).

Equation (28.1) reflects the proportionality of the material response. The linearity, however, implies more:

$$\begin{array}{ccc} \text{If} & & \\ \text{stress history } \sigma_a(t) & \text{corresponds to} & \text{strain history } \varepsilon_a(t) \\ & \text{and} & \\ \text{stress history } \sigma_b(t) & \text{corresponds to} & \text{strain history } \varepsilon_b(t) \\ & \text{then} & \\ \sigma(t) = \sigma_a(t) + \sigma_b(t) & \text{corresponds to} & \varepsilon(t) = \varepsilon_a(t) + \varepsilon_b(t) \end{array} \tag{28.2}$$

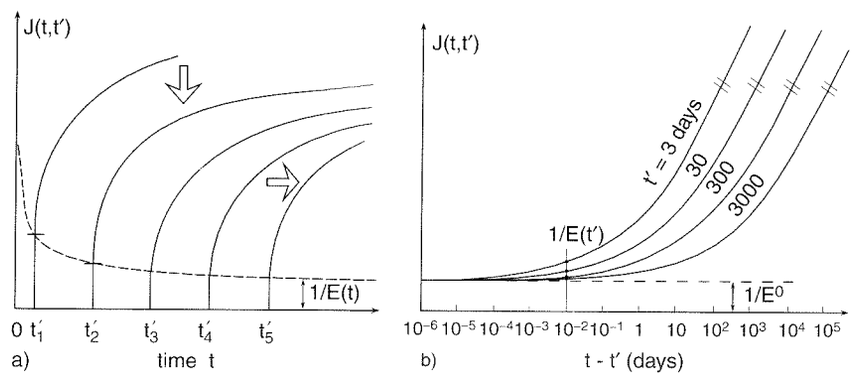


Figure 28.3 Compliance function shown by unit creep curves for different ages at loading

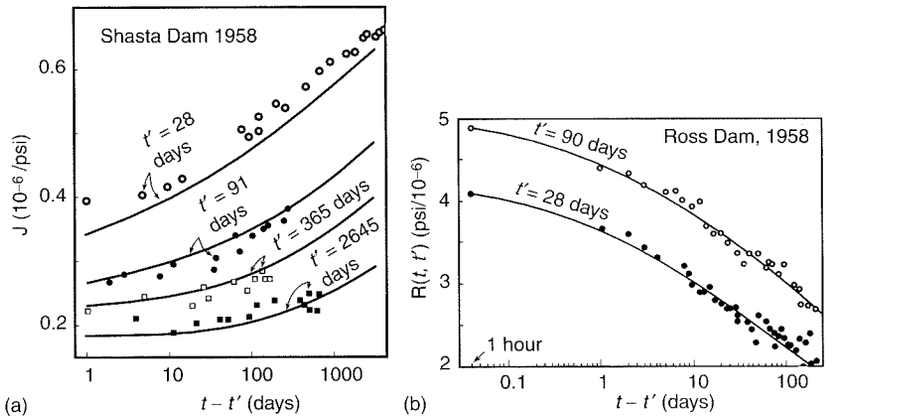


Figure 28.4 a) Compliance function compared to measurements on Shasta Dam concrete (after RILEM, 1988), b) relaxation function predicted by principle of superposition from measured compliance function, compared to relaxation measurements of Hanson (1953) and Harboe *et al.* (1958)

This property represents the *principle of superposition* (in time), which was proposed in general for non-aging phenomena by Boltzmann (1874) and for aging phenomena by Volterra (1909).

Every stress history may be decomposed into infinitesimal steps, $\mathrm{d}\sigma(t')$, applied at various times t' less than the current time t. According to (28.1), each step history produces infinitesimal strain history $\mathrm{d}\varepsilon(t) = J(t,t')\,\mathrm{d}\sigma(t')$ (Figure 28.5(a)). Summing all these infinitesimal contributions and adding the initial strain $\varepsilon_0(t)$, one obtains

$$\varepsilon(t) = \int_0^t J(t,t')\,\mathrm{d}\sigma(t') + \varepsilon_0(t) \tag{28.3}$$

The *initial strain* ε_0 in general includes not only the shrinkage strain ε_{sh}, but also the thermal strain and cracking strain. Only the shrinkage strain will be considered here, in which case $\varepsilon_0(t) = \varepsilon_{sh}(t)$.

Equation (28.3) is not only a consequence of the principle of superposition but its equivalent alternative statement. Indeed, (28.2) follows from (28.3). Using superposition, one may further generalize equation (28.3) for the case of multiaxial stress and strain; see Section 29.3.1.

Note that the integral in (28.3) is not the standard (Riemann) integral but a generalized type of integral called the *Stieltjes integral*, in which piece-wise continuous histories $\sigma(t)$ (with jumps) are admitted. So, if there is a jump from σ^- to σ^+ at time t_j, then the contribution to the integral (28.3) from the jump is $(\sigma^+ - \sigma^-)J(t,t_j)$,

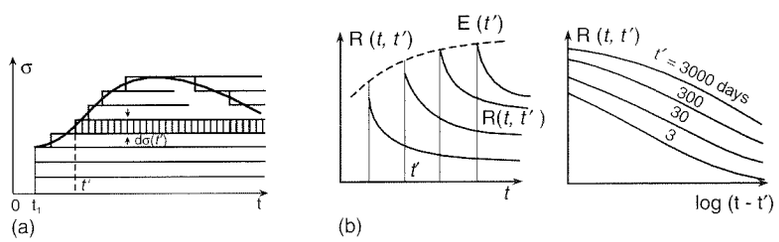


Figure 28.5 a) Decomposition of stress history in infinitesimal stress steps, b) relaxation function of concrete

which follows from (28.3) by integrating the increments $\mathrm{d}\sigma(t_j)$ from σ^- to σ^+ at constant time t_j.

For the frequently assumed case that the concrete remains stress-free until, at some age t_1, a finite stress σ_1 is applied suddenly (by a jump) and subsequently the stress $\sigma(t)$ varies continuously, (28.3) may be written in terms of the Riemann integral as follows:

$$\varepsilon(t) = \sigma_1 J(t,t_1) + \int_{t_1^+}^{t} J(t,t') \frac{\mathrm{d}\sigma(t')}{\mathrm{d}t'} \, \mathrm{d}t' + \varepsilon_0(t) \tag{28.4}$$

where t_1^+ denotes the beginning of integration just after the initial jump.

When the strain history $\varepsilon(t)$ is prescribed, then (28.3) or (28.4) represent an integral equation for the stress history $\sigma(t)$. It is of a well known type – *the Volterra integral* equation, a special case of integral equations of the second kind. The compliance function $J(t,t')$ is called the *kernel* of the integral equation.

For realistic forms of $J(t,t')$ (Figure 28.4(a)), this integral equation cannot be solved analytically. But accurate numerical solutions, in which the integral is approximated by a discrete sum, are easy. In this manner one can, for instance, solve the relaxation function $R(t,t')$, which represents the uniaxial stress history $\sigma(t)$ caused by a constant strain $\varepsilon = 1$ enforced at any age t' ($t' \leq t$); see Figure 28.5(b). The solution of relaxation curves by the principle of superposition agrees very well with experiments, but only in absence of severe drying; see Figure 28.4(b). If the relaxation function is known, the superposition principle can be written in its inverse form

$$\sigma(t) = \int_0^t R(t,t') \, \mathrm{d}\left[\varepsilon(t') - \varepsilon_0(t')\right] \tag{28.5}$$

which gives the stress history in terms of the strain history.

In general, when neither $\sigma(t)$ nor $\varepsilon(t)$ is prescribed, (28.3) or (28.5) represent a uniaxial stress-strain relation for creep. In the special case that the compliance function depends only on the time lag $t-t'$, i.e. $J(t,t') \equiv J_0(t-t')$, one has the general stress-strain relation of *classical linear viscoelasticity* of polymers which exhibit no aging.

The fact that $J(t,t')$ does not depend merely on the time lag $t-t'$, but on t and t' separately, is a consequence of the aging of concrete – the creep curves of concrete for different ages at loading t' are very different (Figure 28.3). In that case the continuum mechanics theory based on (28.3) or (28.5) is called the *aging linear viscoelasticity*. Because of aging, the Laplace transform methods, very effective for non-aging linear viscoelasticity, cannot, unfortunately, be transplanted. Thus, the aging is a major complication of concrete creep theory. It requires that accurate solutions be obtained numerically, by step-by-step integration in time.

When the stress exceeds about one half of the strength limit, creep becomes nonlinear, which is due to time-dependent growth of microcracks. The nonlinear creep behavior is best shown by plotting the so-called stress-strain isochrones, obtained by conducting constant load creep tests at different stress levels and connecting the points reached at the same time. The typical isochrones are shown in Figure 28.6.

28.3 SIMPLE MODELS FOR CREEP AND SHRINKAGE

28.3.1 Basic Creep

Figures 28.3 and 28.4(a) show the typical unit creep curves (compliance curves) of sealed concrete specimens (basic creep) for unit stress $\sigma = 1$, applied at various

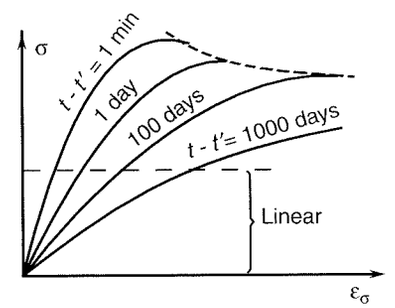


Figure 28.6 Creep isochrones at high stress

ages at loading, $t' = t'_1, t'_2, t'_3, \ldots$. If the first reading is taken several minutes after loading and the initial deformation is regarded as elastic, as has often been done, then the initial strain $1/E(t)$ decreases with the age a loading, at a gradually decreasing rate (Figure 28.3(a)). This is one consequence of the long-time progress of cement hydration. Another consequence is the decrease of creep with an increase of age at loading. For loading of young concrete, the effect of loading age increase is predominantly to shift the creep curve down vertically, while for loading of old concrete, the effect is predominantly to shift the creep curve horizontally to the right (Figure 28.3(a)).

Analysis and design of structures necessitates a realistic model for predicting the compliance function and shrinkage function of the given concrete. This is a difficult problem that has engendered polemics for a long time. The problem calls for careful analysis of extensive data from long-time tests conducted on different concretes and in different environments, generally exhibiting large statistical scatter.

A realistic prediction model for creep and shrinkage properties of concrete is *model B3* (Bažant and Baweja, 1995, 2000a), summarized in Appendix E. However, such a model is inevitably more involved and its details are unimportant for explaining the essentials of creep and shrinkage analysis of structures, the main goal of this section. Therefore, it is preferred to present here the following simple and for the most part empirical, yet not too unrealistic, creep prediction formula, called the *short form of B3 model* (Bažant and Chern, 1985; Bažant and Baweja, 1996, 2000b):

$$J(t,t') = \frac{1}{E_0} + q_s \ln\left[1 + \psi\left(t'^{-m} + \alpha\right)(t-t')^n\right] + C_d(t,t',t_0) \tag{28.6}$$

Here, all the ages are in days, $t \geq t'$, and $C_d(t,t',t_0)$ is the additional creep compliance due to simultaneous drying (cross coupling with drying shrinkage), to be defined in Section 28.3.3. The typical values of the model parameters are $n = 0.1$, $m = 0.5$ and, if the times are expressed in days, $\alpha = 0.001$ and $\psi = 0.3$. As crude empirical estimates, $E_0 = E_{28}/0.6$ and $q_s = 11.4/E_{28}$, where E_{28} is the conventional elastic modulus of concrete at age 28 days. For basic creep, for which $C_d(t,t',t_0) = 0$, formula (28.6) is called the *log-double-power law*, which is a generalization of the *double-power law* (Bažant and Osman, 1976).

Parameter E_0 represents the so-called *asymptotic instantaneous modulus*, which corresponds to extrapolating the creep curve to loading durations orders of magnitude shorter than ever measured. The reason for introducing E_0 with such a definition is that concrete happens to exhibit non-negligible creep, even for extremely short load durations, all the way down to the durations of the sound vibration periods (governing

the sound velocity in measurements of the dynamic modulus E_{dyn}). As experimental data have shown (Bažant and Baweja, 1995, 2000a), E_0 can be considered as age-independent, a constant, whereas an elastic modulus corresponding to any load duration not much shorter than the sound vibration periods is age-dependent. The usual *static modulus of elasticity*, including its age-dependence, is obtained from (28.6) as follows:

$$E(t') = \frac{1}{J(t' + \Delta t, t')} \tag{28.7}$$

where $\Delta t \approx 0.01$ day gives (for $t' = 28$) good agreement with the *conventional elastic modulus* E_{28} defined in the codes. The *dynamic modulus* E_{dyn}, including its age-dependence, is obtained by substituting $\Delta t \approx 10^{-7}$ day.

Since concrete creep is significant for at least ten orders of magnitude of load duration, the unit creep curves plotted in the actual time scale, as in Figure 28.3(a), can show the creep trend for only a small portion of the full time range – only for up to one to two orders of magnitude of t and t'. The plots in the actual time scale obscure either long times or short times, or both. It is much more informative to plot the creep curves in the logarithmic time scale (comparisons with test data made in the actual time scale should be distrusted).

A typical such plot in the logarithmic scale of load duration $t - t'$ is shown in Figure 28.3(b). Now we see the reason for the term asymptotic modulus; since $\log 0 = -\infty$, the value of $1/E_0$ corresponds to $\log t' \to -\infty$, and so the curves for all the ages at loading approach the horizontal dashed line $1/E_0$ asymptotically from the left. The slope of the creep curves in the log-time initially increases, but at a certain creep strain there is a transition to a straight line, and this line has about the same slope for all the ages at loading. For a higher age t' at loading, this transition occurs at a longer duration $t - t'$ but at about the same strain. Note in Figure 28.4 the very large effect of the age at loading, t'.

One basic property of the creep curves for different t' is that they always approach each other with increasing t (i.e. that $\partial^2 J(t,t')/\partial t \partial t' \geq 0$). This is called the non-divergence property, which is satisfied by the full B3 model (Appendix E) but mildly violated by its short form (28.6). When a creep formula with divergent creep curves is used in the superposition principle to predict creep recovery after unloading, the recovery curve is nonmonotonic. This feature is thermodynamically questionable and may inhibit the convergence of a computer solution. Therefore, the short form is not suitable for large-scale computer analysis; exclusively the full B3 model should be used for that purpose.

Creep Coefficient

For structural creep analysis, it is conceptually convenient to characterize creep by a dimensionless *creep coefficient*, $\varphi(t,t')$, which is defined as the ratio of the creep strain for the load duration $t - t'$ to the initial 'elastic' strain at age t':

$$\varphi(t,t') = \frac{\text{creep strain at } t \text{ for stress applied at age } t'}{\text{initial 'elastic' strain at age } t'} = \frac{J(t,t') - 1/E(t')}{1/E(t')} = E(t')J(t,t') - 1 \tag{28.8}$$

Most codes and most authors have in the past defined creep by the creep coefficient instead of the compliance function, which is a sum of initial elastic strain and creep strain. However, if one considers, as typical in design, only the response for load durations at least an order of magnitude longer than the load duration corresponding to the definition of E, then it is only the sum (i.e. the total compliance J) that matters, the subdivision between the creep and elastic parts being irrelevant. Since the ratio of the J-values for load durations 1 second and 1 hour is quite large (typically between 1.50 and 1.75), the range of freedom in the definition of the initial deformation is considerable. However, this carries a danger.

The danger is that the code makers specify only the creep coefficient but not the corresponding value of E that gives the correct compliance J, in agreement with the measurement used to calibrate the code. The designer may then combine this creep coefficient with some different elastic modulus, for example with the value specified in the code that has been calibrated by tests involving load cycles and is very different from the moduli corresponding to the initial deformations in the creep tests behind the code. Therefore, the designer should always first establish the values of the compliance function J and determine φ from them, using (28.8). If that is done, the choice of the particular definition of elastic modulus has no appreciable effect on the calculated long-time response at variable stress (and no effect at all at constant stress). For example, one can predict E at age 28 days from the uniaxial compression strength (standard cylinder strength) $\bar{f}_c$ using the empirical formula (ACI, 1999)

$$\frac{E_{28}}{1 \text{ ksi}} = 57\sqrt{\frac{\bar{f}_c}{1 \text{ psi}}} \qquad \text{or, in SI units,} \qquad \frac{E_{28}}{1 \text{ GPa}} = 4.733\sqrt{\frac{\bar{f}_c}{1 \text{ MPa}}} \tag{28.9}$$

The actual code formula uses the reduced strength, f'_c, but to obtain an estimate of the mean E, $\bar{f}_c$ must be used because that is how this empirical formula was calibrated (Pauw, 1960).

In the case of extrapolation of short-time measurements on a given concrete to long time, however, it is important to use a compliance function that is correct also for very short times.

28.3.2 Mean Cross Section Shrinkage

Since shrinkage evolves differently at different points of the cross section of a concrete beam or plate, its effects should properly be analyzed two- or three-dimensionally, e.g. by subdividing the cross section or the structure into many finite elements. In that case one needs a local constitutive relation for shrinkage, which gives the shrinkage strain rate $\dot{\varepsilon}_{sh}$ as a function of the rate of pore humidity $\dot{h}_p$ at any point of the specimen (the pore humidity h_p is the relative vapor pressure in the pores, i.e. the partial pressure of water vapor divided by the saturation vapor pressure of water at the given temperature). Such a constitutive relation is relatively simple; see Appendix E.3.

For the sake of simplicity, however, structural analysis is normally conducted on the basis of an 'average' or 'effective' shrinkage of the cross section as a whole. Its mathematical formulation is not as simple as that for local shrinkage because the cross section is typically in a highly nonuniform state of pore humidity and residual stress. But the use of such 'average' shrinkage takes one or two dimensions away from the structural analysis, which is a great simplification.

According to model B3 (Bažant and Baweja, 1995, 2000a), the average longitudinal shrinkage of a cross section of a long beam or plate may be approximately calculated as

$$\varepsilon_{sh}(t) = -\varepsilon_{sh\infty} k_h S(t - t_0) \tag{28.10}$$

where t is the current age of concrete, t_0 is the age at the start of drying, $\varepsilon_{sh\infty}$ is the magnitude of the final shrinkage strain (typically 0.0002 to 0.0008), k_h is a coefficient depending on the average environmental humidity h (relative vapor pressure), and $S(\hat{t})$ is an increasing function of the duration of drying, $\hat{t} = t - t_0$. This function describes the evolution of normalized shrinkage strain $|\varepsilon_{sh}|/\varepsilon_{sh\infty}$ in a perfectly dry environment, starting from its initial value 0 at $\hat{t} = 0$ and approaching asymptotically 1 as $\hat{t} \to \infty$. A suitable formula is

$$S(\hat{t}) = \tanh \sqrt{\frac{\hat{t}}{\tau_{sh}}} \tag{28.11}$$

where τ_{sh} is called the shrinkage half-time, because it roughly indicates the time at which ε_{sh} reaches one half of its final value. The shrinkage half-time can be estimated as

$$\tau_{sh} = k_t (k_s D)^2 \tag{28.12}$$

in which k_s is a cross section shape factor ($k_s = 1$ for an infinite slab, 1.15...infinite cylinder, 1.25...infinite square prism, 1.30...sphere, 1.55...cube, based on solutions of the nonlinear diffusion equation for drying of concrete (Bažant and Najjar, 1972)); and $D = 2v/s$ = effective cross section thickness (v and s are the volume and exposed surface of the concrete part). The definition of D is such that, for an infinite slab, it represents the actual thickness. Factor $k_t \approx 0.03/C_1$ where C_1 is the diffusivity of pore water in saturated concrete, about 0.03 mm^2/day (concrete diffusivity is very low, as manifested for instance by the fact that the core of a standard 6-inch cylinder takes about 20 years to dry to the environmental humidity).

The fact that τ_{sh} is proportional to the square of thickness is a basic property of all kinds of diffusion processes. Another consequence of the diffusion origin of drying shrinkage is that the initial shrinkage curve must evolve as $\sqrt{\hat{t}}$. Function $S(\hat{t})$ in (28.11) is used because it satisfies this property asymptotically for small $\hat{t}$, and because it also exhibits the correct asymptotic form for long times $\hat{t}$, as required by the diffusion theory (Bažant, 1999).

The final shrinkage strain can be estimated from the empirical formula (Bažant and Baweja, 1996)

$$\varepsilon_{sh\infty} = \left[26 \left(\frac{w}{1 \text{ lb/ft}^3} \right)^{2.1} \left(\frac{\bar{f}_c}{1 \text{ psi}} \right)^{-0.28} + 270 \right] \times 10^{-6} \tag{28.13}$$

or, in SI units,

$$\varepsilon_{sh\infty} = \left[0.019 \left(\frac{w}{1 \text{ kg/m}^3} \right)^{2.1} \left(\frac{\bar{f}_c}{1 \text{ MPa}} \right)^{-0.28} + 270 \right] \times 10^{-6} \tag{28.14}$$

where w = specific water content of concrete mix, and $\bar{f}_c$ = average compression strength measured on cylinders of diameter 6 inches and length 12 inches at age 28 days.

The final shrinkage strain $\varepsilon_{sh\infty}$ corresponds to a perfectly dry environment ($h = 0$). In an environment of relative humidity $h > 0$ it is reduced by the factor

$$k_h = 1 - h^3 \qquad \text{if } h \le 0.98 \tag{28.15}$$

This formula is used only up to $h = 0.98$, which is the value recommended for sealed concrete (it is not 100%, the initial relative humidity in the pores of concrete, because cement hydration causes self-desiccation). With $k_h = 1 - 0.98^3 = 0.0588$, formula (28.10) gives for normal concretes (but not high-strength concretes) reasonable approximate values for the autogeneous shrinkage. To approximately model the swelling of concrete immersed in water, i.e. at $h = 100\%$, one may substitute $k_h = -0.2$ in (28.10).

The nature of the pore water diffusion process, governed by the diffusion equation, is illustrated by Figure 28.7(a), which portrays the typical profiles at subsequent times of the relative pore humidity h_p over the cross section. The free shrinkage of a material point is roughly proportional to the loss of water from concrete, which in turn is roughly proportional to the relative humidity h_p in the pores. The typical shapes of the shrinkage curves, proportional to function $S(\hat{t})$, are shown in Figure 28.8. Figure 28.9 compares the curves of shrinkage ε_{sh}, as well as the additional creep due to drying, for various relative environmental humidities h and for various thicknesses D. A change of h causes a vertical scaling of the shrinkage curve (or drying creep curve), while a change of thickness causes a horizontal shift of the curve in the logarithmic time scale.

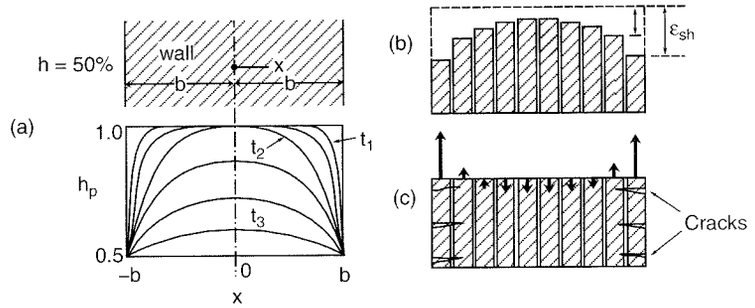


Figure 28.7 a) Pore humidity distributions during drying, b) free shrinkage of slices imagined to be cut out, and c) shrinkage stresses

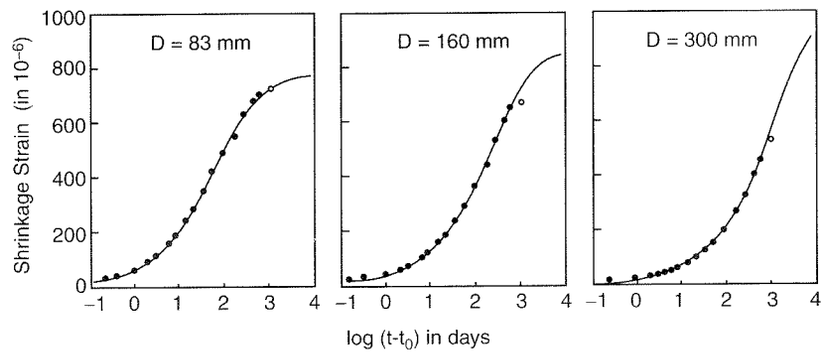


Figure 28.8 Shrinkage curves for different thicknesses and their comparison with measurements (after Wittmann *et al.*, 1987)

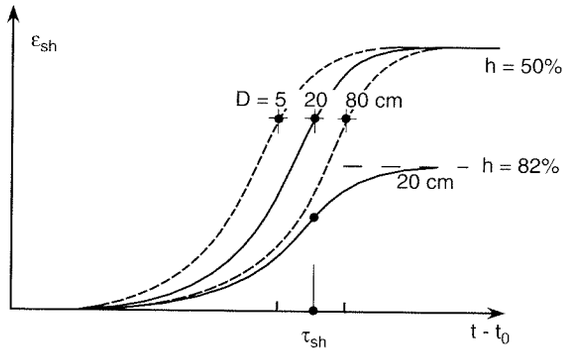


Figure 28.9 Thickness and humidity effects on shrinkage curves and drying creep curves

Note that shrinkage induces large self-equilibrated stresses, called shrinkage stresses, within the cross sections of drying concrete parts. This is clarified by Figure 28.7(b),(c), which shows what would happen if the cross section were cut in slices but drying proceeded the same way as in a contiguous cross section. Each slice would shrink independently, according to its current water content, as shown. In the middle of the drying process, this would cause the slices to develop different shortenings. In a long concrete beam or slab, however, the cross section must remain plane and, as a consequence, stresses required to achieve equal lengths of all the slices must develop. This obviously produces longitudinal tensile stresses near the surface and longitudinal compressive stresses in the core (but at the end of drying, this picture gets reversed, because of creep and unequal aging). The tensile stresses produced by shrinkage are normally large enough to cause microcracking.

28.3.3 Mean Drying Creep in the Cross Section

The additional mean cross sectional compliance caused by simultaneous drying, expressing a coupling between creep and shrinkage, can be estimated from the formula (Bažant and Baweja, 1996)

$$C_d(t,t',t_0) = q_5 \left(\mathrm{e}^{-g(t-t_0)} - \mathrm{e}^{-g(t'-t_0)}\right)^{1/2}, \qquad g(\hat{t}) \;=\; 3\left[1-(1-h)S(\hat{t})\right] \quad (28.16)$$

valid for $t \geq t' \geq t_0$. The formula can be used as a rough approximation (for long times only) even if $t_0 > t'$, in which case t' must be replaced by t_0. As an empirical estimate, $q_5 = 0.0006/\bar{f}_c$. Like shrinkage, and unlike basic creep, the drying creep is bounded. It depends on humidity and cross section thickness through the shrinkage shape function $S(\hat{t})$ defined in (28.11). Recall that, in the short form of B3 model, the additional compliance C_d is added to the log-double-power law describing the basic creep; see equation (28.6).

28.4 AGE-ADJUSTED EFFECTIVE MODULUS METHOD (AAEM)

Since the solutions of structural creep problems with a realistic compliance function such as the log-double-power law (28.6) cannot be performed analytically and require a step-by-step numerical integration, simplified approximate methods have been favored by designers. Besides, it makes little sense to embark on a very accurate linear creep analysis of a structure if the nonlinear effects of drying with the inherent

cracking and the large statistical uncertainty of creep and shrinkage are not taken into account.

The oldest simplified method is the *effective modulus method* (McMillan, 1916), in which the creep solution for time t is obtained by elastic structural analysis based on the so-called *effective modulus*

$$E_{ef} \;=\; \frac{1}{J(t,t')} \;=\; \frac{E(t')}{1+\varphi(t,t')} \quad (28.17)$$

Obviously, this method is exact only if the loads and stresses in a structure have a single-step history (i.e. are constant from the moment of first loading), which is far from true in statically indeterminate systems with significant stress redistributions induced by creep or drying, or if there are changes of the structural system during the construction, or if the permanent loads are not applied at once. The degree of deviation from a single-step history of load and stress governs the error magnitude. The error can be very large, especially for long-time response of structures loaded at a young age (Bažant and Najjar, 1973).

Two more sophisticated methods, which attempt to simplify the creep analysis of structures by introducing a simplified form of the compliance function $J(t,t')$, were historically important:

1. One such method was the *Glanville–Dischinger method* (Glanville, 1933; Dischinger, 1937, 1939), also called the rate-of-creep method or the theory of aging, and its refinement known as the improved Dischinger method (Rüsch, Jungwirth and Hilsdorf, 1973) or rate-of-flow method (England and Illston, 1965), in which the compliance function was simplified to a form that allows reducing structural creep problems to ordinary linear differential equations in transformed time, with constant coefficients, one first-order equation for each static or kinematic unknown.
2. Another such method was the *Maslov–Arutyunyan method* (Maslov, 1941; Arutyunian, 1952), which used another (purportedly better) simplification of $J(t,t')$ allowing the creep problem to be reduced to ordinary linear first-order differential equations with time-dependent coefficients, again one equation for each static or kinematic unknown; in the case of one unknown, this equation can be solved analytically in terms of an incomplete gamma function. This method, which became popular with mathematically inclined Soviet researchers, is obviously much more complicated than the first, but the first is substantially more complicated than the effective modulus method.

When more extensive test data and data of long duration became available and were systematically analyzed (Bažant, 1975; Bažant and Panula, 1978), it turned out that the afore-mentioned two methods leading to differential equations are overall not more accurate than the effective modulus method, which leads to algebraic linear equations with respect to time. The afore-mentioned two methods were thus rendered pointless. None of them is sufficiently accurate compared to the computer solutions for a realistic (unsimplified) compliance function based on long-time measurements with a broad range of ages at loading. A remedy that is sufficiently accurate in most basic situations was found in the *age-adjusted effective modulus method*, proposed and mathematically proven by Bažant (1972b), as a modification and refinement of the *relaxation method*, semi-empirically developed by Trost (1967).

The age-adjusted effective modulus (AAEM) method is formulated for a one-step loading history – a load that is applied suddenly at age t_1 and then is either constant

until the current time t or varies monotonically at a gradually decreasing rate. The response to multistep load histories can, of course, be obtained by superposing the solutions for several one-step histories. While the effective modulus method takes one step from the unstressed state of structure at time t_1^- *just prior* to the first loading to the current state at time t, the AAEM takes one step from the initial stressed state at time t_1^+ *just after* application of the load to the current state at time t. So, the initial state just after loading, which plays no role in the effective modulus method, must be calculated separately, which is accomplished by standard elastic analysis of the structure based on the initial elastic modulus $E(t_1)$. The reader who is not interested in the theoretical background can skip the details of the derivation and proceed directly to the final formulae (28.28)–(28.31).

The AAEM method may be derived as follows. The history of stress and strain between the initial and the current state is approximated by a linear combination of creep at constant stress and relaxation at constant strain. Recall that the strain α, suddenly applied at time t_1 and subsequently kept constant, produces stress history $\sigma(t) = \alpha R(t, t_1)$, where R is the relaxation function. Also, the stress β, suddenly applied at time t_1 and subsequently kept constant, produces strain history $\varepsilon(t) = \beta J(t, t_1)$, where J is the compliance function. According to the principle of superposition, the strain and stress histories

$$\varepsilon(t) = \alpha + \beta J(t, t_1) \tag{28.18}$$

$$\sigma(t) = \alpha R(t, t_1) + \beta \qquad (t \geq t_1) \tag{28.19}$$

correspond to each other (i.e. satisfy the viscoelastic constitutive equations). The coefficients α and β can be expressed in terms of the stress values $\sigma(t_1^+) = \sigma_1$, just after load application, and $\sigma(t) = \sigma_1 + \Delta\sigma$, at current time t. Substituting these values into (28.19) and recalling that $R(t_1, t_1) = 1/J(t_1, t_1) = E(t_1) =$ elastic modulus at age t_1, we obtain a set of two linear equations for α and β,

$$\alpha E(t_1) + \beta = \sigma_1 \tag{28.20}$$

$$\alpha R(t, t_1) + \beta = \sigma_1 + \Delta\sigma \tag{28.21}$$

from which it is possible to calculate the coefficients

$$\alpha = \frac{\Delta\sigma}{R(t, t_1) - E(t_1)} \tag{28.22}$$

$$\beta = \frac{\Delta\sigma\, E(t_1)}{E(t_1) - R(t, t_1)} + \sigma_1 \tag{28.23}$$

Using this result, it is easy to evaluate from (28.18) the initial (elastic) strain,

$$\varepsilon_1 = \varepsilon(t_1^+) = \alpha + \beta J(t_1, t_1) = \alpha + \frac{\beta}{E(t_1)} = \frac{\sigma_1}{E(t_1)} \tag{28.24}$$

and the strain increment,

$$\begin{aligned} \Delta\varepsilon &= \varepsilon(t) - \varepsilon(t_1^+) = \beta\,[J(t, t_1) - J(t_1, t_1)] \\ &= \Delta\sigma \frac{E(t_1)J(t, t_1) - 1}{E(t_1) - R(t, t_1)} + \sigma_1 \left[J(t, t_1) - \frac{1}{E(t_1)} \right] \end{aligned} \tag{28.25}$$

The expression $E(t_1)J(t, t_1) - 1$ is recognized as the creep coefficient $\varphi(t, t_1)$; see (28.8). So, if we introduce the so-called *age-adjusted effective modulus* (Bažant, 1972b),

$$E''(t, t_1) = \frac{E(t_1) - R(t, t_1)}{E(t_1)J(t, t_1) - 1} = \frac{E(t_1) - R(t, t_1)}{\varphi(t, t_1)} \tag{28.26}$$

and replace $\sigma_1/E(t_1)$ by ε_1 according to (28.24), formula (28.25) assumes a convenient form

$$\Delta\varepsilon = \frac{\Delta\sigma}{E''(t, t_1)} + \varepsilon_1\, \varphi(t, t_1) \tag{28.27}$$

This is the fundamental equation of AAEM, stating that the increment of strain over the interval $(t_1, t]$ is equal to the increment of stress divided by the effective modulus plus the initial (elastic) strain multiplied by the creep coefficient.

To keep the previous derivation simple, we have neglected the effect of shrinkage. However, this effect can be automatically incorporated by replacing $\varepsilon(t)$ with the mechanical strain $\varepsilon_0(t) = \varepsilon(t) - \varepsilon_{sh}(t)$, which means that the increment of shrinkage strain must be added to the right-hand side of (28.27). The final form of the *basic equation of AAEM* (in the case of uniaxial stress) is then

$$\Delta\varepsilon(t) = \frac{\Delta\sigma(t)}{E''(t, t_1)} + \varepsilon(t_1^+)\, \varphi(t, t_1) + \Delta\varepsilon_{sh}(t), \qquad \varepsilon(t_1^+) = \frac{\sigma(t_1^+)}{E(t_1)} \tag{28.28}$$

where

$$\Delta\varepsilon(t) = \varepsilon(t) - \varepsilon(t_1^+), \quad \Delta\sigma(t) = \sigma(t) - \sigma(t_1^+), \quad \Delta\varepsilon_{sh}(t) = \varepsilon_{sh}(t) - \varepsilon_{sh}(t_1) \tag{28.29}$$

For convenience, the age-adjusted effective modulus, whose primary definition is (28.26), can be expressed in the form

$$E''(t, t_1) = \frac{E(t_1)}{1 + \chi(t, t_1)\varphi(t, t_1)} \tag{28.30}$$

which represents a correction of the effective modulus. This has the advantage that the so-called *aging coefficient*,

$$\chi(t, t_1) = \frac{E(t_1)}{E(t_1) - R(t, t_1)} - \frac{1}{\varphi(t, t_1)} \tag{28.31}$$

varies relatively little (usually from 0.5 to 1.0, with 0.8 as the most typical value).

The value $\chi = 1$ characterizes the limiting case of a non-aging material well. Indeed, if there is no aging, as in the case of shorter creep durations for concrete loaded at old age, the optimal value of χ is close to 1 (about 0.992), and E'' is nearly equal to E_{ef}. This fact (which would not be true if the exponent n in creep formula (28.6) were much larger than 0.1, say 0.5) clarifies that χ accounts principally for the effect of aging. Hence the name.

Tables of χ, computed for certain compliance functions, have been included in ACI Committee 209 design recommendations. To avoid a computer solution of the relaxation function $R(t, t')$ for a given compliance function $J(t, t')$, one may use the following semi-empirical approximate formula with correct asymptotic properties (Bažant and Kim, 1979a):

$$R(t, t_1) = \frac{0.992}{J(t, t_1)} - \frac{0.115}{J(t, t-1)} \left[\frac{J(t - \Delta, t_1)}{J(t, t_1 + \Delta)} - 1 \right], \qquad \Delta = \frac{t - t_1}{2} \tag{28.32}$$

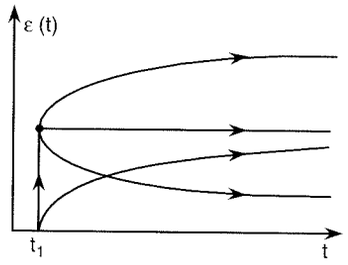


Figure 28.10 Strain histories for which AAEM is exact

For normal concretes, the error of this formula (recommended for calculating χ in CEB-FIP Model Code 1990, Eqs. 5.8-7 and 5.8-3) is generally under 1%; t and t_1 must be given in days.

The basic equation (28.28) of AAEM is exact for all the strain histories representing a linear transformation of the creep coefficient curve, i.e.

$$\varepsilon(t) - \varepsilon_{sh}(t) = a + b\varphi(t, t_1) \qquad (t > t_1) \tag{28.33}$$

where a and b are arbitrary constants. This includes a broad range of strain histories illustrated in Figure 28.10, for all of which the AAEM is exact. The actual strain histories in structures under permanent load are usually well approximated by (28.33). This is the reason for the good accuracy of AAEM in a broad range of problems. By contrast, the effective modulus method is exact only for one history, namely the creep at constant stress.

The AAEM has been generalized in a matrix form for the vectors of force and displacement components (Bažant, 1987), and also for bending creep of composite cross sections (Lazić and Lazić, 1984, 1985) and for loads varying in time monotonically at a decreasing rate (Khazanov, private communication 1990).

29

Creep and Shrinkage Effects in Concrete Structures

29.1 HOMOGENEOUS AND NONHOMOGENEOUS STRUCTURES

Typically, the stresses in concrete structures vary in time even if the load is constant. An exception is a structure with homogeneous linear creep properties.

In a homogeneous elastic structure, the stresses produced by applied forces do not depend on the value of the elastic modulus. Since the basic equation (28.28) of AAEM is a quasi-elastic stress-strain relation and creep is taken into account by a change of modulus E'', it follows that the stress distribution in a structure is not affected by creep, provided that modulus $E''(t, t_1)$ is uniform, i.e. the creep properties are homogeneous throughout the structure. If the applied forces remain constant, all the strains and displacements in a homogeneous structure grow in proportion to the compliance function $J(t, t_1)$.

By a similar argument, the stresses produced in a homogeneous structure by an enforced constant displacement will grow in proportion to the relaxation function and the strains and displacements will remain constant. If the enforced displacement varies in time, the strains and displacements in a homogeneous structure will vary in proportion.

The previously described properties of homogeneous aging viscoelastic structures can be established rigorously on the basis of the integral-type stress-strain relation (28.3) or (28.5) (McHenry, 1943; Bažant, 1975, 1982).

The stress redistributions caused by creep in statically indeterminate structures generally consist of a gradual transfer of stress from the parts that creep more (e.g. a younger concrete, a lower-strength concrete, a thinner cross section, a drying member) to the parts that creep less. Steel parts in composite steel-concrete structures do not creep (except possibly in fire), and so the effect of creep is, in general, a gradual transfer of stress into the steel parts.

Unprestressed steel reinforcement, likewise, does not creep (except in fire). Because of their very high stress, the prestressing tendons creep, but very differently from concrete. Their creep causes prestress relaxation.

The reinforcement thus introduces an inhomogeneity into the creep properties of the structure, and leads to stress redistributions between steel and concrete. A rigorous approach is to consider the reinforcement and the concrete as separate parts of the structure.

In prestressed concrete structures, though, the effect of reinforcement stiffness on the creep deformations of beams is generally small, for two reasons: (1) because the

ratio of the cross section area of reinforcement to the total cross section area is small, compared to unprestressed structure, and (2) the cross sections do not undergo cracking. In the examples that follow, this effect is therefore neglected.

The restraining effect of reinforcement on creep becomes significant in structures that develop distributed cracking, which includes all unprestressed reinforced concrete beams and plates. This effect is nonlinear and its approximate linearized treatment will be discussed in Section 29.2.3.

29.2 STRUCTURAL ANALYSIS ACCORDING TO AAEM

29.2.1 Statically Indeterminate Beam Systems

Defining a Volterra integral operator for creep on the basis of (28.3), one can approach the structural creep analysis in a powerful, general and elegant manner. In this concise exposition, however, it is preferrable to use the simplest possible approach.

The equations of AAEM governing the creep effects in structures may simply be based on the following general rule for the analysis of changes in the structure state from t_1 to t, which follows from the quasi-elastic nature of the stress-strain relation (28.28):

$$\begin{pmatrix}\text{deformation}\\ \text{increment}\end{pmatrix} = \begin{pmatrix}\text{elastic}\\ \text{deformation}\\ \text{increment based}\\ \text{on modulus } E''\end{pmatrix} + \begin{pmatrix}\text{initial}\\ \text{elastic}\\ \text{deformation}\end{pmatrix} \varphi(t,t') + \begin{pmatrix}\text{shrinkage}\\ \text{increment}\end{pmatrix} \tag{29.1}$$

The excellent agreement of the results obtained by AAEM with the exact solutions is documented in Figure 29.1.

Some typical examples, dealing mostly with statically indeterminate beam systems (frames), will now be shown. The statically indeterminate internal forces will be denoted as X_i $(i = 1, 2, \ldots s)$, where s is the degree of statical indeterminacy (redundancy). Considering elastic homogeneous structures in which the deformations due to axial forces in the beams are zero or negligible, the flexibility δ_{ij} of the primary

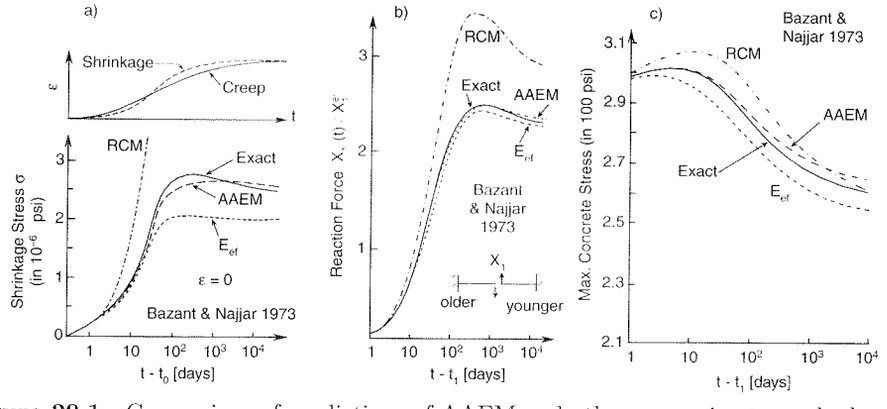


Figure 29.1 Comparison of predictions of AAEM and other approximate methods with exact solutions: a) shrinkage stress evolution, b) reaction force between two pin-joined cantilevers of different age, c) evolution of maximum stress in concrete in steel-concrete composite girder

structure (a statically determinate structure obtained by releasing all X_i), which represent the generalized displacement in the sense of X_i caused in the primary structure by applying force $X_j = 1$, may be expressed according to the principle of virtual work as follows:

$$\delta_{ij} = \frac{1}{E} f_{ij}, \qquad f_{ij} = \int_{(x)} \frac{\bar{M}_i(x)\bar{M}_j(x)}{I(x)}\,\mathrm{d}x, \qquad i,j = 1,2,\ldots s \tag{29.2}$$

Here x is the length coordinate of the beams and the integration is carried out over the length of all beams, $I(x)$ is the centroidal moment of inertia of the cross sections of the beams, and $\bar{M}_i(x)$ is the distribution of bending moment M caused in the primary structure by $X_i = 1$; f_{ij} represents the flexibility for a unit value of elastic modulus $(E = 1)$ and will be called the *unit flexibility*. Such flexibilities are normally not introduced for elastic or viscoelastic analysis but are convenient for aging creep analysis.

The values of the generalized displacement in the sense of X_i caused in the primary structure by the applied load may similarly be expressed as

$$\delta_{i,L} = \frac{1}{E} f_{i,L}, \qquad f_{i,L} = \int_{(x)} \frac{\bar{M}_i(x) M_L(x)}{I(x)}\,\mathrm{d}x, \qquad i = 1,2,\ldots s \tag{29.3}$$

where $M_L(x)$ is the distribution of the bending moment M caused in the primary structure by the applied load L. For elastic behavior, the compatibility conditions governing the values of X_i may be written as

$$\sum_{j=1}^{s} \frac{1}{E} f_{ij} X_j + \frac{1}{E} f_{i,L} + \delta_{i,sh} = 0, \qquad i = 1,2,\ldots s \tag{29.4}$$

where $\delta_{i,sh}$ is the generalized displacement in the sense of X_i caused in the primary structure by shrinkage (analogous to the displacement due to thermal expansion).

In the examples that follow, these elastic compatibility conditions will be generalized to creep. Generalizations for the effect of axial forces N, shear forces Q and, in the case of three dimensions, torques T and bending moments about different axes, are obvious. The purpose of the examples will be primarily didactic. Accordingly, solutions leading to simple formulae will be favored. Therefore, the statically indeterminate (redundant) structures in the examples will be limited to the case of a single statically indeterminate internal force, $X_1(t)$. However, generalization to many $X_i(t)$ would be easy.

29.2.2 Typical Creep and Shrinkage Problems Solved via AAEM

Example 29.1: Deflection of a Statically Determinate Beam

The deflection of a statically determinate beam in the sense of force F (Figure 29.2) caused by permanent load $p(x)$ applied at age t_1 is

$$w_F(t) = f_{F,L} \frac{1 + \varphi(t,t_1)}{E(t_1)} = f_{F,L} J(t,t_1), \qquad f_{F,L} = \int_{(x)} \frac{\bar{M}(x) M_L(x)}{I(x)}\,\mathrm{d}x \tag{29.5}$$

where $\bar{M}(x)$ = distribution of bending moments caused by a fictitious load $F = 1$, and $M_L(x)$ = distribution of bending moments caused by the permanent load. For a

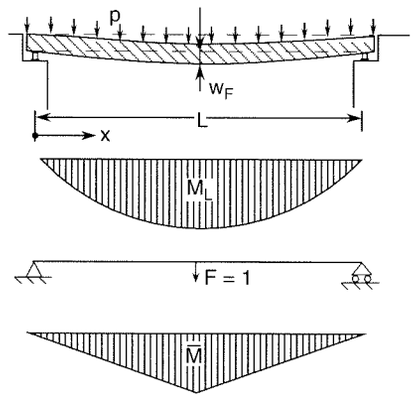


Figure 29.2 Long-time deflection of an uncracked beam

uniform permanent load p and a beam of a uniform cross section, we obtain

$$M_L(x) = \frac{1}{2}px(L-x), \qquad f_{F,L} = \frac{5}{384}\frac{pL^4}{I} \tag{29.6}$$

For concrete with compressive strength $\bar{f}_c = 30$ MPa, the conventional Young's modulus can be estimated from (28.9) as $E_{28} = \sqrt{30} \times 4.733$ GPa $= 25.9$ GPa and the parameters of the compliance function (28.6) are $E_0 = E_{28}/0.6 = 43.2$ GPa, $q_s = 11.4/E_{28} = 0.44$/GPa (the drying creep is neglected in this example). Consider a one-way slab of span $L = 10$ m and depth $h = 0.5$ m, made of concrete with specific weight $\gamma = 25$ kN/m^3, and loaded by its own weight and additional dead load $p^* = 2.5$ kN/m^2. All the subsequent calculations refer to a strip of width $b = 1$ m, which has the cross sectional moment of inertia $I = bh^3/12 = 0.01042$ m^4 and is subjected to the total dead load $p = \gamma bh + p^*b = 15$ kN/m. The midspan deflection evaluated from (29.6), $f_{F,L} = 187.5$ MN/m, refers to a beam with unrestrained deformations in the lateral direction and with unit Young's modulus. For a slab, which experiences cylindrical bending, the deflections are reduced by a factor $1-\nu^2$ (cf. Chapter 18). A typical value of Poisson's ratio for concrete is $\nu = 0.18$; so $1-\nu^2 = 0.9676$, and the deflection reduction due to this effect is only 3%. If the permanent load is applied 14 days after casting, the deflection after 1 day of loading (i.e. at $t = 15$ days) can be evaluated as

$$\begin{aligned} w_F(15) &= (1-\nu^2) f_{F,L} J(15,14) = 0.9676 \times 187.5 \text{ MNm}^{-1} \times 0.0572 \text{ (GPa)}^{-1} \\ &= 10.4 \text{ mm} \end{aligned} \tag{29.7}$$

where the compliance value

$$\begin{aligned} J(15,14) &= \left\{ \frac{1}{43.2} + 0.44 \ln \left[1 + 0.3 \left(14^{-0.5} + 0.001 \right) (15-14)^{0.1} \right] \right\} \text{(GPa)}^{-1} \\ &= 0.0572 \text{ (GPa)}^{-1} \end{aligned} \tag{29.8}$$

has been calculated from (28.6). The elastic deflection corresponding to the conventional Young's modulus E_{28} is $w_e = (1-\nu^2) f_{F,L}/E_{28} = 7$ mm. Deflections calculated for different ages t_1 at the beginning of loading and for different durations of loading $t - t_1$ are summarized in Table 29.1. It is clear that the deflection is strongly

Table 29.1 Slab deflection (in mm) for different ages at the beginning of loading and for different durations of loading

Load duration	Age at loading 14 days	1 month	1 year
1 sec	6.2	5.6	4.6
1 min	7.2	6.3	4.8
1 hour	8.7	7.3	5.1
1 day	10.4	8.5	5.5
1 month	12.8	10.2	6.0
1 year	15.0	11.8	6.5
50 years	19.7	15.1	7.5
500 years	23.3	17.8	8.4

affected by both parameters. Even though the compliance function (28.6) contains a logarithm and is unbounded, its actual growth for extremely long times is very slow. Note in Table 29.1 that the difference between the deflections after 50 years and after 500 years is less than 20%.

Example 29.2: Shrinkage Stress in a Restrained Slab

A pavement slab is prevented from shrinkage in all directions (Figure 29.3). The curing terminates and exposure to drying atmosphere begins at age t_0. By virtue of symmetry, the normal strains and stresses in all horizontal directions must be equal. There is no initial stress before shrinkage starts, i.e. $\sigma(t_0) = 0$. According to (29.1), taking into account creep Poisson effect, $\Delta\varepsilon(t) = \sigma(t)/E''(t,t_0) - \nu\sigma(t)/E''(t,t_0) + 0 + \varepsilon_{sh}(t)$. Since shrinkage is prevented, $\Delta\varepsilon(t) = 0$. Solution yields the shrinkage stress

$$\sigma(t) = -\frac{E''(t,t_0)}{1-\nu}\,\varepsilon_{sh}(t) \tag{29.9}$$

The magnitude of shrinkage strain $\varepsilon_{sh}(t)$ increases in time while the age-adjusted effective modulus $E''(t,t_0)$ decreases (for a fixed t_0). Consequently, the shrinkage-induced stress is relaxed by creep, which means that creep in this case has a beneficial effect (in view of extensive tensile cracking, the optimal value of ν is in this case much less than 0.18, and is sometimes taken as 0).

Example 29.3: Shrinkage Force in a Frame

Consider now the effect of homogeneous shrinkage in a homogeneous portal frame shown in Figure 29.4, supported on two hinges. The frame has one statically indeterminate internal force, $X_1(t)$, which may be defined, for instance, as the horizontal thrust between the hinges. The primary structure is a simply supported portal frame (a frame with one hinge allowed to slide horizontally). The displacement

Figure 29.3 Restrained shrinkage of pavement slab

in the sense of X_1 caused by shrinkage, representing the change of distance between the support points, is $\delta_{1,sh} = L\varepsilon_{sh}(t)$ where L is the distance between the support points.

We need to express the displacement of the primary structure in the sense of X_1 caused by X_1 and by shrinkage (the load effect being relegated to separate solution which may be superposed). Applying the general rule (29.1) to the displacement of the primary structure in the sense of X_1, we obtain the compatibility condition at time t:

$$\frac{f_{11}}{E''(t,t_0)}X_1(t) + \frac{f_{11}X_1(t_0)}{E(t_0)}\varphi(t,t_0) + L\varepsilon_{sh}(t) = 0 \tag{29.10}$$

where $f_{11} = \int_{(x)}(\bar{M}_1^2(x)/I(x))\ \mathrm{d}x$, the integration being carried out over the horizontal beam and both legs of the frame. At the start of drying, $X_1(t_0) = 0$. From the above equation, the shrinkage force in the frame is

$$X_1(t) = -E''(t,t_0)L\varepsilon_{sh}(t)/f_{11} \tag{29.11}$$

Example 29.4: Stress Relaxation After Initial Deformation

Consider that a concrete cylinder of length L is installed in a perfectly rigid testing frame, and at age t_1 is suddenly forced to shorten by w (Figure 29.5). This produces the initial elastic strain $\varepsilon_e(t_1) = -w/L$ and the corresponding stress $\sigma(t_1) = E(t_1)\varepsilon_e(t_1) = -E(t_1)w/L$. Subsequently, the length of the specimen is held constant, i.e. $\Delta\varepsilon(t) = 0$. According to (28.28),

$$\Delta\varepsilon(t) = \frac{\Delta\sigma(t)}{E''(t,t_1)} - \frac{w}{L}\varphi(t,t_1) = 0 \tag{29.12}$$

from which the stress relaxation at time t is

$$\Delta\sigma(t) = \frac{w}{L}E''(t,t_1)\varphi(t,t_1) \tag{29.13}$$

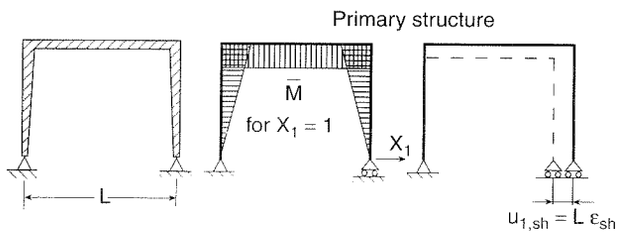


Figure 29.4 Shrinkage force in a portal frame

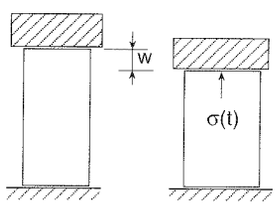


Figure 29.5 Relaxation after forced deformation of a specimen

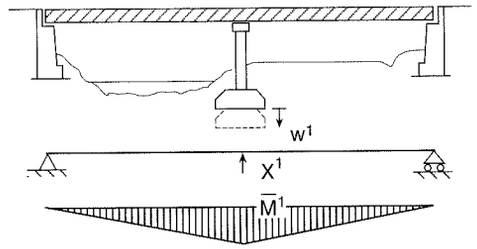


Figure 29.6 Effect of support settlement

Adding $\sigma(t_1)$ and substituting for $E''(t,t_1)$ from (28.26), we obtain

$$\sigma(t) = \frac{\sigma(t_1)}{E(t_1)}R(t,t_1) \tag{29.14}$$

As expected, in this case, which corresponds to stress relaxation at constant strain, the AAEM method provides the exact solution.

Example 29.5: Force Relaxation in a Structure

Consider now that the middle support of a continuous bridge girder with two spans, shown in Figure 29.6, undergoes a sudden vertical settlement w_1 at age t_1. After that, there is no more settlement. The girder has one statically indeterminate force $X_1(t)$, which may be taken as the vertical reaction in the middle support (considered here as positive if pointing downwards), and the primary structure is then a simply supported beam obtained by removing the middle support (Figure 29.6). Only the changes of internal forces caused by support settlement are analyzed here. To get a complete description of the structural behavior, the creep solution for the permanent load would have to be superposed. The initial elastic value of X_1 right after the sudden settlement is

$$X_1(t_1) = w_1/\delta_{11} = w_1E(t_1)/f_{11} \tag{29.15}$$

The condition that the deflection increment from t_1 to t of the primary structure at the middle support must be compatible with the pier settlement increment $\Delta w_1(t) = 0$ may be written, according to the general rule in (29.1), as follows:

$$\Delta w_1(t) = \frac{f_{11}}{E''(t,t_1)}\Delta X_1(t) + w_1\varphi(t,t_1) = 0 \tag{29.16}$$

From this, the relaxation of the statically indeterminate reaction from time t_1 to time t is

$$\Delta X_1(t) = -\frac{w_1}{f_{11}}E''(t,t_1)\varphi(t,t_1) = \left(\frac{R(t,t_1)}{E(t_1)} - 1\right)X_1(t_1) \tag{29.17}$$

Similar to the previous example, the structural behavior corresponds to stress relaxation at constant strain, which is described by AAEM exactly. Note that this is the case only if the structure is homogeneous, in the sense that the creep properties are the same everywhere. The reaction computed in this example is only the part of the total reaction that is caused by the support settlement; it must be added to the reaction due to other loads.

Example 29.6: Cantilever End Placed on Support

In cantilever construction of prestressed concrete box girder bridges, the balancing of the pier is made easier if the end span is cast from the pier toward the abutment as a cantilever (Figure 29.7(a)). To simplify the analysis, we assume the whole cantilever to have the same age t, representing the average of the actual ages of the casting segments of the cantilever, and the dead load to be applied when this average age is $t = t_1$. At age $t = t_2$ ($t_2 > t_1$), a bearing is placed under the end of the cantilever and is made to fit tightly (which may induce an initial vertical reaction at the support). This changes the structural system from statically determinate (system I, for $t < t_2$) to statically indeterminate, or redundant (system II, for $t \geq t_2$). Assuming, for the sake of simplicity, the forces acting at the end of the opposite, counterbalancing cantilever erected from the same pier to have a negligible effect on the vertical reaction $X_1(t)$ at the abutment, $X_1(t)$ is a single statically indeterminate force (Figure 29.7(b)).

In the case of a change of structural system from determinate to indeterminate, the general equation of AAEM used as a compatibility condition requires an adjustment. For $t > t_2$, the evolution of stress and strain during the interval (t_1, t) is not smooth, and its approximation by a linear combination of creep and relaxation would be too crude. It is much better to decompose the problem into two loading histories: (1) a cantilever subjected to the permanent load, acting over the entire time interval (t_1, t), and (2) a cantilever loaded at its end by a variable force $X_1(t)$, which is nonzero only in the time interval (t_2, t). The first case leads to creep at constant stress, and the second case can accurately be described by AAEM. The deflection of the initially free end is obtained by summing the contributions of both loading cases, and the sum should remain constant starting from t_2, i.e. from the time at which the right end is placed on a support. This provides the condition

$$\Delta w_1 = \frac{f_{1,L}}{E(t_1)}\left[\varphi(t,t_1) - \varphi(t_2,t_1)\right] + \frac{\Delta X_1(t) f_{11}}{E''(t,t_2)} + \frac{X_1(t_2) f_{11}}{E(t_2)}\varphi(t,t_2) = 0 \qquad (29.18)$$

Assuming further that there is no jacking up nor pulling down of the cantilever end before the bearing at the abutment is installed, we have the initial condition $X_1(t_2) = 0$. Therefore, the change from t_2 to t is $\Delta X_1(t) = X_1(t)$. With these conditions, the solution of the last equation is

$$X_1(t) = -\frac{E''(t,t_2)}{E(t_1)}\frac{f_{1,L}}{f_{11}}\left[\varphi(t,t_1) - \varphi(t_2,t_1)\right] \qquad (29.19)$$

The negative sign indicates that X_1 actually acts upwards.

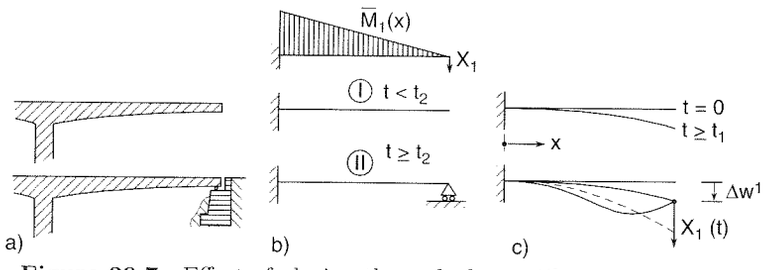


Figure 29.7 Effect of placing the end of a cantilever on a support

The effect of the gradual build-up of reaction X_1 at the abutment is to produce positive bending moments within the span and reduce the negative bending moment peak at the pier. The deflection curves at subsequent times have a uniformly negative curvature up to time t_2, and after that the right portion of the span develops a progressively increasing positive curvature (Figure 29.7(c)). Note that if the bearing on the abutment were not installed, the end of cantilever would continue deflecting down as shown in the figure, and thus it is obvious that an upward reaction X_1 must develop to offset this deflection.

A general property of the response to a change from structural system I to structural system II is that the internal force distributions gradually change from an elastic distribution corresponding to system I toward an elastic distribution corresponding to system II. However, the latter distribution is never attained; it is approached only partially.

The previous calculation yields only the reaction $X_1(t)$ due to own weight. The additional reaction due to permanent loads (such as the weight of the roadway) placed after installation of the bearing, is not changed by creep (assuming the creep properties to be homogeneous throughout the structure).

Example 29.7: Simply Supported Beams Made Continuous

Continuous bridge beams are often assembled from precast prestressed concrete beams, which are transported from a plant to the site, raised on the supports and then joined at time t_2 to create a continuous beam (this is more efficient than simply supported beams and has a better seismic performance); see Figure 29.8(a). The joining can, for instance, be achieved by placing a continuous reinforced concrete slab on top of the beams, by the welding of steel bars, or by installing additional prestressing tendons running from one span to the next.

Let the age of concrete at placing the beams on the supporting piers be denoted as t_1. Initially, each span acts as a simply supported beam and is statically determinate. So, the deflections due to own weight, acting since time t_1, increase freely, which causes the beam ends to gradually undergo rotations $\theta(t)$ (slope increase, Figure 29.8(b)). This continues until age t_2 at which the beam ends are joined.

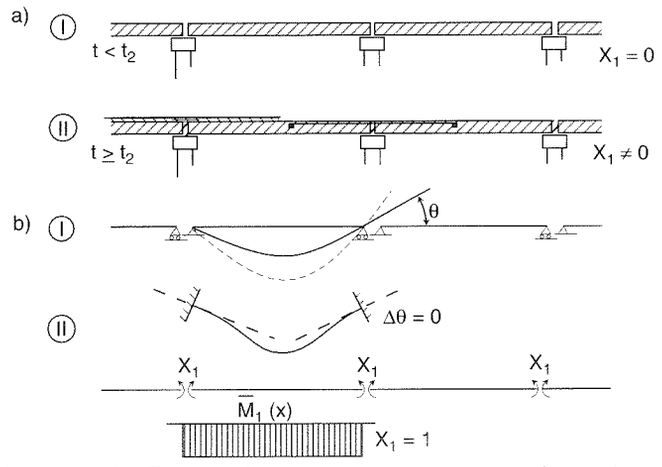


Figure 29.8 Precast simply supported beams made continuous

The joining at age t_2 changes the structural system from statically determinate (system I) to statically indeterminate (system II). After that, the additional end rotations, $\Delta\theta(t)$, of the adjacent beams are forced to be the same. For the sake of simplicity, we will consider a continuous beam of infinitely many identical spans. Then, by virtue of symmetry conditions, the additional rotations above the support must be zero, $\Delta\theta(t) = 0$, and we have only one statically indeterminate internal force $X_1(t)$, the bending moment on top of the support (Figure 29.8(b)). Since the continuing deflection is opposed by this restraint, negative moments above the supports gradually build up.

Applying the general AAEM equation (29.1) to the condition of compatibility of rotations above the supports, $\Delta\theta(t) = 0$, we obtain the same equation as (29.18) but with a different meaning of its variables and coefficients. The resulting formula (29.19) again applies, as it does for all changes of structural system leading to one unknown $X_1(t)$.

Example 29.8: Stress Redistributions Due to Differences in Age of Concrete

Since concrete loaded at young age creeps more than concrete loaded at old age, large differences in age lead to significant redistributions of internal forces in structures. As an example, consider again a box girder bridge cast by the cantilever method. In the interest of economy, the pair of steel trusses at each pier that support new segments during erection is used repeatedly to erect one cantilever pair after another. Consider, therefore, that in Figure 29.9(a) the left cantilever is older by Δ than the right cantilever (a difference Δ of 3–6 months is not unusual).

Let t be the age of the left (older) cantilever, and t_1 be the age of each cantilever at which the own weight is assumed to be applied (this is of course a simplification; in reality the bending moments due to own weight grow gradually as the cantilever is being erected). As long as the opposite cantilevers of the pair in one span are separated, they deflect independently of each other. At time $t = t_2$ they are joined. For the sake of simplicity, assume that their joining is done by installing a horizontally sliding hinge, which only provides one statically indeterminate internal force $X_1(t)$ in each span (a full continuity at midspan is a better design, but it provides three statically indeterminate forces at midspan).

The age difference makes the structure inhomogeneous, but in view of our simplifying assumptions, each cantilever is homogeneous. It is convenient to split each flexibility coefficient into two parts, each of them associated with the left or

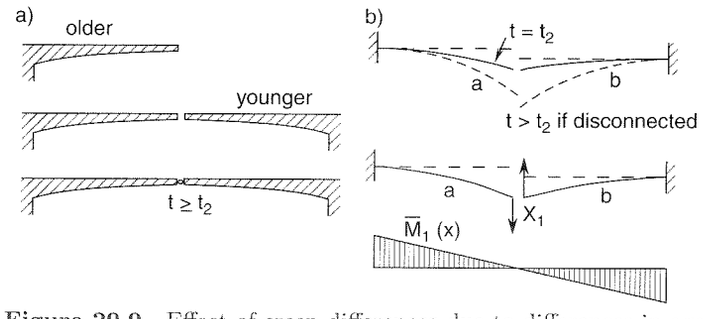


Figure 29.9 Effect of creep differences due to differences in age

right cantilever, labeled by superscripts ^{a} and ^{b}. Thus,

$$f_{11} = f_{11}^a + f_{11}^b, \qquad f_{1,L} = f_{1,L}^a + f_{1,L}^b \tag{29.20}$$

If the left and right cantilevers were allowed to deflect independently, the left (older) one would deflect after time t_2 less than the right (younger) one because more of creep has occurred before t_2 in the left cantilever. Therefore, if the differences in deflection are prevented, the pair of vertical forces representing X_1 will act on the left cantilever downward, and on the right cantilever upward (Figure 29.9(b)). Similar to the preceding two examples, we decompose the problem into two loading cases, one corresponding to the permanent load applied during the entire time interval (t_1, t), and the other corresponding to the statically indeterminate force X_1 first applied at t_2 and evolving with time. The general AAEM equation (29.1) applied to both cases provides the following condition of compatibility (zero vertical relative displacement in the joint):

$$\begin{aligned}\delta_1 = {} & \frac{f_{11}^a \Delta X_1(t)}{E''(t,t_2)} + \frac{f_{11}^b \Delta X_1(t)}{E''(t-\Delta, t_2-\Delta)} + \frac{f_{11}^a X_1(t_2)}{E(t_2)}\varphi(t,t_2) \\ & + \frac{f_{11}^b X_1(t_2)}{E(t_2-\Delta)}\varphi(t-\Delta, t_2-\Delta) + \frac{f_{1,L}^a[\varphi(t,t_1)-\varphi(t_2,t_1)]}{E(t_1)} \\ & + \frac{f_{1,L}^b[\varphi(t-\Delta,t_1)-\varphi(t_2-\Delta,t_1)]}{E(t_1)} = 0\end{aligned} \tag{29.21}$$

If the initial value $X_1(t_2)$ is specified (which depends on the technological procedure used to join the cantilever ends), the increment $\Delta X_1(t)$ may be solved from (29.21). The high accuracy of the AAEM solution is illustrated in Figure 29.1(b).

Example 29.9: Stress Redistributions in Beams of Composite Cross Section

In steel-concrete composite beams, as well as composite beams with cross sections consisting of parts of very different age or quality, creep causes stress redistributions within the cross sections, which generally transfer stress to the part creeping less, that is, to the steel part or the concrete part of higher age or higher quality (Bažant, 1975, 1982; Bažant and Buyukozturk, 1988). As illustrated in Figure 29.1(c), the solutions of these redistributions by AAEM are very accurate; they were shown to differ from the exact solutions according to the principle of superposition generally by less than 1% (Bažant and Najjar, 1973).

Consider a general composite beam consisting of n parts (Figure 29.10), each of which exhibits different linear creep properties (a steel part is a special case with zero

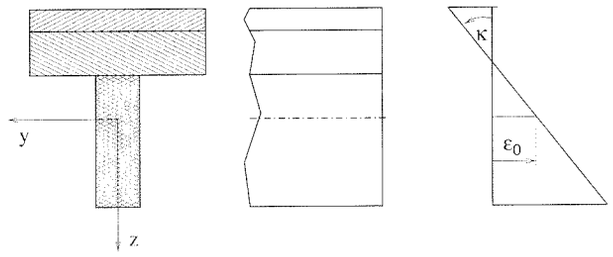


Figure 29.10 Cross-section of a composite beam

creep). Similar to the elastic theory of beam bending, we assume the cross sections to remain plane and normal to the deformed beam axis. The initial elastic state at time t_1 is calculated, as usual, by the method of transformed cross section. The increment of deformation from time t_1 to t can be characterized on the level of the cross section by the increment of curvature, $\Delta\kappa$, and the increment of normal strain at point O on the chosen reference axis, $\Delta\varepsilon_0$. From the assumption of a plane section it follows that the increments of normal strain are linearly distributed and described by

$$\Delta\varepsilon(z) = \Delta\varepsilon_0 + z\,\Delta\kappa \tag{29.22}$$

where z is the coordinate measured from point O along the depth of the cross section. The strain increments must satisfy the basic equation (28.28) of AAEM,

$$\Delta\varepsilon = \frac{\Delta\sigma}{E''} + \varepsilon_1\,\varphi + \Delta\varepsilon_{sh} \tag{29.23}$$

from which we can express the stress increments

$$\Delta\sigma = E''\,(\Delta\varepsilon - \varepsilon_1\,\varphi - \Delta\varepsilon_{sh}) = E''\,(\Delta\varepsilon_0 + z\,\Delta\kappa - \varepsilon_1\,\varphi - \Delta\varepsilon_{sh}) \tag{29.24}$$

For the sake of brevity we have omitted the arguments z, t_1 and t. Integration over the cross section provides the increments of normal force and bending moment,

$$\Delta N = \int_A \Delta\sigma\,\mathrm{d}A = R''_N\,\Delta\varepsilon_0 + R''_{NM}\,\Delta\kappa + \Delta N'' \tag{29.25}$$

$$\Delta M = \int_A \Delta\sigma\,z\,\mathrm{d}A = R''_{NM}\,\Delta\varepsilon_0 + R''_M\,\Delta\kappa + \Delta M'' \tag{29.26}$$

where

$$R''_N = \int_A E''\,\mathrm{d}A, \qquad R''_{NM} = \int_A E''\,z\,\mathrm{d}A, \qquad R''_M = \int_A E''\,z^2\,\mathrm{d}A \tag{29.27}$$

are the age-adjusted effective cross sectional stiffnesses, and

$$\Delta N'' = -\int_A E''(\varepsilon_1\varphi + \Delta\varepsilon_{sh})\,\mathrm{d}A \tag{29.28}$$

$$\Delta M'' = -\int_A E''(\varepsilon_1\varphi + \Delta\varepsilon_{sh})\,z\,\mathrm{d}A \tag{29.29}$$

are the inelastic increments of the internal forces due to creep and shrinkage. Note that, in a composite beam, the age-adjusted effective modulus E'' is variable across the cross section and thus cannot be brought in front of the integrals. Of course, if the cross section consists of n homogeneous parts, we can write

$$R''_N = \sum_{i=1}^{n} E''_i A_i, \qquad R''_M = \sum_{i=1}^{n} E''_i A_i\,(r_i^2 + e_i^2) \tag{29.30}$$

where E''_i is the age-adjusted effective modulus of part number i, A_i is the area of that part, r_i is its radius of inertia, and e_i is the z-coordinate of its centroid.

Equations (29.25)–(29.26) generalize the well-known relations between cross sectional generalized strains (axial strain and curvature) and internal forces (normal force and moment). In linear elasticity, such equations have the simple form $N = EA\varepsilon_0$ and $M = EI\kappa$, provided that the reference point O is placed at the centroid of the homogeneous cross section. For composite elastic beams, it is necessary to take the centroid of the transformed cross section, i.e. to consider areas of individual parts weighted by their relative elastic moduli. In the general case treated here (composite beams with creep and shrinkage), the position of such a generalized centroid must be determined using the age-adjusted effective moduli, and so it depends on the times t_1 and t. If O is placed at such a generalized centroid, we obtain $R''_{NM} = 0$ and the basic equations (29.25)–(29.26) become decoupled, which facilitates their inversion. The increments of axial strain and curvature are then given by

$$\Delta\varepsilon_0 = \frac{\Delta N - \Delta N''}{R''_N}, \qquad \Delta\kappa = \frac{\Delta M - \Delta M''}{R''_M} \tag{29.31}$$

If the composite beam is statically determinate, the distribution of internal forces depends only on the applied loads but not on the stiffness variation along the beam. Under constant load, the increments of internal forces must be zero, and the increments of axial strain and curvature directly follow from (29.31) with $\Delta N = 0$, $\Delta M = 0$, and with $\Delta N''$ and $\Delta M''$ evaluated from (29.28)–(29.29). If needed, the stress distribution in different parts of the composite cross section can be determined by substituting into (29.24), and the deflections can be obtained from the distribution of curvatures using the principle of virtual work.

If the composite beam is statically indeterminate, one must first use R''_N, R''_M, $\Delta N''$ and $\Delta M''$ in a virtual work integral over the structure, in order to determine the flexibilities based on $E''(t, t_1)$ and the inelastic displacements of the primary structure in the sense of X_i. Solution of the compatibility equations in the sense of X_i then yields the changes in X_i caused by creep and shrinkage. From these one obtains $\Delta N(t)$ and $\Delta M(t)$ in the cross section of interest, and the subsequent procedure is the same as before. For a generalization of this procedure to step-by-step integration in time, see Bažant and Najjar (1973).

Example 29.10: Relaxation of Prestressing Tendon

Over time, the prestress force $P(t)$ in a concrete beam relaxes due to creep and shrinkage of concrete as well as creep of steel tendon, manifested as prestress relaxation. This problem is normally handled by simple formulae based on the effective modulus, whose error is about 15%. Although this is not a serious error compared to all the other uncertainties, a more accurate calculation according to AAEM could easily be incorporated into 'black-box' design programs. In view of the recent suspicion that excessive long-time deflections of many large bridges might be caused by underprediction of prestress losses, a more accurate calculation might be worthwhile.

The problem is a special case of the earlier problem of a composite cross section, with $n = 2$ parts, one of which (prestressing steel) does not exhibit shrinkage (Figure 29.11(a)). Instead of substituting into the general equations derived in Example 29.9, let us set up the solution directly. This will elucidate the physical meaning of the basic formulae. We neglect the bending stiffness of the tendon (or group of tendons). The prestress as well as the dead load are introduced at concrete age t_1, and the initial distribution of strains and stresses can be calculated from standard equations valid for elasticity. Now consider the changes due to creep and shrinkage of concrete and loss of prestress in steel (at constant internal forces) in

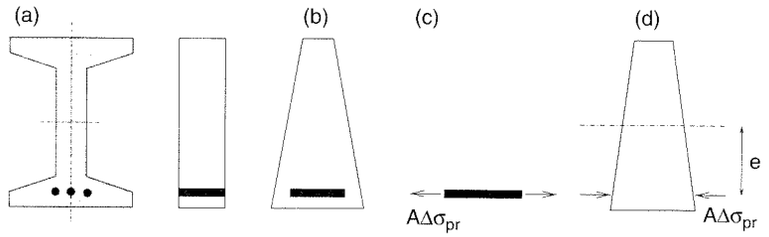


Figure 29.11 Cross-section of concrete beam with a prestressed tendon

a time interval (t_1, t). First we imagine that these changes take place at constant stress. The prestressing tendon would experience a strain increase $\Delta\varepsilon''_{pr}$ at constant stress, which may be figured out from the formula for steel relaxation, provided by the tendon manufacturer. The strain change in concrete under constant stress is $\varepsilon_1\varphi(t,t_1) + \Delta\varepsilon_{sh}(t)$. However, the compatibility condition requires that the change of strain in the tendon from t_1 to t be the same as the change of strain in concrete at the level of the centroid of tendon. The incompatibility that would appear if both parts deformed independently (Figure 29.11(b)) leads to a stress redistribution.

We can imagine that the prestressing tendon is loaded by an additional force $A_{pr}\,\Delta\sigma_{pr}$, where A_{pr} is the tendon area and $\Delta\sigma_{pr}$ is the change of prestress. This force produces an additional strain $\Delta\sigma_{pr}/E_{pr}$ in the tendon, where E_{pr} is the elastic modulus of prestressing steel. An opposite force must be applied on the concrete cross section, which produces nonuniform stress and strain changes in concrete.

Recall that a force F with eccentricity e applied on an elastic cross section produces axial strain $\varepsilon_0 = F/EA$ and curvature $\kappa = Fe/EI = Fe/EAr^2$, where E = Young's modulus, A = cross sectional area, I = cross sectional moment of inertia, and $r = \sqrt{I/A}$ = radius of inertia. The strain at the point where F is applied is then $\varepsilon = \varepsilon_0 + \kappa e = (F/EA)(1 + e^2/r^2)$.

In the present problem, the force F applied on the concrete cross section is $-A_{pr}\,\Delta\sigma_{pr}$; Young's modulus must be replaced by the age-adjusted effective modulus of concrete; parameters A, I and r refer to the concrete cross section and will be denoted by subscripts c; and e is the eccentricity of the tendon with respect to the centroid of the concrete cross section. Combining the previous results, we can set up the compatibility condition

$$\frac{\Delta\sigma_{pr}}{E_{pr}} + \Delta\varepsilon''_{pr} = -\frac{A_{pr}\Delta\sigma_{pr}}{E''A_c}\left(1 + \frac{e^2}{r_c^2}\right) + \varepsilon_1\,\varphi + \Delta\varepsilon_{sh} \tag{29.32}$$

from which it is possible to find the loss of prestress due to creep, shrinkage and steel relaxation combined,

$$\Delta\sigma_{pr} = E_{pr}\left(\varepsilon_1\,\varphi + \Delta\varepsilon_{sh} - \Delta\varepsilon''_{pr}\right)\left[1 + \frac{E_{pr}A_{pr}}{E''A_c}\left(1 + \frac{e^2}{r_c^2}\right)\right]^{-1} \tag{29.33}$$

Note that the elastic strain ε_1 and the shrinkage strain increment $\Delta\varepsilon_{sh}$ that appear in (29.33) are evaluated for the concrete at the level of the prestressing tendon.

The previous derivation is based on the simplifying assumption that the creep strain in steel depends only on the initial prestress in steel, but not on the loss of prestress. This would be exact only if the strain in the tendon were constant, which is not the case. Strictly speaking the creep strain in steel should be solved from the highly

nonlinear differential equation for steel creep. This cannot be incorporated into the present solution. However, the error caused by the simplified approach presented is probably small (doubtless much smaller than the error of the simple standard formula currently used in codes).

Example 29.11: Restrained Shrinkage of Top Slab in Composite Beam

Consider now the effect of shrinkage of a top slab of a simply supported steel-concrete composite beam (Figure 29.12(a)) in which the bending moments from permanent load are large enough to prevent shrinkage cracking of the slab. In that case the full drying shrinkage and drying creep (or stress-induced shrinkage) takes place. The slab is assumed to be so thin that the bending moment in the slab is negligible.

The problem is a special case of the general composite cross section analyzed in Example 29.9. For the sake of simplicity we assume that the drying and loading start at the same age $t_1 = t_0$. If we imagine the concrete slab to be separated from the steel beam (Figure 29.12(b)), its free shrinkage and creep during the time interval (t_1, t) gets cancelled by tensile force

$$\Delta N''(t) = -A_c E''(t,t_1)[\varepsilon_1\varphi(t,t_1) + \varepsilon_{sh}(t)] \tag{29.34}$$

in the slab (Figure 29.12(c)). This force, which restores compatibility of the slab and the steel beam, must then be cancelled by applying the opposite force $-\Delta N''(t)$ on the entire cross section (Figure 29.12(d)). For the purpose of elastic analysis, the composite cross section can be replaced by a homogeneous one, with concrete modulus $E''(t,t_0)$ and with the area of steel amplified by the ratio $E_s/E''(t,t_0)$, where E_s is the elastic modulus of steel. If $A'' = A_c + A_sE_s/E''$ denotes the area of the transformed cross section, r'' its radius of inertia and e_c the distance of the slab centroid from the centroid of the transformed cross section, then the stress in the concrete slab produced by force $-\Delta N''(t)$ applied on the entire cross section is equal to $-[\Delta N''(t)/A''][1 + (e_c/r'')^2]$. Adding this to the stress produced by force $\Delta N''(t)$ applied on the slab separated from the steel beam, we obtain the resulting stress change in the concrete slab due to redistributed creep and shrinkage.

$$\begin{aligned}\Delta\sigma_c(t) &= \frac{\Delta N''(t)}{A_c} - \frac{\Delta N''(t)}{A''}\left[1 + \left(\frac{e_c}{r''}\right)^2\right] \\ &= -E''(t,t_1)[\varepsilon_1\varphi(t,t_1) + \varepsilon_{sh}(t)]\left\{1 - \frac{A_c}{A''}\left[1 + \left(\frac{e_c}{r''}\right)^2\right]\right\}\end{aligned} \tag{29.35}$$

Note that if the steel area is much larger than the concrete area, we have $A_c/A'' \approx 0$ and the stress change in concrete is the largest possible, i.e. the same as if the concrete slab were fully restrained. On the other hand, if the steel area is much smaller than the concrete area, we have $A_c/A'' \approx 1$ and $e_c \approx 0$, and then there is no stress change in concrete, the same as in the case of free shrinkage.

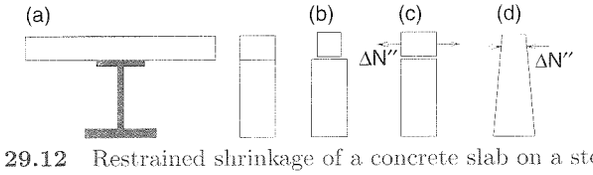


Figure 29.12 Restrained shrinkage of a concrete slab on a steel beam

If the beam is statically indeterminate (a continuous beam), then the curvatures caused by applying $-\Delta N''(t)$ in each cross section will produce statically indeterminate reactions ΔX_i, $i = 1, 2, \ldots s$, which can be computed from the compatibility conditions

$$\int_{(x)} \frac{\bar{M}_i}{E''I''} \left(\sum_{j=1}^{s} \bar{M}_j \, \Delta X_j - e_c \, \Delta N'' \right) \mathrm{d}x = 0, \qquad i = 1, 2, \ldots s \tag{29.36}$$

Note that the moment of inertia $I'' = A''r''^2$ is variable in time.

Example 29.12: Creep Buckling

Another important effect of creep is to reduce the long-time buckling strength of slender columns, thin plates and shells; see Bažant and Cedolin (1991, Chapter 9). As a prototype creep buckling problem, consider a simply supported reinforced concrete column (Figure 29.13) with initial sinusoidal curvature given by initial ordinates

$$z_0(x) = a \sin \frac{\pi x}{L} \tag{29.37}$$

where x = axial coordinate and L = column length. Applying at age t_1 axial load P causes an immediate increase of the deflection ordinates to

$$z(x, t_1) = \mu(t_1)\, a \sin \frac{\pi x}{L} \tag{29.38}$$

where

$$\mu(t_1) = \frac{1}{1 - P/P_{cr}(t_1)}, \qquad P_{cr}(t_1) = \frac{\pi^2}{L^2}[E_c(t_1)I_c + E_s I_s] \tag{29.39}$$

Here I_c, I_s = centroidal moments of inertia of the cross sections of concrete and of the group of steel reinforcing bars (their centroids being assumed coincident), $E_c(t_1)$ = elastic modulus of concrete at age t_1, E_s = elastic modulus of steel, $P_{cr}(t_1)$ = short-time critical load at age t_1; and μ_1 = magnification factor at age t_1 (Bažant and Cedolin, 1991).

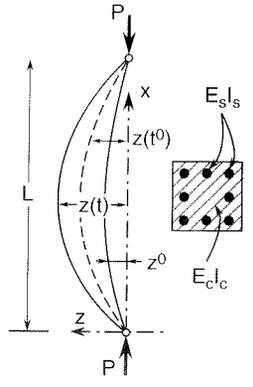


Figure 29.13 Creep buckling of slender column with initial curvature

The question is how much the deflections increase over time if the axial load is held constant. We can again use AAEM for that purpose. The increase of curvature at midlength may be expressed according to (29.1) as

$$\Delta z''(t) = \frac{\Delta M_c(t)}{E''(t, t_1)I_c} + \frac{M_c(t_1)}{E(t_1)I_c}\varphi(t, t_1) \tag{29.40}$$

where M = bending moment at midlength. The curvature increase causes in steel reinforcement the bending moment increase $E_s I_s \Delta z''(t)$. Equilibrium of the cross section at column midlength requires that $\Delta M(t) = -P\Delta z(t)$, and so we have the condition

$$\Delta M_c(t) + E_s I_s \Delta z''(t) = -P\Delta z(t) \tag{29.41}$$

In view of the linearity of creep and the assumption of sinusoidal initial curvature, $z(x,t) = \mu(t)z_0(x)$, and at midlength $\Delta z(t) = z(t) - z(t_1) = [\mu(t) \quad \mu(t_1)]a$. Introducing into (29.41) these expressions and the expression for $\mu(t_1)$ from (29.39), as well as the expression for $\Delta M_c(t)$ obtained by solving (29.40), we get

$$E''(t, t_1)I_c \left\{ -\frac{\pi^2}{L^2}[\mu(t) - \mu(t_1)] + \frac{\pi^2}{L^2}\frac{P}{[P_{cr}(t_1) - P]}\varphi(t, t_1) \right\}$$
$$-E_s I_s \frac{\pi^2}{L^2}[\mu(t) - \mu(t_1)] = -P[\mu(t) - \mu(t_1)] \tag{29.42}$$

Solving for the magnification factor $\mu(t)$ at time t, and substituting (29.39) for $\mu(t_1)$, we get the result (Bažant and Cedolin, 1991, Section 9.4):

$$\mu(t) = \frac{1}{1 - P/P_{cr}(t_1)} \left(1 + \frac{P''_{cr,c}(t, t_1)\varphi(t, t_1)}{P''_{cr,c}(t, t_1) + P_{cr,s} - P} \frac{P}{P_{cr}(t_1)} \right) \tag{29.43}$$

where

$$P''_{cr,c}(t, t_1) = \frac{\pi^2}{L^2}E''_c(t, t_1)I_c, \qquad P_{cr,s} = \frac{\pi^2}{L^2}E_s I_s \tag{29.44}$$

are the critical loads of the concrete and steel parts alone based on moduli $E''_c(t, t_1)$ and E_s.

A live load P_L may be superposed on dead load P at time t, which further magnifies the deflection. To calculate the magnification, we may imagine that the column is at time t instantly unloaded and immediately after that instantly reloaded with the combined load $P + P_L$ (Bažant and Cedolin, 1991, Section 9.5). The instantaneous unloading reduces the magnification factor to $\mu(t)[1 - P/P_{cr}(t)]$, where

$$P_{cr}(t) = \frac{\pi^2}{L^2}(E_c(t)I_c + E_s I_s) \tag{29.45}$$

is the instantaneous critical load at time t. The subsequent reloading increases the magnification factor to

$$\mu_u(t) = \mu(t)\frac{1 - P/P_{cr}(t)}{1 - (P + P_L)/P_{cr}(t)} \tag{29.46}$$

In view of design safety requirements, P and P_L, of course, must not be the actual loads but must include the dead and live load safety factors (called load factors).

It was shown that for loads P within the service stress range the calculation of long-time buckling deflections based on AAEM agrees almost exactly with accurate computer solutions based on approximating the history integral in (28.3) with a sum. Since the statistical variability of the dead load is quite small, this covers most design situations.

The assumption of sinusoidal curvature might seem too restrictive but it is not. Similar to elastic buckling, what matters near the critical load is only the first Fourier component of the given arbitrary initial deflection profile. For small loads, other Fourier components also matter, but buckling is not an important consideration at small loads (Bažant and Cedolin, 1991, Chapter 1). Therefore, formulae (29.44) and (29.46) have an almost universal applicability for design, similar to elastic columns. They are equally valid when the initial deflection before the application of axial load is caused by initial bending moments. Furthermore, it can be shown that the effective length of columns is not affected by linear creep, and therefore these formulae can be applied to columns with arbitrary end support and to columns in frames, provided that the actual column length L is replaced by the effective buckling length l, i.e. by the distance between the inflexion points of deflection curve (as in elastic analysis).

Formulae (29.44) and (29.46) give poor accuracy only in the case when the permanent load is larger than about 50% of the short-time critical load, based on $E(t_1)$. Such loads cannot actually occur in practice (they can be considered only in the sense of factored loads). Figure 29.14 shows the growth of the deflection as a function of time (in semi-logarithmic scale) for two ratios of the permanent load to the elastic critical load. The agreement between the exact solution and AAEM solution is very good up to $P/P_{cr}(t_1) = 0.25$ and still quite good for $P/P_{cr}(t_1) = 0.4$. In this last case, AAEM overpredicts the actual deflection, and thus provides a safe estimate.

The critical loads for long-time buckling of perfect columns, with no imperfections, are a more subtle problem (Bažant and Cedolin, 1991). This question, however, is practically unimportant. ACI Standard 318 (ACI, 1999, Art. 10.12.3.2) requires a certain minimum initial bending moment to be considered in design. The design requirement is that the magnified bending moment, which is here obtained by magnifying the initial bending moments (at no axial load) with the factors in (29.43)

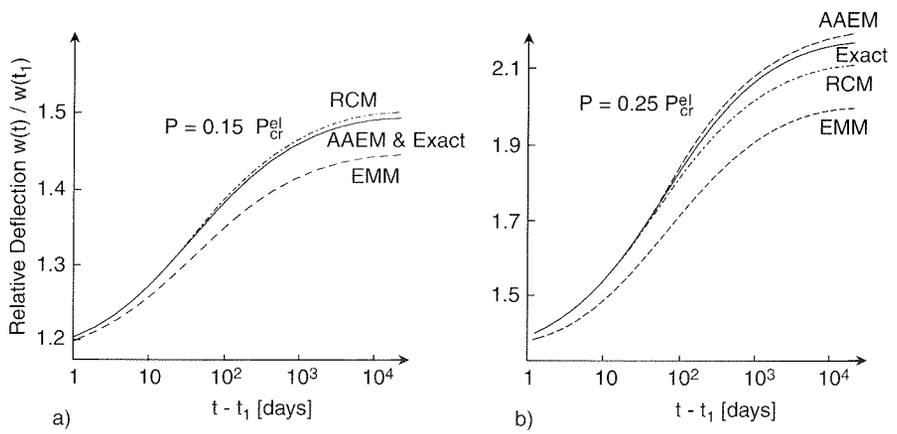


Figure 29.14 Evolution of relative deflection of slender column with initial curvature for different levels of the applied compressive load (adapted from Bažant and Najjar, 1973)

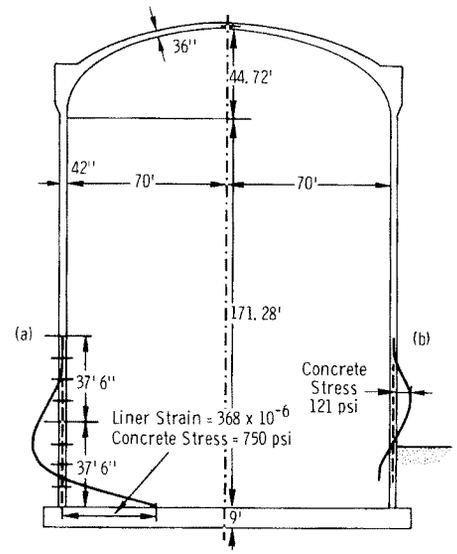


Figure 29.15 Stresses produced in a prestressed concrete containment shell by (a) differential shrinkage, and (b) differential creep (after Bažant, Carreira and Walser, 1975)

and (29.46), be safely carried by the cross section, and the present analysis can in this way be applied in the design practice.

Creep buckling is very important for thin concrete shells. It has been responsible for some collapses of large roof shells. The analysis may again be approximately carried out on the basis of equation (29.1). A thorough presentation of concrete creep buckling may be found in Bažant and Cedolin (1991, Chapter 9).

Example 29.13: Shrinkage and Prestress Effects in Reactor Containment Shell

The base of a cylindrical shell of a nuclear reactor containment is typically connected to a very massive foundation buried in the ground (Figure 29.15). The deformations of this foundation are almost nil. If the base of shell were disconnected, the free shrinkage $\varepsilon_{sh}(t)$ of the shell and the creep deformations of the shell due to circumferential prestress $\sigma_{y,pr}$ and to Poisson effect of vertical prestress $\sigma_{z,pr}$ would cause a decrease in the radius of the shell;

$$\Delta u(t) = R\,\varepsilon_{sh}(t) \;+\; \frac{1}{E(t_1)}(\sigma_{y,pr} - \nu\sigma_{z,pr})\;\varphi(t,t_1) \tag{29.47}$$

where R = radius of the middle surface of the shell.

This gradually increasing displacement as well as any rotation at base is prevented by the rigid foundation. This is equivalent to imposing at the base of the shell an outward displacement equal to $\Delta u(t)$, at zero base rotation, in order to restore compatibility. According to (29.1), the effect of this displacement on the shell can be calculated elastically using modulus $E''(t,t_0)$. The age at prestressing, t_1, is here assumed to be equal to t_0; if it is not, then the prestressing effect and the effect of shrinkage must be considered separately using different moduli $E''(t,t_1)$ and $E''(t,t_0)$.

Based on the well-known solution of the differential equation for axisymmetric bending of a cylindrical shell with a remote top, fixed to a rigid base (e.g. Billington, 1965, p. 91), it was shown by Bažant, Carreira and Walser (1975) that the bending

moments about horizontal axis and the associated shear forces corresponding to elastic modulus equal to $E''(t,t_0)$ are

$$\Delta M(z,t) = 2\Delta u(t) D''(t,t_0)\beta^2 \mathrm{e}^{-\beta z}(\sin\beta z - \cos\beta z) \tag{29.48}$$

$$\Delta Q(z,t) = 4\Delta u(t) D''(t,t_0)\beta^3 \mathrm{e}^{-\beta z}\cos\beta z \tag{29.49}$$

where

$$\beta^4 = \frac{3(1-\nu^2)}{R^2h^2}, \qquad D'' = \frac{E''h^3}{12(1-\nu^2)} \tag{29.50}$$

Here $D''(t,t_0)$ represents the cylindrical stiffness of the shell corresponding to modulus $E''(t,t_0)$; h = shell thickness; and t_0 = age of shell at the start of drying exposure.

Figure 29.15 shows a typical containment shell built in Illinois in the 1970s (with dimensions in feet and inches). The distribution of maximum normal stress produced by shrinkage at the shell surface at time $t = 30$ years is plotted at the bottom left of the shell (Bažant, Carreira and Walser, 1975).

If the shell is buried in the ground up to a certain height, it is protected from drying and does not shrink up to that height. Then the shell may first be considered to be disconnected at the ground level, the top part of shell shrinking freely, and a displacement canceling the shrinkage displacement from t_0 to t must then be imposed to restore compatibility of the shell at the ground level. The stresses produced by this, which are plotted in Figure 29.15 at bottom right of shell, are much smaller than in the former case.

Example 29.14: Deflection of a Statically Indeterminate Structure

We want to calculate the deflection $w_F(t)$ of a prestressed concrete beam structure (a planar frame) in the sense of force F, caused by permanent load applied at age t_1, initial prestress applied at age t_0, loss of prestress, and shrinkage that is nonuniform throughout the structure. The shrinkage strain $\varepsilon_{sh}(x,y,z)$ depends not only on the coordinate x measured along the beam axis and identifying the cross section, but also on the coordinates y and z measured along the beam thickness and depth. Let $X_1(t),\dots X_s(t)$ = statically indeterminate (redundant) forces, $\bar{M}(x)$ = bending moments caused by $F = 1$ in a chosen primary structure (statically determinate), $\bar{\sigma}(x,z)$ = stresses caused by $F = 1$ in the primary structure, $M_L(x)$ = bending moments caused in the primary structure by the given permanent load applied at age t_1, $M_{pr}(x)$ = given bending moments caused in the primary structure by the initial prestress applied at age t_0, λ_{pr} = long-time loss of prestress from t_0 to t as a fraction of the initial prestress, and $\varepsilon_{sh}(x,y,z,t)$ = shrinkage strains from the start of drying to t. We assume that the changes $\Delta X_{i,L}(t)$ and $\Delta X_{i,pr}(t)$ due to given load and to prestress have been solved. Using (29.1) and the principle of superposition, and noting the expressions (29.2) for the unit flexibilities, we obtain

$$w_F(t) = f_{F,L}\frac{1+\varphi(t,t_1)}{E(t_1)} + f_{F,pr}\frac{1+\varphi(t,t_0)}{E(t_0)} - f_{F,pr}\frac{\lambda_{pr}}{E''(t,t_0)} + \sum_{i=1}^{s} f_{F,i}\left(\frac{\Delta X_{i,L}(t)}{E''(t,t_1)} + \frac{\Delta X_{i,pr}(t)}{E''(t,t_0)}\right) + w_{F,sh}(t) \tag{29.51}$$

where $f_{F,L}$ is given by (29.6) and

$$f_{F,pr} = \int_{(x)} \frac{\bar{M}(x)}{I(x)}\, M_{pr}(x)\,\mathrm{d}x, \qquad f_{F,i} = \int_{(x)} \frac{\bar{M}(x)}{I(x)}\, \bar{M}_i(x)\,\mathrm{d}x \tag{29.52}$$

$$w_{F,sh}(t) = \int_{(x)}\int_A \bar{\sigma}(x,z)\varepsilon_{sh}(x,y,z,t)\,\mathrm{d}A\,\mathrm{d}x \tag{29.53}$$

Other Practical Problems

The AAEM as well as general computational approaches based on the history integral have also been used to solve the following creep problems:

1. Effect of differences in water loss and free shrinkage across the thickness of containment shell (Bažant, Carreira and Walser, 1975).
2. Effect of temporary shoring in construction of concrete frames of high-rise buildings (Aguinaga-Zapata and Bažant, 1986).
3. Effects of nonuniform drying of the walls in a prestressed concrete box girder (Bažant, Křístek and Vítek, 1992).
4. Effect of shear lag in box girder bridges (Křístek and Bažant, 1987).
5. Effects of random environment on containment shells, analyzed by spectral approach (Bažant and Wang, 1984).
6. Relaxation of shrinkage and thermal stresses in nuclear reactor vessel or containment (Bažant *et al.*, 1984a, 1985).
7. Stresses produced by hydration heat in dams or nuclear reactor containments, which get strongly relaxed due to creep.
8. Stresses due to nonuniform drying and caused by random environment treated by spectral approach.
9. Bridge deflections due to additional nonlinear creep caused by cyclic nature of loading (Bažant, 1968).

29.2.3 Structures with Distributed Cracking

Unprestressed reinforced concrete beams and plates are designed to exhibit extensive distributed cracking under permanent loads in service. The tensile stresses in the structure are then totally transmitted by steel reinforcement with an adhering layer of concrete. The reinforcement does not creep, which causes an overall reduction of creep effects in the structure. Because of this reduction, it is not important to perform a rigorous analysis, which is a complex nonlinear problem. For many applications, it suffices to treat the effect of cracking and tensile reinforcement in an approximate linearized manner, and the classical effective modulus method appears to suffice here.

The complexity arises from the fact that, despite distributed cracking, the concrete in tension still contributes to bending stiffness, because of two distinct mechanisms:

1. Between each two parallel tensile cracks, the concrete adheres to the bar, due to bond stresses, and consequently, the region adjacent to the bar (shaded in Figure 29.16(a)), transmits significant tensile stresses and opposes the extension of the steel bar. This mechanism is known as the *tension stiffening*.
2. The tensile cracks are *cohesive cracks*, i.e. they transmit crack-bridging stresses (marked as σ_t in Figure 29.16(b)) which decrease as the crack opening increases and vanish only when the opening becomes quite large.

The former mechanism is important. The latter mechanism is less important, and probably unimportant for long-time loading. The reason for this is that, since the crack-bridging stresses rapidly relax with time (Bažant and Li, 1997), they can be neglected

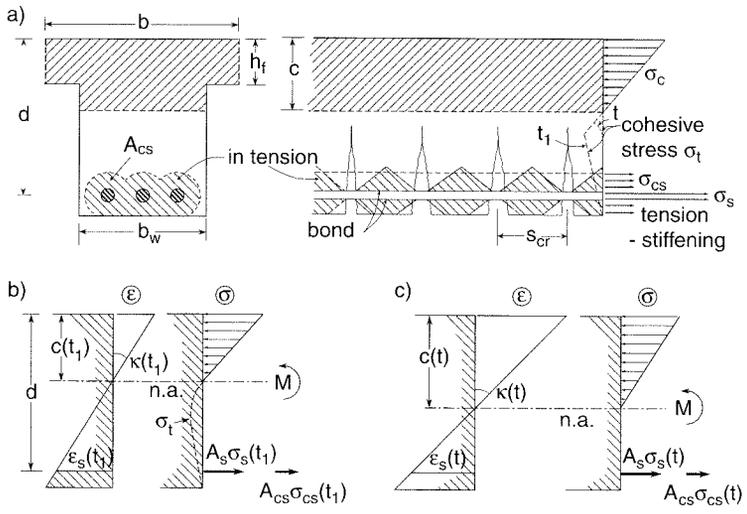


Figure 29.16 Cracked reinforced concrete beam

for long-time loading. In the former mechanism, the tensile stresses in the region parallel to the bar are in the hardening (pre-peak) range, and the creep strains they engender should roughly follow the compliance function $J(t,t')$.

So, for long-time loading, we only need to consider the tension stiffening mechanism. For this purpose, we will imagine the concrete between the cracks that deforms together with the extension of the steel bar to be equivalent to an imagined 'tension-stiffening bar' having, as a rough approximation, the cross section area (Figure 29.16(a))

$$A_{cs} \approx \alpha_{cs} A_s \tag{29.54}$$

where α_{cs} is an empirical constant factor greater than 1 (and of the order of 10). More accurately, A_{cs} can be figured out as a union of overlapping areas such that its points are no farther from the closest steel bar than $s_{cr}/3$, where s_{cr} is the estimated average crack spacing; $s_{cr} \approx 3$ to 6 maximum aggregate sizes; see Figure 29.16(a). If σ_{cs} is the average tensile stress in this bar, the term in AAEM equation (29.1) that gives the creep strain at constant stress in the imagined tension-stiffening bar is $\sigma_{cs}\varphi(t,t_1)/E(t_1)$.

Under the usual assumption of cross sections remaining planar, the deformation of the cross section is fully characterized by the distance of the neutral axis from the top face, c, and the curvature, κ. First we solve the initial state after suddenly applying M at age t_1. The distributions of strains ε and stresses σ in the cross section are plotted in Figure 29.16(b). The extreme compressive stress in concrete is $-\kappa c$, and the average tensile strain in steel and in the surrounding tensile-stiffening bar is $\kappa(d-c)$. In the compressed part of the cross section (above the neutral axis), the concrete deforms according to elastic modulus $E_c = E(t_1)$; in the tensile part of the cross section (below the neutral axis), the concrete is considered to transmit no tensile stresses, except for those in the assumed tension-stiffening bar of area A_{cs}.

To keep the derivation simple, we analyze a rectangular cross section, i.e. we consider $b = b_w$ (Figure 29.16(a)). The more general case of a T-shaped cross section (concrete slab supported by a reinforced concrete beam) is left to the reader as an exercise (see Problem 29.4). In the case of pure bending, the total normal force is

zero, and so the force resultant of compressive stresses in concrete must be equal to the resultant of the tensile stresses in steel and the tension-stiffening bar:

$$bc \times \tfrac{1}{2} E_c \kappa c = A_s \left(E_s + \alpha_{cs} E_c\right) \kappa (d-c) \tag{29.55}$$

Canceling the curvature and denoting $n = E_s/E_c$, we obtain from (29.55) a quadratic equation for the unknown position of neutral axis,

$$\frac{b}{2A_s(n+\alpha_{cs})} c^2 + c - d = 0 \tag{29.56}$$

This equation has only one meaningful (positive) root,

$$c = (n+\alpha_{cs}) \frac{A_s}{b} \left[\sqrt{1 + \frac{2bd}{A_s(n+\alpha_{cs})}} - 1 \right] \tag{29.57}$$

Once the position of the neutral axis is known, the bending moment can be expressed as the product of the compressive resultant and the distance between the two resultants,

$$M = \tfrac{1}{2} E_c b c^2 \kappa \left(d - \tfrac{c}{3}\right) = \tfrac{1}{6} E_c b c^2 (3d-c)\kappa \tag{29.58}$$

The term $I_{ef} = bc^2(3d-c)/6$ can be interpreted as the effective moment of inertia of the cross section, and (29.58) can be written as $M = E_c I_{ef}\, \kappa$.

The preceding derivation applies both to the initial state at time t_1 and to the current state at time $t > t_1$. In the former case, we have to substitute $E_c = E(t_1) =$ elastic modulus at time t_1, while in the latter case we set $E_c = E_{ef}(t,t_1) = 1/J(t,t_1)$ $= E(t_1)/[1+\varphi(t,t_1)]$, for both the compressed concrete above the neutral axis and the concrete in the tension-stiffening bar. Due to the change of concrete modulus, the stiffness ratio n, position of the neutral axis c and the curvature κ produced by a constant moment M will all vary in time. We can now define the effective creep coefficient for bending in a cross section with distributed cracking, $\bar{\varphi}(t,t_1)$, such that the increment of curvature under constant moment be

$$\Delta\kappa(t) = \kappa(t_1)\bar{\varphi}(t,t_1) \tag{29.59}$$

To achieve this, we have to set

$$\bar{\varphi}(t,t_1) = \frac{\kappa(t) - \kappa(t_1)}{\kappa(t_1)} = \frac{E(t_1) I_{ef}(t_1)}{E_{ef}(t,t_1) I_{ef}(t)} - 1 = [1+\varphi(t,t_1)] \frac{I_{ef}(t_1)}{I_{ef}(t)} - 1 \tag{29.60}$$

where $I_{ef}(t_1)$ and $I_{ef}(t)$ are the effective moments of inertia computed for the positions of the neutral axis $c(t_1)$ and $c(t)$, respectively. Note that if there is no shifting of the neutral axis, $\bar{\varphi}(t,t_1)$ is equal to the standard creep coefficient $\varphi(t,t_1)$.

It may seem strange that we have used the simple effective modulus instead of the age-adjusted effective modulus. However, for the problem treated in this example, the difference between both techniques was found to be very small (see Bažant and Najjar 1973, Figure 10). The reason is that, at a constant bending moment, the stress in concrete remains almost constant, and creep under constant stress is described by both methods exactly. AAEM can be expected to provide superior results in more complicated cases with variable loading or with shrinkage. Due to the

nonlinear character of the problem (the tensile response of concrete is not elastic), the principle of superposition does not apply and all the effects must be analyzed simultaneously.

In Examples 29.1–29.12, it was tacitly assumed that the beams subjected to bending are prestressed (fully). If they are not, it suffices, as an approximation, to replace $\varphi(t,t_1)$ in these examples with $\bar{\varphi}(t,t_1)$. In some examples, the long-time moments M in the cross sections are indeterminate and thus unknown in advance, and they also vary in time. In these cases, one must assume a certain average M value before the computations, and then possibly correct it, iterating the calculations.

For rough approximations it suffices to calculate $\bar{\varphi}(t,t_1)$ only for the cross section with the maximum bending moment. It is more accurate, however, to calculate it for several cross sections along the beam and interpolate; this in general yields a smaller deflection. The long-time flexibilities then need to be calculated using an effective creep coefficient $\bar{\varphi}(t,t')$ that varies along the beam.

While the neglect of the cohesive stresses in the cracks is doubtless justified for long-time loading, it might be non-negligible for the initial curvature $\kappa(t_1)$. However, this does not appear to be a serious problem for long-time creep analysis because the subdivision of the total compliance into the elastic and creep compliances is unimportant for long-time predictions as long as the total long-time compliance is correct. The previous discussion of the choice of the short-time (or elastic) modulus is pertinent to this point and need not be repeated here.

Crude though the present approximation might be, larger percentage errors are nevertheless acceptable because (1) the reduced creep coefficient makes the magnitude of creep effects generally small, and (2) unprestressed structures are generally more massive then the prestressed ones, which makes not only the elastic deformations, but also the creep deformations relatively small.

29.3 COMPUTATIONAL APPROACHES

29.3.1 Multiaxial Constitutive Equation for Creep and Shrinkage

So far, we have discussed simple engineering solutions of elementary creep and shrinkage problems. Analysis of sensitive real structures with complicated shapes and loading histories must be done by numerical techniques, usually based on the finite element method. Since the stress states are often multiaxial, we need an appropriate generalization of the uniaxial constitutive relations.

The generalization is simplified by the fact that the Poisson ratio for creep of concrete is approximately time-independent and about the same as the elastic Poisson ratio ν, both being about 0.18 (for basic creep). Thus, in view of isotropy of concrete, and in analogy to the elastic relation $\boldsymbol{\varepsilon} = \boldsymbol{C}_\nu \boldsymbol{\sigma}$, the multiaxial generalization of the history integral (28.3) based on the principle of superposition may be written as

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{C}_\nu \int_0^t J(t,t')\, \mathrm{d}\boldsymbol{\sigma}(t') + \boldsymbol{\varepsilon}_0(t) \tag{29.61}$$

Here, $\boldsymbol{\varepsilon}$ is the column matrix of engineering strain components defined in Section D.7, $\boldsymbol{\sigma}$ is the column matrix of stress components, $\boldsymbol{\varepsilon}_0$ is the column matrix of shrinkage

strains (possibly combined with thermal strains), and

$$\boldsymbol{C}_\nu = \begin{bmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ & 1 & -\nu & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & 2(1+\nu) & 0 & 0 \\ \text{sym.} & & & & 2(1+\nu) & 0 \\ & & & & & 2(1+\nu) \end{bmatrix} \tag{29.62}$$

is the unit elastic compliance matrix, corresponding to Young's modulus $E = 1$; cf. (D.110). The inverse form of (29.61), which is the multiaxial generalization of (28.5), reads

$$\boldsymbol{\sigma}(t) = \boldsymbol{D}_\nu \int_0^t R(t,t')\, \mathrm{d}\left[\boldsymbol{\varepsilon}(t') - \boldsymbol{\varepsilon}_0(t')\right] \tag{29.63}$$

where $\boldsymbol{D}_\nu = \boldsymbol{C}_\nu^{-1}$ is the unit elastic stiffness matrix, which can be deduced from (D.109) with $E = 1$.

29.3.2 AAEM via Elastic Finite Element Program

In general, the effects of linear aging creep can easily be calculated with an elastic finite element program. According to (29.1), two runs of the elastic finite element program are necessary (their results must then be superposed):

1. **Initial Elastic State.** The elastic material stiffness matrix at each integration point (of each finite element) is set to $E(t_1)\boldsymbol{D}_\nu$, where $E(t_1)$ is the elastic modulus for the age t_1 at first loading of the structure. For the given initial loads, the program yields the stresses, strains and displacements at time t_1^+, just after the load application.
2. **Change from t_1 to t.** The elastic material stiffness matrix at each integration point is set to $E''(t,t_1)\boldsymbol{D}_\nu$ where $E''(t,t_1)$ is the age-adjusted effective modulus. In addition to this, one must consider at each integration point the inelastic strain increments

$$\Delta\boldsymbol{\varepsilon}'' = \frac{\varphi(t,t_1)}{E(t_1)} \boldsymbol{C}_\nu \boldsymbol{\sigma}(t_1) + \Delta\boldsymbol{\varepsilon}_0(t) \tag{29.64}$$

If the finite element program does not have the capability to handle general inelastic strains, and if all the stresses are uniaxial (as in layered beam theory), one can prescribe in each finite element an equivalent temperature change $\Delta T = \Delta\varepsilon''/\alpha$, where α is the thermal expansion coefficient.

29.3.3 Step-by-Step Solution Based on Hereditary Integral

If discrete times t_n $(n = 0, 1, 2, \ldots N_{step})$ are introduced and the history integral in (29.61) is approximated by a sum according to the trapezoidal rule or the midpoint rule, one obtains the incremental stress-strain relation (Bažant, 1972a, 1975)

$$\Delta\boldsymbol{\varepsilon}^{(n)} = \frac{1}{\bar{E}} \boldsymbol{C}_\nu\, \Delta\boldsymbol{\sigma}^{(n)} + \Delta\boldsymbol{\varepsilon}'' \tag{29.65}$$

where

$$\bar{E} = \frac{1}{J_{n,n-1/2}}, \qquad \Delta\boldsymbol{\varepsilon}'' = \boldsymbol{C}_\nu \sum_{m=1}^{n-1} \Delta J_{n,m} \Delta\boldsymbol{\sigma}^{(m)} + \Delta\boldsymbol{\varepsilon}_0^{(n)} \tag{29.66}$$

$$\Delta J_{n,m} = J_{n,m-1/2} - J_{n-1,m-1/2} \tag{29.67}$$

Here, $\Delta(...)^{(n)}$ refers to the increments during the time interval $[t_{n-1}, t_n]$ and

$$J_{n,m-1/2} = \begin{cases} \frac{1}{2}[J(t_n, t_{m-1}) + J(t_n, t_m)] & \text{for the trapezoidal rule} \\ J(t_n, \frac{1}{2}[t_{m-1} + t_m]) & \text{for the midpoint rule} \end{cases} \tag{29.68}$$

Equations (29.67)–(29.68) are suitable for the theories in which the compliance function has a simple expression. For the solidification theory and in particular the B3 model, they should be replaced by

$$\Delta J_{n,m} = \Delta t_n \dot{J}(\tfrac{1}{2}(t_{n-1} + t_n), \tfrac{1}{2}(t_{m-1} + t_m)) \tag{29.69}$$

where $\Delta t_n = t_n - t_{n-1}$.

In each time step, $\Delta\boldsymbol{\varepsilon}''$ can be evaluated from the previous stress history, and thus is known. Therefore, (29.65) is equivalent to an elastic stress-strain relation with inelastic strains, and the response of the structure can be analyzed using the following algorithm (Bažant, 1972a, 1975).

Algorithm 29.1:

1. For a given time step number n, assemble the elastic stiffness matrix of the structure using the elastic material stiffness matrix $\bar{E}\boldsymbol{D}_\nu$.
2. Assemble the equivalent load vector corresponding to inelastic strain increments $\Delta\boldsymbol{\varepsilon}''$, evaluated at each integration point of the finite element model on the basis of the previous stress history at that point up to time t_{n-1} and of the non-mechanical strain increment during step number n.
3. Solve the set of linear equations to obtain the displacement increments.
4. Update the displacements and stresses, return to 1 and start the analysis of the next time step.

The solution converges quadratically (Bažant, 1972a) in the sense that, as the time step is refined, the error decreases proportionally to the square of the time step.

Step-by-step approximation of a hereditary integral analogous to (29.65)–(29.68) can also be used for a structural model based on the flexibility method. First, we replace strains by curvatures, stresses by bending moments, and the unit compliance matrix $\boldsymbol{C}_\nu$ by the reciprocal value of the cross sectional moment of inertia. This leads to the equation

$$\Delta\kappa^{(n)} = \frac{1}{\bar{E}I}\,\Delta M^{(n)} + \frac{1}{I}\sum_{m=1}^{n-1} \Delta J_{n,m} \Delta M^{(m)} + \Delta\kappa_0^{(n)} \tag{29.70}$$

where $\bar{E}$ and $\Delta J_{n,m}$ have exactly the same meaning as in (29.66)–(29.67), and $\Delta\kappa_0^{(n)}$ is the increment of the curvature due to shrinkage and thermal strain in the time interval $[t_{n-1}, t_n]$.

Next, we substitute these relations into the incremental compatibility conditions of the flexibility method, $\int_{(x)} \bar{M}_i\,\Delta\kappa^{(n)}\,\mathrm{d}x = 0$, expressing the moments as a superposition of moments from applied loading on the primary structure and moments from the redundant forces, $M = M_L + \sum_j \bar{M}_j X_j$. The resulting equations read

$$\frac{1}{\bar{E}}\left(\Delta f_{i,L}^{(n)} + \sum_{j=1}^{s} f_{ij}\,\Delta X_j^{(n)}\right) + \sum_{m=1}^{n-1} \Delta J_{n,m}\left(\Delta f_{i,L}^{(m)} + \sum_{j=1}^{s} f_{ij}\,\Delta X_j^{(m)}\right) + \int_{(x)} \bar{M}_i\,\Delta\kappa_0^{(n)}\,\mathrm{d}x = 0, \qquad i = 1, 2, \ldots s \tag{29.71}$$

From these incremental conditions of compatibility, we can in each step calculate the increments of redundant forces, $\Delta X_j^{(n)}$, $j = 1, 2, \ldots s$. For example, for a frame with degree of redundancy 1, we would get

$$\Delta X_1^{(n)} = -\frac{\Delta f_{1,L}^{(n)}}{f_{11}} - \bar{E}\sum_{m=1}^{n-1} \Delta J_{n,m}\,\Delta X_1^{(m)} - \frac{\bar{E}}{f_{11}}\left(\sum_{m=1}^{n-1} \Delta J_{n,m}\,\Delta f_{1,L}^{(m)} + \int_{(x)} \bar{M}_1\,\Delta\kappa_0^{(n)}\,\mathrm{d}x\right) \tag{29.72}$$

29.3.4 Rate-Type Creep Law and Internal Variables

In very large structural systems, the storage of the entire stress history at each integration point of each finite element and the evaluation of history integrals according to (29.65) can be prohibitively expensive, even with the most powerful computers. Note that the storage requirements and the number of operations needed to assemble *one* structure stiffness matrix and *one* right-hand side are proportional to the number of preceding time steps. If the total number of time steps in the analysis, N_{step}, is large, the CPU time becomes proportional to N_{step}^2 (or worse, if the memory capacity is exceeded and external storage must be used). For example, the total number of assembly-related operations in an analysis with 1000 time steps is 100 times larger than in an analysis with 100 time steps (on the same finite element mesh).

The computation can be made much more efficient by approximating the integral-type stress-strain relation (29.61) or (29.63) with a rate-type stress-strain relation based on Kelvin chain or Maxwell chain spring-dashpot model of aging viscoelasticity. In such a differential formulation, the previous history is characterized by the current values of several internal variables. Since the number of internal variables is fixed, the computational cost per step is constant, independent of the number of preceding steps. The rate-type formulation is also advantageous for extensions to variable temperature and for fracture analysis, because it makes it easier to incorporate various physical concepts. For example, the concept of temperature activation energy can be applied to the viscosities, and separately to the rate of aging as a chemical reaction (Bažant and Wu, 1974).

Due to space limitations, we will present here only the approach leading to a *Kelvin chain*. The compliance function is approximated by the *Dirichlet series* (Hardy and Riesz, 1915; Schapery, 1962)

$$J(t, t') \approx \frac{1}{E_0} + \sum_{i=1}^{M} \frac{1}{D_i(t')}\left[1 - \exp\left(-\frac{t - t'}{\tau_i}\right)\right] \tag{29.73}$$

where τ_i, $i = 1, 2, \ldots M$, are fixed parameters called the *retardation times*, E_0 is the asymptotic modulus, and $D_i(t')$, $i = 1, 2, \ldots M$, are age-dependent moduli which can be determined, for example, by least-square fitting of the 'exact' compliance function. Alternatively, and in fact preferably, these moduli may be determined from the so-called continuous retardation spectrum; see Bažant and Xi (1995) for details. In either case, the retardation times must be selected in advance (an attempt to determine them by the least square method would lead to an ill-conditioned problem); they must cover the entire time range of interest and must not be too far apart in the logarithmic time scale (or the creep curves would look 'bumpy') nor too close (or else the calculation of D_i would be ill-conditioned). It is recommended to distribute the retardation times in a geometric progression with quotient 10, taking $\tau_1 \leq 0.3\tau_{min}$ and $\tau_M \geq 0.5\tau_{max}$, where τ_{min} and τ_{max} are the shortest and longest time delay after an instantaneous load application or a change for which the response should be accurately reproduced. For example, if we want to study the creep of a structure during 50 years of service life, and we want to correctly resolve the short-term response already five minutes after a sudden change of loading, it is necessary to use as many as 8 terms of the Dirichlet series, with $\tau_1 = 1.5$ min $= 0.001$ day, $\tau_2 = 10\ \tau_1 = 0.01$ day, $\tau_3 = 10\ \tau_2 = 0.1$ day, $\ldots\ \tau_8 = 10^7\tau_1 = 10,000$ days ≈ 27 years. Figure 29.17 shows the compliance function given by the log-double-power law (28.6) and its approximation by a Dirichlet series with $M = 8$ and with E_0 considered as a free parameter determined by fitting rather than as the asymptotic modulus. If we are only interested in the long-term behavior, but not in the details of the response shortly after load application or change, it is sufficient to use only a few Dirichlet terms, say four, with retardation times from 10 days to 10,000 days. The corresponding approximation of the compliance function is also shown in Figure 29.17.

To simplify the derivations, we will first analyze a uniaxial model, for which stress and strain are described by scalars and the matrices $\boldsymbol{D}_\nu$ and $\boldsymbol{C}_\nu$ can be omitted. Also, shrinkage is not considered, but its incorporation would be straightforward.

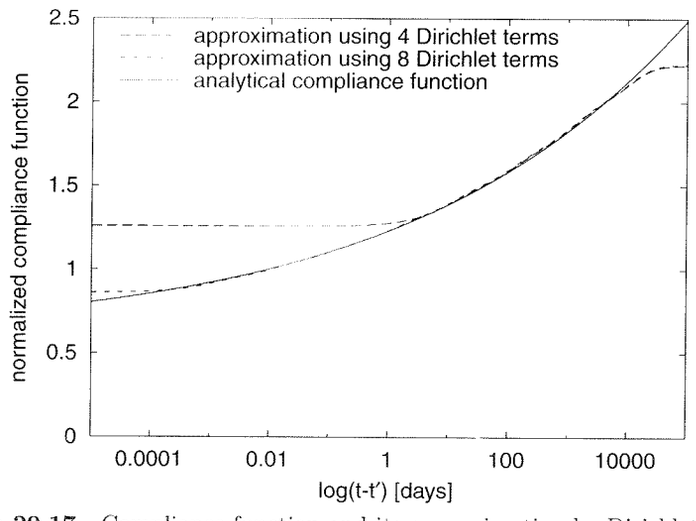


Figure 29.17 Compliance function and its approximation by Dirichlet series

Substituting the degenerate kernel (29.73) into the integral relation (28.3), we obtain

$$\varepsilon(t) = \sum_{i=0}^{M} \varepsilon_i(t) \tag{29.74}$$

where

$$\varepsilon_0(t) = \int_0^t \frac{1}{E_0}\, \mathrm{d}\sigma(t') = \frac{1}{E_0}\sigma(t) \tag{29.75}$$

$$\varepsilon_i(t) = \int_0^t \frac{1}{D_i(t')} \left[1 - \exp\left(-\frac{t-t'}{\tau_i}\right)\right] \mathrm{d}\sigma(t'), \qquad i = 1, 2, \ldots M \tag{29.76}$$

Relation (29.74) means that the total strain is the sum of contributions of $M + 1$ units coupled in series. According to (29.75), unit number 0 is an elastic spring of stiffness E_0. To find the rheological model of the other units, we differentiate (29.76) and express the strain rate

$$\dot{\varepsilon}_i(t) = \int_0^t \frac{1}{\tau_i D_i(t')} \exp\left(-\frac{t-t'}{\tau_i}\right) \mathrm{d}\sigma(t') \tag{29.77}$$

Combining this with the original relation (29.76), we find that the following differential equation is satisfied;

$$\varepsilon_i(t) + \tau_i\dot{\varepsilon}_i(t) = \int_0^t \frac{\mathrm{d}\sigma(t')}{D_i(t')} \tag{29.78}$$

Now consider the non-aging case, in which $D_i(t') = D_i = \text{const}$. The integral in (29.78) is then equal to $\sigma(t)/D_i$, and we obtain a differential equation

$$D_i\varepsilon_i(t) + \tau_i D_i\dot{\varepsilon}_i(t) = \sigma(t) \tag{29.79}$$

The total stress is expressed here as a sum of two terms. The first term, $D_i\varepsilon_i$, corresponds to the stress in an elastic spring of stiffness D_i. The second term, $\tau_i D_i\dot{\varepsilon}_i$, is the stress generated by the strain rate $\dot{\varepsilon}_i$ in a linear damper of viscosity $\eta_i = \tau_i D_i$. This means that in the non-aging case the approximation of the compliance function by the Dirichlet series (29.73) corresponds to the Kelvin chain shown in Figure 29.18(a), with constant properties of individual elements.

The general case of an aging material is somewhat more involved. To eliminate the integral from (29.78), we must differentiate again, arriving at

$$\dot{\varepsilon}_i(t) + \tau_i\ddot{\varepsilon}_i(t) = \frac{\dot{\sigma}(t)}{D_i(t)} \tag{29.80}$$

Intuitively, we expect this equation to describe a Kelvin unit (parallel coupling of a spring and damper) with age-dependent properties. However, the equation for an aging spring cannot be written as $\sigma_i(t) = D_i(t)\varepsilon_i(t)$ because certain thermodynamic restrictions due to the solidification nature of aging would be violated (Bažant, 1975, 1979). Rather the equation for the spring must be written as $\dot{\sigma}_i(t) = D_i(t)\dot{\varepsilon}_i(t)$. Therefore, if we consider a Kelvin unit characterized by a variable stiffness $E_i(t)$ and a variable viscosity $\eta_i(t)$, the corresponding differential equation is

$$E_i(t)\dot{\varepsilon}_i(t) + \frac{\mathrm{d}}{\mathrm{d}t}\left[\eta_i(t)\dot{\varepsilon}_i(t)\right] = \dot{\sigma}(t) \tag{29.81}$$

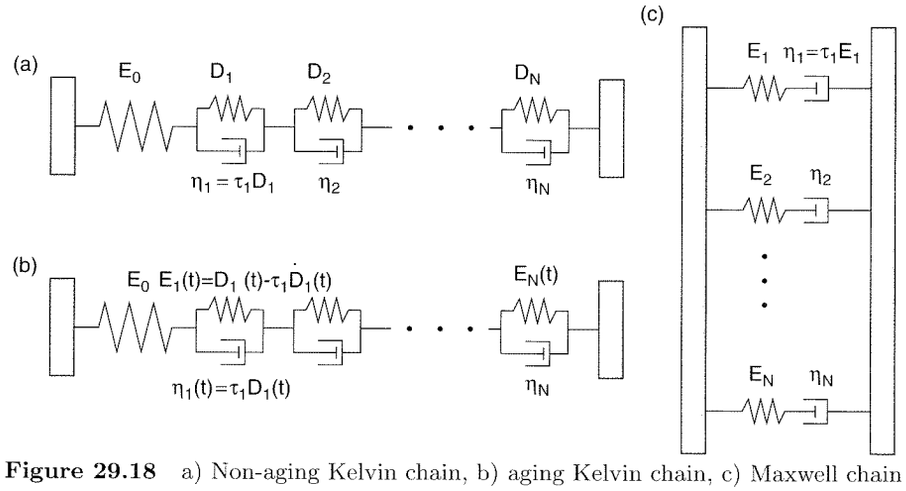


Figure 29.18 a) Non-aging Kelvin chain, b) aging Kelvin chain, c) Maxwell chain

which can be expanded into

$$E_i(t)\dot{\varepsilon}_i(t) + \dot{\eta}_i(t)\dot{\varepsilon}_i(t) + \eta_i(t)\ddot{\varepsilon}_i(t) = \dot{\sigma}(t) \tag{29.82}$$

Comparing this with (29.80) we realize that the properties of the rheological model must be identified as (Bažant, 1996)

$$\eta_i(t) = \tau_i D_i(t) \tag{29.83}$$

$$E_i(t) = D_i(t) - \dot{\eta}_i(t) = D_i(t) - \tau_i \dot{D}_i(t) \tag{29.84}$$

The corresponding Kelvin chain is shown in Figure 29.18(b).

The previous derivations are easily generalized to multiaxial stress. The total strain is given by the sum

$$\boldsymbol{\varepsilon}(t) = \frac{1}{E_0}\boldsymbol{C}_\nu\boldsymbol{\sigma}(t) + \sum_{i=1}^{M}\boldsymbol{\varepsilon}_i(t) \tag{29.85}$$

where partial strains $\boldsymbol{\varepsilon}_i$, $i = 1, 2, \ldots M$, are governed by a set of differential equations. For non-aging materials we obtain equations of the first order,

$$D_i\boldsymbol{\varepsilon}_i(t) + \tau_i D_i\dot{\boldsymbol{\varepsilon}}_i(t) = \boldsymbol{C}_\nu\boldsymbol{\sigma}(t) \tag{29.86}$$

but for aging materials the resulting equations

$$D_i(t)\dot{\boldsymbol{\varepsilon}}_i(t) + \tau_i D_i(t)\ddot{\boldsymbol{\varepsilon}}_i(t) = \boldsymbol{C}_\nu\dot{\boldsymbol{\sigma}}(t) \tag{29.87}$$

are of the second order. The initial conditions for the partial strains are $\boldsymbol{\varepsilon}_i(0) = \mathbf{0}$. For the second-order differential equations (29.87) we need another set of initial conditions specifying the partial strain rates. Since the time is measured from the set of concrete, there is always an initial period of time during which the stress is zero, same as the strain and the strain rate. The response is studied starting from some time $t_1 > 0$ at which the first load is applied. Just before that, the strain rate is still zero, and this provides the second initial condition.

In the differential (rate-type) formulation, the effect of history is described by a finite number of internal variables that characterize individual units of the

corresponding rheological chain. To update the solution, it is sufficient to know the values of the internal variables at the end of the previous time step. This greatly reduces the storage requirements and increases the efficiency of the numerical solution.

From the physical point of view, it would be most natural to assign the role of internal variables directly to the partial strains $\boldsymbol{\varepsilon}_i$, $i = 1, 2, \ldots M$. However, their evolution would be governed by second-order differential equations (29.87) and the numerical integration scheme would become more demanding. A more practical choice is to select internal variables $\boldsymbol{\gamma}_i$ proportional to the partial strain rates (they are not truly internal variables in the sense of thermodynamics, but we will keep referring to them like that). More specifically, we set

$$\boldsymbol{\gamma}_i(t) = \tau_i\dot{\boldsymbol{\varepsilon}}_i(t) \tag{29.88}$$

Equation (29.87) is now replaced by a first-order differential equation

$$\dot{\boldsymbol{\gamma}}_i(t) + \frac{1}{\tau_i}\boldsymbol{\gamma}_i(t) = \frac{\boldsymbol{C}_\nu\dot{\boldsymbol{\sigma}}(t)}{D_i(t)} \tag{29.89}$$

with the initial condition $\boldsymbol{\gamma}_i(0) = \mathbf{0}$.

An alternative approach can be based on the Dirichlet approximation of the relaxation function $R(t, t')$. This leads to a system of first-order differential equations corresponding to the *Maxwell chain* in Figure 29.18c.

29.3.5 Exponential Algorithm

In general, equations (29.89) have to be integrated numerically for $i = 1, 2, \ldots M$. The solution proceeds in a sequence of time steps from $t = t_n$ to $t = t_{n+1}$. At the beginning of each step, the value

$$\boldsymbol{\gamma}_i(t_n) = \boldsymbol{\gamma}_i^{(n)} \tag{29.90}$$

is known from the previous step (or, in the first step, from the initial condition). Regarding the selection of time steps Δt for long-time creep analysis of structures under constant loads, note that at the beginning the response varies fast and Δt needs to be very short, e.g. 0.001 day, while at the end the response varies slowly, and Δt can be very long, e.g. $\Delta t = 100$ days. So, Δt must be greatly increased for efficiency of computation. The usual integration algorithms for ordinary differential equations would become either unstable or grossly inaccurate if $\Delta t \gg \tau_1$. This is overcome by using the *exponential algorithm*, originally proposed for non-aging materials by Zienkiewicz, Watson and King (1968) and extended to aging materials by Bažant (1971b). For efficiency of calculations, it is best to select the time steps Δt as constant in the scale of $\log(t - t_1)$, which means that the intervals $t_n - t_1$ form a geometric progression. Several steps per decade usually suffice.

The key idea of the exponential algorithm, which allows $\Delta t \gg \tau_1$, is that (29.89) can be integrated analytically if the right-hand side remains constant over the time step. For simplicity, we first develop the algorithm for the uniaxial case and then extend it to multiple dimensions. Denoting $\Delta t = t_{n+1} - t_n$, $\Delta\sigma = \sigma(t_{n+1}) - \sigma(t_n)$, and $D_i^{(n+1/2)} = D_i(t_n + \Delta t/2)$, we can approximate the right-hand side of (29.89) by $\Delta\sigma/\Delta t\, D_i^{(n+1/2)}$. The ordinary differential equation

$$\dot{\gamma}_i(t) + \frac{1}{\tau_i}\gamma_i(t) = \frac{\Delta\sigma}{\Delta t\, D_i^{(n+1/2)}} \tag{29.91}$$

has a general solution

$$\gamma_i(t) = \frac{\tau_i \Delta\sigma}{\Delta t\, D_i^{(n+1/2)}} + K \exp\left(-\frac{t}{\tau_i}\right) \tag{29.92}$$

where K is an arbitrary constant. From condition (29.90) it follows that

$$K = \left(\gamma_i^{(n)} - \frac{\tau_i \Delta\sigma}{\Delta t\, D_i^{(n+1/2)}}\right) \exp\left(\frac{t_n}{\tau_i}\right) \tag{29.93}$$

and the corresponding particular solution reads

$$\gamma_i(t) = \frac{\tau_i \Delta\sigma}{\Delta t\, D_i^{(n+1/2)}} \left[1 - \exp\left(-\frac{t - t_n}{\tau_i}\right)\right] + \gamma_i^{(n)} \exp\left(-\frac{t - t_n}{\tau_i}\right) \tag{29.94}$$

The value of γ_i at the end of the step is given by

$$\gamma_i^{(n+1)} = \gamma_i(t_{n+1}) = \lambda_i \frac{\Delta\sigma}{D_i^{(n+1/2)}} + \beta_i \gamma_i^{(n)} \tag{29.95}$$

where, for convenience, we have denoted

$$\beta_i = \exp\left(-\frac{\Delta t}{\tau_i}\right), \qquad \lambda_i = \frac{\tau_i}{\Delta t}(1 - \beta_i) \tag{29.96}$$

The strain increment over the time step can be evaluated by integrating (29.88). This can be done analytically, and the result is

$$\Delta\varepsilon_i = \frac{1}{\tau_i} \int_{t_n}^{t_{n+1}} \gamma_i(t)\, \mathrm{d}t = \frac{1 - \lambda_i}{D_i^{(n+1/2)}} \Delta\sigma + (1 - \beta_i)\gamma_i^{(n)} \tag{29.97}$$

Note that $0 \le \lambda_i \le 1$, $\lambda_i \to 1$ if $\Delta t \to 0$ and $\lambda_i \to 0$ if $\Delta t \to \infty$. When $\Delta t/\tau_i$ is too large, e.g. > 30, the first expression in (29.96) causes an underflow and must be replaced by $\beta_i = 0$. On the other hand, when $\Delta t/\tau_i$ is much smaller than 1, β_i approaches 1 and the evaluation of λ_i from the second expression in (29.96) is inaccurate; it can be replaced by the first few terms of the Taylor series, $\lambda_i = 1 - \frac{1}{2}(\Delta t/\tau_i) + \frac{1}{6}(\Delta t/\tau_i)^2 - \ldots$.

So far, all the considerations have been done on the level of a single Kelvin unit. Substituting (29.97) into the incremental form of (29.74), we obtain the total strain increment

$$\Delta\varepsilon = \sum_{i=0}^{M} \Delta\varepsilon_i = \Delta\sigma \left(\frac{1}{E_0} + \sum_{i=1}^{M} \frac{1 - \lambda_i}{D_i^{(n+1/2)}}\right) + \sum_{i=1}^{M} (1 - \beta_i)\gamma_i^{(n)} \tag{29.98}$$

Note that we have taken into account the 'zero-th' chain unit, which is an elastic spring of stiffness E_0 (without any damper). The first term on the right-hand side of (29.98) can be interpreted as the strain increment due to a change of the applied stress. In this term, the stress increment is multiplied by the effective compliance

$$\frac{1}{\bar{E}} = \frac{1}{E_0} + \sum_{i=1}^{M} \frac{1 - \lambda_i}{D_i^{(n+1/2)}} \tag{29.99}$$

Note that λ_i is nearly 1 if $\Delta t/\tau_i \le 0.01$, and nearly 0 if $\Delta t/\tau_i \ge 100$. In the former case, the ith Kelvin unit acts as almost rigid, while in the latter case it acts almost as an elastic spring (with no viscous damping). The second term on the right-hand side of (29.98) is the strain increment due to creep at constant stress,

$$\Delta\varepsilon'' = \sum_{i=1}^{M} (1 - \beta_i)\gamma_i^{(n)} \tag{29.100}$$

With this notation, (29.98) can be written in the compact form

$$\Delta\varepsilon = \frac{\Delta\sigma}{\bar{E}} + \Delta\varepsilon'' \tag{29.101}$$

In a strain-driven material subroutine of a finite element code, the strain increment is prescribed and the stress increment has to be computed; see Section 22.1. Equation (29.101) is easily inverted, and the stress increment is evaluated as

$$\Delta\sigma = \bar{E}(\Delta\varepsilon - \Delta\varepsilon'') \tag{29.102}$$

Finally, substituting this into (29.95), we obtain a formula for the update of the internal variables,

$$\gamma_i^{(n+1)} = \frac{\lambda_i \bar{E}}{D_i^{(n+1/2)}} (\Delta\varepsilon - \Delta\varepsilon'') + \beta_i \gamma_i^{(n)} \tag{29.103}$$

The algorithm has been presented for stress and strain as scalars (uniaxial case) but it is easily extensible to the general multiaxial case. The only important modification is that the matrix $\boldsymbol{D}_\nu$ or $\boldsymbol{C}_\nu$ must be inserted in the relation between stresses $\boldsymbol{\sigma}$ and strain-like quantities ($\boldsymbol{\varepsilon}_i$, $\boldsymbol{\gamma}_i$, etc.). The resulting algorithm can be summarized as follows.

Algorithm 29.2:

1. At the beginning of the simulation (time $t = t_0$), set all components of internal variables $\boldsymbol{\gamma}_i^{(0)}$ to zero.
2. For a given time step from t_n to $t_{n+1} = t_n + \Delta t$, compute the mid-step stiffness moduli $D_i^{(n+1/2)}$, the factors β_i and λ_i given by (29.96), and the effective Young's modulus $\bar{E}$ given by (29.99).
3. Compute the strain increment due to creep,

$$\Delta\boldsymbol{\varepsilon}'' = \sum_{i=1}^{M} (1 - \beta_i)\boldsymbol{\gamma}_i^{(n)} \tag{29.104}$$

4. For a given strain increment $\Delta\boldsymbol{\varepsilon}$, compute the stress increment

$$\Delta\boldsymbol{\sigma} = \bar{E}\boldsymbol{D}_\nu(\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}'') \tag{29.105}$$

5. Update the internal variables using the formula

$$\boldsymbol{\gamma}_i^{(n+1)} = \frac{\lambda_i \bar{E}}{D_i^{(n+1/2)}} (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}'') + \beta_i \boldsymbol{\gamma}_i^{(n)} \tag{29.106}$$

Note that step 1 is done only once, before the start of the time incrementation loop. Steps 2 and 3 are done only once per time step, and steps 4 and 5 have to be repeated in every equilibrium iteration. If the entire structure is linearly viscoelastic and the effective material stiffness $\bar{E}\boldsymbol{D}_\nu$ is used in setting up the structural stiffness matrix, equilibrium is restored immediately after the first iteration. However, if certain parts of the model are nonlinear, additional iterations may be needed.

29.3.6 Solidifying Kelvin Chain

The form of the compliance function used by the full model B3 (Appendix E) has been motivated by the *solidification theory* (Bažant, 1977; Bažant and Prasannan, 1989; Carol and Bažant, 1993), which states that the changes of apparent rheological properties (such as the stiffnesses and viscosities in a Kelvin chain) are due to solidification of a non-aging constituent (e.g. due to the hydration reaction in concrete, which leads to the formation of cement gel). The growth of the relative volume of the solidified material is described by a function $v(t)$ that starts from $v(0) = 0$ and asymptotically tends to 1. The compliance function of the solidifying material is then expressed as (Bažant and Prasannan, 1989)

$$J(t,t') = \int_{t'}^{t} \frac{1}{v(\theta)} \dot{\Phi}(\theta - t')\, \mathrm{d}\theta \tag{29.107}$$

where Φ is the compliance function of the non-aging constituent. If this function is approximated by a Dirichlet series

$$\Phi(t) = \sum_{i=1}^{M} \frac{1}{D_i} \left[1 - \exp\left(-\frac{t}{\tau_i}\right)\right] \tag{29.108}$$

it is possible to show that the total strain is a sum of partial strains, ε_i, that are governed by the second-order partial differential equation

$$\tau_i v(t) \ddot{\varepsilon}_i(t) + [v(t) + \tau_i \dot{v}(t)]\, \dot{\varepsilon}_i(t) = \frac{\dot{\sigma}(t)}{D_i} \tag{29.109}$$

After a suitable choice of internal variables

$$\gamma_i(t) = \tau_i v(t) \dot{\varepsilon}_i(t) \tag{29.110}$$

equation (29.109) can be converted into a first-order differential equation

$$\dot{\gamma}_i(t) + \frac{1}{\tau_i}\gamma_i(t) = \frac{\dot{\sigma}(t)}{D_i} \tag{29.111}$$

Formally, this is the same equation as for a non-aging Kelvin chain. The dependence on the maturity of the material is hidden in the definition of γ_i. Consequently, the exponential algorithm developed for a Kelvin unit can easily be adapted for the present model after two small modifications. First of all, moduli D_i are taken as constants. Second, to express the strain increment we must integrate equation (29.110). An approximation of the solidification function $v(t)$ in the interval (t_n, t_{n+1}) by the mid-point value $v^{(n+1/2)} = v(t_n + \Delta t/2)$ leads to

$$\Delta\varepsilon_i = \frac{1}{\tau_i} \int_{t_n}^{t_{n+1}} \frac{\gamma_i(t)}{v(t)}\, \mathrm{d}t \approx \frac{1 - \lambda_i}{D_i v^{(n+1/2)}} \Delta\sigma + \frac{1 - \beta_i}{v^{(n+1/2)}} \gamma_i^{(n)} \tag{29.112}$$

This formula replaces relation (29.97). If needed, the accuracy could be increased by a better approximation of $v(t)$ over each time step. This is usually not necessary, unless the concrete is loaded at a very young age.

Closing Comments

Creep and shrinkage of concrete is a rich and vast subject that has only been touched in this brief exposition. During the last quarter century, various sophisticated formulations have been advanced for the thermodynamics of creep and shrinkage mechanism, interaction with capillary and adsorbed water, coupling with macroscopic diffusion through concrete, effects of elevated and very high temperature, coupling with damage, fracture and size effect, nonlinear creep due to growth of cracking, randomness of creep properties, stochastic process aspects of creep, randomness of environmental effects, statistical prediction of creep and shrinkage from material composition and curing, etc. (e.g. Bažant and Liu, 1985; Planas and Elices, 1993; Bažant and Xi, 1993; Bažant and Kaplan, 1996; Bažant *et al.*, 1997; Baweja, Dvorak and Bažant, 1998).

PROBLEMS

Problem 29.1: Consider that a cantilever placed on a support at time t_2 (Example 29.6) is jacked up. This introduces an initial force $X_1(t_2) \neq 0$. Derive the equation for the time evolution of this force.

Problem 29.2: Find the distribution of bending moments due to restrained shrinkage in the frame in Figure 29.4.

Problem 29.3: Generalize the solution of Example 29.13 to the case when the prestressing is applied at age t_1 different from the age at the onset of shrinkage, t_0.

Problem 29.4: Analyze the stress redistribution due to creep in a cracked reinforced concrete cross section from Section 29.2.3 in the general case when $b \neq b_w$ (a T section). First derive general formulae for the position of the neutral axis and for the effective creep parameter, and then substitute the following specific data: $b = 1$ m, $b_w = 0.3$ m, $h_f = 0.2$ m, $d = 0.5$ m, $E_s = 210$ GPa, $A_s = 2500$ mm^2, $\alpha_{cs} = 10$, and concrete properties same as in Example 29.1. Consider a bending moment M applied at concrete age $t_1 = 30$ days and remaining constant until $t = 1000$ days. How much does the neutral axis shift, and how does the effective creep parameter $\varphi(t, t_1)$ differ from the usual creep parameter $\varphi(t, t_1)$?

Problem 29.5: Based on the step-by-step approximation of the hereditary integral by a sum, set up the algorithm for computer solution of the relaxation curve from $J(t,t')$. Run the algorithm for 100, 1000 and 10 000 time steps and observe the increase of the execution time.

Problem 29.6: Plot the evolution of shrinkage-induced stress in a drying slab analyzed in Example 29.2, using the following specific data: compressive strength 30 MPa, specific water content of concrete mix 150 kg/m^3, Poisson's ratio 0.18, environmental humidity 50%, slab thickness 0.2 m. Consider three different ages at start of drying: 7 days, 14 days and 28 days. For each case determine the maximum stress and the time at which it occurs.

Problem 29.7: Solve the relaxation of prestressing tendon using the general approach outlined in Example 29.9, and show that the result is the same as in Example 29.10. Discuss the physical meaning of terms that appear in the general equations.

Problem 29.8: Derive the differential equations corresponding to the Dirichlet approximation of the relaxation function in the form

$$R(t,t') = \sum_{i=1}^{M} E_i(t') \exp\left(-\frac{t-t'}{\tau_i}\right) \tag{29.113}$$

where τ_i are the *relaxation times* and $E_i(t')$ are time-dependent partial stiffness moduli. Discuss the physical meaning of individual terms.

Problem 29.9: Select appropriate internal variables for the model from Problem 29.8 and develop the exponential algorithm for numerical integration of the set of differential equations.

Appendix A

Linear Elastic Trusses

It is not the aim of this appendix to replace basic textbooks on matrix structural analysis. We give only a brief review of the material needed for the purpose of this book. For more details and exercises, the reader is referred elsewhere (Hibbler, 1994; West 1993; Norris, Wilbur and Utku, 1991). Our attention is restricted to static equilibrium of elastic structures under small displacements.

A.1 FUNDAMENTAL VARIABLES

An arbitrary state of a given elastic structure is completely described by three groups of variables specifying the displacements, strains and stresses. In general, we will refer to these variables as to the *generalized displacements, generalized strains* and *generalized stresses* (internal forces). Their specific nature depends on the type of structural model under consideration. For a truss model, the displacements at an arbitrary point immediately follow from the joint displacements, the strains are uniquely determined by the bar extensions (changes of length compared to the initial stress-free state), and the stress state is represented by the axial forces (also called the normal forces).

The external effects on the structure are described by the applied forces, as well as other external effects (not considered here) such as the settlements of supports and temperature variations.

As an example, consider the simple planar truss in Figure A.1(a), located in the coordinate plane (x, z). Joints 1 and 3 are attached to the supports, and so their displacements are prescribed in advance (we assume that they are equal to zero). If we know the displacement of joint 2, represented by its horizontal component, u_2, and vertical component, w_2, we can determine the displacement of any other point of the structure. The state of strain is represented by extensions e_1 and e_2. A more common notation for the bar extension would be Δl but we reserve the symbol Δ for increments during individual loading steps of incremental analysis. Finally, the state of stress is characterized by axial forces S_1 and S_2. Subscripts 1 and 2 refer here to the respective bars. Sometimes it is more convenient to identify a bar by the numbers of the joints that it connects. In this alternative notation, the extensions may be denoted as e_{12} and e_{23}, and the axial forces as S_{12} and S_{23}. Finally, the applied (external) forces are represented by the components F_{2x} and F_{2z} of the force acting on the unsupported joint (Figure A.1(b)).

Note the sign conventions that we have tacitly adopted:

- The extensions and axial forces are positive if they correspond to tension.

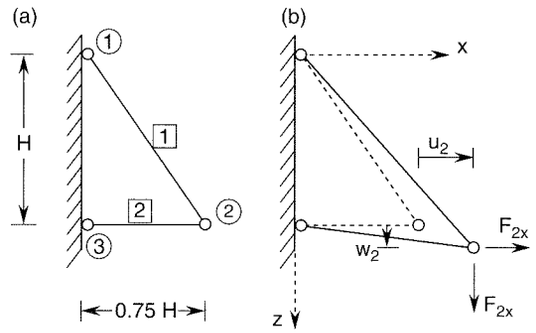


Figure A.1 a) Two-bar truss, b) displacements and external loads

- The components of joint displacements and external forces are positive if they have the same orientation as the coordinate axes, i.e. the horizontal components are positive 'to the right', and the vertical components are positive 'downwards'.

A.2 FUNDAMENTAL EQUATIONS

The static structural models are usually described by three fundamental sets of equations. The kinematic (or geometric) equations link the displacements with the strains (therefore they are sometimes called the strain-displacement equations), the constitutive equations describe the material behavior by a stress-strain law, and the equations of equilibrium relate the external and internal forces (stresses).

For a truss, the *kinematic equations* express the simple fact that each bar extension can be calculated from the displacements of the joints connected by the bar (called the end joints). Although generalization to three dimensions would be straightforward, consider a two-dimensional truss located in the coordinate plane (x, z). Assuming the displacements to remain very small compared to the bar length, we can neglect second-order geometrical effects and express the extension of a bar connecting joints i and j as

$$e_{ij} = (u_j - u_i)\cos\alpha_{ij} + (w_j - w_i)\sin\alpha_{ij} \tag{A.1}$$

where α_{ij} is the angle between the bar axis and the coordinate axis x, measured as shown in Figure A.2(a). The smallness of displacements is the key assumption of the geometrically linear theory, which causes the kinematic equations – and, due to duality, also the equations of equilibrium – to be linear. For the truss in Figure A.1, we have $\cos\alpha_{12} = 0.6$, $\sin\alpha_{12} = 0.8$, $\cos\alpha_{23} = -1$, and $\sin\alpha_{23} = 0$, and taking into account that the displacements of the supported joints are prescribed as zero, we get

$$e_1 \equiv e_{12} = 0.6u_2 + 0.8w_2 \tag{A.2}$$

$$e_2 \equiv e_{23} = u_2 \tag{A.3}$$

In the original sense, the *constitutive equations* describe the behavior of the material at one point of a continuum, i.e. they relate the stresses to the strains. However, since the structural model is discrete, the basic 'building block' is not an infinitely small volume element, but a finite structural element such as a bar. Thus, we need the generalized constitutive equations relating the internal forces (generalized stresses) to the deformation variables (generalized strains). The generalized constitutive equations for linear elastic trusses can be derived by combining the uniaxial Hooke's law

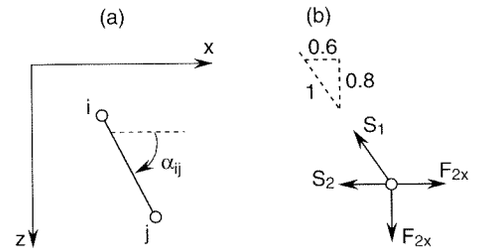


Figure A.2 a) Definition of angle α_{ij}, b) free-body diagram

$\sigma = E\varepsilon$ (that is, the constitutive equation of one-dimensional linear elasticity) with the elementary relations $\sigma = S/A$ and $\varepsilon = e/\ell$, where σ = stress, ε = strain, E = Young's modulus of elasticity, S = axial force, A = cross sectional area (assumed to be constant along the bar), e = bar extension, and ℓ = bar length. The resulting relation between the axial force and the bar extension reads

$$S = \frac{EA}{\ell}e \tag{A.4}$$

The proportionality factor $k = EA/\ell$ is the *axial stiffness* of the bar. When we write the generalized constitutive equation for a specific bar connecting joints i and j, we add subscripts ij:

$$S_{ij} = k_{ij}e_{ij}, \quad \text{where } k_{ij} = \frac{EA_{ij}}{\ell_{ij}} \tag{A.5}$$

Finally, the *equilibrium equations* can be obtained by writing the conditions of equilibrium for the unsupported joints and for the partially supported joints in unsupported directions (the equilibrium conditions at the supported joints are useful for evaluating the reactions but they do not enter the fundamental equations needed to compute the displacements, strains and stresses). For example, the free-body diagram in Figure A.2(b) provides

$$-0.6S_1 - S_2 + F_{2x} = 0 \tag{A.6}$$

$$-0.8S_1 + F_{2z} = 0 \tag{A.7}$$

A.3 MATRIX NOTATION

The theoretical derivations as well as programming are greatly facilitated by introducing matrix notation for the fundamental groups of variables and sets of equations. Let us collect the generalized displacements, strains and stresses into vectors $\boldsymbol{d}$, $\boldsymbol{e}$ and $\boldsymbol{s}$, respectively, and denote by $\boldsymbol{f}$ the vector containing the external forces (by vectors we mean column matrices in this context). The fundamental equations can then be written as follows:

$$\boldsymbol{e} = \boldsymbol{B}\boldsymbol{d} \quad \ldots \quad \text{kinematic equations} \tag{A.8}$$

$$\boldsymbol{s} = \boldsymbol{D}\boldsymbol{e} \quad \ldots \quad \text{constitutive equations} \tag{A.9}$$

$$\boldsymbol{f} = \boldsymbol{B}^T\boldsymbol{s} \quad \ldots \quad \text{equilibrium equations} \tag{A.10}$$

Here, $\boldsymbol{B}$ is the kinematic matrix, $\boldsymbol{D}$ is the generalized material stiffness matrix, and $\boldsymbol{B}^T$ is the static matrix. Whenever the vectors $\boldsymbol{d}$ and $\boldsymbol{f}$ and the vectors $\boldsymbol{e}$ and $\boldsymbol{s}$

are work-conjugate,[1] the static matrix is the transpose of the kinematic matrix. This statement, in general known as the *duality of kinematic and equilibrium equations*, follows from the principle of virtual work. Indeed, $\boldsymbol{f}^T\delta\boldsymbol{d} = \boldsymbol{s}^T\delta\boldsymbol{e}$ must hold for any virtual generalized displacements $\delta\boldsymbol{d}$ and the corresponding virtual generalized strains $\delta\boldsymbol{e} = \boldsymbol{B}\ \delta\boldsymbol{d}$, which implies that $\boldsymbol{f}^T = \boldsymbol{s}^T\boldsymbol{B}$. We will always select the fundamental variables such that they form work-conjugate pairs, which permits us to denote the static matrix immediately as $\boldsymbol{B}^T$ instead of introducing a special symbol.

For our simple example, the fundamental equations in matrix notation read

$$\begin{Bmatrix} e_1 \\ e_2 \end{Bmatrix} = \begin{bmatrix} 0.6 & 0.8 \\ 1 & 0 \end{bmatrix} \begin{Bmatrix} u_3 \\ w_3 \end{Bmatrix} \tag{A.11}$$

$$\begin{Bmatrix} S_1 \\ S_2 \end{Bmatrix} = \begin{bmatrix} EA/1.25H & 0 \\ 0 & EA/0.75H \end{bmatrix} \begin{Bmatrix} e_1 \\ e_2 \end{Bmatrix} \tag{A.12}$$

$$\begin{Bmatrix} F_{2x} \\ F_{2z} \end{Bmatrix} = \begin{bmatrix} 0.6 & 1 \\ 0.8 & 0 \end{bmatrix} \begin{Bmatrix} S_1 \\ S_2 \end{Bmatrix} \tag{A.13}$$

If the external forces are given, we can calculate the axial forces from the equations of equilibrium, then the extensions from the constitutive equations, and finally the displacements from the kinematic equations. However, this is possible only because the structure in our example is statically determinate. For statically indeterminate structures, the static matrix has more columns than rows (the number of unknown axial forces is larger than the number of available equilibrium equations), and so the fundamental sets of equations cannot be solved separately.

A.4 STIFFNESS METHOD

The most common approach to analysis of statically indeterminate structures is the *stiffness method* (also called the *displacement method*), which combines the fundamental sets of equations in the following manner:

$$\boldsymbol{f} = \boldsymbol{B}^T\boldsymbol{s} = \boldsymbol{B}^T\boldsymbol{D}\boldsymbol{e} = \boldsymbol{B}^T\boldsymbol{D}\boldsymbol{B}\boldsymbol{d} \tag{A.14}$$

For properly supported structures, the *stiffness matrix* $\boldsymbol{K} = \boldsymbol{B}^T\boldsymbol{D}\boldsymbol{B}$ is regular, and so the governing equations of the displacement method,

$$\boldsymbol{K}\boldsymbol{d} = \boldsymbol{f} \tag{A.15}$$

have a unique solution for any applied load. After solving for the generalized displacements, the remaining unknown vectors $\boldsymbol{e}$ and $\boldsymbol{s}$ are easily obtained by substituting the solution $\boldsymbol{d}$ into the kinematic and constitutive equations:

$$\boldsymbol{e} = \boldsymbol{B}\boldsymbol{d}, \quad \boldsymbol{s} = \boldsymbol{D}\boldsymbol{e} = \boldsymbol{D}\boldsymbol{B}\boldsymbol{d} \tag{A.16}$$

For our example, the stiffness matrix could be evaluated as

$$\boldsymbol{K} = \boldsymbol{B}^T\boldsymbol{D}\boldsymbol{B} = \begin{bmatrix} 0.6 & 1 \\ 0.8 & 0 \end{bmatrix} \frac{EA}{H} \begin{bmatrix} 0.8 & 0 \\ 0 & 1.333 \end{bmatrix} \begin{bmatrix} 0.6 & 0.8 \\ 1 & 0 \end{bmatrix} = \frac{EA}{H} \begin{bmatrix} 1.621 & 0.384 \\ 0.384 & 0.512 \end{bmatrix} \tag{A.17}$$

[1] Generalized displacements $\boldsymbol{d}$ are said to be work-conjugate to external forces $\boldsymbol{f}$ if the scalar product $\boldsymbol{f}^T\dot{\boldsymbol{d}}$ has the meaning of external power, i.e. if each component of $\boldsymbol{f}$ does work on the corresponding component of $\boldsymbol{d}$. Similarly, $\boldsymbol{e}$ and $\boldsymbol{s}$ are work-conjugate if the scalar product $\boldsymbol{s}^T\dot{\boldsymbol{e}}$ has the meaning of internal power. As usual, dots over symbols denote differentiation with respect to time.

Suppose that the structure is loaded by a vertical force, i.e. $\boldsymbol{f} = \{0\ F\}^T$. Solving the set of equations

$$\frac{EA}{H} \begin{bmatrix} 1.621 & 0.384 \\ 0.384 & 0.512 \end{bmatrix} \begin{Bmatrix} u_2 \\ w_2 \end{Bmatrix} = F \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \tag{A.18}$$

we get the displacements of the unsupported joint

$$\begin{Bmatrix} u_2 \\ w_2 \end{Bmatrix} = \frac{FH}{EA} \begin{Bmatrix} -0.563 \\ 2.375 \end{Bmatrix} \tag{A.19}$$

After a substitution into the kinematic and constitutive equations, the axial forces are finally obtained:

$$\begin{Bmatrix} S_1 \\ S_2 \end{Bmatrix} = \frac{EA}{H} \begin{bmatrix} 0.8 & 0 \\ 0 & 1.333 \end{bmatrix} \begin{bmatrix} 0.6 & 0.8 \\ 1 & 0 \end{bmatrix} \frac{FH}{EA} \begin{Bmatrix} -0.563 \\ 2.375 \end{Bmatrix} = F \begin{Bmatrix} 1.25 \\ 0.75 \end{Bmatrix} \tag{A.20}$$

The reader can verify that these forces satisfy the equations of equilibrium.

A.5 ASSEMBLY OF STIFFNESS MATRIX

For large structures it would be excessively laborious to construct the stiffness matrix by directly evaluating the product $\boldsymbol{B}^T\boldsymbol{D}\boldsymbol{B}$ of three large matrices. A much more efficient procedure is the assembly of the stiffness matrix from the contributions of individual bars.

A truss structure can be decomposed by cuts at the end sections of individual bars into a collection of joints and bars (Figure A.3). Interactions among the components are represented by the so-called member *(bar) end forces*, which are considered as positive when acting *on the bars* in the positive directions. The positive directions are defined by the orientation of the global coordinate axes. With the coordinate system shown in Figure A.1(b), the horizontal components are positive to the right, and the vertical components are positive downwards. The equations of joint equilibrium can formally be written as

$$\sum_j X_{ij} = F_{ix} \tag{A.21}$$

$$\sum_j Z_{ij} = F_{iz} \tag{A.22}$$

where the left-hand side represents the sum of the horizontal or vertical components of the end forces acting on joint number i, and the right-hand side is the horizontal or vertical component of the prescribed external force acting on that joint. The sum

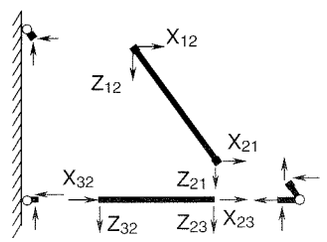


Figure A.3 Truss decomposed into bars and joints

on the left-hand side is taken over all the bars connected to joint i, each of which is distinguished by the number of the opposite joint, j.

To obtain the governing equations of the stiffness approach (A.14), the end forces must be expressed in terms of the joint displacements. This can be achieved by combining the three fundamental sets of equations written for a single bar, i.e. the equilibrium equations

$$X_{ij} = -S_{ij}\cos\alpha_{ij} \tag{A.23}$$
$$Z_{ij} = -S_{ij}\sin\alpha_{ij} \tag{A.24}$$
$$X_{ji} = S_{ij}\cos\alpha_{ij} \tag{A.25}$$
$$Z_{ji} = S_{ij}\sin\alpha_{ij} \tag{A.26}$$

the constitutive equation

$$S_{ij} = \frac{EA_{ij}}{\ell_{ij}} e_{ij} \tag{A.27}$$

and the kinematic equation

$$e_{ij} = (u_j - u_i)\cos\alpha_{ij} + (w_j - w_i)\sin\alpha_{ij} \tag{A.28}$$

The result in matrix form reads

$$\begin{Bmatrix} X_{ij} \\ Z_{ij} \\ X_{ji} \\ Z_{ji} \end{Bmatrix} = \frac{EA_{ij}}{\ell_{ij}} \begin{bmatrix} c_{ij}^2 & c_{ij}s_{ij} & -c_{ij}^2 & -c_{ij}s_{ij} \\ c_{ij}s_{ij} & s_{ij}^2 & -c_{ij}s_{ij} & -s_{ij}^2 \\ -c_{ij}^2 & -c_{ij}s_{ij} & c_{ij}^2 & c_{ij}s_{ij} \\ -c_{ij}s_{ij} & -s_{ij}^2 & c_{ij}s_{ij} & s_{ij}^2 \end{bmatrix} \begin{Bmatrix} u_i \\ w_i \\ u_j \\ w_j \end{Bmatrix} \tag{A.29}$$

where $c_{ij} = \cos\alpha_{ij}$ and $s_{ij} = \sin\alpha_{ij}$. The 4×4 matrix

$$\boldsymbol{K}_{ij} = \frac{EA_{ij}}{\ell_{ij}} \begin{bmatrix} c^2 & cs & -c^2 & -cs \\ cs & s^2 & -cs & -s^2 \\ -c^2 & -cs & c^2 & cs \\ -cs & -s^2 & cs & s^2 \end{bmatrix}_{ij} \tag{A.30}$$

is the *member axial stiffness matrix* of bar ij. In (A.30) we have simplified the notation by omitting the subscripts ij at c and s and mentioning them only once, after the right matrix bracket.

The previous relations between the end forces and joint displacements for a single bar make it possible to construct the stiffness matrix of the entire structure without multiplying the potentially large matrices in $\boldsymbol{K} = \boldsymbol{B}^T\boldsymbol{D}\boldsymbol{B}$. Note that each equation of the stiffness method is one of the equations of joint equilibrium expressed in terms of joint displacements. For example, for our simple truss we can proceed as follows:

$$\cos\alpha_{12} = 0.6,\ \sin\alpha_{12} = 0.8,\ EA_{12}/\ell_{12} = EA/1.25H = 0.8EA/H \tag{A.31}$$

$$\boldsymbol{K}_{12} = \frac{EA}{H} \begin{bmatrix} 0.288 & 0.384 & -0.288 & -0.384 \\ 0.384 & 0.512 & -0.384 & -0.512 \\ -0.288 & -0.384 & 0.288 & 0.384 \\ -0.384 & -0.512 & 0.384 & 0.512 \end{bmatrix} \tag{A.32}$$

$$\cos\alpha_{23} = -1,\ \sin\alpha_{23} = 0,\ EA_{23}/\ell_{23} = EA/0.75H = 1.333EA/H \tag{A.33}$$

$$\boldsymbol{K}_{23} = \frac{EA}{H} \begin{bmatrix} 1.333 & 0 & -1.333 & 0 \\ 0 & 0 & 0 & 0 \\ -1.333 & 0 & 1.333 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{A.34}$$

The rows of the member stiffness matrices contain the coefficients that appear in the expressions for member end forces in terms of joint displacements; cf. (A.29). They are useful for assembling the joint equilibrium equations. We start with the equilibrium conditions for the horizontal forces acting on joint 2, $X_{21} + X_{23} = F_{2x}$. Substituting

$$X_{21} = \frac{EA}{H}(-0.288u_1 - 0.384w_1 + 0.288u_2 + 0.384w_2) \tag{A.35}$$
$$X_{23} = \frac{EA}{H}(1.333u_2 + 0 \cdot w_2 - 1.333u_3 + 0 \cdot w_3) \tag{A.36}$$

we obtain

$$\frac{EA}{H}(0.288 + 1.333)u_2 + \frac{EA}{H}(0.384 + 0)w_2 = F_{2x} \tag{A.37}$$

Note that we have substituted the prescribed joint displacements $u_1 = 0$, $w_1 = 0$, $u_3 = 0$ and $w_3 = 0$. In a similar fashion, we set up the equation expressing equilibrium of vertical forces acting on joint 2, $Z_{21} + Z_{23} = F_{2z}$. Substituting

$$Z_{21} = \frac{EA}{H}(-0.384u_1 - 0.512w_1 + 0.384u_2 + 0.512w_2) \tag{A.38}$$
$$Z_{23} = \frac{EA}{H}(0 \cdot u_2 + 0 \cdot w_2 + 0 \cdot u_3 + 0 \cdot w_3) \tag{A.39}$$

we obtain

$$\frac{EA}{H}(0.384 + 0)u_2 + \frac{EA}{H}(0.512 + 0)w_2 = F_{2z} \tag{A.40}$$

Equations (A.37) and (A.40) are the basic equations of the displacement method, and the coefficients at the unknown generalized displacements u_2 and w_2 represent entries of the global stiffness matrix, cf. (A.17).

In practical calculations, it is more efficient to proceed in a loop over all the bars, and for each bar evaluate its contribution to the four equations of equilibrium that are associated with its end joints. It is very useful to introduce the notion of *dof numbers* (degree-of-freedom numbers), which are obtained by numbering the unknown joint displacements consecutively. Each degree of freedom (i.e. unknown displacement component) has an equilibrium equation associated with it, and so the dof numbers give at the same time the ordering of governing equations of the displacement method. The displacements of the supported joints are not among the unknowns (and the corresponding equations of equilibrium are not needed to compute the unknown displacements); accordingly, they will be assigned zero dof numbers as described by the following algorithm.

Algorithm A.1: Assigning the dof numbers

1. Set NUMBER_OF_UNKNOWNS = 0.
2. Loop over joints.
   For each joint:

(a) If the joint is supported in the horizontal direction then set its first dof number to zero, else increase NUMBER_OF_UNKNOWNS by one and set the first dof number to NUMBER_OF_UNKNOWNS.
(b) If the joint is supported in the vertical direction then set its second dof number to zero, else increase NUMBER_OF_UNKNOWNS by one and set the second dof number to NUMBER_OF_UNKNOWNS.

3. End of loop over joints.
4. The final value of NUMBER_OF_UNKNOWNS gives the total number of unknown displacement components, which is equal to the number of rows and columns of the stiffness matrix.

Once the dof numbers have been assigned to the horizontal and vertical displacement component at each joint, the contribution of an arbitrary bar to the stiffness matrix can be evaluated. Each column of the member stiffness matrix is associated with a certain joint displacement, which yields the column dof number, and each row of the member stiffness matrix is associated with an end force entering a certain equilibrium equation, which yields the row dof number. Now, each stiffness coefficient found in the member stiffness matrix will appear on the global (structural) level in the equilibrium equation given by its row dof number, at the unknown given by its column dof number. If at least one of the dof numbers is zero, the coefficient does not contribute to the structure stiffness matrix. This leads to the following assembly procedure:

Algorithm A.2: Assembling the stiffness matrix

1. Initialize the stiffness matrix, i.e. set all its entries to zero.
2. Loop over bars.
   For each bar:
   (a) Evaluate the member stiffness matrix (A.30).
   (b) Assign dof numbers to the columns (and rows) of the member stiffness matrix: Dof numbers of the 'left joint', i, are assigned to columns (and rows) 1 and 2, dof numbers of the 'right joint', j, are assigned to columns (and rows) 3 and 4.
   (c) Loop over rows of the member stiffness matrix.
   (d) If the row dof number is zero, go to next row.
   (e) Loop over columns of the member stiffness matrix.

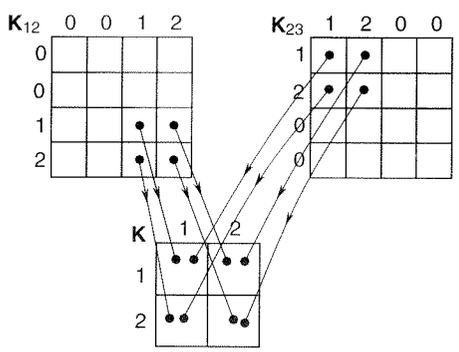


Figure A.4 Assembly procedure

Table A.1 Assembly procedure

$\boldsymbol{K}_{12}$	0	0	1	2
0	.	.	.	.
0	.	.	.	.
1	.	.	0.288	0.384
2	.	.	0.384	0.512

$\boldsymbol{K}_{23}$	1	2	0	0
1	1.333	0	.	.
2	0	0	.	.
0	.	.	.	.
0	.	.	.	.

$\boldsymbol{K}$	1	2
1	0.288+1.333	0.384+0
2	0.384+0	0.512+0

(f) If the column dof number is not zero, add the corresponding stiffness coefficient to the entry of the structure stiffness matrix in the row and column given by the respective dof numbers.
(g) End of loop over columns.
(h) End of loop over rows.

3. End of loop over bars.

For our simple example, the dof numbers of joints 1, 2, and 3 are (0,0), (1,2), and (0,0), respectively, and thus the dof numbers of bars 1-2 and 2-3 are (0,0,1,2) and (1,2,0,0), respectively. Figure A.4 illustrates the assembly procedure. The numerical values of the stiffness coefficients are given in Table A.1; cf. (A.32) and (A.34). The common multiplier EA/H has been left out. The dots denote the entries that are not needed by the assembly procedure. Note that if we changed the order of the end joints of a bar, the order of the bar dof numbers and the signs of c_{ij} and s_{ij} would change but the contribution to the structure stiffness matrix would remain the same.

Appendix B

Linear Elastic Beams and Frames

In contrast to trusses consisting of pin-ended bars, the beams and frames contain members that are connected to joints in a rigid manner. A rigid connection forces the rotations of the joint and the attached end cross section to be the same. Of course, not all of the connections have to be rigid – beams and frames sometimes contain internal hinges. For the sake of simplicity, let us assume that the loads are applied only at the joints. The generalization to loads acting on the members is not difficult – the state of the frame within the beam length is equal to the state where the load is replaced by the end reactions at the joints, plus the state of beam acted upon by the load and the reactions. However, the basic ideas will be more transparent if we omit the terms corresponding to this type of loading. We also neglect the effects of support settlement, temperature variation and fabrication error. To keep this overview relatively simple, we restrict our attention to planar frames with prismatic straight members.

B.1 FUNDAMENTAL VARIABLES

In addition to joint displacements u_i and w_i, the fundamental variables describing the *displacement* of a beam or frame must also include the joint rotations, ϕ_i. The quantities work-conjugate with the generalized displacements u_i, w_i and ϕ_i are the horizontal and vertical components of the external force applied at joint i, F_{ix} and F_{iz}, and the external moment applied at that joint, M_i, respectively.

Consider a generic member of a planar frame connecting joints i and j, and further referred to as member ij. The local coordinate system of member ij has its origin at joint i, and its axis x' passes through joint j (Figure B.1(a)). The member can stretch and bend, and so its *deformation* is described by the extension, e_{ij}, and by two bending angles, τ_{ij} and τ_{ji}, representing the angles of rotation of the end cross sections with respect to the line perpendicular to the rotated chord (straight line connecting the centroids of the end cross sections – see Figure B.1(b)). As for trusses, the generalized *stress* that is work-conjugate to the extension is the axial force, S_{ij}. In frame analysis it is customary to call it the normal force, N_{ij}. The variables work-conjugate to the bending angles τ_{ij} and τ_{ji} are the moments M_{ij} and M_{ji} acting on the member at its ends (Figure B.2(a)). The order of subscripts i and j is irrelevant for the normal force ($N_{ij} \equiv N_{ji}$) but it is important for the end moments (M_{ji} is the moment acting on member ij from joint j). Note that the end moments are considered to be positive when acting on the member counterclockwise. This sign convention is objective in the sense that it is independent of the chosen 'orientation' of the member, i.e. it does not matter which joint has been considered as the 'left' one, i, and which

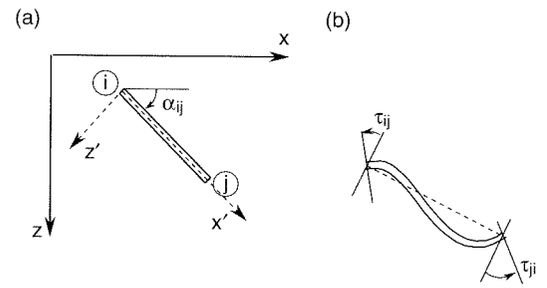


Figure B.1 a) Global and local coordinate system, b) bending angles

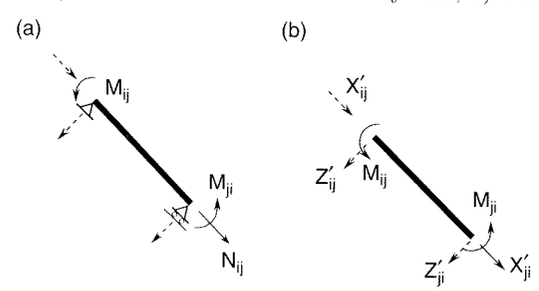


Figure B.2 a) Generalized stresses, b) end forces and moments

as the 'right' one, j. The sign of the normal force is objective as well (a tensile force is positive, a compressive force is negative).

When we isolate a member by fictitious cuts at its end sections, the action of the surrounding structure on the member is represented by the end forces and moments (Figure B.2(b)). The end moments are equal to the generalized stresses M_{ij} and M_{ji} already introduced, and the longitudinal end force at the right end, X'_{ji}, is identical to the normal force, N_{ij}. The remaining end forces, X'_{ij}, Z'_{ij}, and Z'_{ji}, are uniquely determined by the equilibrium conditions of the member. Note that the transversal end force Z'_{ji} corresponds to the shear force. The primes at X and Z remind us that the components of the end forces are expressed with respect to the local coordinate system. For planar frames, the values of end moments M_{ij} and M_{ji} are not affected by the transformation between the local and global coordinate system, and so we do not use primes at M.

The *loading* applied on the structure consists of concentrated forces F_{ix} and F_{iz} and moments M_i, acting on the joints.

The fundamental variables are collected into vectors $\boldsymbol{d}$, $\boldsymbol{e}$, $\boldsymbol{s}$ and $\boldsymbol{f}$, which are generalizations of the vectors introduced in Appendix A (see Table B.1). It is important to bear in mind that the components must be ordered in a systematic

Table B.1 Fundamental variables

vector	truss	frame	beam (bending)
$\boldsymbol{d}$	u_i, w_i	u_i, w_i, ϕ_i	w_i, ϕ_i
$\boldsymbol{e}$	e_{ij}	$e_{ij}, \tau_{ij}, \tau_{ji}$	τ_{ij}, τ_{ji}
$\boldsymbol{s}$	$S_{ij} \equiv N_{ij}$	N_{ij}, M_{ij}, M_{ji}	M_{ij}, M_{ji}
$\boldsymbol{f}$	F_{ix}, F_{iz}	F_{ix}, F_{iz}, M_i	F_{iz}, M_i

way to make sure that $(\boldsymbol{d}, \boldsymbol{f})$ and $(\boldsymbol{e}, \boldsymbol{s})$ are work-conjugate pairs, i.e. that $\boldsymbol{f}^T\dot{\boldsymbol{d}}$ is the external power and $\boldsymbol{s}^T\dot{\boldsymbol{e}}$ is the internal power. The components of $\boldsymbol{d}$ and $\boldsymbol{f}$ are usually ordered according to the joint numbers, and the components of $\boldsymbol{e}$ and $\boldsymbol{s}$ according to the member numbers.

Vectors $\boldsymbol{d}$ and $\boldsymbol{f}$ contain only the components corresponding to the unsupported directions. This means that the prescribed displacements and unknown reaction forces in the supports are not included in these vectors.

For straight beams satisfying certain simple conditions on the type and location of supports, any general state can be decomposed into bending (accompanied by shear) and axial stretching. The corresponding problems are completely uncoupled, and so the fundamental variables can be divided into two groups: u_i, e_{ij}, N_{ij} and F_{ix} enter the equations that describe axial stretching while w_i, ϕ_i, τ_{ij}, M_{ij}, F_{iz} and M_i enter the equations that describe bending; see Table B.1.

We will take advantage of this decomposition and first develop the fundamental equations related to beam bending. A general description of a frame will be obtained later as a combination of the equations for trusses and beams (rotated into a general position with respect to the global coordinate system).

Example B.1: Set up the vectors of fundamental variables for the beam in Figure B.3(a).

Solution: For the purpose of analysis we introduce computational joints at the supports and at the points of application of external loads. Such artificial joints are often referred to as nodes of the structural model. Fictitious cuts at the nodes decompose the structure into a collection of members, in a general context referred to as elements of the model. In our case, we introduce four nodes, and the beam is then divided into three elements; see Figure B.3(b).

The vector of unknown generalized displacements,

$$\boldsymbol{d} = \{w_2, \phi_2, \phi_3, \phi_4\}^T \tag{B.1}$$

collects the free nodal displacements and rotations (i.e. those not restrained by the supports). The vector of generalized strains,

$$\boldsymbol{e} = \{\tau_{12}, \tau_{21}, \tau_{23}, \tau_{32}, \tau_{34}, \tau_{43}\}^T \tag{B.2}$$

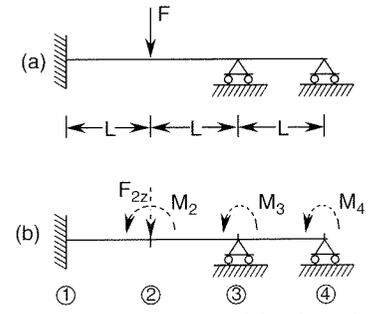


Figure B.3 Continuous beam: a) geometry and loading, b) nodes, external forces and moments

and the vector of generalized stresses (internal forces),

$$\boldsymbol{s} = \{M_{12}, M_{21}, M_{23}, M_{32}, M_{34}, M_{43}\}^T \tag{B.3}$$

each contain two variables per member. Finally, the vector of external forces,

$$\boldsymbol{f} = \{F_{2z}, M_2, M_3, M_4\}^T \tag{B.4}$$

is composed of the forces and moments that are work-conjugate to the generalized displacements. In systematic notation, this vector consists of all the external forces and moments that could be applied at the nodes. In the specific loading case in Figure B.3(a) we have $\boldsymbol{f} = \{F, 0, 0, 0\}^T$.

Of course, it is not necessary to introduce a node at every point of application of an external force or moment. We could solve the problem with a smaller number of unknowns by leaving out node 2 and transforming the force F into equivalent forces and moments applied at nodes 1 and 3. Such an approach is indeed routinely used in linear elastic analysis. However, in incremental plastic analysis we must consider the points of application of concentrated loads as potential plastic hinges, and so it is algorithmically convenient to treat them as nodes already from the first stage of analysis, in which the structure responds in an elastic manner. This has the advantage that, at later stages in which plastic hinges form at some of these nodes, we do not have to change the topological description of the structure (connectivity of nodes and elements).

B.2 FUNDAMENTAL EQUATIONS

B.2.1 Kinematic Equations

Kinematic equations express the generalized strains in terms of the generalized displacements. They follow from simple geometric considerations. It is clear from Figure B.4(a) that

$$\tau_{ij} = \phi_i - \psi_{ij} \tag{B.5}$$

$$\tau_{ji} = \phi_j - \psi_{ij} \tag{B.6}$$

where ψ_{ij} is the rotation of the member chord. In general, ψ_{ij} would depend on the joint displacements in a nonlinear manner. According to Figure B.4(b), the chord rotation satisfies the relation

$$\sin\psi_{ij} = \frac{w_i - w_j}{\ell_{ij} + e_{ij}} \tag{B.7}$$

In the geometrically linear theory based on the assumption of small displacements, we can replace $\sin\psi_{ij}$ by ψ_{ij} and neglect e_{ij} with respect to ℓ_{ij}, which leads to a linearized expression

$$\psi_{ij} = \frac{w_i - w_j}{\ell_{ij}} \tag{B.8}$$

Substituting (B.8) into (B.5)–(B.6), we get the desired kinematic relations

$$\tau_{ij} = \phi_i + \frac{w_j - w_i}{\ell_{ij}} \tag{B.9}$$

$$\tau_{ji} = \phi_j + \frac{w_j - w_i}{\ell_{ij}} \tag{B.10}$$

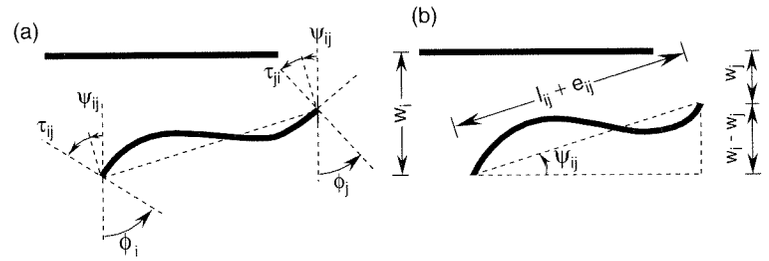


Figure B.4 Kinematic quantities

Example B.2: Set up the kinematic equations for the beam in Figure B.3(a).

Solution: For member 1-2 we have

$$\tau_{12} = \phi_1 + \frac{w_2 - w_1}{\ell_{12}} = \frac{w_2}{L} \tag{B.11}$$

$$\tau_{21} = \phi_2 + \frac{w_2 - w_1}{\ell_{12}} = \phi_2 + \frac{w_2}{L} \tag{B.12}$$

because the displacement w_1 and the rotation ϕ_1 are prescribed as zero. Analogous substitutions lead to the kinematic equations for members 2-3 and 3-4. The final result can be presented in the matrix form as

$$\boldsymbol{e} = \boldsymbol{B}\boldsymbol{d} \tag{B.13}$$

where the components of the vector of generalized strains $\boldsymbol{e}$ and of the vector of generalized displacements $\boldsymbol{d}$ were listed in Example B.1, and

$$\boldsymbol{B} = \begin{bmatrix} 1/L & 0 & 0 & 0 \\ 1/L & 1 & 0 & 0 \\ -1/L & 1 & 0 & 0 \\ -1/L & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{B.14}$$

is the kinematic matrix of the structure.

B.2.2 Constitutive Equations

Generalized constitutive equations are the relations between the generalized stresses and the generalized strains. They can be derived from the elastic properties of the material and from the geometry of the beam member (the length and sectional characteristics). For the simplest beam theory, based on the *Bernoulli–Navier hypothesis* (cross sections remain planar and normal to the deformed centerline), elementary analysis of a simply supported beam of constant cross section provides the following expressions for the bending angles in terms of the end moments:

$$\begin{Bmatrix} \tau_{ij} \\ \tau_{ji} \end{Bmatrix} = \frac{\ell_{ij}}{6EI_{ij}} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{Bmatrix} M_{ij} \\ M_{ji} \end{Bmatrix} \tag{B.15}$$

where I_{ij} is the moment of inertia of the cross section of member ij with respect to the centroidal axis perpendicular to the plane of bending. By inversion we get

$$\begin{Bmatrix} M_{ij} \\ M_{ji} \end{Bmatrix} = \frac{2EI_{ij}}{\ell_{ij}} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{Bmatrix} \tau_{ij} \\ \tau_{ji} \end{Bmatrix} \tag{B.16}$$

This is the generalized constitutive equation for bending (with the effect of shear distortion neglected).

Later on we will see that it is useful to combine the kinematic and constitutive equations, and derive expressions for the end moments in terms of the joint displacements and rotations. Substituting (B.5)–(B.6) into (B.16), we get

$$M_{ij} = \frac{2EI_{ij}}{\ell_{ij}} (2\phi_i + \phi_j - 3\psi_{ij}) \tag{B.17}$$

$$M_{ji} = \frac{2EI_{ij}}{\ell_{ij}} (\phi_i + 2\phi_j - 3\psi_{ij}) \tag{B.18}$$

The chord rotation ψ_{ij} is only an auxiliary variable that could be eliminated from the above formulae by substitution from (B.8). However, we prefer to keep the formulae in their present form because (B.17)–(B.18) will remain valid for the frame members, and only the expression for the chord rotation will have to be generalized.

B.2.3 Equilibrium Equations

Equilibrium equations are set up for the unsupported joints (and the partially supported joints in free directions). They can easily be constructed from the free-body diagrams for individual joints. A systematic approach is based on the auxiliary notion of vertical end forces, Z_{ij} and Z_{ji} (Figure B.5; see also Figure B.2(b) and note that, in the present case, the local and global coordinate systems coincide). These forces in principle correspond to the shear force, but they are considered as positive when acting in the direction of the positive z-axis, i.e. downwards. The vertical end forces can be expressed in terms of two independent generalized stresses – the end moments. From the equilibrium conditions for the isolated member we get

$$Z_{ij} = -\frac{M_{ij} + M_{ji}}{\ell_{ij}} \tag{B.19}$$

$$Z_{ji} = \frac{M_{ij} + M_{ji}}{\ell_{ij}} \tag{B.20}$$

With the previous relations at hand, it is easy to set up the equilibrium equations at an arbitrary joint i,

$$\sum_j Z_{ij} = F_{iz} \tag{B.21}$$

Figure B.5 End moments and vertical end forces on a beam member

$$\sum_j M_{ij} = M_i \tag{B.22}$$

Here the sums are taken over all the members framing into joint i. The final equilibrium equations are obtained by expressing the vertical end forces in terms of the end moments according to (B.19)–(B.20).

Example B.3: Set up the equilibrium equations for the beam in Figure B.3(a).

Solution: The equilibrium conditions for the unsupported joints read

$$\begin{aligned} Z_{21} + Z_{23} &= F_{2z} \\ M_{21} + M_{23} &= M_2 \\ M_{32} + M_{34} &= M_3 \\ M_{43} &= M_4 \end{aligned}$$

Substituting

$$Z_{21} = \frac{M_{12} + M_{21}}{L} \tag{B.23}$$

$$Z_{23} = -\frac{M_{23} + M_{32}}{L} \tag{B.24}$$

we obtain the desired equations

$$\begin{bmatrix} 1/L & 1/L & -1/L & -1/L & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} M_{12} \\ M_{21} \\ M_{23} \\ M_{32} \\ M_{34} \\ M_{43} \end{Bmatrix} = \begin{Bmatrix} F_{2z} \\ M_2 \\ M_3 \\ M_4 \end{Bmatrix} \tag{B.25}$$

or

$$\boldsymbol{B}^T \boldsymbol{s} = \boldsymbol{f} \tag{B.26}$$

where, as implied by duality, the static matrix $\boldsymbol{B}^T$ is the transpose of the kinematic matrix $\boldsymbol{B}$ constructed in Example B.2; cf. (B.19).

B.3 STIFFNESS MATRIX

The general form of the governing equations of the stiffness (displacement) method,

$$\boldsymbol{K}\boldsymbol{d} = \boldsymbol{f} \tag{B.27}$$

is exactly the same for trusses, beams and frames, and in fact for any other discrete structural model. The stiffness matrix of the structure is defined as $\boldsymbol{K} = \boldsymbol{B}^T\boldsymbol{D}\boldsymbol{B}$, where $\boldsymbol{B}^T$ is the static matrix, $\boldsymbol{D}$ is the generalized material stiffness matrix, and $\boldsymbol{B}$ is the kinematic matrix. As already explained in Appendix A, the stiffness matrix can be constructed efficiently by assembling the contributions of individual members, which are extracted from the member stiffness matrices. To derive the general formula for the stiffness matrix of a beam member, we rewrite the fundamental equations for

an individual member in the matrix form. From (B.9)–(B.10), we get the kinematic equations

$$\begin{Bmatrix} \tau_{ij} \\ \tau_{ji} \end{Bmatrix} = \begin{bmatrix} -1/\ell & 1 & 1/\ell & 0 \\ -1/\ell & 0 & 1/\ell & 1 \end{bmatrix}_{ij} \begin{Bmatrix} w_i \\ \phi_i \\ w_j \\ \phi_j \end{Bmatrix} \tag{B.28}$$

In compact notation, (B.28) can be written as

$$\boldsymbol{e}_{ij} = \boldsymbol{B}_{ij}\boldsymbol{d}_{ij} \tag{B.29}$$

where $\boldsymbol{e}_{ij} = \{\tau_{ij}, \tau_{ji}\}^T$ is the vector of member generalized strains, $\boldsymbol{B}_{ij}$ is the member kinematic matrix, and $\boldsymbol{d}_{ij} = \{w_i, \phi_i, w_j, \phi_j\}^T$ is the vector of member generalized displacements (end displacements and rotations).

For an isolated member, the end forces and moments play the role of applied loads. Using expressions (B.19)–(B.20) and realizing that the end moments represent the generalized stresses, we can construct the equilibrium equations of a beam member,

$$\begin{Bmatrix} Z_{ij} \\ M_{ij} \\ Z_{ji} \\ M_{ji} \end{Bmatrix} = \begin{bmatrix} -1/\ell & -1/\ell \\ 1 & 0 \\ 1/\ell & 1/\ell \\ 0 & 1 \end{bmatrix}_{ij} \begin{Bmatrix} M_{ij} \\ M_{ji} \end{Bmatrix} \tag{B.30}$$

The corresponding member static matrix is recognized as the transpose of the member kinematic matrix $\boldsymbol{B}_{ij}$ from (B.28), and so we can write the static equations as

$$\boldsymbol{f}_{ij} = \boldsymbol{B}_{ij}^T \boldsymbol{s}_{ij} \tag{B.31}$$

where $\boldsymbol{f}_{ij} = \{Z_{ij}, M_{ij}, Z_{ji}, M_{ji}\}^T$ is the vector of generalized member end forces, and $\boldsymbol{s}_{ij} = \{M_{ij}, M_{ji}\}^T$ is the vector of member generalized stresses (end moments).

Finally, the generalized constitutive equations (B.16) can be written in compact notation as

$$\boldsymbol{s}_{ij} = \boldsymbol{D}_{ij}\boldsymbol{e}_{ij} \tag{B.32}$$

where

$$\boldsymbol{D}_{ij} = \frac{2EI_{ij}}{\ell_{ij}} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{B.33}$$

is the member material stiffness matrix.

Now it is easy to combine the fundamental equations for a beam member in the same manner as it has been done before for the entire structure. The resulting equations

$$\boldsymbol{K}_{ij}\boldsymbol{d}_{ij} = \boldsymbol{f}_{ij} \tag{B.34}$$

relate the end forces and moments to the end displacements and rotations. The matrix

$$\boldsymbol{K}_{ij} = \boldsymbol{B}_{ij}^T \boldsymbol{D}_{ij} \boldsymbol{B}_{ij} = \frac{EI_{ij}}{\ell_{ij}} \begin{bmatrix} 12 & -6\ell & -12 & -6\ell \\ -6\ell & 4\ell^2 & 6\ell & 2\ell^2 \\ -12 & 6\ell & 12 & 6\ell \\ -6\ell & 2\ell^2 & 6\ell & 4\ell^2 \end{bmatrix}_{ij} \tag{B.35}$$

is the *beam member stiffness matrix.*

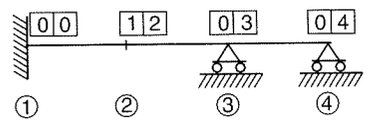


Figure B.6 Dof numbers assigned to joints of a continuous beam

Table B.2 Assembly of stiffness matrix

$\boldsymbol{K}_{12}$	0	0	1	2
0	.	.	.	.
0	.	.	.	.
1	.	.	12	6L
2	.	.	6L	$4L^2$

$\boldsymbol{K}_{23}$	1	2	0	3
1	12	−6L	.	−6L
2	−6L	$4L^2$	.	$2L^2$
0	.	.	.	.
3	−6L	$2L^2$	.	$4L^2$

$\boldsymbol{K}_{34}$	0	3	0	4
0	.	.	.	.
3	.	$4L^2$	.	$2L^2$
0	.	.	.	.
4	.	$2L^2$	.	$4L^2$

$\boldsymbol{K}$	1	2	3	4
1	12+12	6L−6L	−6L	0
2	6L−6L	$4L^2$+$4L^2$	$2L^2$	0
3	−6L	$2L^2$	$4L^2$+$4L^2$	$2L^2$
4	0	0	$2L^2$	$4L^2$

Example B.4: Set up the governing equations of the stiffness method for the beam in Figure B.3(a).

Solution: The algorithms given in Appendix A are applicable to beam analysis after a few simple modifications. Each node still has two dof numbers but the first dof number now corresponds to the vertical displacement, w_i, and the second dof number corresponds to the rotation, ϕ_i. In our case, we assign the dof numbers as shown in Figure B.6. For each member we evaluate the beam stiffness matrix according to formula (B.35), and then add its components to appropriate entries of the global stiffness matrix, which incidentally has four rows and four columns, same as the member stiffness matrices. The assembly procedure is recorded in Table B.2. In all matrices, the common factor EI/L^3 has been left out.

The resulting equations describing equilibrium of the structure in terms of the generalized displacements read

$$\frac{EI}{L^3} \begin{bmatrix} 24 & 0 & -6L & 0 \\ 0 & 8L^2 & 2L^2 & 0 \\ -6L & 2L^2 & 8L^2 & 2L^2 \\ 0 & 0 & 2L^2 & 4L^2 \end{bmatrix} \begin{Bmatrix} w_2 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{Bmatrix} = \begin{Bmatrix} F \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{B.36}$$

B.4 MEMBERS CONNECTED BY HINGES

So far we have tacitly assumed that the beam does not contain any internal hinge. If a beam member is connected to one of its end joints by a hinge, the generalized constitutive equations have to be modified. For a member connected by a hinge at its

right end, we have $M_{ji} = 0$, and (B.15) simplifies to

$$\tau_{ij} = \frac{\ell_{ij}}{3EI_{ij}} M_{ij} \tag{B.37}$$

$$\tau_{ji} = -\frac{\ell_{ij}}{6EI_{ij}} M_{ij} \tag{B.38}$$

In this case, the nonvanishing end moment M_{ij} can be expressed in terms of a single generalized strain as

$$M_{ij} = \frac{3EI_{ij}}{\ell_{ij}} \tau_{ij} \tag{B.39}$$

Combining this with the kinematic equation (B.5) we obtain

$$M_{ij} = \frac{3EI_{ij}}{\ell_{ij}} (\phi_i - \psi_{ij}) \tag{B.40}$$

which is a formula analogous to (B.17) but valid for a beam member connected by a hinge at its right end.

Consequently, neither M_{ji} nor τ_{ji} need to be considered as fundamental variables. M_{ji} is *a priori* known to be zero, and $\tau_{ji} = -\tau_{ij}/2$ as follows from (B.37)–(B.38).

To handle members with hinge-connected and rigidly connected ends by the same algorithm, we will write the generalized constitutive equations in formal analogy with (B.16) as

$$\begin{Bmatrix} M_{ij} \\ M_{ji} \end{Bmatrix} = \frac{3EI_{ij}}{\ell_{ij}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} \tau_{ij} \\ \tau_{ji} \end{Bmatrix} \tag{B.41}$$

Substituting

$$\boldsymbol{D}_{ij}^{rh} = \frac{3EI_{ij}}{\ell_{ij}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{B.42}$$

into the formula for the member stiffness matrix, we obtain

$$\boldsymbol{K}_{ij}^{rh} = \boldsymbol{B}_{ij}^T \boldsymbol{D}_{ij}^{rh} \boldsymbol{B}_{ij} = \frac{3EI_{ij}}{\ell_{ij}^3} \begin{bmatrix} 1 & -\ell & -1 & 0 \\ -\ell & \ell^2 & \ell & 0 \\ -1 & \ell & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}_{ij} \tag{B.43}$$

Superscript ^{rh} means that the left end section of the beam is connected in a rigid manner while the right end section is connected by a hinge.

By analogy, for a member connected by a hinge at its left end we have $M_{ij} = 0$ and

$$M_{ji} = \frac{3EI_{ij}}{\ell_{ij}} \tau_{ji} \tag{B.44}$$

Combining this with the kinematic equation (B.6), we obtain

$$M_{ji} = \frac{3EI_{ij}}{\ell_{ij}} (\phi_j - \psi_{ij}) \tag{B.45}$$

which is a formula analogous to (B.18) but valid for a beam member connected by a hinge at its left end. Substituting

$$\boldsymbol{D}_{ij}^{hr} = \frac{3EI_{ij}}{\ell_{ij}} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{B.46}$$

into the formula for the member stiffness matrix, we obtain

$$\boldsymbol{K}_{ij}^{hr} = \boldsymbol{B}_{ij}^T \boldsymbol{D}_{ij}^{hr} \boldsymbol{B}_{ij} = \frac{3EI_{ij}}{\ell_{ij}^3} \begin{bmatrix} 1 & 0 & -1 & -\ell \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & \ell \\ -\ell & 0 & \ell & \ell^2 \end{bmatrix}_{ij} \tag{B.47}$$

The rotation of a hinge joint has to be excluded from the list of fundamental variables, and the corresponding moment equation of equilibrium is not to be included among the equilibrium equations.

Example B.5: Solve the beam in Figure B.7(a) by the stiffness method.

Solution: The only unknown generalized displacement is the deflection of the internal hinge, w_2. We assign the dof numbers as shown in Figure B.7(b). The global stiffness matrix contains only one element,

$$k = \frac{3EI}{L^3} + \frac{3EI}{(L/2)^3} = \frac{27EI}{L^3} \tag{B.48}$$

which is obtained as the sum of entry (3,3) from $\boldsymbol{K}_{12}^{rh}$ and entry (1,1) from $\boldsymbol{K}_{23}^{hr}$. This stiffness coefficient has the physical meaning of the force that must be applied at joint 2 in order to enforce a unit deflection w_2.

From the governing equation of the stiffness method,

$$kw_2 = F \tag{B.49}$$

we compute the deflection

$$w_2 = \frac{FL^3}{27EI} \tag{B.50}$$

Now it is easy to evaluate the chord rotations,

$$\psi_{12} = \frac{w_1 - w_2}{\ell_{12}} = -\frac{w_2}{L} = -\frac{FL^2}{27EI} \tag{B.51}$$

$$\psi_{23} = \frac{w_2 - w_3}{\ell_{23}} = \frac{w_2}{L/2} = \frac{2FL^2}{27EI} \tag{B.52}$$

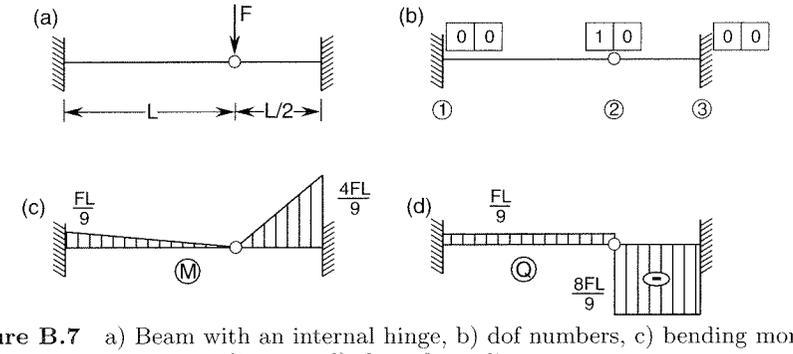


Figure B.7 a) Beam with an internal hinge, b) dof numbers, c) bending moment diagram, d) shear force diagram

and the end moments,

$$M_{12} = \frac{3EI_{12}}{\ell_{12}}(\phi_1 - \psi_{12}) = -\frac{3EI}{L}\psi_{12} = \frac{FL}{9} \tag{B.53}$$

$$M_{32} = \frac{3EI_{23}}{\ell_{23}}(\phi_3 - \psi_{23}) = -\frac{3EI}{L/2}\psi_{23} = -\frac{4FL}{9} \tag{B.54}$$

Based on these results we can plot the moment and shear force diagrams in Figures B.7(c),(d). □

In Example B.4, the right end section of member 3-4 could have been considered as connected to joint 4 by a hinge. In that case, the rotation ϕ_4 would not be considered as an unknown generalized displacement, and the stiffness matrix $\boldsymbol{K}_{34}$ would be replaced by $\boldsymbol{K}_{34}^{rh}$. The reader can set up the resulting equations of the stiffness method and verify that the same set of equations is obtained from (B.36) by subtracting 0.5 times the fourth equation from the third equation. This process is known as *static condensation*, and its application on the level of a single member transforms the standard member stiffness matrix $\boldsymbol{K}_{ij}$ in to the modified member stiffness matrix $\boldsymbol{K}_{ij}^{rh}$, or $\boldsymbol{K}_{ij}^{hr}$.

B.5 FRAMES

Based on what we know about trusses and beams, it is relatively easy to describe planar frames. It is sufficient to combine the equations for axial stretching with those for bending, and then transform them such that they correspond to a general inclination of the member axis with respect to the global coordinate system. The choice of fundamental variables for a planar frame has already been explained; cf. Table B.1. The fundamental equations for beams have been derived under the assumption that the beam centerline is parallel to the global axis x. They remain valid for members in a general position if we work in the local coordinate system with axes x' and z'; see Figure B.2(b). In local coordinates, the kinematic equations read

$$e_{ij} = u'_j - u'_i \tag{B.55}$$

$$\tau_{ij} = \phi_i + \frac{w'_j - w'_i}{\ell_{ij}} \tag{B.56}$$

$$\tau_{ji} = \phi_j + \frac{w'_j - w'_i}{\ell_{ij}} \tag{B.57}$$

and the equilibrium equations read

$$X'_{ij} = -N_{ij} \tag{B.58}$$

$$Z'_{ij} = -\frac{M_{ij} + M_{ji}}{\ell_{ij}} \tag{B.59}$$

$$M_{ij} = M_{ij} \tag{B.60}$$

$$X'_{ji} = N_{ij} \tag{B.61}$$

$$Z'_{ji} = \frac{M_{ij} + M_{ji}}{\ell_{ij}} \tag{B.62}$$

$$M_{ji} = M_{ji} \tag{B.63}$$

To be able to sum the contributions of members with different orientations of their centerlines, we have to transform the above relations into the global coordinate system. The transformation rules

$$u'_i = c_{ij}u_i + s_{ij}w_i \tag{B.64}$$

$$w'_i = -s_{ij}u_i + c_{ij}w_i \tag{B.65}$$

readily follow from Figure B.8(a). These rules apply in the same form to the right member end as well. Substituting (B.64)–(B.65) and analogous expressions for u'_j and w'_j into the kinematic equations (B.55)–(B.57) we get

$$\begin{Bmatrix} e_{ij} \\ \tau_{ij} \\ \tau_{ji} \end{Bmatrix} = \begin{bmatrix} -c & -s & 0 & c & s & 0 \\ s/\ell & -c/\ell & 1 & -s/\ell & c/\ell & 0 \\ s/\ell & -c/\ell & 0 & -s/\ell & c/\ell & 1 \end{bmatrix}_{ij} \begin{Bmatrix} u_i \\ w_i \\ \phi_i \\ u_j \\ w_j \\ \phi_j \end{Bmatrix} \tag{B.66}$$

or, in compact notation,

$$\boldsymbol{e}_{ij} = \boldsymbol{B}_{ij}\boldsymbol{d}_{ij} \tag{B.67}$$

The equilibrium equations (B.58)–(B.63) must be transformed using the inverse transformation law (see Figure B.8(c)),

$$X_{ij} = c_{ij}X'_{ij} - s_{ij}Z'_{ij} \tag{B.68}$$

$$Z_{ij} = s_{ij}X'_{ij} + c_{ij}Z'_{ij} \tag{B.69}$$

and analogous expressions for X_{ji} and Z_{ji}. The resulting equations

$$\begin{Bmatrix} X_{ij} \\ Z_{ij} \\ M_{ij} \\ X_{ji} \\ Z_{ji} \\ M_{ji} \end{Bmatrix} = \begin{bmatrix} -c & s/\ell & s/\ell \\ -s & -c/\ell & -c/\ell \\ 0 & 1 & 0 \\ c & -s/\ell & -s/\ell \\ s & c/\ell & c/\ell \\ 0 & 0 & 1 \end{bmatrix}_{ij} \begin{Bmatrix} N_{ij} \\ M_{ij} \\ M_{ji} \end{Bmatrix} \tag{B.70}$$

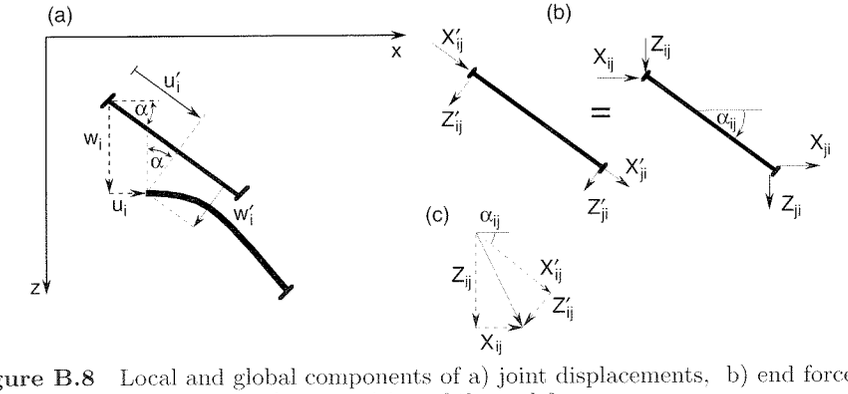


Figure B.8 Local and global components of a) joint displacements, b) end forces; c) decomposition of the end force

or

$$\boldsymbol{f}_{ij} = \boldsymbol{B}_{ij}^T \boldsymbol{s}_{ij} \tag{B.71}$$

are dual to (B.67) in the sense that the static matrix $\boldsymbol{B}_{ij}^T$ is the transpose of the kinematic matrix $\boldsymbol{B}_{ij}$.

To formulate the generalized constitutive equations for a planar frame member, it is convenient to introduce the axial stiffness $k_{a,ij} = EA_{ij}/\ell_{ij}$ and the bending stiffness $k_{b,ij} = EI_{ij}/\ell_{ij}$. Depending on the type of member, the generalized material stiffness matrix can assume the following forms:

1. For a member rigidly connected at both ends,

$$\boldsymbol{D}_{ij} = \begin{bmatrix} k_a & 0 & 0 \\ 0 & 2k_b & k_b \\ 0 & k_b & 2k_b \end{bmatrix}_{ij} \tag{B.72}$$

2. For a member rigidly connected at end i and hinged at end j,

$$\boldsymbol{D}_{ij}^{rh} = \begin{bmatrix} k_a & 0 & 0 \\ 0 & 3k_b & 0 \\ 0 & 0 & 0 \end{bmatrix}_{ij} \tag{B.73}$$

3. For a member hinged at end i and rigidly connected at end j,

$$\boldsymbol{D}_{ij}^{hr} = \begin{bmatrix} k_a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3k_b \end{bmatrix}_{ij} \tag{B.74}$$

4. For a member hinged at both ends,

$$\boldsymbol{D}_{ij}^{hh} = \begin{bmatrix} k_a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{ij} \tag{B.75}$$

A standard procedure leads to the frame member stiffness matrix $\boldsymbol{K}_{ij} = \boldsymbol{B}_{ij}^T \boldsymbol{D}_{ij} \boldsymbol{B}_{ij}$. It is useful to separate the contribution of axial stiffness (terms containing EA) from the contribution of the bending stiffness (terms containing EI), and write

$$\boldsymbol{K}_{ij} = \boldsymbol{K}_{a,ij} + \boldsymbol{K}_{b,ij} \tag{B.76}$$

The contribution of axial stiffness to the member stiffness matrix,

$$\boldsymbol{K}_{a,ij} = \frac{EA_{ij}}{\ell_{ij}} \begin{bmatrix} c^2 & cs & 0 & -c^2 & -cs & 0 \\ cs & s^2 & 0 & -cs & -s^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -c^2 & -cs & 0 & c^2 & cs & 0 \\ -cs & -s^2 & 0 & cs & s^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{ij} \tag{B.77}$$

is the same for any type of member. The contribution of bending stiffness depends on the type of member end connection:

1. For a member rigidly connected at both ends,

$$\boldsymbol{K}_{ij} = \frac{EI_{ij}}{\ell_{ij}^3} \begin{bmatrix} 12s^2 & -12cs & 6s\ell & -12s^2 & 12cs & 6s\ell \\ -12cs & 12c^2 & -6c\ell & 12cs & -12c^2 & -6c\ell \\ 6s\ell & -6c\ell & 4\ell^2 & -6s\ell & 6c\ell & 2\ell^2 \\ -12s^2 & 12cs & -6s\ell & 12s^2 & -12cs & -6s\ell \\ 12cs & -12c^2 & 6c\ell & -12cs & 12c^2 & 6c\ell \\ 6s\ell & -6c\ell & 2\ell^2 & -6s\ell & 6c\ell & 4\ell^2 \end{bmatrix}_{ij} \tag{B.78}$$

2. For a member rigidly connected at joint i and hinged at joint j,

$$\boldsymbol{K}_{ij}^{rh} = \frac{3EI_{ij}}{\ell_{ij}^3} \begin{bmatrix} s^2 & -cs & s\ell & -s^2 & cs & 0 \\ -cs & c^2 & -c\ell & cs & -c^2 & 0 \\ s\ell & -c\ell & \ell^2 & -s\ell & c\ell & 0 \\ -s^2 & cs & -s\ell & s^2 & -cs & 0 \\ cs & -c^2 & c\ell & -cs & c^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{ij} \tag{B.79}$$

3. For a member hinged at joint i and rigidly connected at joint j,

$$\boldsymbol{K}_{ij}^{hr} = \frac{3EI_{ij}}{\ell_{ij}^3} \begin{bmatrix} s^2 & -cs & 0 & -s^2 & cs & s\ell \\ -cs & c^2 & 0 & cs & -c^2 & -c\ell \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -s^2 & cs & 0 & s^2 & -cs & -s\ell \\ cs & -c^2 & 0 & -cs & c^2 & c\ell \\ s\ell & -c\ell & 0 & -s\ell & c\ell & \ell^2 \end{bmatrix}_{ij} \tag{B.80}$$

4. For a member hinged at both ends,

$$\boldsymbol{K}_{ij}^{hh} = \boldsymbol{O} \tag{B.81}$$

It is worth noting that the coefficients coming from the axial stiffness are typically several orders of magnitude larger than those coming from the bending stiffness. This fact becomes obvious if we evaluate the ratio of the stiffness coefficients,

$$\frac{\frac{EI}{\ell^3}}{\frac{EA}{\ell}} = \frac{I}{A\ell^2} = \frac{r^2}{\ell^2} = \frac{1}{\lambda^2} \tag{B.82}$$

Here, $r = \sqrt{I/A}$ is the radius of inertia of the cross section, and $\lambda = \ell/r$ is the slenderness ratio of the member, which must exceed about 20 for the bending theory to be valid. In structural engineering, λ ranges up to about 120 but for $\lambda > 40$ to 50, the second-order geometric effects, which cause coupling of rotations to the axial forces, need to be included (see e.g. Bažant and Cedolin 1991, Chapters 1 and 2). It is often possible to assume the axial stiffness to be infinite and, compared to the effects of bending, neglect the effect of axial shortening (or elongation) of the columns and beams on the distribution of internal forces. This assumption of *axial inextensibility* leads to a formulation with a smaller number of unknowns than the fully general formulation of the stiffness method.

After solving for the nodal displacements and rotations, we must evaluate the generalized stresses – the end moments and normal forces. Now we appreciate that

formulae (B.17)–(B.18), (B.40) and (B.45) can still be used, and only expression (B.8) for the chord rotation needs to be transformed into

$$\psi_{ij} = \frac{w_i' - w_j'}{\ell_{ij}} = \frac{s_{ij}}{\ell_{ij}}(u_j - u_i) - \frac{c_{ij}}{\ell_{ij}}(w_j - w_i) \tag{B.83}$$

In the standard stiffness method, the normal forces would be evaluated from the axial extensions

$$e_{ij} = c_{ij}(u_j - u_i) + s_{ij}(w_j - w_i) \tag{B.84}$$

via the generalized constitutive equations

$$N_{ij} = \frac{EA_{ij}}{\ell_{ij}} e_{ij} \tag{B.85}$$

In the simplified method, however, this formula cannot be applied since the assumption of axial inextensibility means that $EA_{ij}/\ell_{ij} = \infty$ and $e_{ij} = 0$. The normal forces must be calculated from the appropriate joint equilibrium conditions.

Example B.6: Construct the internal force diagrams for the frame in Figure B.9(a).

Solution: The general formulation would lead to a set of six equations for six unknown generalized displacements specified by the dof numbers in Figure B.9(b). Note that the rotation ϕ_4 is not among the fundamental unknowns because the right end of member 3-4 will be treated as hinge-connected. We will present a simplified solution based on the assumption of axial inextensibility. If the axial stiffness is considered as infinitely large, the axial extensions

$$e_{12} = w_1 - w_2 \tag{B.86}$$

$$e_{23} = u_3 - u_2 \tag{B.87}$$

$$e_{34} = w_4 - w_3 \tag{B.88}$$

must vanish. The resulting conditions $w_1 = w_2$, $u_2 = u_3$ and $w_3 = w_4$ lead to a reduction of the number of unknown generalized displacements. Realizing that the support displacements w_1 and w_4 are prescribed as zero, we obtain $w_2 = 0$, $w_3 = 0$, and $u_2 = u_3$. This means that only one of the nodal displacements, $u_2 \equiv u_3$, is an independent unknown. In addition, we still have two unknown nodal rotations, ϕ_2 and ϕ_3. The fundamental unknowns resulting from this simplified approach are described by the dof numbers in Figure B.9(c). Note that the reduced number of unknowns (three) is equal to the full number of unknowns (six) minus the number of members (three) because each member provides one inextensibility constraint.

The structure stiffness matrix is now assembled in the usual way, however, only the contributions of bending stiffness matrices $\boldsymbol{K}_{b,ij}$ are taken into account. From the governing equations of the (simplified) stiffness method

$$\frac{EI}{L^3}\begin{bmatrix} 15 & 6L & 3L \\ 6L & 8L^2 & 2L^2 \\ 3L & 2L^2 & 7L^2 \end{bmatrix} \begin{Bmatrix} u_2 \\ \phi_2 \\ \phi_3 \end{Bmatrix} = \begin{Bmatrix} F \\ 0 \\ 0 \end{Bmatrix} \tag{B.89}$$

we compute the generalized displacements

$$u_2 \equiv u_3 = 0.09848\frac{FL^3}{EI}, \qquad \phi_2 = -0.06818\frac{FL^2}{EI}, \qquad \phi_3 = -0.02272\frac{FL^2}{EI} \tag{B.90}$$

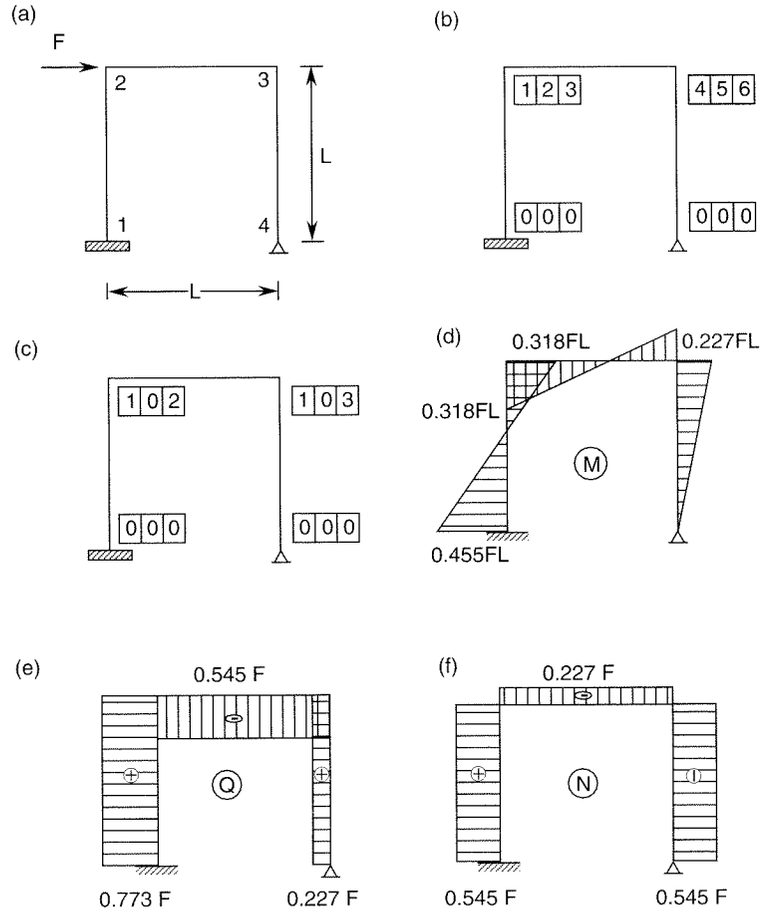


Figure B.9 Elastic analysis of a portal frame: a) geometry and loading, b) dof numbers for standard stiffness method, c) dof numbers for simplified stiffness method based on assumption of axial inextensibility, d) bending moment diagram, e) shear force diagram, f) normal force diagram

and substitute them into the formulae for the chord rotations and end moments,

$$\psi_{21} = \psi_{34} = -\frac{u_2}{L} = -0.09848\frac{FL^2}{EI} \tag{B.91}$$

$$M_{21} = \frac{2EI}{L}(2\phi_2 - 3\psi_{21}) = 0.318FL \tag{B.92}$$

$$M_{12} = \frac{2EI}{L}(\phi_2 - 3\psi_{21}) = 0.455FL \tag{B.93}$$

$$M_{34} = \frac{3EI}{L}(\phi_3 - \psi_{34}) = 0.227FL \tag{B.94}$$

The normal forces are calculated from the appropriate joint equilibrium conditions. For example, normal force N_{12} follows from the force equilibrium condition of joint 2 in the vertical direction. The resulting internal forces are plotted in the diagrams in Figures B.9(d)–(f). □

Appendix C

Linear Programming

Linear programming (LP) is a mathematical theory of minimization of a linear function subject to linear constraints. The constraints can have the form of equalities or inequalities. The general form of a *linear programming problem* (LPP) is

$$\text{Minimize } f(\boldsymbol{x}) \equiv \boldsymbol{c}^T \boldsymbol{x} \tag{C.1}$$

subject to

$$\boldsymbol{A}_1 \boldsymbol{x} = \boldsymbol{b}_1 \tag{C.2}$$
$$\boldsymbol{A}_2 \boldsymbol{x} \geq \boldsymbol{b}_2 \tag{C.3}$$

Vector $\boldsymbol{x}$ contains the unknown variables while matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ and vectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ and $\boldsymbol{c}$ consist of given constants. Any vector $\boldsymbol{x}$ satisfying the constraints is called a *feasible solution*. The function f to be minimized is called the *objective function* (or *cost function*). The inequality sign between vectors in (C.3) is somewhat unusual but it simplifies notation, replacing a set of like inequalities between all the individual components.

The above formulation also covers maximization problems because maximizing $f(\boldsymbol{x})$ is equivalent to minimizing $-f(\boldsymbol{x})$, and at the same time it covers problems with inequality constraints of the form $\boldsymbol{A}_3 \boldsymbol{x} \leq \boldsymbol{b}_3$, which can be rewritten as $-\boldsymbol{A}_3 \boldsymbol{x} \geq -\boldsymbol{b}_3$.

The most popular method of LP is the *simplex method* (Dantzig, 1949), which is applicable to LPPs in the so-called *standard form*, i.e. to problems cast as

$$\text{Minimize } f(\boldsymbol{x}) \equiv \boldsymbol{c}^T \boldsymbol{x} \tag{C.4}$$

subject to

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{C.5}$$
$$\boldsymbol{x} \geq \boldsymbol{0} \tag{C.6}$$

with the vector on the right-hand side of (C.5) having non-negative components, $\boldsymbol{b} \geq \boldsymbol{0}$. Note the characteristic features of the standard form:

1. the unknowns are non-negative, and
2. all the other constraints are equalities.

It is easy to show that any LPP in the form (C.1)–(C.3) can be converted to the standard form (C.4)–(C.6), however, additional unknowns must be introduced in the

course of transformation. For example, those inequality constraints in (C.3) that do not happen to have the form $x_i \geq 0$ are replaced by equality constraints $\boldsymbol{A}_2\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{b}_2$ with additional non-negative unknowns $\boldsymbol{y} \geq \boldsymbol{0}$ called the *slack variables*. It is therefore advantageous if, by a suitable choice of fundamental unknowns, the physical problem under consideration can be cast directly in the standard form.

The simplex method relies heavily on the notion of a *basic feasible solution*. To explain its definition, let us focus on the equality constraints (C.5). It is natural to assume that the constraints are linearly independent, and that their number is smaller than the number of unknowns. Note that

1. if the constraints were linearly dependent, either no feasible solution would exist, or the dependent constraints could be removed without any effect on the solution;
2. if the number of linearly constraints were larger than the number of unknowns, the constraints could not be linearly independent;
3. if the number of linearly independent constraints were equal to the number of unknowns, the problem would have only one feasible solution, and so it would not be a genuine minimization problem.

Let us denote the number of equality constraints by m and the number of unknown variables by n, with $m < n$. The matrix $\boldsymbol{A}$ is a full-rank matrix of type (m, n), and so it must be possible to find m linearly independent columns. Let us arrange these columns into a regular square matrix $\boldsymbol{A}_B$ of type (m, m) and the remaining columns into a matrix $\boldsymbol{A}_N$ of type $(m, n-m)$. The corresponding unknown variables can be arranged into vectors $\boldsymbol{x}_B$ and $\boldsymbol{x}_N$, referred to as the *basic variables* and the *nonbasic variables*, respectively. The equality constraints (C.5) can now be written as

$$\boldsymbol{A}_B\boldsymbol{x}_B + \boldsymbol{A}_N\boldsymbol{x}_N = \boldsymbol{b} \tag{C.7}$$

and the objective function as

$$f(\boldsymbol{x}_B, \boldsymbol{x}_N) = \boldsymbol{c}_B^T\boldsymbol{x}_B + \boldsymbol{c}_N^T\boldsymbol{x}_N \tag{C.8}$$

where $\boldsymbol{c}_B$ and $\boldsymbol{c}_N$ represent the coefficients of the objective function associated with the basic and nonbasic variables, respectively. Since the matrix $\boldsymbol{A}_B$ is regular, the basic variables can be expressed from (C.7) in terms of the nonbasic variables as

$$\boldsymbol{x}_B = \boldsymbol{A}_B^{-1}\boldsymbol{b} - \boldsymbol{A}_B^{-1}\boldsymbol{A}_N\boldsymbol{x}_N \tag{C.9}$$

Now suppose that the vector $\boldsymbol{A}_B^{-1}\boldsymbol{b}$ happens to have all components non-negative. In that case, the solution $\boldsymbol{x}_N = \boldsymbol{0}$ and $\boldsymbol{x}_B = \boldsymbol{A}_B^{-1}\boldsymbol{b}$ is feasible, and it is called a basic feasible solution. In general, there are many ways of selecting the set of basic variables, and each of them may or may not lead to a basic feasible solution (depending on the signs of components of vector $\boldsymbol{A}_B^{-1}\boldsymbol{b}$). The theory of LP shows that if the problem has an optimal solution, then one of the basic feasible solutions is an optimal one (we avoid saying '*the* optimal one' because the solution does not need to be unique). In other words, when looking for an optimal solution it suffices to check the basic feasible ones. This is very important because the number of basic feasible solutions is finite, unlike the number of all feasible solutions.

An obvious but naive solution technique would be to construct all the basic feasible solutions and find that which minimizes the objective function. However, the basic variables can be selected in $\binom{n}{m}$ different ways, and for each selection we would have

to invert a matrix of type (m, m). This would be prohibitively expensive for large problems. The simplex method starts from an initial basic feasible solution, and it systematically modifies the choice of the basic variables such that the objective function monotonically decreases. The solution for which a further decrease of the objective function turns out to be impossible is optimal.

Suppose that we have constructed an initial basic feasible solution. The relation (C.9) is fully equivalent with the original equality constraints (C.5), and so it must hold for any feasible solution. Therefore, we can substitute (C.9) into (C.8) and express the objective function

$$\begin{aligned} f(\boldsymbol{x}_B, \boldsymbol{x}_N) &= \boldsymbol{c}_B^T\boldsymbol{x}_B + \boldsymbol{c}_N^T\boldsymbol{x}_N = \boldsymbol{c}_B^T(\boldsymbol{A}_B^{-1}\boldsymbol{b} - \boldsymbol{A}_B^{-1}\boldsymbol{A}_N\boldsymbol{x}_N) + \boldsymbol{c}_N^T\boldsymbol{x}_N \\ &= \boldsymbol{c}_B^T\boldsymbol{A}_B^{-1}\boldsymbol{b} + (\boldsymbol{c}_N - \boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B)^T\boldsymbol{x}_N \end{aligned} \tag{C.10}$$

in terms of the nonbasic variables only. For the current basic feasible solution we have $\boldsymbol{x}_N = \boldsymbol{0}$, $\boldsymbol{x}_B = \boldsymbol{A}_B^{-1}\boldsymbol{b}$ and $f(\boldsymbol{x}) = \boldsymbol{c}_B^T\boldsymbol{A}_B^{-1}\boldsymbol{b}$.

Now look what happens if we depart from the current basic feasible solution, i.e. if we change the value of one or more components in $\boldsymbol{x}_N$. These variables can only increase, since their current values vanish and the new values must be non-negative. If some component of the vector $\boldsymbol{c}_N - \boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B$ is negative, then an increase of the corresponding nonbasic variable results in a decrease of the objective function, and the current solution can be improved. If, on the other hand, all the components of $\boldsymbol{c}_N - \boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B$ are non-negative, there is no way to decrease the objective function while keeping the solution feasible, and so the current solution is optimal.

If the current basic feasible solution is found not to be optimal, we can construct an improved basic feasible solution with a lower value sof the objective function. To this end, we select one of the currently nonbasic variables for which the corresponding component of $\boldsymbol{c}_N - \boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B$ is negative, and we decide that this variable will become a basic one (we say that it will *enter the basis*). Of course, the total number of basic variables must remain unchanged, and so one of the basic variables must *leave the basis* and become a nonbasic one. The choice of this variable is dictated by the condition that, after we set it to zero, the other basic variables must remain non-negative.

Look again at equation (C.9), which must hold for any feasible solution. The vector $\boldsymbol{A}_B^{-1}\boldsymbol{b}$ represents the current values of the basic variables, which are usually positive (sometimes the value of a basic variable happens to vanish but this is the so-called degenerate case that we leave out of our considerations for a while). The column of $\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$ corresponding to the variable that enters the basis is called the *key column*. The coefficients in the key column describe the evolution of the variables in $\boldsymbol{x}_B$ during an increase of the variable entering the basis. Positive values of these coefficients indicate that the corresponding variable from $\boldsymbol{x}_B$ will decrease and eventually reach zero. The variable for which this happens first leaves the basis while the others remain in the basis with values that are still positive. So the variable to leave the basis can be detected by looking at the ratios between the coefficients in the key column and the elements of $\boldsymbol{A}_B^{-1}\boldsymbol{b}$.

Now we have all the necessary ingredients of the simplex method. The actual calculations can conveniently be performed in a special table known as the *simplex tableau*. The initial simplex tableau for the LPP in the standard form (C.4)–(C.6) is shown in Table C.1(a). The core of the simplex tableau is the coefficient matrix of the equality constraints. This matrix is augmented on the right by the right-hand side of

the equality constraints, at the bottom by the coefficients of the objective function, and on the left by a zero column. The element in the lower left corner is (forever) set to one, and the element in the lower right corner is (initially) set to zero. The top line only lists the variables associated with individual columns, and so we do not count it when referring to the row numbers. Similarly, we do not count the leftmost column when referring to the column number, and so the row and column numbering in the simplex tableau is the same as in the coefficient matrix. The first m rows obviously represent the equality constraints $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. The bottom row records the definition of the objective function written in the form of an equality,

$$-f + \boldsymbol{c}^T\boldsymbol{x} = 0 \tag{C.11}$$

This is the reason why we had to add the leftmost column that corresponds to minus the objective function. Since the objective function does not enter the equality constraints, the first m coefficients in this column must be zero. In fact, the leftmost column is never modified by the simplex algorithm, and so it does not have to be actually represented in the numerical implementation.

Once we make an initial choice of the basic variables, the tableau can be partitioned as shown in Table C.1(b). Of course, the blocks corresponding to the basic and nonbasic variables will in reality not be contiguous but rather spread through the tableau. Now we can construct the initial feasible solution by transforming the equality constraints from (C.7) to (C.9), for the present purpose written as

$$\boldsymbol{I}\boldsymbol{x}_B + \boldsymbol{A}_B^{-1}\boldsymbol{A}_N\boldsymbol{x}_N = \boldsymbol{A}_B^{-1}\boldsymbol{b} \tag{C.12}$$

At the same time, the objective function must be transformed from (C.8) to (C.10), with (C.10) written for the present purpose as

$$-f + \boldsymbol{0}^T\boldsymbol{x}_B + (\boldsymbol{c}_N - \boldsymbol{A}_N^T\boldsymbol{A}_B^{-T}\boldsymbol{c}_B)^T\boldsymbol{x}_N = -\boldsymbol{c}_B^T\boldsymbol{A}_B^{-1}\boldsymbol{b} \tag{C.13}$$

The transformation from (C.7) to (C.12) consists of multiplying both sides of the original equation by $\boldsymbol{A}_B^{-1}$. This is equivalent to applying the Gauss elimination procedure so that the block corresponding to the basic variables is transformed into a unit matrix. Similarly, the transformation from (C.8) to (C.13) is equivalent to eliminating the bottom row coefficients that correspond to the basic variables. The transformed simplex tableau is shown in Table C.2. If, after these manipulations, the first m coefficients on the right-hand side turn out to be positive (or at least non-negative), the solution is feasible. The element in the right lower corner then gives the opposite value of the objective function. This is verified by substituting $\boldsymbol{x}_N = \boldsymbol{0}$ into (C.13).

Table C.1 Initial simplex tableau

(a)

$-f$	$\boldsymbol{x}^T$	RHS
$\boldsymbol{0}$	$\boldsymbol{A}$	$\boldsymbol{b}$
1	$\boldsymbol{c}^T$	0

(b)

$-f$	$\boldsymbol{x}_B^T$	$\boldsymbol{x}_N^T$	RHS
$\boldsymbol{0}$	$\boldsymbol{A}_B$	$\boldsymbol{A}_N$	$\boldsymbol{b}$
1	$\boldsymbol{c}_B^T$	$\boldsymbol{c}_N$	0

Table C.2 Simplex tableau representing a basic solution

$-f$	$\boldsymbol{x}_B^T$	$\boldsymbol{x}_N^T$	RHS
$\boldsymbol{0}$	$\boldsymbol{I}$	$\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$	$\boldsymbol{A}_B^{-1}\boldsymbol{b}$
1	$\boldsymbol{0}^T$	$\boldsymbol{c}_N^T - \boldsymbol{c}_B^T\boldsymbol{A}_B^{-1}\boldsymbol{A}_N$	$-\boldsymbol{c}_B^T\boldsymbol{A}_B^{-1}\boldsymbol{b}$

Example C.1: Consider the following LPP in the standard form:

$$\text{Minimize } f(x_1, x_2, x_3) = x_1 - x_2 + x_3 \tag{C.14}$$

subject to

$$x_1 + x_2 + 2x_3 = 5 \tag{C.15}$$

$$-2x_1 + x_3 = 2 \tag{C.16}$$

$$x_i \geq 0, \quad i = 1, 2, 3 \tag{C.17}$$

Set up the initial simplex tableau and transform it into a tableau corresponding to an initial feasible solution.

Solution: The initial simplex tableau reflecting the problem formulation is given in Table C.3(a). Let us try to select x_1 and x_2 as the basic variables. First, we add twice the first row to the second row, and we subtract the first row from the third row in order to eliminate the coefficients in the column corresponding to the first basic variable, x_1. The intermediate state of the simplex tableau is given in Table C.3(b). Next, we divide the second row by 2, we subtract it from the first row, and we add twice the second row to the third row in order to eliminate the coefficients in the column corresponding to the second basic variable, x_2. The resulting tableau in Table C.3(c) has the general structure from Table C.2 but the corresponding solution is not feasible because of the presence of a negative value -1 on the right-hand side.

Let us try to select x_1 and x_3 as the basic variables. After elimination of the coefficients in the first and the third row, we get the tableau in Table C.4(c). The

Table C.3 Transformations of simplex tableau leading to an infeasible solution

(a)

$-f$	x_1	x_2	x_3	RHS
0	1	1	2	5
0	−2	0	1	2
1	1	−1	1	0

(b)

$-f$	x_1	x_2	x_3	RHS
0	1	1	2	5
0	0	2	5	12
1	0	−2	−1	−5

(c)

$-f$	x_1	x_2	x_3	RHS
0	1	0	−0.5	−1
0	0	1	2.5	6
1	0	0	4	7

right-hand side is non-negative, and so the corresponding solution is feasible. The elements of the modified right-hand side give values of the basic variables, $x_1 = 0.2$, $x_3 = 2.4$, while the element in the lower right corner gives the opposite value of the objective function, $-f = -2.6$, i.e. $f = 2.6$. The nonbasic variable x_2 has of course a zero value. The reader can verify that the solution indeed satisfies all the constraints, and that $f(0.2,\ 0,\ 2.4) = 0.2 - 0 + 2.4 = 2.6$. □

The construction of an initial feasible solution is in general not an easy task. Special techniques have been developed for this purpose. However, when solving LPPs which correspond to actual problems, we can often be guided by the physical meaning of the variables. For example, in LPPs representing the kinematic approach to plastic limit analysis (see Chapter 7), any kinematically admissible mechanism with one degree of freedom corresponds to a basic feasible solution. Therefore, we will only briefly explain the most universal starting procedure. Instead of the original problem (C.4)–(C.6), we consider a modified problem

$$\text{Minimize } g(\boldsymbol{x}, \boldsymbol{y}) = y_1 + y_2 + \ldots y_m \tag{C.18}$$

subject to

$$\boldsymbol{A}\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{b} \tag{C.19}$$

$$\boldsymbol{x} \geq \boldsymbol{0}, \quad \boldsymbol{y} \geq \boldsymbol{0} \tag{C.20}$$

where $\boldsymbol{y}$ is a vector of m additional (artificial) variables. The modified problem has a basic feasible solution that is very easy to construct, namely $\boldsymbol{x} = \boldsymbol{0}$ and $\boldsymbol{y} = \boldsymbol{b}$. The auxiliary objective function g is minimized for $\boldsymbol{y} = \boldsymbol{0}$. Consequently, if we find the optimal solution of the modified problem by the simplex algorithm, we obtain a vector $\boldsymbol{x}$ that can be used as an initial basic feasible solution of the original problem.

Once the initial feasible solution has been constructed, we can check it for optimality. As explained before, the solution is optimal if the vector $\boldsymbol{c}_N - \boldsymbol{A}_N^T \boldsymbol{A}_B^{-T} \boldsymbol{c}_B$ has all components non-negative. Since this vector appears in the bottom row of the transformed simplex tableau (see Table C.2), the optimality condition can easily be checked. If some of the coefficients in the bottom row have a negative sign, any of the corresponding nonbasic variables can be selected as the variable that enters the basis. In practice, the variable corresponding to the negative coefficient with the

Table C.4 Transformations of simplex tableau leading to a feasible solution

(a)

$-f$	x_1	x_2	x_3	RHS
0	1	1	2	5
0	−2	0	1	2
1	1	−1	1	0

(b)

$-f$	x_1	x_2	x_3	RHS
0	1	1	2	5
0	0	2	5	12
1	0	−2	−1	−5

(c)

$-f$	x_1	x_2	x_3	RHS
0	1	0.2	0	0.2
0	0	0.4	1	2.4
1	0	−1.6	0	−2.6

largest absolute value is usually selected. The column in which this coefficient is located is the key column. Next, we have to find the basic variable that will leave the basis. It follows from the previous discussion that this variable corresponds to the row in which the ratio of the right-hand side value to the coefficient in the key column is minimized, where only rows with positive coefficients in the key column are considered. The row determined by this criterion is the *key row*, and the corresponding basic variable (that which appears in the equation represented by the key row) will leave the basis.

An important point is that the next basic feasible solution obtained after the exchange of one basic variable can easily be constructed from the current feasible solution, i.e. we do not have to start from the initial simplex tableau. Indeed, we only need to transform the key column into a column with a unit coefficient in the key row and zero coefficients everywhere else. This can be achieved by Gauss elimination, using the element in the key row and key column as the pivot. If the key column and key row have been selected properly (and if the solution is not a degenerate one – see the discussion following the next example), the transformed tableau is guaranteed to correspond to a feasible solution with a smaller value of the objective function than the original solution. The change of basis just described is repeated until we arrive at an optimal solution.

Example C.2: Solve the problem from Example C.1 by the simplex method.

Solution: In Example C.1 we have constructed an initial feasible solution described by the tableau in Table C.4(a), which is reproduced in Table C.5(a). This solution is not optimal as there is a negative coefficient in the bottom row. Column 2, in which this negative coefficient appears, is the key column, and the corresponding variable x_2 will enter the basis. Now we have to determine the key row. The ratio of the right-hand side value to the value in the key column is $0.2/0.2 = 1$ for the first row and $2.4/0.4 = 6$ for the second row. The smallest value of this ratio is obtained for the first row, which therefore becomes the key row. The basic variable below which we find the unit coefficient in the key row will leave the basis; in our case, this is variable x_1. The tableau must be transformed such that the coefficient in the first row and second column assumes a unit value while all other coefficients in the second column vanish. So we divide the first row by 0.2 (the pivot) and then we subtract its 0.4-multiple from the second row and add its 1.6-multiple to the third row. The resulting simplex tableau is displayed in Table C.5(b). We note that the value of the objective function has decreased to 1 (this follows from the fact that $-f$ = value in the lower right corner = -1). As all the coefficients in the bottom row are non-negative, this solution is optimal. The values

Table C.5 Transformation of simplex tableau from the initial feasible solution to the optimal solution

(a)

$-f$	x_1	x_2	x_3	RHS
0	1	0.2	0	0.2
0	0	0.4	1	2.4
1	0	−1.6	0	−2.6

(b)

$-f$	x_1	x_2	x_3	RHS
0	5	1	0	1
0	−2	0	1	2
1	8	0	0	−1

that appear on the right-hand side correspond to the basic variables $x_2 = 1$ and $x_3 = 2$ while the nonbasic variable x_1 has a zero value. □

So far we have tacitly assumed that we deal with a 'well-behaved' problem, for which a straightforward application of the simplex algorithm leads to a unique optimal solution. Now it is time to mention some of the special situations that might arise.

First, the problem may have no feasible solution simply because the constraints are too strict. On the other hand, the constraints might be so weak that the objective function is not bounded from below in the feasible domain. In this case, no optimal solution exists either. The above cases can be excluded for LPPs describing a real physical problem for which we know that an optimal solution must exist.

Even if an optimal solution exists, it does not have to be unique. It is possible that the same minimum value of the objective function is obtained for more than one feasible solution. This may easily happen, for instance, for problems with symmetry. The *lack of uniqueness* is easy to detect and it does not present any serious complication. Let us return to the fundamental equations (C.9) and (C.10), and recall that the solution is optimal if all components of the vector $\boldsymbol{c}_N - \boldsymbol{A}_N^T \boldsymbol{A}_B^{-T} \boldsymbol{c}_B$ are non-negative. However, the solution is unique only if all these components are strictly positive. If any of the components vanishes, then the corresponding variable can change its value without any effect on the objective function. This means that additional optimal solutions can be constructed by changing the value of this nonbasic variable and modifying the values of the basic variables according to (C.9). The set of all optimal solutions is the convex hull of a finite number of basic optimal solutions, which can be constructed by appropriate simplex transformations.

A more severe complication is the so-called *degeneracy*. A basic feasible solution is said to be degenerate if one or more basic variables computed from (C.9) vanish. In that case, the same basic feasible solution is obtained for various distinct choices of the basic set of variables, which might lead to the so-called *cycling*. By cycling we mean that a closed loop of simplex tableaux is generated, each of them corresponding to the same basic feasible solution, naturally with the same value of the objective function (which does not have to be the optimal one). Degeneracy manifests itself by the presence of one or more zero values on the right-hand side of the simplex tableau. If the corresponding coefficients in the key column are positive, one of the rows with the zero right-hand side becomes the key row (because the ratio of the right-hand side to the coefficient in the key row is zero, which is clearly the minimum possible value taken into account when looking for the key row), and the next simplex transformation does not change the value of the objective function. The algorithm may continue towards the optimal solution in the subsequent steps, but it might also get stuck in a closed loop. The danger of cycling can be eliminated by a special rule for the selection of the key row in cases where the same minimum ratio is obtained for more than one row. For details, see standard textbooks on linear programming, some of which are listed at the end of this appendix.

In the non-degenerate case, the basic steps of the simplex algorithm are as follows.

Algorithm C.1: Simplex Method

1. Construct an initial basic feasible solution.
2. Check for optimality:

   (a) If all the coefficients in the bottom row are non-negative, *the current solution is optimal*. In that case, if all the coefficients in the bottom row that correspond to the nonbasic variables are positive, then the optimal solution is unique.
   (b) If at least one of the coefficients in the bottom row is negative, the solution is not optimal. Find the negative coefficient with the largest absolute value, and select the corresponding column as the key column.
3. Find the variable leaving the basis:
   (a) If all the coefficients in the key column are nonpositive, then the objective function is unbounded on the feasible domain, and *no optimal solution exists*.
   (b) If at least one of the coefficients in the key column is positive, find the row for which the ratio of the right-hand side value to the coefficient in the key column is minimal, taking into account only the rows with a positive coefficient in the key column. If the same minimum ratio is obtained for more than one column, there is a danger of cycling, and a special rule for the selection of the key row should be used in order to prevent it. Otherwise, the row with the minimum ratio is selected as the key row.
4. Transform the table:
   (a) Divide the key row by the pivot, i.e. by the coefficient at the intersection of the key row and key column.
   (b) Subtract appropriate multiples of the key row from the other rows in order to eliminate the coefficients in the key column.
5. Repeat from step 2.

Finally, it is appropriate to mention that the standard simplex method presented here is not the only solution technique of linear programming by far. The simplex technique itself has many variants and modifications. The *revised simplex method* (Dantzig and Orchard-Hays, 1953) is basically a different implementation of the standard simplex method that requires less storage space but more operations per step. The simplex algorithm can be adapted to problems with bounded variables, i.e. to problems where the variables must remain inside given intervals instead of being just nonnegative (Charnes and Lemke, 1954). Another family of approaches, including the *dual simplex method* (Lemke, 1954) and the *primal-dual method* (Dantzig, Ford and Fulkerson, 1956), makes use of the so-called duality. Among the recent advances of optimization algorithms, the most notable one is *Karmarkar's method* (Karmarkar, 1984), which has been developed mainly for nonlinear problems but is applicable to linear programming problems as well.

For further reading, see Bazaraa, Jarvis and Sherali (1990), Gass (1984), Goldfarb and Todd (1989), Luenberger (1973), Sakarovitch (1983), Strayer (1989) and Walsh (1985). Perhaps the best introduction to the simplex algorithm and to Karmarkar's method is given in Chapter 8 of the excellent textbook on applied mathematics by Strang (1986).

Appendix D

Cartesian Tensors and Elasticity

D.1 CARTESIAN TENSORS

In this book, we denote first-order tensors (representing vectors) by lower-case Latin letters, second-order tensors by lower-case Greek or Latin letters, and fourth-order tensors by capital Latin letters. Bold face letters represent the entire tensors (in compact notation). Referring to the Cartesian components of tensors (in indicial notation) we use lower-case Latin subscripts $i, j, k, l, ...$, which can take values 1, 2 and 3 corresponding to the Cartesian coordinate axes x_1, x_2 and x_3. For example, $\boldsymbol{u}$ is the displacement vector (first-order tensor) with components u_i, $i = 1, 2, 3$; $\boldsymbol{\varepsilon}$ is the strain tensor (which is of the second order) with components ε_{ij}, $i = 1, 2, 3$ and $j = 1, 2, 3$; $\boldsymbol{\sigma}$ is the stress tensor (which is also of the second order) with components σ_{ij}, $i = 1, 2, 3$ and $j = 1, 2, 3$; and $\boldsymbol{D}$ is the stiffness tensor (which is a fourth-order tensor) with components D_{ijkl}, all subscripts running again from 1 to 3. When the tensor itself has a subscript, we raise it to a superscript in the indicial notation in order to avoid confusion with the subscripts referring to individual components. For example, the components of the elastic stiffness tensor $\boldsymbol{D}_e$ are denoted as D^e_{ijkl}.

The transpose $\boldsymbol{\varepsilon}^T$ of a second-order tensor $\boldsymbol{\varepsilon}$ is obtained by switching the order of subscripts, i.e. $(\boldsymbol{\varepsilon}^T)_{ij} = \varepsilon_{ji}$. If the transposed tensor is equal to the original one ($\varepsilon_{ij} = \varepsilon_{ji}$ for any i and j), it is said to be symmetric. The symmetric part of an arbitrary second-order tensor $\boldsymbol{f}$ is defined as $\boldsymbol{f}_{\text{sym}} = (\boldsymbol{f}+\boldsymbol{f}^T)/2$. Fourth-order tensors can exhibit minor symmetry ($D_{ijkl} = D_{jikl} = D_{ijlk} = D_{jilk}$ for any i, j, k, l), or major symmetry ($D_{ijkl} = D_{klij}$ for any i, j, k, l).

In expressions involving components of tensors we adhere to *Einstein's summation convention*. Repeated subscripts in a product-like expression imply summation from 1 to 3 (unless explicitly stated to the contrary), e.g. $\boldsymbol{u} \cdot \boldsymbol{v} = u_i v_i$ means in the 'full' notation $\boldsymbol{u} \cdot \boldsymbol{v} = \sum_{i=1}^{3} u_i v_i$, and $\sigma_{ij} = D_{ijkl}\varepsilon_{kl}$ means $\sigma_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} D_{ijkl}\varepsilon_{kl}$. The summation convention is used because sums over repeated subscripts arise almost as a rule in tensor analysis. The basic tensorial operations used in this book are as follows:

1. The dot product (or inner product, or contraction, or single contraction) of two first-order tensors, $\boldsymbol{u}\cdot\boldsymbol{v} = u_i v_i$, which produces a scalar, and therefore is sometimes called the scalar product;
2. the double-dot product (inner product, double contraction) of two second-order tensors, $\boldsymbol{\sigma} : \boldsymbol{\varepsilon} = \sigma_{ij}\varepsilon_{ij}$, which also produces a scalar (and therefore could also be called the scalar product);
3. the dot product of two second-order tensors, $\boldsymbol{f} \cdot \boldsymbol{g}$, which produces a second-order tensor with components $(\boldsymbol{f} \cdot \boldsymbol{g})_{ij} = f_{ik}g_{kj}$;

4. the dot products of a second-order tensor with a first-order tensor, $\boldsymbol{\sigma} \cdot \boldsymbol{n}$, or $\boldsymbol{n} \cdot \boldsymbol{\sigma}$, which produce first-order tensors with components $(\boldsymbol{\sigma} \cdot \boldsymbol{n})_i = \sigma_{ij} n_j$, or $(\boldsymbol{n} \cdot \boldsymbol{\sigma})_i = n_j \sigma_{ji}$;
5. the double-dot products of a fourth-order tensor with a second-order tensor, $\boldsymbol{D} : \boldsymbol{\varepsilon}$, or $\boldsymbol{\varepsilon} : \boldsymbol{D}$, which produce second-order tensors with components $(\boldsymbol{D} : \boldsymbol{\varepsilon})_{ij} = D_{ijkl}\varepsilon_{kl}$, or $(\boldsymbol{\varepsilon} : \boldsymbol{D})_{ij} = \varepsilon_{kl} D_{klij}$; and
6. the direct product (or outer product, or tensorial product) of two second-order tensors, $\boldsymbol{f} \otimes \boldsymbol{g}$, sometimes denoted simply $\boldsymbol{fg}$, which produces a fourth-order tensor with components $(\boldsymbol{fg})_{ijkl} = f_{ij} g_{kl}$.

The scalar products are commutative (i.e. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$ and $\boldsymbol{\sigma} : \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon} : \boldsymbol{\sigma}$); the third operation is not; the fourth operation is commutative if the second-order tensor is symmetric ($\boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{n} \cdot \boldsymbol{\sigma}$ if $\sigma_{ij} = \sigma_{ji}$ for any i, j); the fifth operation is commutative if the fourth-order tensor exhibits major symmetry ($\boldsymbol{D} : \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon} : \boldsymbol{D}$ if $D_{ijkl} = D_{klij}$ for any i, j, k, l); and the direct product is noncommutative. Readers who do not wish to memorize the above rules may carry out all tensor operations in the indicial notation, for which the summation convention is all that is needed. For brevity of writing, the result of the operations may then be translated into the compact notation.

An important example of a second-order tensor is the *Kronecker delta*, $\boldsymbol{\delta}$, with components $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$. It plays the role of the unit element with respect to the (single) dot product of second order tensors ($\boldsymbol{f} \cdot \boldsymbol{\delta} = \boldsymbol{\delta} \cdot \boldsymbol{f} = \boldsymbol{f}$ for any $\boldsymbol{f}$), and therefore it is also called the unit second-order tensor. For our purpose, a more important property of Kronecker delta is that all its principal values are equal to 1, and so the stress tensor $\boldsymbol{\sigma} = \sigma_V \boldsymbol{\delta}$ represents a volumetric stress state, i.e., a state in which the shear stress on any plane is zero. Note that $\boldsymbol{\delta} : \boldsymbol{\delta} = \delta_{ij}\delta_{ij} = \delta_{ii} = 3$ (in a three-dimensional space).

Let us also introduce the unit fourth-order tensor, $\boldsymbol{I}$, with components $I_{ijkl} = \delta_{ik}\delta_{jl}$, which has the important property that $\boldsymbol{I} : \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon} : \boldsymbol{I} = \boldsymbol{\varepsilon}$ for any second-order tensor $\boldsymbol{\varepsilon}$. This tensor exhibits major symmetry but not minor symmetry. It is therefore useful to introduce another tensor $\boldsymbol{I}_s$ with components $I^s_{ijkl} = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})/2$, which has both major and minor symmetry. The product $\boldsymbol{I}_s : \boldsymbol{f}$ is equal to the symmetric part of $\boldsymbol{f}$, since $(\boldsymbol{I}_s : \boldsymbol{f})_{ij} = (\delta_{ik}\delta_{jl}f_{kl} + \delta_{il}\delta_{jk}f_{kl})/2 = (f_{ij} + f_{ji})/2 = (\boldsymbol{f}_{\text{sym}})_{ij}$.

Finally, let us introduce the notation $\partial\boldsymbol{u}/\partial\boldsymbol{x}$ for the tensor produced by differentiation of a tensor-valued function $\boldsymbol{u}$ with respect to its tensorial argument $\boldsymbol{x}$. For example, if both $\boldsymbol{u}$ and $\boldsymbol{x}$ are first-order tensors (i.e. vectors), then $\partial\boldsymbol{u}/\partial\boldsymbol{x}$ is a second-order tensor with components

$$\left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}\right)_{ij} = \frac{\partial u_i}{\partial x_j} \tag{D.1}$$

Similarly, if $\boldsymbol{\sigma}$ is a second-order tensor dependent on another second-order tensor $\boldsymbol{\varepsilon}$, the differentiation yields a fourth-order tensor with components

$$\left(\frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}}\right)_{ijkl} = \frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} \tag{D.2}$$

Note that, for any first-order tensor $\boldsymbol{u}$, we have

$$\left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{u}}\right)_{ij} = \boldsymbol{\delta} \quad \text{because } \frac{\partial u_i}{\partial u_j} = \delta_{ij} \tag{D.3}$$

and for any second-order tensor $\boldsymbol{\sigma}$ we have

$$\left(\frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\sigma}}\right)_{ijkl} = \boldsymbol{I} \quad \text{because } \frac{\partial \sigma_{ij}}{\partial \sigma_{kl}} = \delta_{ik}\delta_{jl} = I_{ijkl} \tag{D.4}$$

The chain rule can also be generalized to tensor-valued functions of tensorial arguments, but attention must be paid to the nature of product operations between individual terms. The derivation is most safely done in the indicial notation. For example, if a second-order tensor $\boldsymbol{f}$ depends on a second-order tensor $\boldsymbol{u}$ and a scalar v, each of which depends on a second-order tensor $\boldsymbol{x}$, we derive the chain rule in indicial notation as

$$\frac{\partial f_{ij}}{\partial x_{kl}} = \frac{\partial f_{ij}}{\partial u_{mn}} \frac{\partial u_{mn}}{\partial x_{kl}} + \frac{\partial f_{ij}}{\partial v} \frac{\partial v}{\partial x_{kl}} \tag{D.5}$$

and then rewrite the formula in compact notation as

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} : \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}} + \frac{\partial \boldsymbol{f}}{\partial v} \otimes \frac{\partial v}{\partial \boldsymbol{x}} \tag{D.6}$$

The first product operation is a double-dot product of two fourth-order tensors while the second one is a direct product of two second-order tensors.

Sometimes it is convenient to use a special differential operator $\boldsymbol{\nabla}$ (read 'nabla') having the character of a first-order tensor with components $\nabla_i \equiv \partial/\partial x_i$. The direct product of $\boldsymbol{\nabla}$ with a tensor field gives the *gradient* of that field, and the dot product of $\boldsymbol{\nabla}$ with a tensor field gives the so-called *divergence* of that field. For example, the gradient of the scalar temperature field $T(\boldsymbol{x})$ is

$$\text{grad } T = \boldsymbol{\nabla} T = \frac{\partial T}{\partial \boldsymbol{x}} \tag{D.7}$$

and the divergence of the heat flux vector field $\boldsymbol{q}(\boldsymbol{x})$ is

$$\text{div } \boldsymbol{q} = \boldsymbol{\nabla} \cdot \boldsymbol{q} = \frac{\partial \boldsymbol{q}}{\partial \boldsymbol{x}} : \boldsymbol{\delta} \tag{D.8}$$

However, if the result is a tensor of a higher order than 1, one has to be careful about the sequence of the subscripts. For example, the gradient of the displacement field is a second-order tensor $\partial\boldsymbol{u}/\partial\boldsymbol{x}$ with components $(\partial\boldsymbol{u}/\partial\boldsymbol{x})_{ij} = \partial u_i/\partial x_j$, i.e. the first subscript determines the component u_i that is to be differentiated, and the second subscript determines the spatial coordinate x_j with respect to which the partial derivative should be taken. To get the same tensor using the symbol $\boldsymbol{\nabla}$, we would have to write $\partial\boldsymbol{u}/\partial\boldsymbol{x} = \boldsymbol{u}\boldsymbol{\nabla}$, but this might lead to confusion in more complicated expressions where it is not clear that the operator $\boldsymbol{\nabla}$ acts 'to the left'. It is therefore safer to work with the transposed gradient, $\boldsymbol{\nabla}\boldsymbol{u} = (\partial\boldsymbol{u}/\partial\boldsymbol{x})^T$. The divergence of a vector field is a scalar, and so it is irrelevant whether we write $\boldsymbol{\nabla} \cdot \boldsymbol{q}$ or $\boldsymbol{q} \cdot \boldsymbol{\nabla}$. The divergence of a second- or higher-order tensor field is not a scalar, and one has to be careful again. For example, the divergence of the stress field $\boldsymbol{\sigma}(\boldsymbol{x})$ (which will later appear in the differential equations of equilibrium (D.79)) is a first-order tensor div $\boldsymbol{\sigma}$ with components $(\text{div } \boldsymbol{\sigma})_i = \partial\sigma_{ij}/\partial x_j$. This can be written as

$$\text{div } \boldsymbol{\sigma} = \boldsymbol{\sigma} \cdot \boldsymbol{\nabla} = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{x}} : \boldsymbol{\delta} \tag{D.9}$$

and the operator $\boldsymbol{\nabla}$ again acts 'to the left'.

D.2 LINEAR ELASTIC STRESS-STRAIN RELATIONS

General anisotropic linear elastic stress-strain relations may be written as

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} \qquad \left[\text{or } \sigma_{ij} = D^e_{ijkl}\varepsilon_{kl}\right] \tag{D.10}$$

in which $\boldsymbol{\sigma}$ is the stress tensor, $\boldsymbol{\varepsilon}$ is the strain tensor and $\boldsymbol{D}_e$ is the *elastic material stiffness tensor* (or elastic moduli tensor) with components that are called the *elastic moduli*. Due to the symmetry of the stress tensor and strain tensor, the relation must not change when we interchange subscripts i, j or subscripts k, l. Therefore, the elastic stiffness tensor always satisfies the minor symmetry relations $D^e_{ijkl} = D^e_{jikl}$ and $D^e_{ijkl} = D^e_{ijlk}$. Furthermore, if an elastic potential, W, exists, it must be expressed as

$$W = \tfrac{1}{2}\boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{\varepsilon} = \tfrac{1}{2}D^e_{ijkl}\varepsilon_{ij}\varepsilon_{kl} \tag{D.11}$$

Then, the stress is equal to the gradient (with respect to $\boldsymbol{\varepsilon}$) of the elastic potential,

$$\boldsymbol{\sigma} = \frac{\partial W}{\partial \boldsymbol{\varepsilon}} \qquad \left[\text{or } \sigma_{ij} = \frac{\partial W}{\partial \varepsilon_{ij}}\right] \tag{D.12}$$

and the elastic stiffness tensor may be obtained from the elastic potential as

$$\boldsymbol{D}_e = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}} = \frac{\partial^2 W}{\partial \boldsymbol{\varepsilon}^2} \qquad \left[\text{or } D^e_{ijkl} = \frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} = \frac{\partial^2 W}{\partial \varepsilon_{ij}\partial \varepsilon_{kl}}\right] \tag{D.13}$$

Since the mixed derivatives $\partial^2 W/\partial\varepsilon_{ij}\partial\varepsilon_{kl}$ and $\partial^2 W/\partial\varepsilon_{kl}\partial\varepsilon_{ij}$ are equal, we must have $D^e_{ijkl} = D^e_{klij}$. These are the major symmetry relations valid for all *hyperelastic materials*, i.e. for materials with an elastic potential.

Often it is convenient to separate the effects of a volume change from those of a shape change. The *volumetric part* of a second-order tensor is obtained by projecting it onto the direction given by Kronecker delta, which represents a unit volumetric state. For example, the volumetric part of the strain tensor is $\varepsilon_V\boldsymbol{\delta}$, where $\varepsilon_V = \boldsymbol{\varepsilon} : \boldsymbol{\delta}/(\boldsymbol{\delta} : \boldsymbol{\delta}) = \varepsilon_{ii}/3$ is the mean (or volumetric) strain (summation implied due to repetition of subscript i). Similarly, the volumetric part of the stress tensor is $\sigma_V\boldsymbol{\delta}$, where $\sigma_V = \sigma_{ii}/3$ is the mean (or hydrostatic, or volumetric) stress. The difference between a tensor and its volumetric part is called the *deviatoric part* (deviator). The deviators of the stress and strain tensors are, respectively, defined as

$$\boldsymbol{s} = \boldsymbol{\sigma} - \sigma_V\boldsymbol{\delta} \tag{D.14}$$

$$\boldsymbol{e} = \boldsymbol{\varepsilon} - \varepsilon_V\boldsymbol{\delta} \tag{D.15}$$

The following relations are often useful when tensor expressions have to be simplified:

$$\delta_{ii} = \delta_{ij}\delta_{ij} = \boldsymbol{\delta} : \boldsymbol{\delta} = 3 \tag{D.16}$$

$$\sigma_{ii} = \sigma_{ij}\delta_{ij} = \boldsymbol{\sigma} : \boldsymbol{\delta} \tag{D.17}$$

$$\varepsilon_{ii} = \varepsilon_{ij}\delta_{ij} = \boldsymbol{\varepsilon} : \boldsymbol{\delta} \tag{D.18}$$

$$\begin{aligned} s_{ii} = s_{ij}\delta_{ij} &= \boldsymbol{s} : \boldsymbol{\delta} = (\boldsymbol{\sigma} - \sigma_V\boldsymbol{\delta}) : \boldsymbol{\delta} = \boldsymbol{\sigma} : \boldsymbol{\delta} - \sigma_V\boldsymbol{\delta} : \boldsymbol{\delta} \\ &= \sigma_{ii} - 3\sigma_V = 0 \end{aligned} \tag{D.19}$$

$$\begin{aligned} e_{ii} = e_{ij}\delta_{ij} &= \boldsymbol{e} : \boldsymbol{\delta} = (\boldsymbol{\varepsilon} - \varepsilon_V\boldsymbol{\delta}) : \boldsymbol{\delta} = \boldsymbol{\varepsilon} : \boldsymbol{\delta} - \varepsilon_V\boldsymbol{\delta} : \boldsymbol{\delta} \\ &= \varepsilon_{ii} - 3\varepsilon_V = 0 \end{aligned} \tag{D.20}$$

The identities $\boldsymbol{s} : \boldsymbol{\delta} = 0$ and $\boldsymbol{e} : \boldsymbol{\delta} = 0$ imply that the scalar product of any purely deviatoric tensor with any purely volumetric tensor (a multiple of $\boldsymbol{\delta}$) is equal to zero. In this sense, the volumetric-deviatoric split is a decomposition of the tensor into two mutually orthogonal components.

If an elastic material is *isotropic*, i.e. has the same properties in all directions, its elastic stress-strain relation may be given by independent equations for the volumetric components and the deviatoric components,

$$\sigma_V = 3K\varepsilon_V = K\boldsymbol{\delta} : \boldsymbol{\varepsilon} \tag{D.21}$$

$$\boldsymbol{s} = 2G\boldsymbol{e} \tag{D.22}$$

in which K is the *bulk modulus* and G is the *shear modulus*.

The elastic material is often characterized by its *Young's modulus*, E, and its *Poisson's ratio*, ν. The relations between K, G, E, and ν can easily be established by considering a state of unit hydrostatic stress $\boldsymbol{\sigma} = \boldsymbol{\delta}$ (Figure D.1(a)), for which the normal strain may be expressed as $\varepsilon_{11} = 1/3K = 1/E - 2\nu/E$, and a unit deviatoric stress state $\sigma_{11} = -\sigma_{22} = 1$ and $\sigma_{33} = 0$ (Figure D.1(b)), for which we may write $\varepsilon_{11} = 1/2G = 1/E + \nu/E$. From the preceding equalities we get the well-known relations

$$K = \frac{E}{3(1-2\nu)}, \qquad G = \frac{E}{2(1+\nu)} \tag{D.23}$$

It is also possible to derive an explicit expression for the isotropic elastic stiffness tensor $\boldsymbol{D}_e$. Combining constitutive equations (D.21)–(D.22) with relations (D.14)–(D.15) describing the volumetric-deviatoric split, we obtain

$$\begin{aligned} \boldsymbol{\sigma} &= \sigma_V\boldsymbol{\delta} + \boldsymbol{s} = 3K\varepsilon_V\boldsymbol{\delta} + 2G\boldsymbol{e} = 3K\varepsilon_V\boldsymbol{\delta} + 2G(\boldsymbol{\varepsilon} - \varepsilon_V\boldsymbol{\delta}) \\ &= (3K - 2G)\boldsymbol{\delta}\varepsilon_V + 2G\boldsymbol{\varepsilon} = \left(K - \tfrac{2}{3}G\right)\boldsymbol{\delta}\otimes\boldsymbol{\delta} : \boldsymbol{\varepsilon} + 2G\boldsymbol{I} : \boldsymbol{\varepsilon} \end{aligned} \tag{D.24}$$

The unit fourth-order tensor $\boldsymbol{I}$ with components $I_{ijkl} = \delta_{ik}\delta_{jl}$ does not exhibit minor symmetry. As $\boldsymbol{\varepsilon}$ is a symmetric second-order tensor, it is possible to replace $\boldsymbol{I}$ by $\boldsymbol{I}_s$ without affecting the value of the last term in (D.24). The elastic stiffness tensor can then be presented in a compact form

$$\boldsymbol{D}_e = \lambda\boldsymbol{\delta}\otimes\boldsymbol{\delta} + 2\mu\boldsymbol{I}_s, \qquad \left[\text{or } D^e_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\right] \tag{D.25}$$

where

$$\lambda = K - \frac{2}{3}G = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = G = \frac{E}{2(1+\nu)} \tag{D.26}$$

are the so-called *Lamé coefficients*. The formulae for conversions among three frequently used pairs of elastic constants, (E, ν), (K, G) and (λ, μ), are summarized in Table D.1. The table also presents expressions for three typical components of the elastic stiffness tensor, D^e_{1111}, D^e_{1122} and D^e_{1212}. The other nonzero components can be obtained by cyclic permutation of the subscripts, taking into account the rules of major and minor symmetry.

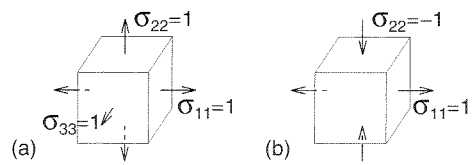


Figure D.1 a) Unit hydrostatic stress state, b) unit deviatoric stress state

Table D.1 Conversion formulae for elastic constants

	E, ν	K, G	λ, μ
E	E	$\dfrac{9KG}{3K+G}$	$\dfrac{(3\lambda+2\mu)\mu}{\lambda+\mu}$
ν	ν	$\dfrac{3K-2G}{6K+2G}$	$\dfrac{\lambda}{2(\lambda+\mu)}$
K	$\dfrac{E}{3(1-2\nu)}$	K	$\lambda+\dfrac{2}{3}\mu$
G	$\dfrac{E}{2(1+\nu)}$	G	μ
$\lambda \equiv D^e_{1122}$	$\dfrac{E\nu}{(1+\nu)(1-2\nu)}$	$K-\dfrac{2}{3}G$	λ
$\mu \equiv D^e_{1212}$	$\dfrac{E}{2(1+\nu)}$	G	μ
D^e_{1111}	$\dfrac{E(1-\nu)}{(1+\nu)(1-2\nu)}$	$K+\dfrac{4}{3}G$	$\lambda+2\mu$

D.3 STRESS INVARIANTS

The principal stresses may be defined as the normal stresses on planes of such orientation that the associated shear stresses vanish. This condition for principal stresses can be stated without reference to any particular coordinate system. Therefore, the principal stresses must be the same regardless of the coordinate system used, i.e. they must be independent of coordinate transformations. Realizing that the components of the traction vector $\boldsymbol{t}$ acting on any plane may be expressed as $\boldsymbol{t} = \boldsymbol{\sigma}\cdot\boldsymbol{n}$, where $\boldsymbol{\sigma}$ is the stress tensor and $\boldsymbol{n}$ is the unit vector normal to the plane considered (its components being the direction cosines), a condition characterizing the principal stress σ may be written as $\boldsymbol{\sigma}\cdot\boldsymbol{n} = \sigma\boldsymbol{n}$, which means that the traction must be perpendicular to the plane; σ represents the magnitude of the traction vector, i.e. the value of the principal stress. Substituting the identity $\boldsymbol{n} = \boldsymbol{\delta}\cdot\boldsymbol{n}$, we may rewrite this equation as

$$(\boldsymbol{\sigma} - \sigma\,\boldsymbol{\delta})\cdot\boldsymbol{n} = \boldsymbol{0} \qquad [\text{or } (\sigma_{ij} - \sigma\delta_{ij})n_j = 0] \tag{D.27}$$

This represents a system of three linear algebraic equations for the unknown direction cosines n_j, $j = 1, 2, 3$. The equations are homogeneous, and so a nonzero solution exists if and only if the determinant of the equation system vanishes, i.e.

$$\det(\boldsymbol{\sigma} - \sigma\boldsymbol{\delta}) = 0 \qquad [\text{or } \det(\sigma_{ij} - \sigma\delta_{ij}) = 0] \tag{D.28}$$

Expanding the determinant, we obtain a cubic equation for the magnitude of the principal stress σ, traditionally written in the form

$$-\sigma^3 + I_1\sigma^2 - I_2\sigma + I_3 = 0 \tag{D.29}$$

The coefficients I_1, I_2 and I_3 are given by

$$I_1 = \sigma_{11} + \sigma_{22} + \sigma_{33} = \sigma_{ii} = \boldsymbol{\sigma} : \boldsymbol{\delta} \tag{D.30}$$

$$I_2 = \begin{vmatrix}\sigma_{22} & \sigma_{23}\\ \sigma_{32} & \sigma_{33}\end{vmatrix} + \begin{vmatrix}\sigma_{11} & \sigma_{13}\\ \sigma_{31} & \sigma_{33}\end{vmatrix} + \begin{vmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{vmatrix} \tag{D.31}$$

$$I_3 = \begin{vmatrix}\sigma_{11} & \sigma_{12} & \sigma_{13}\\ \sigma_{21} & \sigma_{22} & \sigma_{23}\\ \sigma_{31} & \sigma_{32} & \sigma_{33}\end{vmatrix} \tag{D.32}$$

where $|\boldsymbol{A}|$ denotes the determinant of matrix $\boldsymbol{A}$.

The roots σ of (D.27) are in general called the eigenvalues (or characteristic values), and in algebra texts it is proven that they are real if the matrix corresponding to $\boldsymbol{\sigma}$ is symmetric (and positive, if the matrix is positive definite). This means in particular that, due to the symmetry of tensor $\boldsymbol{\sigma}$, the cubic equation (D.29) has three real roots, σ_1, σ_2 and σ_3 (eigenvalues of $\boldsymbol{\sigma}$), which are in solid mechanics called the *principal stresses*. Vectors $\boldsymbol{n}_1$, $\boldsymbol{n}_2$ and $\boldsymbol{n}_3$ associated with the three roots (in general called the eigenvectors) represent the *principal directions* and are mutually orthogonal.

Equation (D.29) can now be written in the equivalent form

$$-(\sigma - \sigma_1)(\sigma - \sigma_2)(\sigma - \sigma_3) = 0 \tag{D.33}$$

or

$$-\sigma^3 + (\sigma_1 + \sigma_2 + \sigma_3)\sigma^2 - (\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1)\sigma + \sigma_1\sigma_2\sigma_3 = 0 \tag{D.34}$$

Comparing this to (D.29), we obtain the following expressions for the coefficients I_1, I_2, and I_3:

$$I_1 = \sigma_1 + \sigma_2 + \sigma_3,\ I_2 = \sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1,\ I_3 = \sigma_1\sigma_2\sigma_3 \tag{D.35}$$

Now, since we concluded that the principal stress must be invariant due to a physical argument, it is clear that the coefficients of the cubic equation (D.29) must also be invariants. The quantities I_1, I_2, I_3 are called the *basic invariants* of the stress tensor. Any function of these invariants is, of course, an invariant too. The invariance of the expressions in (D.30)–(D.32) can be formally proven by coordinate transformations.

Let us now consider the invariants of the stress deviator $\boldsymbol{s}$ defined by (D.14). The equation for the principal deviatoric stresses must again have the same form as (D.29), however, the term corresponding to I_1 vanishes, since $s_{ii} = \boldsymbol{s} : \boldsymbol{\delta} = 0$; see (D.19). Thus, the equation may be written in the form

$$s^3 - J_2 s - J_3 = 0 \tag{D.36}$$

in which s is the principal deviatoric stress, and J_2 and J_3 are the same as I_2 and I_3 (see (D.30)) except that they are expressed in terms of s_{ij} rather than σ_{ij}, and that the sign of J_2 compared to I_2 has been changed. In particular, for the *second deviatoric invariant* we have, according to (D.31),

$$\begin{aligned} J_2 &= -\begin{vmatrix}s_{22} & s_{23}\\ s_{32} & s_{33}\end{vmatrix} - \begin{vmatrix}s_{11} & s_{13}\\ s_{31} & s_{33}\end{vmatrix} - \begin{vmatrix}s_{11} & s_{12}\\ s_{21} & s_{22}\end{vmatrix}\\ &= -s_{11}s_{22} - s_{22}s_{33} - s_{33}s_{11} + s_{12}^2 + s_{23}^2 + s_{31}^2 \end{aligned} \tag{D.37}$$

Since $s_{11} + s_{22} + s_{33} = \boldsymbol{s} : \boldsymbol{\delta} = 0$, the expression for J_2 may be transformed to

$$\begin{aligned} J_2 &= \tfrac{1}{2}\left[-2s_{11}s_{22} - 2s_{22}s_{33} - 2s_{33}s_{11} + (s_{11} + s_{22} + s_{33})^2\right]\\ &\quad + s_{12}^2 + s_{23}^2 + s_{31}^2\\ &= \tfrac{1}{2}(s_{11}^2 + s_{22}^2 + s_{33}^2) + s_{12}^2 + s_{23}^2 + s_{31}^2 = \tfrac{1}{2}s_{ij}s_{ij} = \tfrac{1}{2}\boldsymbol{s} : \boldsymbol{s} \end{aligned} \tag{D.39}$$

We see that the second deviatoric invariant can be expressed as a sum of squares, and so it is always non-negative. Note that an analogous expression in terms of $\boldsymbol{\sigma}$, namely $\boldsymbol{\sigma} : \boldsymbol{\sigma}/2$, is not a correct representation of I_2. Another useful expression for the second deviatoric invariant is

$$J_2 = \tfrac{1}{6}\left[(\sigma_{11}-\sigma_{22})^2 + (\sigma_{22}-\sigma_{33})^2 + (\sigma_{33}-\sigma_{11})^2\right] + \sigma_{12}^2 + \sigma_{23}^2 + \sigma_{31}^2 \tag{D.40}$$

or, in terms of principal stresses,

$$J_2 = \tfrac{1}{6}\left[(\sigma_1-\sigma_2)^2 + (\sigma_2-\sigma_3)^2 + (\sigma_3-\sigma_1)^2\right] \tag{D.41}$$

To prove (D.40), we realize that the bracketed term in (D.40) may be expressed as

$$\begin{aligned}
&\tfrac{1}{6}[(s_{11}-s_{22})^2 + (s_{22}-s_{33})^2 + (s_{33}-s_{11})^2] \\
&= \tfrac{1}{6}[2s_{11}^2 + 2s_{22}^2 + 2s_{33}^2 - 2s_{11}s_{22} - 2s_{22}s_{33} - 2s_{33}s_{11} + (s_{11}+s_{22}+s_{33})^2] \\
&= \tfrac{1}{6}(3s_{11}^2 + 3s_{22}^2 + 3s_{33}^2) = \tfrac{1}{2}(s_{11}^2 + s_{22}^2 + s_{33}^2)
\end{aligned} \tag{D.42}$$

The relations between the invariants can easily be obtained by considering their expressions in terms of principal stresses. For example, for the second deviatoric invariant we have from (D.37)

$$\begin{aligned}
J_2 &= -s_1 s_2 - s_2 s_3 - s_3 s_1 \\
&= -(\sigma_1-\sigma_V)(\sigma_2-\sigma_V) - (\sigma_2-\sigma_V)(\sigma_3-\sigma_V) - (\sigma_3-\sigma_V)(\sigma_1-\sigma_V) \\
&= 2\sigma_V(\sigma_1+\sigma_2+\sigma_3) - 3(\sigma_V)^2 - (\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_1\sigma_3) \\
&= 6\,(\sigma_V)^2 - 3\,(\sigma_V)^2 - I_2 = \tfrac{1}{3}I_1^2 - I_2
\end{aligned} \tag{D.43}$$

A useful tensorial expression for the third deviatoric invariant is

$$J_3 = \tfrac{1}{3}s_{ij}s_{jk}s_{ki} = \tfrac{1}{3}(\boldsymbol{s}\cdot\boldsymbol{s}) : \boldsymbol{s} \tag{D.44}$$

Recall that J_3 was defined as the determinant of $\boldsymbol{s}$. The equivalence of expression (D.44) with this definition is easily verified in the principal stress coordinates, in which only the normal components s_{11}, s_{22} and s_{33} are in general nonzero. Since $s_{33} = -s_{11} - s_{22}$, expression (D.44) gives $J_3 = [s_{11}^3 + s_{22}^3 - (s_{11}+s_{22})^3]/3$, which is equal to the determinant $\det \boldsymbol{s} = -s_{11}s_{22}(s_{11}+s_{22})$.

Since J_2 is always non-negative, we may expect it to be related to energy. The *strain energy density* (energy stored per unit volume of the linear elastic material) can be expressed as

$$\begin{aligned}
W &= \tfrac{1}{2}\boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{\varepsilon} = \tfrac{1}{2}\boldsymbol{\sigma} : \boldsymbol{\varepsilon} = \tfrac{1}{2}(\sigma_V\boldsymbol{\delta} + \boldsymbol{s}) : (\varepsilon_V\boldsymbol{\delta} + \boldsymbol{e}) \\
&= \tfrac{1}{2}(\sigma_V\varepsilon_V\boldsymbol{\delta}:\boldsymbol{\delta} + \sigma_V\boldsymbol{\delta}:\boldsymbol{e} + \varepsilon_V\boldsymbol{s}:\boldsymbol{\delta} + \boldsymbol{s}:\boldsymbol{e}) \\
&= \tfrac{3}{2}\sigma_V\varepsilon_V + \tfrac{1}{2}\boldsymbol{s}:\boldsymbol{e}
\end{aligned} \tag{D.45}$$

We see that the strain energy may be separated into two parts. The first part,

$$W_V = \frac{3}{2}\sigma_V\varepsilon_V = \frac{9}{2}K\varepsilon_V^2 = \frac{1}{2K}\sigma_V^2 \tag{D.46}$$

obviously represents the energy due to the change of volume. The second part,

$$W_D = \frac{1}{2}\boldsymbol{s}:\boldsymbol{e} = G\boldsymbol{e}:\boldsymbol{e} = \frac{1}{4G}\boldsymbol{s}:\boldsymbol{s} = \frac{1}{2G}J_2 \tag{D.47}$$

represents the energy due to the change of shape, called the *distortional energy*. The second deviatoric invariant J_2 is equal to the distortional energy times twice the shear modulus.

In plastic analysis we often have to differentiate a scalar function f (the yield function or the plastic potential) with respect to the stress tensor $\boldsymbol{\sigma}$ where the functional dependence of f on σ_{ij} is not given explicitly, but through the invariants, e.g. as $f(I_1, J_2)$. Let us demonstrate the application of the chain rule and the differentiation of invariants:

$$\frac{\partial f}{\partial \boldsymbol{\sigma}} = \frac{\partial f}{\partial I_1}\frac{\partial I_1}{\partial \boldsymbol{\sigma}} + \frac{\partial f}{\partial J_2}\frac{\partial J_2}{\partial \boldsymbol{\sigma}} \tag{D.48}$$

$$\frac{\partial I_1}{\partial \boldsymbol{\sigma}} = \frac{\partial}{\partial \boldsymbol{\sigma}}(\boldsymbol{\sigma}:\boldsymbol{\delta}) = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\sigma}} : \boldsymbol{\delta} = \boldsymbol{I}:\boldsymbol{\delta} = \boldsymbol{\delta} \tag{D.49}$$

$$\frac{\partial J_2}{\partial \boldsymbol{\sigma}} = \frac{\partial J_2}{\partial \boldsymbol{s}} : \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{\sigma}} \tag{D.50}$$

$$\begin{aligned}
\frac{\partial J_2}{\partial \boldsymbol{s}} &= \frac{\partial}{\partial \boldsymbol{s}}\left(\frac{1}{2}\boldsymbol{s}:\boldsymbol{s}\right) = \frac{1}{2}\left(\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{s}} : \boldsymbol{s} + \boldsymbol{s} : \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{s}}\right) \\
&= \tfrac{1}{2}\left(\boldsymbol{I}:\boldsymbol{s} + \boldsymbol{s}:\boldsymbol{I}\right)\tfrac{1}{2}(\boldsymbol{s}+\boldsymbol{s}) = \boldsymbol{s}
\end{aligned} \tag{D.51}$$

$$\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{\sigma}} = \frac{\partial}{\partial \boldsymbol{\sigma}}\left(\boldsymbol{\sigma} - \frac{1}{3}I_1\boldsymbol{\delta}\right) = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\sigma}} - \frac{1}{3}\frac{\partial I_1}{\partial \boldsymbol{\sigma}} \otimes \boldsymbol{\delta} = \boldsymbol{I} - \frac{1}{3}\boldsymbol{\delta}\otimes\boldsymbol{\delta} \tag{D.52}$$

Substituting the partial results back into (D.48) we finally get

$$\begin{aligned}
\frac{\partial f}{\partial \boldsymbol{\sigma}} &= \frac{\partial f}{\partial I_1}\frac{\partial I_1}{\partial \boldsymbol{\sigma}} + \frac{\partial f}{\partial J_2}\frac{\partial J_2}{\partial \boldsymbol{\sigma}} = \frac{\partial f}{\partial I_1}\boldsymbol{\delta} + \frac{\partial f}{\partial J_2}\boldsymbol{s} : \left(\boldsymbol{I} - \frac{1}{3}\boldsymbol{\delta}\otimes\boldsymbol{\delta}\right) \\
&= \frac{\partial f}{\partial I_1}\boldsymbol{\delta} + \frac{\partial f}{\partial J_2}\boldsymbol{s}
\end{aligned} \tag{D.53}$$

Note that the first term represents the volumetric part of $\partial f/\partial\boldsymbol{\sigma}$, while the second term is the deviatoric part. In the indicial form, the previous transformations are longer to write, but the advantage is that one does not need to remember the rules of symbolic tensor operations.

Differentiating a function that depends on the third deviatoric invariant, we also need an expression for $\partial J_3/\partial\boldsymbol{\sigma}$. The chain rule gives

$$\frac{\partial J_3}{\partial \boldsymbol{\sigma}} = \frac{\partial J_3}{\partial \boldsymbol{s}} : \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{\sigma}} \tag{D.54}$$

where $\partial\boldsymbol{s}/\partial\boldsymbol{\sigma}$ is already known from (D.52). The first term is easily evaluated using (D.44). This time, it is more convenient to work in the indicial notation, and write

$$\frac{\partial J_3}{\partial s_{pq}} = \frac{1}{3}\frac{\partial}{\partial s_{pq}}(s_{ij}s_{jk}s_{ki}) = \frac{1}{3}\left(s_{qk}s_{kp} + s_{ip}s_{qi} + s_{qj}s_{jp}\right) = s_{pk}s_{kq} = \boldsymbol{s}\cdot\boldsymbol{s} \tag{D.55}$$

Substituting this back into (D.54) we obtain

$$\frac{\partial J_3}{\partial \boldsymbol{\sigma}} = (\boldsymbol{s}\cdot\boldsymbol{s}) : \left(\boldsymbol{I} - \frac{1}{3}\boldsymbol{\delta}\otimes\boldsymbol{\delta}\right) = \boldsymbol{s}\cdot\boldsymbol{s} - \frac{1}{3}(\boldsymbol{s}:\boldsymbol{s})\boldsymbol{\delta} = \boldsymbol{s}\cdot\boldsymbol{s} - \frac{2}{3}J_2\boldsymbol{\delta} \tag{D.56}$$

Sometimes it is necessary to differentiate a function defined in terms of the principal stresses, σ_I, $I = 1, 2, 3$, and find its gradient with respect to the stress components.

The chain rule gives

$$\frac{\partial f}{\partial \boldsymbol{\sigma}} = \sum_{I=1}^{3} \frac{\partial f}{\partial \sigma_I} \frac{\partial \sigma_I}{\partial \boldsymbol{\sigma}} \tag{D.57}$$

Note that the summation convention does not apply to the capital subscript I, because I does not refer to a Cartesian component of a tensor but it labels three scalars, σ_1, σ_2 and σ_3. The gradients $\partial\sigma_I/\partial\boldsymbol{\sigma}$ could be found by applying the chain rule to the expressions for principal stresses in terms of stress invariants, to be explained in the next section. However, a more elegant and much simpler procedure is based on differentiation of the eigenproblem from which the principal stresses are computed. Let $\boldsymbol{n}_I$, $I = 1, 2, 3$, be the unit eigenvectors (principal directions) associated with principal stresses σ_I, $I = 1, 2, 3$. The principal stress σ_1 and principal direction $\boldsymbol{n}_1$ are one of the solutions of the eigenvalue problem

$$(\sigma_{ij} - \sigma_1 \delta_{ij}) n_{1j} = 0, \qquad i = 1, 2, 3 \tag{D.58}$$

We may regard the stress tensor $\boldsymbol{\sigma}$ as an independent variable and σ_1 and $\boldsymbol{n}_1$ as functions of $\boldsymbol{\sigma}$ implicitly defined by (D.58). Now we differentiate (D.58) with respect to $\boldsymbol{\sigma}$ and multiply the result from the left by $\boldsymbol{n}_1$. This leads to

$$n_{1i}\left(\delta_{ik}\delta_{jl} - \frac{\partial \sigma_1}{\partial \sigma_{kl}}\delta_{ij}\right) n_{1j} + n_{1i}(\sigma_{ij} - \sigma_1\delta_{ij})\frac{\partial n_{1j}}{\partial \sigma_{kl}} = 0, \qquad k, l = 1, 2, 3 \tag{D.59}$$

The second term on the left-hand side vanishes because of (D.58). Realizing that $n_{1i}\delta_{ik}\delta_{jl}n_{1j} = n_{1k}n_{1l}$ and $n_{1i}\delta_{ij}n_{1j} = n_{1i}n_{1i} = 1$, we can convert (D.59) to

$$n_{1k}n_{1l} - \frac{\partial \sigma_1}{\partial \sigma_{kl}} = 0 \tag{D.60}$$

Analogous relations hold for the other principal stresses, σ_2 and σ_3. Putting them all together, we finally have the result

$$\frac{\partial \sigma_I}{\partial \boldsymbol{\sigma}} = \boldsymbol{n}_I \otimes \boldsymbol{n}_I, \qquad I = 1, 2, 3 \qquad \text{(no sum on } I\text{)} \tag{D.61}$$

D.4 OCTAHEDRAL STRESS AND HAIGH-WESTERGAARD COORDINATES

The stress invariants have a useful interpretation in terms of stress components acting on the so-called *octahedral plane*, also called the *π-projection* (Figure D.2(a)). This is

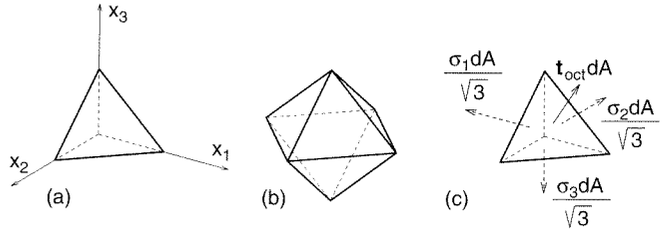


Figure D.2 a) Principal directions and octahedral plane, b) regular octahedron, c) equilibrium between principal stresses and octahedral stress

the plane which forms equal angles with the axes of principal stresses, and the unit normal is therefore given by

$$\boldsymbol{n} = \frac{1}{\sqrt{3}} \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \tag{D.62}$$

The term 'octahedral' stems from the fact that all the planes that have equal angles with the three Cartesian axes number eight in total and form a regular octahedron (Figure D.2(b)). The intersection of the octahedral plane with the coordinate planes forms an equilateral triangle, and denoting the area of this triangle by $\mathrm{d}A$, the projections of the triangle area on the coordinate planes are right-angle triangles of areas $\mathrm{d}A/\sqrt{3}$. Considering the conditions of equilibrium of forces acting on the element in Figure D.2(c) in the directions of x_1, x_2 and x_3, we find the vector of the traction acting on the octahedral plane to be

$$\boldsymbol{t}_{\mathrm{oct}} = \frac{1}{\sqrt{3}} \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{Bmatrix} \tag{D.63}$$

The normal component of the stress on the octahedral plane may be obtained as the scalar product of vectors $\boldsymbol{t}_{\mathrm{oct}}$ and $\boldsymbol{n}$, i.e.

$$\sigma_{\mathrm{oct}} = \boldsymbol{t}_{\mathrm{oct}} \cdot \boldsymbol{n} = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3} = \frac{1}{3} I_1 = \sigma_V \tag{D.64}$$

The magnitude τ_{oct} of the shear stress acting on the octahedral plane is obtained from the Pythagorean theorem;

$$\tau_{\mathrm{oct}}^2 = |\boldsymbol{t}_{\mathrm{oct}}|^2 - \sigma_{\mathrm{oct}}^2 = \frac{\sigma_1^2}{3} + \frac{\sigma_2^2}{3} + \frac{\sigma_3^2}{3} - \frac{(\sigma_1 + \sigma_2 + \sigma_3)^2}{9} \tag{D.65}$$

This expression may be further transformed as

$$\begin{aligned} 9\tau_{\mathrm{oct}}^2 &= 3\sigma_1^2 + 3\sigma_2^2 + 3\sigma_3^2 - \sigma_1^2 - \sigma_2^2 - \sigma_3^2 - 2\sigma_1\sigma_2 - 2\sigma_2\sigma_3 - 2\sigma_1\sigma_3 \\ &= 2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 - \sigma_1\sigma_2 - \sigma_2\sigma_3 - \sigma_1\sigma_3) \\ &= (\sigma_1 - \sigma_2)^2 + (\sigma_1 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 = 6J_2 \end{aligned} \tag{D.66}$$

This shows that

$$J_2 = \frac{3}{2}\tau_{\mathrm{oct}}^2, \quad \tau_{\mathrm{oct}} = \sqrt{\frac{2J_2}{3}} \tag{D.67}$$

Thus, we can see that the first stress invariant and the second stress deviator invariant may be expressed in terms of the normal and shear octahedral stress components, respectively. The direction of the vector of the octahedral shear stress is further related to the third deviatoric invariant, J_3. This direction can be described by the so-called *Lode angle*, θ, defined by the relation

$$\cos 3\theta = \frac{3\sqrt{3}}{2} \frac{J_3}{J_2^{3/2}} \tag{D.68}$$

The Lode angle is the deviatoric projection of the angle between the radius vector of the current stress point and the axis σ_1. In the interval $\langle 0, 2\pi \rangle$ there exist (in a generic

case) six angles θ satisfying (D.68). This is related to the fact that a given stress state can be represented by six different points in the principal stress space, provided that the principal stresses are mutually different.

Some plasticity models use the so-called *Haigh–Westergaard coordinates* in the principal stress space. These coordinates are usually denoted as ξ, ρ, and θ where

$$\xi = \frac{I_1}{\sqrt{3}} = \sqrt{3}\,\sigma_V \tag{D.69}$$

$$\rho = \sqrt{2J_2} = \sqrt{3}\,\tau_{\text{oct}} \tag{D.70}$$

and θ is the Lode angle defined by (D.68). Coordinate ξ is proportional to the mean stress, and the scaling factor is chosen such that ξ equals the distance of the hydrostatic projection of the stress point from the origin. The second coordinate, ρ, is proportional to the octahedral shear stress, and it equals the distance of the stress point from the hydrostatic axis. Finally, θ is the polar angle in the deviatoric projection. The Haigh–Westergaard coordinates are therefore cylindrical coordinates in the principal stress state, with coordinate ξ measured along the hydrostatic axis, coordinate ρ as the radius, and coordinate θ as the polar angle, measured from the projection of the first principal axis onto the deviatoric plane. Consequently, the principal stresses can be expressed as

$$\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\end{Bmatrix} = \frac{\xi}{\sqrt{3}}\begin{Bmatrix}1\\ 1\\ 1\end{Bmatrix} + \sqrt{\frac{2}{3}}\rho\begin{Bmatrix}\cos\theta\\ \cos(\theta-2\pi/3)\\ \cos(\theta+2\pi/3)\end{Bmatrix}, \tag{D.71}$$

If $0 \le \theta \le \pi/3$ then the principal stresses are ordered, $\sigma_1 \ge \sigma_2 \ge \sigma_3$. Other values of θ satisfying (D.68) correspond to the representations of the same stress state by points in other sextants of the principal stress space. Formulae (D.68)–(D.71) provide a closed-form solution of the cubic equation (D.29), and verify that all the roots are real.

D.5 FUNDAMENTAL EQUATIONS OF LINEAR ELASTICITY

Under the classical assumptions of small displacements and a linear relationship between strains and stresses, the three fundamental sets of equations describing a linear elastic body can be written as follows:

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \ldots \text{strain-displacement (kinematic) equations} \tag{D.72}$$

$$\sigma_{ij} = D^e_{ijkl}\varepsilon_{kl} \ldots \text{stress-strain (constitutive) equations} \tag{D.73}$$

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \bar{b}_i = 0 \ldots \text{equilibrium (static) equations} \tag{D.74}$$

These equations must be supplemented by appropriate boundary conditions:

$$u_i = \bar{u}_i \ \text{ on } S_u \quad \ldots \text{essential (kinematic) boundary conditions} \tag{D.75}$$

$$\sigma_{ij}n_j = \bar{t}_i \ \text{ on } S_t \quad \ldots \text{natural (static) boundary conditions} \tag{D.76}$$

The symbols with an overbar denote the given quantities – the body forces $\bar{\boldsymbol{b}}$ prescribed inside the domain V that represents the body under consideration, the surface

tractions $\bar{\boldsymbol{t}}$ prescribed on a part S_t of the boundary of V, and the displacements $\bar{\boldsymbol{u}}$ prescribed on the remaining part S_u of the boundary of V.

Sometimes it is convenient to write the basic equations (D.72)–(D.74) and the boundary conditions (D.75)–(D.76) in a compact notation

$$\boldsymbol{\varepsilon} = (\boldsymbol{\nabla u})_{\text{sym}} \tag{D.77}$$

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} \tag{D.78}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{\nabla} + \bar{\boldsymbol{b}} = \boldsymbol{0} \tag{D.79}$$

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on } S_u \tag{D.80}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = \bar{\boldsymbol{t}} \quad \text{on } S_t \tag{D.81}$$

In (D.77) and (D.79), $\boldsymbol{\nabla}$ is the differential operator introduced in Section D.1, $(\boldsymbol{\nabla u})_{\text{sym}}$ is the symmetric gradient of the displacement field, and $\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$ is the divergence of the stress field.

To emphasize the duality between the kinematic and equilibrium equations, (D.77) and (D.79) could be written in an alternative form

$$\boldsymbol{\varepsilon} = \boldsymbol{\partial u} \tag{D.82}$$

$$\boldsymbol{\partial}^*\boldsymbol{\sigma} = \bar{\boldsymbol{b}} \tag{D.83}$$

where $\boldsymbol{\partial}$ and $\boldsymbol{\partial}^*$ are differential operators defined by

$$\boldsymbol{\partial u} = (\boldsymbol{\nabla u})_{\text{sym}} = \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}\right)_{\text{sym}} = \frac{1}{2}\left[\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}} + \left(\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}\right)^T\right] \tag{D.84}$$

$$\boldsymbol{\partial}^*\boldsymbol{\sigma} = -\boldsymbol{\sigma}\cdot\boldsymbol{\nabla} = -\frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{x}} : \boldsymbol{\delta} \tag{D.85}$$

The relation between differential operators $\boldsymbol{\partial}$ and $\boldsymbol{\partial}^*$ is similar to the relation between algebraic operators represented by the kinematic matrix $\boldsymbol{B}$ and the static matrix $\boldsymbol{B}^T$. This will become clear in the next section.

Equations (D.79) are the conditions of static equilibrium. They can be extended to the general dynamic behavior by adding the inertial forces, $-\rho\ddot{\boldsymbol{u}}$, considered in the sense of the d'Alembert principle as a special type of body forces. Here, ρ is the density (mass per unit volume), and $\ddot{\boldsymbol{u}}$ is the acceleration vector, i.e. the second time derivative of the displacement vector. After this generalization, we obtain from (D.79) the *Euler–Cauchy equations of motion* (also called the *momentum equations* because they express the balance of momentum):

$$\boldsymbol{\sigma}\cdot\boldsymbol{\nabla} + \bar{\boldsymbol{b}} = \rho\ddot{\boldsymbol{u}} \tag{D.86}$$

D.6 CLAPEYRON THEOREM

Consider now an arbitrary (but sufficiently smooth) displacement field $\boldsymbol{u}_k$ and an arbitrary (but sufficiently smooth) stress field $\boldsymbol{\sigma}_s$. Substituting $\boldsymbol{u}_k$ into the kinematic equations, we can find the corresponding strain field $\boldsymbol{\varepsilon}_k$, and substituting $\boldsymbol{\sigma}_s$ into the static equations and static boundary conditions, we can find the corresponding body forces $\boldsymbol{b}_s$ and surface tractions $\boldsymbol{t}_s$. Applying the Gauss theorem to the integral of the

scalar product of $\boldsymbol{\sigma}_s$ and $\boldsymbol{\varepsilon}_k$, and using the kinematic and static equations and the symmetry of the stress tensor, we arrive at the following equality:

$$\begin{aligned}\int_V \boldsymbol{\sigma}_s : \boldsymbol{\varepsilon}_k \,\mathrm{d}V &= \int_V \sigma_{ij}^s \varepsilon_{ij}^k \,\mathrm{d}V = \frac{1}{2}\int_V \sigma_{ij}^s \left(\frac{\partial u_i^k}{\partial x_j} + \frac{\partial u_j^k}{\partial x_i}\right) \mathrm{d}V \\ &= \frac{1}{2}\int_S \left(\sigma_{ij}^s u_i^k n_j + \sigma_{ij}^s u_j^k n_i\right) \mathrm{d}S \\ &\quad -\frac{1}{2}\int_V \left(\frac{\partial \sigma_{ij}^s}{\partial x_j} u_i^k + \frac{\partial \sigma_{ij}^s}{\partial x_i} u_j^k\right) \mathrm{d}V \\ &= \frac{1}{2}\int_S \left(t_i^s u_i^k + t_j^s u_j^k\right) \mathrm{d}S + \frac{1}{2}\int_V \left(b_i^s u_i^k + b_j^s u_j^k\right) \mathrm{d}V \\ &= \int_S t_i^s u_i^k \,\mathrm{d}S + \int_V b_i^s u_i^k \,\mathrm{d}V \\ &= \int_S \boldsymbol{t}_s \cdot \boldsymbol{u}_k \,\mathrm{d}S + \int_V \boldsymbol{b}_s \cdot \boldsymbol{u}_k \,\mathrm{d}V \end{aligned} \tag{D.87}$$

The left-hand side has the meaning of the (totally fictitious) internal work that would be done by the stresses $\boldsymbol{\sigma}_s$ on the strains $\boldsymbol{\varepsilon}_k$, while the right-hand side has the meaning of the (totally fictitious) external work that would be done by the body forces $\boldsymbol{b}_s$ and surface tractions $\boldsymbol{t}_s$ on the displacements $\boldsymbol{u}_k$. The work is fictitious because the kinematic state (displacements and strains) that we consider here is completely independent of the static state (stresses, body forces and surface tractions). The identity

$$\int_V \boldsymbol{\sigma}_s : \boldsymbol{\varepsilon}_k \,\mathrm{d}V = \int_S \boldsymbol{t}_s \cdot \boldsymbol{u}_k \,\mathrm{d}S + \int_V \boldsymbol{b}_s \cdot \boldsymbol{u}_k \,\mathrm{d}V \tag{D.88}$$

valid under the above assumptions is called the *Clapeyron theorem*. This is quite a general statement that can be particularized to a number of useful applications. For example, adding the assumptions that the displacements satisfy homogeneous boundary conditions ($\boldsymbol{u} = \boldsymbol{0}$) on S_u and that the stresses are in equilibrium with the prescribed body forces in V and prescribed surface tractions on S_t (with the actual ones, not with fictitious ones) we get the *virtual work equality*

$$\int_V \boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon} \,\mathrm{d}V = \int_{S_t} \bar{\boldsymbol{t}} \cdot \delta\boldsymbol{u} \,\mathrm{d}S + \int_V \bar{\boldsymbol{b}} \cdot \delta\boldsymbol{u} \,\mathrm{d}V \tag{D.89}$$

As another example, if we add the assumption that the stress field is self-equilibrated (it is in equilibrium with zero body forces and zero surface tractions), we get the identity

$$\int_V \boldsymbol{\sigma}_s : \boldsymbol{\varepsilon}_k \,\mathrm{d}V = 0 \tag{D.90}$$

indicating that the work done by any self-equlibrated stress field on any compatible strain field is zero. Recall that a strain field is called compatible if there exists a displacement field such that the kinematic equations (D.72) hold.

The essential step in the derivation of the Clapeyron theorem is the application of the Gauss theorem, which shifts the differential operator from the displacement field to the stress field. In the compact notation this step could be presented as

$$\int_V \boldsymbol{\sigma} : \boldsymbol{\partial}\boldsymbol{u} \,\mathrm{d}V = \int_S \boldsymbol{n} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{u} \,\mathrm{d}S + \int_V \boldsymbol{\partial}^* \boldsymbol{\sigma} \cdot \boldsymbol{u} \,\mathrm{d}V \tag{D.91}$$

For homogeneous boundary conditions ($\boldsymbol{u} = \boldsymbol{0}$ on S_u and $\boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{0}$ on S_t), the surface integral vanishes and we obtain the identity

$$\int_V \boldsymbol{\sigma} : \boldsymbol{\partial}\boldsymbol{u} \,\mathrm{d}V = \int_V \boldsymbol{\partial}^* \boldsymbol{\sigma} \cdot \boldsymbol{u} \,\mathrm{d}V \tag{D.92}$$

which is analogous to the identity

$$\boldsymbol{s}^T(\boldsymbol{B}\boldsymbol{d}) = (\boldsymbol{B}^T\boldsymbol{s})^T\boldsymbol{d} \tag{D.93}$$

valid for discrete structural models (see Appendices A and B). Operator $\boldsymbol{\partial}^*$ is said to be the *adjoint operator* of $\boldsymbol{\partial}$. Now it is clear why we included the negative sign in the definition of $\boldsymbol{\partial}^*$; see (D.85).

D.7 ENGINEERING NOTATION IN MATRIX FORM

The tensorial notation is certainly elegant and most efficient in theoretical derivations. However, for developing a numerical algorithm to be implemented into a computer code, it is more practical to work with a different type of notation. Even though nowadays there are computer libraries that provide direct access to tensorial operations implemented in an efficient manner (e.g. Jeremić and Sture, 1998), the conventional approach is to store the stress and strain components in one-dimensional arrays and the stiffness moduli in two-dimensional arrays. The development of a computer code is then facilitated if the basic formulae are written with the stress and strain components arranged in column matrices and the stiffness components in square matrices. This is the notation normally used in engineering textbooks, and we will call it the *engineering notation*. The representation of the material stiffness tensor by a square matrix is sometimes called the *Voigt notation*.

Using the engineering notation, one must be careful about the ordering of components. The normal components are usually arranged in the natural order (i.e. σ_x followed by σ_y and σ_z), but for the shear components there are several conventions. In principle it is possible to use any of them, but it is extremely important to select one convention and stick to it throughout the entire project. One possibility is to set

$$\boldsymbol{\sigma} = \begin{Bmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \\ \tau_{yz} \\ \tau_{zx} \\ \tau_{xy} \end{Bmatrix} = \begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{Bmatrix}, \qquad \boldsymbol{\varepsilon} = \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \\ \gamma_{yz} \\ \gamma_{zx} \\ \gamma_{xy} \end{Bmatrix} = \begin{Bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \mathbf{2}\varepsilon_{23} \\ \mathbf{2}\varepsilon_{31} \\ \mathbf{2}\varepsilon_{12} \end{Bmatrix} \tag{D.94}$$

There is no standard and universally accepted way of denoting whether the symbol $\boldsymbol{\sigma}$ means the stress tensor or the column matrix of stress components. The physical content is the same, but from the mathematical point of view these objects are different. This is not only a formal detail but a potential source of serious errors in computer programs, as will be seen later.

Note that the engineering shear component γ_{xy} is twice the tensorial shear strain ε_{12}, etc. The reason for introducing the factor of 2 is not only the fact that strain component γ_{xy} defined in this way is endowed with the physical meaning of the shear angle. More importantly, we want the energy product $\boldsymbol{\sigma} : \boldsymbol{\varepsilon} = \sigma_{ij}\varepsilon_{ij}$ to be replaced

in the engineering notation by a simple scalar product of column matrices, $\boldsymbol{\sigma}^T\boldsymbol{\varepsilon}$. Unfortunately, this simple transcription of the double contraction of two second-order tensors works only if we deal with one stress-like tensor and one strain-like tensor (e.g. with the back stress and plastic strain). To evaluate the tensorial norm of the stress tensor, defined as $\|\boldsymbol{\sigma}\| = \sqrt{\boldsymbol{\sigma} : \boldsymbol{\sigma}}$, we must be careful, and realize that

$$\begin{aligned}\boldsymbol{\sigma} : \boldsymbol{\sigma} = \sigma_{ij}\sigma_{ij} &= \sigma_{11}^2 + \sigma_{22}^2 + \sigma_{33}^2 + \sigma_{23}^2 + \sigma_{32}^2 + \sigma_{31}^2 + \sigma_{13}^2 + \sigma_{12}^2 + \sigma_{21}^2 \\ &= \sigma_{11}^2 + \sigma_{22}^2 + \sigma_{33}^2 + 2\left(\sigma_{23}^2 + \sigma_{31}^2 + \sigma_{12}^2\right)\end{aligned} \tag{D.95}$$

while

$$\boldsymbol{\sigma}^T\boldsymbol{\sigma} = \sigma_{11}^2 + \sigma_{22}^2 + \sigma_{33}^2 + \sigma_{23}^2 + \sigma_{31}^2 + \sigma_{12}^2 \tag{D.96}$$

So it is *not* possible to replace the tensorial norm of the stress tensor by the Euclidean norm of the column matrix of stress components! Instead, we must insert a diagonal scaling matrix

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \tag{D.97}$$

and express the tensorial norm in the engineering notation as

$$\|\boldsymbol{\sigma}\|_\sigma = \sqrt{\boldsymbol{\sigma}^T\boldsymbol{P}\boldsymbol{\sigma}} \tag{D.98}$$

In particular, the second invariant of the deviatoric stress $\boldsymbol{s}$ is in the tensorial notation defined as $J_2 = \boldsymbol{s} : \boldsymbol{s}/2$, but in the engineering notation we have to write

$$J_2 = \tfrac{1}{2}\boldsymbol{s}^T\boldsymbol{P}\boldsymbol{s} \tag{D.99}$$

Another way to understand the discrepancy is to note that $s_{ij}s_{ij}$ can be represented as the scalar product of two nine-dimensional vectors (formed from the nine components of s_{ij}), while the stress matrix in engineering notation belongs to a six-dimensional space.

An even more dangerous situation arises for the gradient of the J_2-invariant. In the tensorial notation we have $\partial J_2/\partial \boldsymbol{s} = \boldsymbol{s}$ (which can be interpreted as the differentiation of J_2 with respect to the nine-dimensional vector of components s_{ij}). However, in the engineering notation we must write $\partial J_2/\partial \boldsymbol{s} = \boldsymbol{P}\boldsymbol{s}$, as is easily verified by differentation of (D.99). The double dot in the product $\boldsymbol{s} : \boldsymbol{s}$ clearly indicates that $\boldsymbol{s}$ stands for a tensor, and this gives us a hint that in the engineering transcription we may have to modify something (insert the scaling matrix $\boldsymbol{P}$). However, the equality $\partial J_2/\partial \boldsymbol{s} = \boldsymbol{s}$ can be interpreted in two different ways, but only one interpretation is correct.

When transcribing the double contraction of two strain-like tensors, it is also necessary to insert a scaling matrix, but not the same one as for two stress-like tensors. Since the shear components have already been doubled, the corresponding scaling factors are now 1/2 instead of 2. So it turns out that the scaling matrix is the inverse of $\boldsymbol{P}$ and the tensorial norm of $\boldsymbol{\varepsilon}$ is in the engineering notation evaluated as $\sqrt{\boldsymbol{\varepsilon}^T\boldsymbol{P}^{-1}\boldsymbol{\varepsilon}}$. This is important, for example, for the evaluation of the equivalent plastic strain rate according to (20.5).

For the purpose of volumetric-deviatoric decomposition we need to establish the engineering counterpart of the unit second-order tensor (Kronecker delta). This is a tensor corresponding to a unit hydrostatic state, represented by a column matrix

$$\boldsymbol{\delta} = \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \tag{D.100}$$

Since the shear components are zero, $\boldsymbol{\delta}$ can have the character of both a stress-like quantity and a strain-like quantity. Fortunately, multiplication by the scaling matrix $\boldsymbol{P}$ or by its inverse does not change $\boldsymbol{\delta}$, so no problems can arise. The volumetric stress and strain can be expressed as

$$\sigma_V = \tfrac{1}{3}\boldsymbol{\sigma}^T\boldsymbol{\delta} = \tfrac{1}{3}\boldsymbol{\delta}^T\boldsymbol{\sigma}, \qquad \varepsilon_V = \tfrac{1}{3}\boldsymbol{\varepsilon}^T\boldsymbol{\delta} = \tfrac{1}{3}\boldsymbol{\delta}^T\boldsymbol{\varepsilon} \tag{D.101}$$

and the volumetric-deviatoric decomposition

$$\boldsymbol{\sigma} = \boldsymbol{\delta}\sigma_V + \boldsymbol{s}, \qquad \boldsymbol{\varepsilon} = \boldsymbol{\delta}\varepsilon_V + \boldsymbol{e} \tag{D.102}$$

has exactly the same form as in the tensorial notation. Substituting from (D.101) we obtain

$$\boldsymbol{s} = \boldsymbol{\sigma} - \boldsymbol{\delta}\sigma_V = \boldsymbol{\sigma} - \tfrac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T\boldsymbol{\sigma} = \left(\boldsymbol{I} - \tfrac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T\right)\boldsymbol{\sigma} = \boldsymbol{I}_D\boldsymbol{\sigma} \tag{D.103}$$

where

$$\boldsymbol{I}_D = \boldsymbol{I} - \tfrac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T \tag{D.104}$$

is the *deviatoric projection matrix*. Similarly, for the strain deviator we can write

$$\boldsymbol{e} = \boldsymbol{I}_D\boldsymbol{\varepsilon} \tag{D.105}$$

Taking into account that $\boldsymbol{\delta}^T\boldsymbol{\delta} = 3$, which is the engineering notation counterpart of (D.16), it is easy to prove that $\boldsymbol{I}_D\boldsymbol{\delta} = \boldsymbol{0}$, and that the matrix $\boldsymbol{I}_D$ is idempotent, i.e. $\boldsymbol{I}_D\boldsymbol{I}_D = \boldsymbol{I}_D$. This is related to the fact that the deviatoric part of stress or strain is not modified by another deviatoric projection and that its volumetric part is zero. We could also denote

$$\boldsymbol{I}_V = \boldsymbol{I} - \boldsymbol{I}_D = \tfrac{1}{3}\boldsymbol{\delta}\boldsymbol{\delta}^T \tag{D.106}$$

and call it the *volumetric projection matrix*. The volumetric-deviatoric split then becomes

$$\boldsymbol{\sigma} = (\boldsymbol{I}_V + \boldsymbol{I}_D)\boldsymbol{\sigma} = \boldsymbol{I}_V\boldsymbol{\sigma} + \boldsymbol{I}_D\boldsymbol{\sigma} = (\boldsymbol{\sigma})_{\text{vol}} + (\boldsymbol{\sigma})_{\text{dev}} = \boldsymbol{\delta}\sigma_V + \boldsymbol{s} \tag{D.107}$$

Another pitfall is hidden in the transcription of the constitutive equations of linear elasticity. The counterpart of (D.21) is $\sigma_V = K\boldsymbol{\delta}^T\boldsymbol{\varepsilon}$. However, even though the multiplication by a scalar in (D.22) seems to be a harmless operation, the relation $\boldsymbol{s} = 2G\boldsymbol{e}$ is not valid in the engineering notation. We need to compensate for the fact that the engineering shear components are twice the tensorial shear components

(i.e. $\tau_{xy} = 2G\varepsilon_{xy} = G\gamma_{xy}$) and write $\boldsymbol{s} = 2G\boldsymbol{P}^{-1}\boldsymbol{e}$. Consequently, the engineering transcription of (D.25) is not $\boldsymbol{D}_e = \lambda\boldsymbol{\delta}\boldsymbol{\delta}^T + 2\mu\boldsymbol{I}$, but

$$\boldsymbol{D}_e = \lambda\boldsymbol{\delta}\boldsymbol{\delta}^T + 2\mu\boldsymbol{P}^{-1} = \begin{bmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \tag{D.108}$$

Substituting from (D.26) we obtain the well-known formula

$$\boldsymbol{D}_e = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \dfrac{1-2\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \dfrac{1-2\nu}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \dfrac{1-2\nu}{2} \end{bmatrix} \tag{D.109}$$

This is the inverse of the matrix of elastic compliances,

$$\boldsymbol{C}_e = \boldsymbol{D}_e^{-1} = \frac{1}{E} \begin{bmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ -\nu & 1 & -\nu & 0 & 0 & 0 \\ -\nu & -\nu & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(1+\nu) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(1+\nu) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(1+\nu) \end{bmatrix} \tag{D.110}$$

The elastic constitutive equation (D.78) is represented in the engineering notation by

$$\boldsymbol{\sigma} = \boldsymbol{D}_e\boldsymbol{\varepsilon} \tag{D.111}$$

The kinematic operator from (D.84) in the engineering notation becomes

$$\boldsymbol{\partial} = \begin{bmatrix} \dfrac{\partial}{\partial x} & 0 & 0 \\ 0 & \dfrac{\partial}{\partial y} & 0 \\ 0 & 0 & \dfrac{\partial}{\partial z} \\ 0 & \dfrac{\partial}{\partial z} & \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial z} & 0 & \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} & 0 \end{bmatrix} \tag{D.112}$$

and its adjoint is $\boldsymbol{\partial}^* = -\boldsymbol{\partial}^T$. It is easy to rewrite the Clapeyron theorem (D.88) in the form

$$\int_V \boldsymbol{\sigma}_s^T\boldsymbol{\varepsilon}_k \,\mathrm{d}V = \int_S \boldsymbol{t}_s^T\boldsymbol{u}_k \,\mathrm{d}S + \int_V \boldsymbol{b}_s^T\boldsymbol{u}_k \,\mathrm{d}V \tag{D.113}$$

and the virtual work equality (D.89) in the form

$$\int_V \boldsymbol{\sigma}^T\delta\boldsymbol{\varepsilon} \,\mathrm{d}V = \int_{S_t} \bar{\boldsymbol{t}}^T\delta\boldsymbol{u} \,\mathrm{d}S + \int_V \bar{\boldsymbol{b}}^T\delta\boldsymbol{u} \,\mathrm{d}V \tag{D.114}$$

For problems formulated in a reduced space, such as that corresponding to plane stress, plane strain or axial symmetry, all the matrices can be reduced accordingly. For example, for a plane stress problem in the (x, y)-plane, we set $\boldsymbol{\varepsilon} = \{\varepsilon_x, \varepsilon_y, \gamma_{xy}\}^T = \{\varepsilon_{11}, \varepsilon_{22}, 2\varepsilon_{12}\}^T$ and $\boldsymbol{\sigma} = \{\sigma_x, \sigma_y, \tau_{xy}\}^T$. The scaling matrix is then

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{D.115}$$

The matrix of stiffness moduli (D.109) is, for plane stress, replaced by

$$\boldsymbol{D}_e = \frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \dfrac{1-\nu}{2} \end{bmatrix} \tag{D.116}$$

and the kinematic operator (D.112) is replaced by

$$\boldsymbol{\partial} = \begin{bmatrix} \dfrac{\partial}{\partial x} & 0 \\ 0 & \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} \end{bmatrix} \tag{D.117}$$

Instead of remembering the rules of symbolic tensor operations and their matrix transcription, it is sometimes preferred, and often safer, to conduct all the derivations in the indicial tensor notation and at the end transcribe the results to the matrix form.

Appendix E

Model B3 for Predicting Concrete Creep and Shrinkage

E.1 FRAMEWORKS FOR MODELING OF DRYING EFFECTS

Among many models for predicting the creep and shrinkage properties from the basic characteristics of concrete (e.g. Bažant, 1988, 1999), one model, called *model B3* (Bažant and Baweja, 1995, 2000a), has been based on a systematic theoretical formulation of the basic physical phenomena involved, and on statistical optimization with regard to most of the test data that exist in the literature. This model will now be briefly outlined and some of its assumptions described.

Model B3 covers creep and shrinkage of concrete, including their coupling. The effects of drying (i.e. drying shrinkage and drying creep) can be taken into account at two different levels of accuracy and sophistication:

1. The *'sectional' approach* only takes into account the mean effects of drying averaged over the cross section of a beam or slab. The environmental humidity must be given, but the time evolution of the pore humidity distribution throughout the structure is not solved. The influence of the cross section size and shape on the kinetics of drying is only incorporated approximately, through semiempirical coefficients. The self-equilibrated stresses caused by nonuniform drying cannot be captured. This approach is useful for simple engineering calculations.
2. The *'material' approach* takes into account the time evolution of the distribution of pore humidity throughout the structure, which must be obtained by solving a nonlinear diffusion equation, with the environmental humidity providing the boundary conditions. Once the evolution of pore humidity distribution is known, its effect on the constitutive behavior of the material can be evaluated from a physical law that links the increments of shrinkage strains to the changes of pore humidity. This approach is useful for more accurate (and more demanding) analyses of shrinkage-sensitive structures.

The 'sectional' approach defines the compliance function of concrete in the form

$$J(t,t') = q_1 + C_0(t,t') + C_d(t,t';t_0) \tag{E.1}$$

where $q_1 = 1/E_0$ is the inverse of the asymptotic elastic modulus (cf. Section 28.3.1), $C_0(t,t')$ is the *basic creep compliance*, and $C_d(t,t';t_0)$ is the *additional compliance due to drying*, which is influenced by the time t_0 at the beginning of drying. The average drying shrinkage strain in a cross section, $\varepsilon_{sh}(t)$, is estimated using formulae (28.10)–(28.15) given in Section 28.3.2. This means that the full model B3 differs from

its short version explained in Chapter 26 only by the expressions for basic creep and drying creep, which will be given in Section E.2.

The 'material' approach drops the compliance C_d from (E.1), i.e. the compliance function only reflects the instantaneous elastic response and basic creep. The shrinkage strain is now defined pointwise (i.e. not as an average over the cross section), and it represents not only the drying shrinkage, but also the stress-induced shrinkage (alias drying creep, also called the Pickett effect).

Having introduced the general ideas, we can now present the specific formulae for creep and the effects of drying.

E.2 MODEL B3

E.2.1 Basic Creep Compliance Function

The basic creep compliance is more conveniently defined by its time rate than its accumulated value:

$$\frac{\partial C_0(t,t')}{\partial t} = \frac{n(q_2 t^{-m} + q_3)}{(t-t') + (t-t')^{1-n}} + \frac{q_4}{t} \tag{E.2}$$

in which t and t' must be in days, $m = 0.5$ and $n = 0.1$ are empirical parameters whose values can be taken the same for all normal concretes, and q_2, q_3 and q_4 are empirical constitutive parameters whose prediction will be discussed later. The total basic creep compliance is obtained by integrating (E.2) with the initial condition $C_0(t',t') = 0$ (because the instantaneous compliance is taken into account by the constant term q_1 in (E.1)). The functions multiplied by q_3 and q_4 can be integrated analytically, but the function multiplied by q_2 leads to a binomial integral which cannot be expressed analytically. So the basic creep compliance function has the form

$$C_0(t,t') = q_2 Q(t,t') + q_3 \ln[1 + (t-t')^n] + q_4 \ln\left(\frac{t}{t'}\right) \tag{E.3}$$

where $Q(t,t')$ is a function that can be obtained by numerical integration or by interpolation from a table computed in Bažant and Baweja (1995, 2000a). The values of $Q(t,t')$ can also be calculated from the approximate explicit formula

$$Q(t,t') = Q_f(t')\left[1 + \left(\frac{Q_f(t')}{Z(t,t')}\right)^{r(t')}\right]^{-1/r(t')} \tag{E.4}$$

in which, for $n = 0.1$ and $m = 0.5$ (with an error under 1%),

$$r(t') = 1.7(t')^{0.12} + 8 \tag{E.5}$$

$$Z(t,t') = (t')^{-m} \ln[1 + (t-t')^n] \tag{E.6}$$

$$Q_f(t') = [0.086(t')^{2/9} + 1.21(t')^{4/9}]^{-1} \tag{E.7}$$

The terms in (E.3) containing q_2, q_3 and q_4 represent the aging viscoelastic compliance, non-aging viscoelastic compliance and flow compliance, respectively, as deduced from the solidification theory (Bažant and Prasannan, 1989; Carol and Bažant, 1993; Bažant *et al.*, 1997).

E.2.2 Drying Creep Compliance Function

The additional compliance due to simultaneous drying, C_d, is evaluated from a formula of the same general form as for the short version of model B3, i.e.

$$C_d(t,t',t_0) = q_5\left(\mathrm{e}^{-g(t-t_0)} - \mathrm{e}^{-g(t'-t_0)}\right) \tag{E.8}$$

However, since the formulae for basic creep compliance in the full version and the short version are different, the optimal agreement of this formula with the experimental data is obtained for a different function $g(\hat{t})$ and different q_5;

$$g(\hat{t}) = 8\left[1 - (1-h)S(\hat{t})\right] \tag{E.9}$$

$$q_5 = 7.57 \times 10^5 \bar{f}_c^{-1} \varepsilon_{sh\infty}^{-0.6} \tag{E.10}$$

Here h, $S(\hat{t})$, $\bar{f}_c$ and $\varepsilon_{sh\infty}$ have exactly the same meaning as in Section 28.3.2. Expressions (E.9) and (E.10) replace those given in Section 28.3.3, which are valid only for the short version.

E.2.3 Prediction of Model Parameters

Ideally, the material parameters should be determined from long-term creep and shrinkage tests performed on the specific concrete that will be used in the particular structure. In practical applications, it is impossible to wait with the analysis until such tests are finished. The parameters can be crudely estimated from empirical formulae that have been established by fitting a large set of experimental data for various types of concrete and correlating the obtained model parameters to the concrete strength and composition. The high uncertainties in predicting creep and shrinkage from the basic characteristics of concrete can be drastically reduced by statistical updating of the model based on short-time tests of creep and shrinkage (Bažant and Baweja, 1995, 2000a). The shrinkage tests must be accompanied by measurements of water loss due to drying.

Since the coefficients that appear in the approximative formulae are not dimensionless, their values depend on the choice of the system of units. We present here alternative formulae valid in inch-pound (in-lb) system units and in SI (metric) units. The entire parameter evaluation must be done in one selected system of units.

The required input data specifying the strength and composition of concrete are summarized in Table E.1. The compression strength $\bar{f}_c$ should be determined as the statistical average of test results on cylinders of diameter 6 inches and length 12

Table E.1 Concrete properties serving as input data

Property	Symbol	in-lb units	SI units
average compression strength	$\bar{f}_c$	psi	MPa
water content	w	lb/ft^3	kg/m^3
cement content	c	lb/ft^3	kg/m^3
aggregate content	a	lb/ft^3	kg/m^3

Table E.2 Prediction formulae for creep parameters

Phenomenon	Parameter	Inch-pound formula	Units	SI formula	Units
basic creep	q_1	$10.5\,\bar{f}_c^{-0.5}$	10^{-6}/psi	$127\,\bar{f}_c^{-0.5}$	10^{-6}/MPa
	q_2	$451.1\,c^{0.5}\,\bar{f}_c^{-0.9}$	10^{-6}/psi	$185.4\,c^{0.5}\,\bar{f}_c^{-0.9}$	10^{-6}/MPa
	q_3	$0.29(w/c)^4 q_2$	10^{-6}/psi	$0.29(w/c)^4 q_2$	10^{-6}/MPa
	q_4	$0.14(a/c)^{-0.7}$	10^{-6}/psi	$20.3(a/c)^{-0.7}$	10^{-6}/MPa
shrinkage	$\varepsilon_{sh\infty}$	$\alpha_1\alpha_2\left(26\,w^{2.1}\,\bar{f}_c^{-0.28}+270\right)$	10^{-6}	$\alpha_1\alpha_2\left(0.019\,w^{2.1}\,\bar{f}_c^{-0.28}+270\right)$	10^{-6}
	k_t	$190.8\,t_0^{-0.08}\,\bar{f}_c^{-1/4}$	days/in^2	$8.5\,t_0^{-0.08}\,\bar{f}_c^{-1/4}$	days/cm^2
drying creep	q_5	$7.57\times10^5\,\bar{f}_c^{-1}\varepsilon_{sh\infty}^{-0.6}$	10^{-6}/psi	$7.57\times10^5\,\bar{f}_c^{-1}\varepsilon_{sh\infty}^{-0.6}$	10^{-6}/MPa

inches at age 28 days. The content of water, cement and aggregates is the mass of the component per unit volume of concrete mix. If the water and cement contents to be used have not yet been decided, they may be estimated from their empirical correlation to the required design strength of concrete (Bažant and Baweja, 2000a).

The formulae for predicting the creep and shrinkage parameters are presented in Table E.2. Coefficients α_1 and α_2 that appear in the formulae for the final shrinkage strain $\varepsilon_{sh\infty}$ are defined as

$$\alpha_1 = \begin{cases} 1.0 & \text{for type I cement} \\ 0.85 & \text{for type II cement} \\ 1.1 & \text{for type III cement} \end{cases} \tag{E.11}$$

$$\alpha_2 = \begin{cases} 0.75 & \text{for steam-curing} \\ 1.2 & \text{for sealed or normal curing in air with initial protection against drying} \\ 1.0 & \text{for curing in water or at 100\% relative humidity} \end{cases} \tag{E.12}$$

E.3 MATERIAL LAW FOR FREE SHRINKAGE AND THERMAL STRAIN

Sensitive structures are analyzed by layered beam finite element programs or by two- and three-dimensional finite element programs. In such programs, the material properties used in each finite element must be the constitutive properties, independent of the cross section dimensions and shape, as well as of the environmental conditions, which represent the boundary conditions of the partial differential equations for drying and heat conduction (Bažant, 1988; Bažant and Kaplan, 1996). At drying, the constitutive properties cannot be measured directly, but they have been identified by fitting with a finite element program the overall deformation measurements on test specimens (Bažant, Hauggaard and Baweja, 1996a; Bažant, 1988). For this kind of analysis, the compliance function of concrete has the form

$$J(t,t') = q_1 + C_0(t,t') \tag{E.13}$$

which means that it only includes the instantaneous response and the basic creep. The additional compliance $C_d(t,t';t_0)$ is deleted and the effects of drying creep are incorporated in the shrinkage strain, ε_s, which is now different from (and simpler to characterize than) the average drying shrinkage strain in a cross section, ε_{sh}. The evolution of shrinkage strain is governed by the following rate-type constitutive relation (Bažant and Baweja, 2000a, 1995; Bažant and Chern, 1985):

$$\dot{\boldsymbol{\varepsilon}}_s = \varepsilon_{0s}\frac{E(t_0)}{E(t)}\left[\boldsymbol{\delta} + s_h(r\boldsymbol{\sigma} + r'\sigma_V\boldsymbol{\delta})\right](\dot{h}_p + a_T\dot{T}) \tag{E.14}$$

where $\dot{\boldsymbol{\varepsilon}}_s$ = shrinkage strain rate tensor; ε_{0s} = final shrinkage (at material points), which has a similar but not exactly the same value as $\varepsilon_{sh\infty}$ for the cross section average; $\boldsymbol{\delta}$ = second-order unit tensor (Kronecker delta); $\boldsymbol{\sigma}$ = stress tensor; σ_V = volumetric (or mean) stress; $\dot{h}_p$ and $\dot{T}$ = rates of local relative humidity and temperature in the pores of concrete (which must be obtained by solving the diffusion equations); s_h = sign of $(\dot{h}_p + a_T\dot{T})$, which is either 1 or -1; r and r' = parameters;

and a_T = coefficient relating the stress-induced thermal strain and shrinkage. When (E.14) is used, the cracking or fracture must also be included in the analysis.

The constitutive relation (E.14) is much simpler than equations (28.10)–(28.16) which it replaces. However, at present there are not enough data to predict the values of ε_{0s}, r, r' and a_T from the composition and strength of the concrete. They must be identified by fitting the measured data for drying creep, shrinkage and thermal expansion.

Appendix F

Softening Inelastic Hinges: Deviations from Plasticity and Size Effect

F.1 SIZE EFFECT ON MOMENT-ROTATION DIAGRAM

Although this book is devoted to plasticity, one should be aware of the limitations of plasticity due to brittleness. We will now expound one important limitation, which concerns the inelastic analysis of beams and frames. Reading of this appendix requires knowledge of the basic concepts of fracture mechanics (e.g. Bažant and Planas, 1998) and stability of damage (Bažant and Cedolin, 1991, Chapter 13).

As pointed out in Chapter 14, post-peak softening of inelastic hinges in reinforced concrete frames cannot be avoided when failure is not caused by the yielding of steel. Since it is important to realize the consequences of softening, we will now briefly discuss a simplified analysis of such behavior, based on a recent study by Bažant (2000). This kind of analysis is applicable not only to concrete but also to other materials, for example to the load carrying capacity of floating sea ice plates (Bažant and Guo, 2000).

In floating ice plates, a softening response of an inelastic hinge is engendered by transverse propagation of a tensile crack. Post-peak softening of inelastic hinges in reinforced concrete beams or plates (Figure F.1(a)) is caused by compression fracture. The idea of describing the post-peak rotation of inelastic hinges in concrete beams by a cohesive stress-displacement relation for a 'compression crack' was suggested by Hillerborg (1990). While the effective length of the yielding zone of plastic hinges in steel beams is proportional to the characteristic size, which may be taken as the beam depth D, the effective hinge length of softening hinges with fracture (Figures F.1(b),(c)) is certainly not proportional to D but is close to being approximately constant, roughly equal to the characteristic length of the material. The cause is the localization instability of strain softening (Bažant and Cedolin, 1991, Chapter 13), and one consequence is the size effect.

Hinge softening due to tensile fracture occurs, for example, in floating ice plates, due to propagation of a crack from the tensile face (Figure F.1(b)). The energy dissipated by the crack per unit area of the cross section is the fracture energy of the material, G_f.

Hinge softening in reinforced concrete beams is caused by compression fracture and occurs if the bending moment is accompanied by a sufficient axial compression force (which is the case for prestressed concrete beams, for columns with a large enough

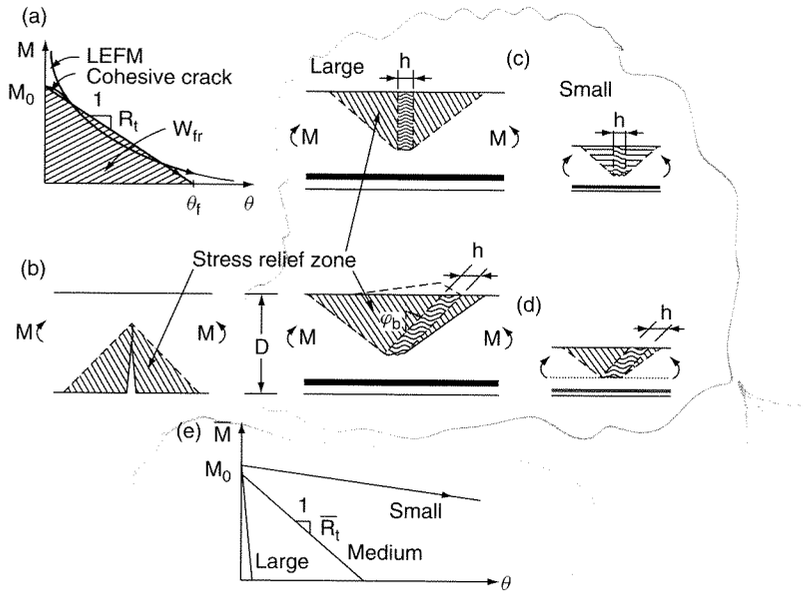


Figure F.1 Softening inelastic hinge: (a) its idealized moment-rotation diagram, (b) hinge in ice plate failing by tensile crack, (c) and (d) hinge in reinforced concrete beams failing by compression fracture band, and (e) size effect on moment-rotation diagram

axial force, and for frames or arches with a large enough horizontal thrust). It also occurs if there is strong enough tensile reinforcement (Figures F.1(c),(d)), as may be the case for concrete beams retrofitted with a fiber laminate bonded to the tensile face.

The compression fracture consists of an inclined band of axial splitting microcracks (a crushing band, which is often regarded as a shear failure in compression, although shear slip becomes possible only after the axial stress gets reduced much below the peak). With respect to a normal cross section, the band can be inclined by angle φ_b in the horizontal plane (Figure F.1(c), where the inclination would be seen by looking down from the top) or in the vertical plane (Figure F.1(d)). The energy dissipated by formation of the band, per unit area of the cross section (unit area of the projection of the compression fracture band onto the critical orthogonal cross section), is G_b; G_b plays the role of the fracture energy of the band, considered to be a material constant.

The assumption of constant G_b is indicated theoretically by studies of localization of softening damage. However, it must be pointed out that this assumption has not yet been systematically tested in the laboratory and at present lacks unambiguous experimental support. On the other hand, it is clear that an assumption of softening damage in compression remaining distributed, with no localization at all, would definitely be wrong (this is the only possibility for which the size effect would be absent). The assumption of no localization, no size effect, would correspond to a fixed post-peak stress-strain curve, with no unloading, which would mean that the area W_{f_r} under the curve would be fixed. Recall that W_f has the meaning of energy dissipated per unit volume, while G_b is the energy dissipated per unit area. The reality might not be a truly constant G_b, but somewhere between the constancy of dissipation per unit area and per unit volume, definitely closer to the former.

According to linear elastic fracture mechanics (LEFM), the diagram of the softening moment-rotation relation for the inelastic hinge is a curve descending from infinity,

as sketched in Figure F.1(a). If the nonlinear cohesive nature of a tensile crack, or the finite size of the fracture process zone of the tensile crack or compression crushing band, is taken into account, the diagram starts its descent from a finite value, M_0, representing the nominal (or maximum) bending moment whose calculation for concrete beams is discussed in Section 14.1.

For an over-reinforced cross section with only tensile reinforcement, the maximum bending moment is obtained when the magnitude of maximum compressive strain in concrete reaches 0.003. The position of the neutral axis follows from the condition that the compressive stress resultant in concrete is equal to the tensile stress resultant in steel, i.e.

$$0.85 f_c' \beta_1 c b \;=\; 0.003 \,\frac{D-c}{c}\, E_s A_s \tag{F.1}$$

where all the symbols are defined in Section 14.1.1, and the beam depth D is now measured as the distance between the reinforcement centroid and the compression face of the beam. For given material parameters and cross sectional dimensions, (F.1) provides a quadratic equation,

$$\frac{0.85 f_c' \beta_1 b D}{0.003 E_s A_s}\gamma^2 + \gamma - 1 = 0 \tag{F.2}$$

for the unknown dimensionless ratio $\gamma \equiv c/D$. For geometrically similar cross sections with a fixed reinforcement ratio ρ, the cross section area of steel is $A_s = \rho b D$, and so the coefficient at γ^2 in (F.2) becomes independent of the structure size. Consequently, the relative position of the neutral axis, γ, is a size-independent parameter, and the maximum moment is

$$M_0 = E_s A_s \varepsilon_s (D - \tfrac{1}{2}a) \;=\; 0.003 E_s \rho \,\frac{1-\gamma}{\gamma}\,(1 - \tfrac{1}{2}\beta_1\gamma)\, b D^2 \tag{F.3}$$

i.e. M_0 is proportional to the width and to the square of the depth of the cross section. This is formally similar to the expression valid for the failure of a plain concrete beam (or an ice plate) driven by tensile cracking,

$$M_0 = \tfrac{1}{6} b D^2 f_r \tag{F.4}$$

in which f_r is the *modulus of rupture* (flexural strength). Defining the flexural strength of an overreinforced cross section failing by compressive crushing as

$$f_r \;=\; 0.018 E_s \rho \,\frac{1-\gamma}{\gamma}\,(1 - \tfrac{1}{2}\beta_1\gamma) \tag{F.5}$$

we can extend the validity of formula (F.4) to this failure mode.

Strictly speaking, the initiation of tensile or compression fracture also causes a size effect on the flexural strength. However, this kind of size effect is of a different type (it is weaker than that to be shown later in Figure F.3) and disappears for larger sizes. Here it will be neglected.

For the sake of simplicity, we will idealize the moment-rotation diagram as linear, i.e. $M = R_t(\theta_f - \theta)$, where $R_t = M_0/\theta_f$ is minus the tangent stiffness of the hinge (representing the slope ot the M–θ diagram) and θ_f is the hinge rotation at complete break (Figure F.1(a)). The energy W_{fr}, dissipated by a total break of the cross section, is given by the area under this diagram (Figure F.1(a));

$$W_{fr} = \tfrac{1}{2} M_0 \theta_f \tag{F.6}$$

We will now restrict attention to rectangular cross sections and plates (although generalization to arbitrary cross sections is not difficult). Then, the energy dissipated over the whole cross- section upon reaching a complete break must be equal to

$$W_{fr} = G_b bD \tag{F.7}$$

where b is the width of a rectangular cross section (in the case of a hinge in a plate, we consider a unit width $b = 1$). In view of energy conservation, we set both expressions for W_{fr} equal, and thus we get

$$\theta_f = \frac{2G_b bD}{M_0} = \frac{12\,G_b}{f_r}\,\frac{1}{D} \tag{F.8}$$

where f_r = flexural strength, whose size dependence is neglected here.

To bring to light the scaling, it is helpful to introduce the dimensionless bending moment and the dimensionless tangential softening stiffness of the softening inelastic hinge:

$$\bar{M} = \frac{M}{EbD^2}, \qquad \bar{R}_t = \frac{R_t}{EbD^2} = \frac{M_0}{EbD^2\theta_f} \tag{F.9}$$

For non-softening elasto-plastic materials, the diagram of $\bar{M}$ versus θ is independent of the beam depth, and so any change in this diagram as a function of structure size reveals a size effect. For our softening material and for a rectangular cross section, we find, upon substituting (F.4) and (F.8) into (F.9), that the dimensionless stiffness is given by

$$\bar{R}_t = \frac{bD^2 f_r/6}{EbD^2 12\,G_b/f_r D} = \frac{f_r^2 D}{72\,EG_b} = \frac{D}{72\,l_{fr}} \tag{F.10}$$

where

$$l_{fr} = \frac{EG_b}{f_r^2} \tag{F.11}$$

represents an Irwin-type characteristic length of the material (see e.g. Bažant and Planas, 1998).

As we see from (F.10) and (F.8), the dimensionless softening stiffness of the hinge increases in proportion to size D, and the rotation at full break decreases at inverse proportion to D (Figure F.1(e)). As the structure size approaches infinity, $\bar{R}_t$ becomes infinite (a vertical drop). As the structure size tends to zero, $\bar{R}_t$ becomes zero (a horizontal line) and the softening hinge becomes equivalent to a plastic hinge. These are size effects with important consequences for structures, which we explore next. In plasticity, by contrast, one cannot speak of any size effect because $\bar{R}_t = R_t = 0$ for any size, and $\theta_f \to \infty$.

F.2 SIZE EFFECT IN BEAMS AND FRAMES FAILING BY SOFTENING HINGES

The size effects of softening hinges are important for: (1) the energy absorption capability of structures (which governs the resistance to earthquake, blast, shock and impact), and (2) the strength of redundant brittle beams and frames, as well as the strength of plates. Here we will deal only with the strength.

Consider a statically indeterminate (redundant) beam structure which requires N inelastic hinges to form in order to collapse (e.g. four hinges in Figure F.2). Let the

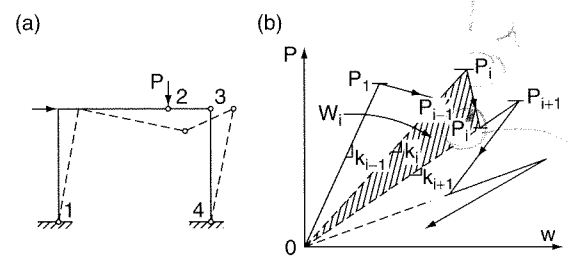


Figure F.2 (a) Frame failing by softening inelastic hinges, and (b) load-deflection diagram when one and only one hinge is softening at a time

inelastic hinges be numbered as $j = 1, 2, ...N$ in the sequence in which they form as the load-point displacement w is increased. Let K_i be the stiffness associated with the applied load P if hinges $j = 1, 2, ...i-1$ have completely softened (i.e. $M = 0$ and $\theta \geq \theta_f$ in these hinges) and hinges $j = i, i+1, ...N$ have not yet started to form (i.e. $M \leq M_0$ and $\theta = 0$ in these hinges). Obviously, $K_1 > K_2 > K_3 > \ldots > K_N > 0$.

If all the hinges $j = 1, 2, ...i-1$ have softened to a zero moment and hinge i has not yet started to form, the load-point deflection w is decided solely by K_i, and the start of softening in the next hinge i is decided by a critical stress, σ_i, representing the stress at the tensile face in the case of tensile fracture, or at the compression face in the case of compression fracture. Therefore

$$P = K_i w, \qquad \sigma_i = S_i w \tag{F.12}$$

where S_i are constants. Restricting again attention to rectangular cross sections, one may introduce dimensionless structure stiffness $\bar{K}_i$ and dimensionless critical stress $\bar{S}_i$ in hinge i, such that

$$K_i = \bar{K}_i Eb, \qquad S_i = \bar{S}_i E/D \tag{F.13}$$

where D is the characteristic size of the structure, considered as two-dimensional, and b is the characteristic thickness in the third dimension. Here K_i is normalized by b because it grows in proportion to the out-of-plane dimension of the cross section, while S_i is normalized by D because the stress in a rectangular cross section depends only on the bending moment per unit width.

Assume now that the hinges form and fully soften one by one, i.e. no two hinges are softening at the same time. Then the load-deflection diagram must look as shown in Figure F.2(b), where the slope of each ray emanating from the origin is K_i and the maximum load on each ray, marked as P_i $(i = 1, 2, ...)$, corresponds to the start of softening of the next hinge. The trough P_i' on each ray is the load at which the softening of each hinge gets completed (i.e. $M = 0$ and $\theta = \theta_f$). Since we assumed the moment-rotation diagram of the hinges to be linear, the load-deflection diagram from P_i to P_i' must be a straight line.

The load peaks P_i are determined from the condition $\sigma_i = f_r$ = modulus of rupture (flexural strength). Noting that $\sigma_i = PS_i/K_i$, we find:

$$P_i = \frac{K_i}{S_i}\,f_r, \tag{F.14}$$

The corresponding nominal stresses are

$$\sigma_{Ni} = \frac{P_i}{bD} = \frac{\bar{K}_i}{\bar{S}_i} f_r \tag{F.15}$$

As may have been expected, peaks σ_{Ni} exhibit only the size effect of fracture initiation, the same as that in the modulus of rupture, which is shown in (Figure F.3) but is neglected in our calculations.

To determine the troughs P_i', consider the area of the shaded triangle in Figure F.2(b) between the rays of slopes K_i and K_{i+1}; the area is

$$W_i = \frac{P_i P_i'}{2}\left(\frac{1}{K_{i+1}} - \frac{1}{K_i}\right) \tag{F.16}$$

W_i represents the work dissipated when hinge i softens from M_0 to 0. Since no other hinge is assumed to be softening simultaneously with hinge i, energy conservation requires W_i to be equal to the work dissipated by hinge i;

$$W_i = G_b b_i D_i = G_b \beta_i \delta_i D^2 \qquad \text{with} \quad \beta_i = b_i/D, \quad \delta_i = D_i/D \tag{F.17}$$

where b_i and D_i are the width and depth of the cross section at hinge i, and β_i and δ_i are constants for structures geometrically similar in three dimensions. Setting this equal to (F.16), we can solve for P_i' and obtain:

$$\frac{\sigma_{Ni}'}{\sigma_{Ni}} = \frac{P_i'}{P_i} = \frac{2EG_b}{\sigma_{Ni}^2}\left(\frac{1}{\bar{K}_{i+1}} - \frac{1}{\bar{K}_i}\right)^{-1}\frac{1}{D} = \frac{\text{const.}}{D} \qquad \text{(for large } D\text{)} \tag{F.18}$$

Factor $1/D$ represents the large-size asymptotic size effect on the troughs (Figure F.3). It is a very strong size effect, much stronger than the LEFM size effect of large cracks, which is of the type $1/\sqrt{D}$.

Figure F.4 shows a sequence of typical load-deflection diagrams for geometrically similar structures of various sizes D. These diagrams are plotted in dimensionless coordinates w/D and σ_N/E (and the size effect on the peaks stemming from crack initiation is neglected). This has the advantage that (1) the slopes corresponding to K_i (and also the points corresponding to the peaks P_i) are independent of the size, and that (2) the diagrams for structures of different sizes would be identical if the material were elasto-plastic (in which case the size effect is known to be absent). Thus, any change in these diagrams with the structure size signifies a size effect.

The scaling of the troughs P_i', as described by (F.18), indicates that when the structure size increases, the response approaches a series of narrow spikes of progressively decreasing inclination (Figure F.4). The descending part of each spike in Figure F.4(e) is unstable for any type of control of P and w, and is known as the *snap-back instability* (e.g. Bažant and Cedolin, 1991). This behavior may have been intuitively expected on the basis of the fact that the diagram of $\bar{M}$ versus θ approaches a vertical drop as $D \to \infty$.

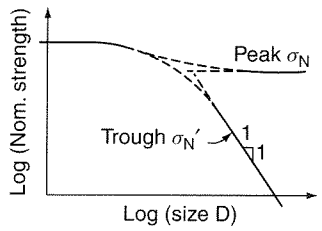


Figure F.3 Size effects on nominal strengths at a peak and at a trough, matched to small-size plastic limit

When the structure size decreases, the line from P_i to P_i' eventually changes its slope from negative to positive, i.e. P_i' ceases to be a trough and P_i a peak. With a further size decrease, P_i' eventually becomes coincident with the next peak P_{i+1}. For still smaller sizes, during the loading process there is a period in which hinges i and $i+1$ soften simultaneously. During such simultaneous softening, the deflection diagram is again linear (lines 67 and 89 in Figure F.4(b)), but different for each softening hinge combination and harder to calculate. When the size is decreased further, more and more hinges, or all the hinges, undergo softening at the same time.

Segments 67 and 89 in Figure F.4(b) correspond to simultaneous softening of two subsequently formed hinges. Since areas 0120 and 0340 must be equal to the fracture energies W_{i-1} and W_i of hinges $i-1$ and i, the area 01678950 must be equal to their sum. Therefore, the two little shaded triangles of each pair (Figure F.4(b)) must have equal areas. This property can be exploited for constructing the response graphically.

In the theoretical asymptotic case of a zero size ($D \to 0$), the slope of the diagram of dimensionless moment $\bar{M}$ versus rotation θ tends to a horizontal asymptote and, consequently, the plastic limit analysis applies. In Figure F.4(a), the zero-size limiting response consists of line segments parallel to the rays emanating from the origin (note

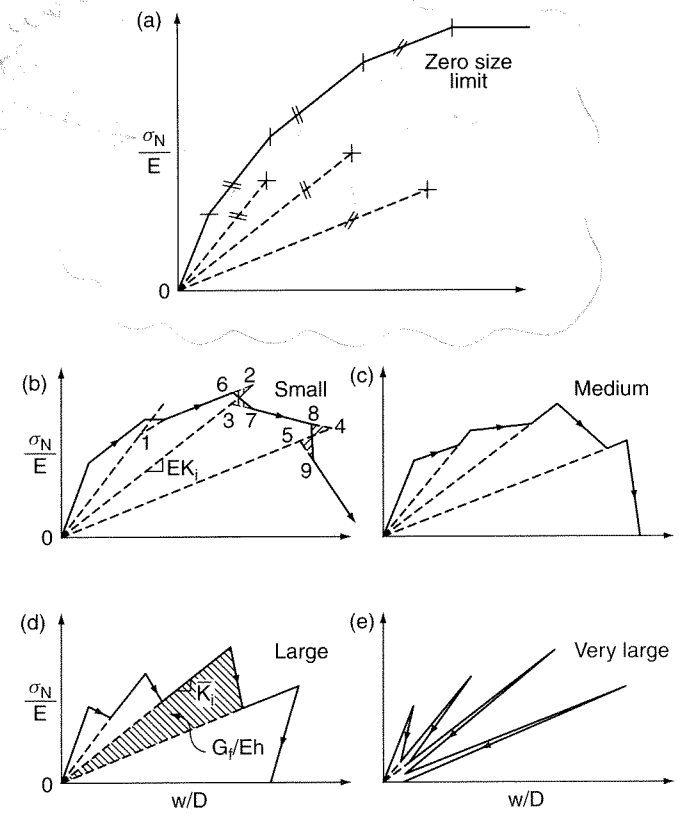


Figure F.4 Evolution of load-displacement diagrams (in terms of dimensionless nominal strength σ_N/E and relative deflection w/D) with increasing size D of a structure failing by softening hinges (in such coordinates, the secant stiffness slopes, as well as the peak points, do not change with the size). Parallel dashes in (a) mark parallel lines

the parallel double dashes marking parallel lines). Such a response is easy to calculate by well-known methods (Section 3.3).

The behavior between this asymptotic case and the case of the smallest structure for which only one hinge is softening at a time (in which $P'_i = P_{i+1}$) can be approximately characterized by interpolation, better regarded as asymptotic matching. Matching of the large-size asymptotic size effect in (F.18) to the horizontal small-size asymptote may be achieved by the following simple approximate formulae for all sizes:

$$\sigma'_{Ni} \approx \frac{\sigma'_{0i}}{1 + D/D'_{0i}} \quad \text{or} \quad \frac{\sigma'_{0i}}{(1 + D^s/D'^{\,s}_{0i})^{1/s}} \tag{F.19}$$

where σ'_{0i}, D'_{0i} and s are positive constants. For small sizes, however, the strongest size effect is caused by the fact that the overall maximum load for different sizes occurs for different i; compare Figures F.4(a) and F.4(b). This aspect is not reflected in Figure F.3.

When the structure is large enough for the load-deflection diagram to involve softening segments, designing for the largest peak, $\max P_i$, becomes questionable because, under load control, the structure is unstable during each softening segment. Such a design is even more questionable when the structure is large enough for snapbacks to occur, in which case it is unstable even under load-point displacement control (Bažant and Cedolin, 1991, Chapter 7). A completely safe design is that for the lowest trough, $\min P'_i$, but then again, the safety margin would often be unnecessarily high by far. A realistic design load P lies somewhere between $\max P_i$ and $\min P'_i$, for a smaller structure closer to $\max P_i$ and for a larger one closer to $\min P'_i$.

To decide the proper design load rationally, one should take into account the imperfections and possible dynamic disturbances. In structures with plastic hinges, a dynamic disturbance in which the peak load is only reached temporarily leads merely to a permanent deflection, but in the case of post-peak snap-back it may trigger stability loss and dynamic failure.

The problem is similar to that of an axially compressed thin cylindrical shell (or a thin spherical dome), for which the critical load of a theoretically perfect shell (unattainable in practice) is followed by a snap-back to a very low residual load. It is now well understood (e.g. Bažant and Cedolin, 1991, Chapter 7) that, due to inevitable imperfections and dynamic disturbances, one must design the shell for that residual load (typically between 1/8 and 1/3 of the theoretical critical load). An effective semi-empirical method has been developed to deal with this problem for shells, and a similar method might have to be developed for large beams and frames with softening hinges.

Another type of size effect in frames with inelastic softening hinges was studied by Jirásek (1997), who considered fixed cross sectional dimensions but a variable number of columns in a one-storey multi-bay frame loaded by a horizontal force. For perfectly plastic hinges ($\theta_f \to \infty$), the limit load is proportional to the number of columns. Therefore, the actual collapse load divided by the plastic limit load can be considered as the dimensionless nominal collapse load. For perfectly brittle 'hinges' ($\theta_f = 0$), the nominal collapse load is 1 for the simple frame with two columns; it drops down to 0.8 for the frame with three columns and then gradually increases with increasing number of columns. For inelastic softening hinges with $0 < \theta_f < \infty$, the dependence of the nominal collapse load on the number of columns is between the plastic and brittle limits, and is controled by the dimensionless ductility parameter l_{fr}/H, where l_{fr} is

the characteristic length defined in (F.11) and H is the column height. The brittle limit is approached for small values and the plastic limit for large values of the ductility parameter.

Finally, it should be noted that brittle materials such as concrete (or ice) are not the only ones that lead to softening inelastic hinges. As shown by Maier and Zavelani (1970), local buckling of thin flanges in the plastic hinge region can cause sharp softening in the moment curvature diagram. The previous analysis applies to that case. Obviously, the local buckling of flanges is much more dangerous in large steel structures than in small ones.

References

Acharya, A. and Bassani, J. L. (2000). Lattice incompatibility and a gradient theory of crystal plasticity, *Journal of the Mechanics and Physics of Solids* **48**: 1565–1595.

ACI (1999). *Building Code Requirements for Reinforced Concrete*, ACI Standard 318-99, American Concrete Institute, Detroit.

Ackroyd, M. H. and Gerstle, K. H. (1983). Elastic stability of flexibly connected frames, *Journal of Structural Engineering, ASCE* **109**: 241–245.

Aguinaga-Zapata, M. and Bažant, Z. P. (1986). Creep deflections in slab buildings and forces in shores during construction, *Journal of the American Concrete Institute* **83**: 719–726.

Aifantis, E. C. (1984). On the microstructural origin of certain inelastic models, *Journal of Engineering Materials and Technology, ASME* **106**: 326–330.

Aifantis, E. C. (1999). Strain gradient interpretation of size effects, *International Journal of Fracture* **95**: 299–314.

AISC (1986). *Load and Resistance Factor Design*, Manual of Steel Construction, American Institute of Steel Construction, Chicago.

Argyris, J. H., Faust, G., Szimmat, J., Warnke, E. P. and Willam, K. J. (1974). Recent developments in the finite element analysis of PCRV, *Nuclear Engineering and Design* **28**: 42–75.

Armstrong, P. J. and Frederick, C. O. (1966). A mathematical representation of the multiaxial Bauschinger effect, Report *RD/B/N 731*, C.E.G.B.

Arutyunian, N. K. (1952). *Some Problems in the Theory of Creep (in Russian)*, Techteorizdat, Moscow. English translation Pergamon Press, 1966.

Asaro, R. J. (1983a). Crystal plasticity, *Journal of Applied Mechanics, ASME* **50**: 921–934.

Asaro, R. J. (1983b). Micromechanics of crystals and polycrystals, *Advances in Applied Mechanics* **23**: 1–115.

Asaro, R. J. and Needleman, A. (1985). Texture development and strain hardening in rate dependent polycrystals, *Acta Metallurgica* **33**: 923–953.

Asaro, R. J. and Rice, J. R. (1977). Strain localization in ductile single crystals, *Journal of the Mechanics and Physics of Solids* **33**: 309–338.

Ashby, M. F. (1970). The deformation of plastically non-homogeneous alloys, *Philosophical Magazine* **21**: 399–424.

Atkinson, J. H. and Bransby, P. L. (1978). *The Mechanics of Soils: An Introduction to Critical State Soil Mechanics*, University series in civil engineering, McGraw-Hill, London.

Auricchio, F. and Taylor, R. L. (1994). A generalized visco-plasticity model and its algorithmic implementation, *Computers and Structures* **53**: 637–647.

Backhaus, G. (1968). Zur Fließgrenze bei allegmeiner Verfestigung, *Zeitschrift für angewandte Mathematik und Mechanik* **48**: 99–108.

Balmer, G. G. (1949). Shearing strength of concrete under high triaxial stress – computation of Mohr's evelope as a curve, *Structural Research Laboratory Report No. SP-23*, United States Department of the Interior, Bureau of Reclamation, Washington, D.C.

Barenblatt, G. I. (1962). The mathematical theory of equilibrium of cracks in brittle fracture, *Advances in Applied Mechanics* **7**: 55–129.

Baron, M. L., Nelson, I. and Sandler, I. S. (1973). Influence of constitutive models on ground motion predictions, *Journal of the Engineering Mechanics Division, ASCE* **99**: 1181–1200.

Batdorf, S. B. and Budianski, B. (1949). A mathematical theory of plasticity based on the concept of slip, *Technical Note 1871*, National Advisory Committee for Aeronautics, Washington, D.C.

Bathe, K.-J. (1996). *Finite Element Procedures*, revised edition, Prentice Hall, Englewood Cliffs, New Jersey.

Bauschinger, J. (1881). Über die Veränderung der Elastizitätsgrenze und des Elastizitätsmoduls verschiedener Metalle, *Civilingenieur* pp. 289–348.

Baweja, S., Dvorak, G. J. and Bažant, Z. P. (1998). Triaxial composite model for basic creep of concrete, *Journal of Engineering Mechanics, ASCE* **124**: 959–966.

Bazaraa, M. S., Jarvis, J. J. and Sherali, H. D. (1990). *Linear Programming and Network Flows*, 2nd edition, Wiley, New York.

Bažant, Z. P. (1966). Phenomenological theories for creep of concrete based on rheological models, *Acta Technica ČSAV* **11**: 82–109.

Bažant, Z. P. (1968). Langzeitige Durchbiegungen von Spannbetonbrücken infolge des Schwingkriechens unter Verkehrslasten, *Beton- und Stahlbetonbau* **63**: 282–285.

Bažant, Z. P. (1971a). Correlation study of formulations of incremental deformations and stability of continuous bodies, *Journal of Applied Mechanics, ASME* **38**: 919–928.

Bažant, Z. P. (1971b). Numerically stable algorithm with increasing time steps for integral-type ageing creep, *Proceedings of the 1st International Conference on Structural Mechanics in Reactor Technology*, Vol. **3**, Berlin, p. H2/3.

Bažant, Z. P. (1972a). Numerical determination of long-range stress history from strain history in concrete, *Materials and Structures* **5**: 135–141.

Bažant, Z. P. (1972b). Prediction of concrete creep effects using age-adjusted effective modulus method, *ACI Journal* **69**: 212–217.

Bažant, Z. P. (1975). Theory of creep and shrinkage in concrete structures: A precis of recent developments, in *Mechanics Today*, S. Nemat-Nasser (ed.), Vol. 2, Pergamon Press, pp. 1–93.

Bažant, Z. P. (1976). Instability, ductility, and size effect in strain-softening solids, *Journal of the Engineering Mechanics Division, ASCE* **102**: 331–344.

Bažant, Z. P. (1977). Viscoelasticity of solidifying porous material – concrete, *Journal of Engineering Mechanics, ASCE* **103**: 1049–1067.

Bažant, Z. P. (1978). Endochronic inelasticity and incremental plasticity, *International Journal of Solids and Structures* **14**: 691–714.

Bažant, Z. P. (1979). Thermodynamics of solidifying or melting viscoelastic material, *Journal of Engineering Mechanics, ASCE* **105**: 933–952.

Bažant, Z. P. (1980). Work inequalities for plastic-fracturing materials, *International Journal of Solids and Structures* **16**: 870–901.

Bažant, Z. P. (1982). Mathematical models for creep and shrinkage of concrete, in *Creep and Shrinkage in Concrete Structures*, Z. P. Bažant and F. H. Wittmann (eds), Wiley, Chichester, chapter 7, pp. 163–256.

Bažant, Z. P. (1983). Comment on orthotropic models for concrete and geomaterials, *Journal of Engineering Mechanics, ASCE* **109**: 849–865.

Bažant, Z. P. (1984a). Design and analysis of concrete reactor vessels: New developments, problems and trends, *Nuclear Engineering and Design* **80**: 181–202.

Bažant, Z. P. (1984b). Imbricate continuum and its variational derivation, *Journal of Engineering Mechanics, ASCE* **110**: 1693–1712.

Bažant, Z. P. (1984c). Microplane model for strain controled inelastic behavior, in *Mechanics of Engineering Materials*, C. S. Desai and R. H. Gallagher (eds), Wiley, Chichester, Chapter 3, pp. 45–59.

Bažant, Z. P. (1985). Mechanics of fracture and progressive cracking in concrete structures, in *Fracture Mechanics of Concrete: Structural Application and Numerical Calculation*, G. C. Sih and A. DiTommaso (eds), Martinus Nijhoff, Dordrecht and Boston, chapter 1, pp. 1–94.

Bažant, Z. P. (1986). Mechanics of distributed cracking, *Applied Mechanics Reviews, ASME* **39**: 675–705.

Bažant, Z. P. (1987). Matrix force-displacement relations in aging viscoelasticity, *Journal of Engineering Mechanics, ASCE* **113**: 1235–1243.

Bažant, Z. P. (1987). Why continuum damage is nonlocal: Justification by quasi-periodic microcrack array, *Mechanics Research Communications* **14**: 407–419.

Bažant, Z. P. (ed) (1988). *Mathematical Modeling of Creep and Shrinkage of Concrete*, Wiley, Chichester and New York.

Bažant, Z. P. (1994). Nonlocal damage theory based on micromechanics of crack interactions, *Journal of Engineering Mechanics, ASCE* **120**: 593–617.

Bažant, Z. P. (1997). Fracturing truss model: Size effect in shear failure of reinforced concrete, *Journal of Engineering Mechanics, ASCE* **123**: 1276–1288.

Bažant, Z. P. (1998). Easy-to-compute tensors with symmetric inverse approximating Hencky finite strain and its rate, *Journal of Engineering Materials and Technology, ASME* **120**: 131–136.

Bažant, Z. P. (1999). Criteria for rational prediction of creep and shrinkage of concrete, *Revue Française de Génie Civil* **3**: 61–89.

Bažant, Z. P. (2001). *Scaling of Structural Strength*, Hermes Science Publications, Oxford and Paris.

Bažant, Z. P. and Baweja, S. (1995). Creep and shrinkage prediction model for analysis and design of concrete structures – model B3, *Materials and Structures* **28**: 357–365. (RILEM Recommendation, in collaboration with RILEM Committee TC 107-GCS, with Errata, Vol. 29, p. 126.)

Bažant, Z. P. and Baweja, S. (1996). Short form of creep and shrinkage prediction model B3 for structures of medium sensitivity, *Materials and Structures* **29**: 587–593. (Addendum to RILEM Recommendation TC 107-GCS.)

Bažant, Z. P. and Baweja, S. (2000a). Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3, in *Adam Neville Symposium: Creep and Shrinkage – Structural Design Effects*, A. AlManasseer (ed), American Concrete Institute, pp. 1–83.

Bažant, Z. P. and Baweja, S. (2000b). Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3–short form, in *Adam Neville Symposium: Creep and Shrinkage – Structural Design Effects*, A. AlManasseer (ed), American Concrete Institute, pp. 85–100.

Bažant, Z. P. and Becq-Giraudon, E. (2000). Size effect in shear design of reinforced concrete, *Project Report C699-00/8s*, Northwestern University.

Bažant, Z. P. and Buyukozturk, O. (1988). Creep analysis of structures, *in* Z. P. Bažant (ed), *Mathematical Modeling of Creep and Shrinkage of Concrete*, Wiley, Chichester and New York, RILEM committee TC-69, chapter 3, pp. 217–273.

Bažant, Z. P. and Cedolin, L. (1983). Finite element modeling of crack band propagation, *Journal of Structural Engineering, ASCE* **109**: 69–92.

Bažant, Z. P. and Cedolin, L. (1991). *Stability of Structures*, Oxford University Press, New York and Oxford.

Bažant, Z. P. and Chern, J. C. (1985). Concrete creep at variable humidity: Constitutive law and mechanism, *Materials and Structures* **18**: 1–20.

Bažant, Z. P. and Frangopol, D. M. (2000). Size effect hidden in excessive dead load factor, *Report 2000-6/A423s*, Departments of Civil Engineering and Materials Science, Northwestern University, Evanston, Illinois.

Bažant, Z. P. and Gambarova, P. (1984). Crack shear in concrete: Crack band microplane model, *Journal of Structural Engineering, ASCE* **110**: 2015–2035.

Bažant, Z. P. and Kaplan, M. F. (1996). *Concrete at High Temperatures: Material Properties and Mathematical Models*, Longman (Addison-Wesley), London.

Bažant, Z. P. and Kim, S. S. (1979a). Approximate relaxation function for concrete, *Journal of the Structural Division, ASCE* **105**: 2695–2705.

Bažant, Z. P. and Kim, S. S. (1979b). Plastic-fracturing theory for concrete, *Journal of the Engineering Mechanics Division, ASCE* **105**: 407–428.

Bažant, Z. P. and Li, Y.-N. (1997). Cohesive crack model with rate-dependent crack opening and viscoelasticity: I. Mathematical model and scaling, *International Journal of Fracture* **86**: 247–265.

Bažant, Z. P. and Lin, F.-B. (1988a). Nonlocal smeared cracking model for concrete fracture, *Journal of Engineering Mechanics, ASCE* **114**: 2493–2510.

Bažant, Z. P. and Lin, F.-B. (1988b). Nonlocal yield-limit degradation, *International Journal for Numerical Methods in Engineering* **26**: 1805–1823.

Bažant, Z. P. and Liu, K.-L. (1985). Random creep and shrinkage in structures: Sampling, *Journal of Structural Engineering, ASCE* **111**: 1113–1134.

Bažant, Z. P. and Najjar, L. J. (1972). Nonlinear water diffusion in nonsaturated concrete, *Materials and Structures* **5**: 3–20.

Bažant, Z. P. and Najjar, L. J. (1973). Comparison of approximate linear methods for concrete creep, *Journal of the Structural Division, ASCE* **99**: 1851–1874.

Bažant, Z. P. and Oh, B.-H. (1983). Crack band theory for fracture of concrete, *Materials and Structures* **16**: 155–177.

Bažant, Z. P. and Oh, B.-H. (1985). Microplane model for progressive fracture of concrete and rock, *Journal of Engineering Mechanics, ASCE* **111**: 559–582.

Bažant, Z. P. and Oh, B.-H. (1986). Efficient numerical integration on the surface of a sphere, *Zeitschrift für angewandte Mathematik und Mechanik* **66**: 37–49.

Bažant, Z. P. and Osman, E. (1976). Double power law for basic creep of concrete, *Materials and Structures* **9**: 3–11.

Bažant, Z. P. and Ožbolt, J. (1990). Nonlocal microplane model for fracture, damage, and size effect in structures, *Journal of Engineering Mechanics, ASCE* **116**: 2485–2505.

Bažant, Z. P. and Panula, L. (1978). Practical prediction of time dependent deformations of concrete, *Materials and Structures* **11**: 307–316, 317–328, 425–434.

Bažant, Z. P. and Pijaudier-Cabot, G. (1988). Nonlocal continuum damage, localization instability and convergence, *Journal of Applied Mechanics, ASME* **55**: 287–293.

Bažant, Z. P. and Planas, J. (1998). *Fracture and Size Effect in Concrete and Other Quasibrittle Materials*, CRC Press, Boca Raton.

Bažant, Z. P. and Prasannan, S. (1989). Solidification theory for concrete creep: I. Formulation, II. Verification and application, *Journal of Engineering Mechanics, ASCE* **115**: 1691–1725.

Bažant, Z. P. and Prat, P. (1988). Microplane model for brittle plastic materials. I: Theory, II: Verification, *Journal of Engineering Mechanics, ASCE* **114**: 1672–1702.

Bažant, Z. P. and Prat, P. C. (1987). Creep of anisotropic clay: New microplane model, *Journal of Engineering Mechanics, ASCE* **113**: 1000–1064.

Bažant, Z. P. and Tsubaki, T. (1979). Concrete reinforcing net: Optimum slip-free limit design, *Journal of the Structural Division, ASCE* **105**: 1375–1383.

Bažant, Z. P. and Tsubaki, T. (1980). Total strain theory and path-dependence of concrete, *Journal of the Engineering Mechanics Division, ASCE* **106**: 1151–1173.

Bažant, Z. P. and Wang, T.-S. (1984). Spectral analysis of random shrinkage stresses in concrete, *Journal of Engineering Mechanics, ASCE* **110**: 173–186.

Bažant, Z. P. and Wu, S. T. (1974). Rate-type creep law of aging concrete based on Maxwell chain, *Materials and Structures* **7**: 45–60.

Bažant, Z. P. and Xi, Y. (1993). Stochastic drying and creep effects in concrete structures, *Journal of Structural Engineering, ASCE* **119**: 301–322.

Bažant, Z. P. and Xi, Y. (1995). Continuous retardation spectrum for solidification theory of concrete creep, *Journal of Engineering Mechanics, ASCE* **121**: 281–288.

Bažant, Z. P., Adley, M. D., Carol, I., Jirásek, M., Akers, S. A., Rohani, B., Cargile, J. D. and Caner, F. C. (2000a). Large-strain generalization of microplane model for concrete and applications, *Journal of Engineering Mechanics, ASCE* **126**: 971–980.

Bažant, Z. P., Caner, F. C., Carol, I, Adley, M. D. and Akers, S. (2000b). Microplane model M4 for concrete: I. Formulation with work-conjugate deviatoric stress, *Journal of Engineering Mechanics, ASCE* **126**: 944–953.

Bažant, Z. P., Carreira, D. J. and Walser, A. (1975). Creep and shrinkage in reactor containment shells, *Journal of the Structural Division, ASCE* **101**: 2117–2131.

Bažant, Z. P., Hauggaard, A. B. and Baweja, S. (1996). Microprestress solidification theory for aging and drying creep of concrete, in *Advances in Building and Materials Science*, A. Gerdes (ed), Aedificatio Publishers, Freiburg, Germany, pp. 111–130.

Bažant, Z. P., Hauggaard, B., Baweja, S. and Ulm, F.-J. (1997). Microprestress-solidification theory for concrete creep. I. Aging and drying effects, II. Algorithm and verification, *Journal of Engineering Mechanics, ASCE* **123**: 1188–1201.

Bažant, Z. P., Kim, J. J. and Brocca, M. (1999). Finite strain tube-squash test of concrete at high pressures and shear angles up to 70 degrees, *ACI Materials Journal* **96**: 580–592.

Bažant, Z. P., Křístek, V. and Vítek, J. L. (1992). Drying and cracking effects in box-girder bridge segment, *Journal of Structural Engineering, ASCE* **118**: 305–321.

Bažant, Z. P., Tsubaki, T. and Belytschko, T. (1980). Concrete reinforcing net: Safe design, *Journal of Engineering Mechanics, ASCE* **106**: 1895–1906.

Bažant, Z.P., Xiang, Y. and Prat, P. C. (1996). Microplane model for concrete. I: Stress-strain boundaries and finite strain, *Journal of Engineering Mechanics, ASCE* **122**: 245–254.

Belytschko, T. and Lasry, D. (1989). Localization limiters and numerical strategies for strain softening materials, in *Cracking and Damage – Strain Localization and Size Effect*, J. Mazars and Z. P. Bažant (eds), Elsevier, London and New York, pp. 269–294.

Belytschko, T., Liu, W. K. and Moran, B. (2000). *Nonlinear Finite Element Analysis for Continua and Structures*, Wiley, Chichester.

Benvenuto, E. (1991). *An Introduction to the History of Structural Mechanics*, Springer, New York.

Bergan, P. G. (1980). Solution algorithms for nonlinear structural problems, *Computers and Structures* **12**: 497–509.

Bertram, A. (1998). An alternative approach to finite plasticity based on material isomorphisms, *International Journal of Plasticity* **52**: 353–374.

Bertsch, P. K. and Findley, W. N. (1962). An experimental study of yield surfaces – corners, normality, Bauschinger and allied effects, *Proceedings of the 4th U.S. National Congress of Applied Mechanics*, ASME, New York, pp. 893–907.

Berveiller, M. and Zaoui, A. (1978). An extension of the self-consistent scheme to plastically flowing polycrystals, *Journal of the Mechanics and Physics of Solids* **26**: 325–344.

Bezukhov, N. I. (1936). On the theory of plastic bending (in Russian), *Vestnik Inzhenerov i Tekhnikov*, Vol. 10, Moskva.

Biezeno, C. B. and Hencky, H. (1928). On the general theory of elastic stability, *Koninklijke Akademie van Wettenschappen te Amsterdam, Proceedings of the Section of Sciences* **31**: 569–592.

Biezeno, C. B. and Hencky, H. (1929). On the general theory of elastic stability, *Koninklijke Akademie van Wettenschappen te Amsterdam, Proceedings of the Section of Sciences* **32**: 444–456.

Billington, D. P. (1965). *Thin Shell Concrete Structures*, McGraw-Hill, New York.

Bingham, E. C. (1919). Paint, a plastic material and not a viscous liquid; the measurement of its mobility and yield value, **19**.

Biot, M. A. (1934). Sur la stabilité de l'équilibre élastique. Equations de l'élasticité d'un milieu soumis à tension initiale, *Annales de la Société Scientifique de Bruxelles* Series B **54**: 18–21.

Biot, M. A. (1955). Variational principles of irreversible thermodynamics with application to viscoelasticity, *Physical Review* **97**: 1463–1469.

Biot, M. A. (1965). *Mechanics of Incremental Deformation*, Wiley, New York.

Bishop, J. F. W. and Hill, R. (1951). A theory of the plastic distortion of a polycrystalline aggregate under combined stress, *Philosophical Magazine* **42**: 414–427.

Bleich, F. (1952). *Buckling Strength of Metal Structures*, McGraw-Hill, New York.

Bleich, H. (1932). Über die Bemessung statisch unbestimmter Stahltragwerke unter Berücksichtigung des elastisch-plastischen Verhaltens des Baustoffes, *Bauingenieur* **13**: 261–267.

Bolton, M. (1979). *A Guide to Soil Mechanics*, MacMillan Educational.

Boltzmann, Z. (1874). Zur Theorie der elastischen Nachwirkungen, *Sitzungsberichte Akademie der Wissenschaften in Wien, Wiener Bericht 70, Wissenschafliche Abhandhungen* **1**: 279–306.

Borino, G., Fuschi, P. and Polizzotto, C. (1999). A thermodynamic approach to nonlocal plasticity and related variational approaches, *Journal of Applied Mechanics, ASME* **66**: 952–963.

Borja, R. (1991). Cam-clay plasticity, Part II, Implicit integration of constitutive equations based on a nonlinear elastic stress predictor, *Computer Methods in Applied Mechanics and Engineering* **88**: 225–240.

Braestrup, M. W. (1974). Plastic analysis of shear in reinforced concrete, *Magazine Concrete Research* **26**, No. 89.

Branson, D. E. (1968). Design procedures for computing deflections, *ACI Journal, Proceedings* **65**: 730–742.

Brocca, M. and Bažant, Z. P. (2000). Microplane constitutive model and metal plasticity, *Applied Mechanics Reviews ASME* **53**: 265–281.

Brocca, M. and Bažant, Z. P. (2001). Microplane finite element analysis of tube-squash test of concrete with shear angles up to 70°, *International Journal for Numerical Methods in Engineering*, in press.

Brocca, M., Bažant, Z. P. and Daniel, I. M. (2001). Microplane model for stiff foams and finite element analysis of sandwich failure by core indentation, *International Journal of Solids and Structures*, in press.

Bronkhorst, C. A., Kalindindi, S. R. and Anand, L. (1992). Polycrystalline plasticity and the evolution of crystallographic texture in FCC metals, *Philosophical Transactions of the Royal Society of London A* **341**: 443–477.

Brown, G. M. (1970). A self-consistent polycrystalline model for creep under combined stress states, *Journal of the Mechanics and Physics of Solids* **18**: 367–382.

Budianski, B. (1959). A reassessment of deformation theories of plasticity, *Journal of Applied Mechanics, ASME* **26**: 259–264.

Budianski, B. and Wu, T. T. (1962). Theoretical prediction of plastic strains of polycrystals, *Proc. Fourth U.S. National Congress of Applied Mechanics*, ASME, New York, pp. 1175–1185.

Budianski, B., Dow, N. F., Peters, R. W. and Ghephard, R. P. (1951). Experimental studies of polyaxial stress-strain laws of plasticity, *Proc. First US National Congress of Applied Mechanics*, ASME, New York, pp. 503–512.

Burland, J. B. (1965). The yielding and dilation of clay, *Géotechnique* **15**: 211–214.

Burzyński, W. (1929). Über die Anstrengungshypothesen, *Schweizerische Bauzeitung, Zürich* **94**: 259.

Butler, G. C. and McDowell, D. L. (1998). Polycrystal constraint and grain subdivision, *International Journal of Plasticity* **14**: 703–717.

Calladine, C. R. (1969). *Engineering Plasticity*, Pergamon Press, Oxford.

Caner, F. C. and Bažant, Z. P. (2000a). Microplane model M4 for concrete: II. Algorithm and calibration, *Journal of Engineering Mechanics, ASCE* **126**: 954–961.

Caner, F. C. and Bažant, Z. P. (2000b). Minimum confinement ensuring ductility of concrete column via tube-squash tests, *Report 2000-6/C808m*, Northwestern University, Evanston, Illinois, submitted to *Journal of Engineering Mechanics, ASCE*.

Caner, F. C., Bažant, Z. P. and Červenka, J. (2001). Vertex effect in strain-softening concrete at rotating principle axes, *Journal of Engineering Mechanics, ASCE*, in press.

Canova, G. R., Kocks, U. F. and Jonas, J. J. (1984). Theory of torsion texture development, *Acta Metall.* **32**: 211.

Carathéodory, C. (1909). Untersuchungen über die Grundlagen der Thermodynamik, *Mathematische Annalen* **67**: 355–386.

Carol, I. and Bažant, Z. P. (1993). Viscoelasticity with aging caused by solidification of nonaging constituent, *Journal of Engineering Mechanics, ASCE* **119**: 2252–2269.

Carol, I. and Bažant, Z. P. (1997). Damage and plasticity in microplane theory, *International Journal of Solids and Structures* **34**: 3807–3835.

Carol, I., Bažant, Z. P. and Prat, P. (1991). Geometric damage tensor based on microplane model, *Journal of Engineering Mechanics, ASCE* **117**: 2429–2448.

Carol, I., Jirásek, M. and Bažant, Z. P. (2001). A thermodynamically consistent approach to microplane theory. Part I: Free energy and consistent microplane stresses, *International Journal of Solids and Structures* **38**: 2921–2931.

Casimir, H. B. G. (1945). On Onsager's principal of microscopic reversibility, *Reviews of Modern Physics* **17**: 343–350.

Cauchy, A.-L. (1828). Sur les équations qui expriment les conditions d'équilibre ou les lois du mouvement intérieur d'un corps solide, élastique ou non élastique, *Ex. de Math.* **3**: 160–187.

Cauchy, A.-L. (1829). Sur l'équilibre et le mouvement intérieur des corps considérés comme des masses continues, *Ex. de Math.* **4**: 293–319.

CEB-FIP (1991). *CEB-FIP Model Code 1990, Design Code*. Thomas Telford, London.

Červenka, J., Červenka, V. and Eligehausen, R. (1998). Fracture-plastic material model for concrete, application to analysis of powder actuated anchors, in *Fracture Mechanics of Concrete Structures*, H. Mihashi and K. Rokugo (eds), Aedificatio Publishers, Freiburg, Germany, pp. 1107–1117.

Chaboche, J. L. (1978). Description thermodynamique et phénoménologique de la viscoplasticité cyclique avec endommagement, *Thése de doctorat*, Université Paris VI, Paris.

Chaboche, J. L. and Nouailhas, D. (1989). A unified constitutive model for cyclic viscoplasticity and its applications to various stainless steels, *Journal of Engineering Materials and Technology, ASME* **111**: 424–430.

Chaboche, J. L. and Rousselier, G. (1983). On the plastic and viscoplastic constitutive equations – Part I: Rules developed with internal variable concept, *Journal of Presure Vessel Technology, ASME* **105**: 153–158.

Chambon, R. (1989). Une classe de lois incrémentalement non linéaires pour les sols non visqueux, résolution de quelques problems de cohérence, *Comptes Rendus de l'Académie des Sciences, Paris, Série II* **308**: 1571–1576.

Chambon, R., Caillerie, D. and Hassan, N. E. (1998). One-dimensional localization studied with a second grade model, *European Journal of Mechanics/A: Solids* **17**: 637–656.

Charlez, P. A. (1991). *Rock Mechanics*, Edition Technip, Paris.

Charnes, A. and Lemke, C. E. (1954). Computational theory of linear programming, I: The bounded variable problem, *ONR Research Memorandum 10*, Carnegie Institute of Technology, Pittsburgh, Pennsylvania.

Chen, A. C. and Chen, W. F. (1975). Constitutive relations for concrete, *Journal of the Engineering Mechanics Division, ASCE* **101**: 465–481.

Chen, W. F. (1982). *Plasticity in Reinforced Concrete*, McGraw Hill, New York.

Chen, W. F. and Baladi, G. Y. (1985). *Soil Plasticity*, Developments in geotechnical engineering, Elsevier, Amsterdam.

Chen, W. F. and Liu, X. L. (1990). *Limit Analysis in Soil Mechanics*, Developments in geotechnical engineering, Elsevier, Amsterdam.

Chen, W. F. and Mizuno, E. (1990). *Nonlinear Analysis in Soil Mechanics*, Developments in geotechnical engineering, Elsevier, Amsterdam.

Christoffersen, J. and Hutchinson, J. W. (1979). A class of phenomenological corner theories in plasticity, *Journal of the Mechanics and Physics of Solids* **27**: 465–487.

Clarke, F. H. (1975). Generalized gradients and applications, *Transactions of the American Mathematical Society* **205**: 247–262.

Clausius, R. (1854). Über eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheorie, *Poggendorffs Annalen* **93**: 481–506.

Clausius, R. (1865). Über verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie, *Poggendorffs Annalen* **125**: 353–400.

Cohn, M. Z., Maier, G. and Grierson, D. (eds) (1979). *Engineering Plasticity by Mathematical Programming*, Pergamon Press, New York.

Coleman, B. D. (1964). Thermodynamics of material with memory, *Archive for Rational Mechanics and Analysis* **17**: 1–46.

Coleman, B. D. and Noll, W. (1963). The thermodynamics of elastic materials with heat conduction and viscosity, *Archive for Rational Mechanics and Analysis* **13**: 167–178.

Collins, M. P. (1978). Towards a rational theory for RC members in shear, *Journal of the Structural Division, ASCE* **104**: 396–408.

Collins, M. P. and Mitchell, D. (1991). *Prestressed Concrete Structures*, Prentice Hall, Englewood Cliffs, New Jersey.

Comi, C. and Corigliano, A. (1991). Dynamic shakedown in elastoplastic structures with general internal variable laws, *International Journal of Plasticity* **7**: 679–692.

Comi, C. and Perego, U. (1996). A generalized variable formulation for gradient dependent softening plasticity, *International Journal for Numerical Methods in Engineering* **39**: 3731–3755.

Considère, A. (1891). Résistance des pièces comprimées, *Congrés international des procédés de construction*, Vol. 3, Paris.

Corradi, L. and Maier, G. (1974). Dynamic non-shakedown theorem for elastic perfectly-plastic continua, *Journal of the Mechanics and Physics of Solids* **22**: 401–413.

Cosserat, E. and Cosserat, F. (1909). *Théorie des corps déformables*, A. Herrman et Fils, Paris.

Cotter, B. A. and Rivlin, R. S. (1955). Tensors associated with time-dependent stress, *Quarterly Journal of Applied Mathematics* **13**: 177–182.

Cottle, R. W., Pang, J.-S. and Stone, R. E. (1992). *The Linear Complementarity Problem*, Academic Press, Boston.

Cottrell, A. H. (1964). *The Mechanical Properties of Matter*, Wiley, New York.

Coulomb, C. A. (1776). Essai sur une application des règles de maximis et minimis à quelques problèmes de statique, relatifs à l'architecture, Mémoires de Mathématique et de Physique presentés à l'Académie Royale des Science par divers savants.

Craig, R. F. (1992). *Soil Mechanics*, 5th edition, Chapman & Hall, London.

Crisfield, M. A. (1981). A fast incremental iterative solution procedure that handles snapthrough, *Computers and Structures* **13**: 55–62.

Crisfield, M. A. (1983). An arc-length method including line searches and accelerations, *Computer Methods in Applied Mechanics and Engineering* **19**: 1269–1289.

Cuellar, V., Bažant, Z. P., Krizek, R. J. and Silver, M. L. (1977). Densification and hysteresis of sand under cyclic shear, *Journal of the Geotechnical Engineering Division, ASCE* **103**: 399–416.

Dafalias, Y. F. (1983). Corotational rates for kinematic hardening at large plastic deformation, *Journal of Applied Mechanics, ASME* **50**: 561–565.

Dantzig, G. B. (1949). Programming of interdependent activities, II, Mathematical model, *Econometrica* **17**: 200–211.

Dantzig, G. B. and Orchard-Hays, W. (1953). Notes on linear programming: Part V–alternate algorithm for the revised simplex method using product form for the inverse, *Research Memorandum RM-1268*, The Rand Corporation, Santa Monica, California.

Dantzig, G. B., Ford, L. R. and Fulkerson, D. R. (1956). A primal-dual algorithm for linear programs, *Annals of Mathematics Study No. 38*, Princeton University Press, Princeton, New Jersey, pp. 171–181.

Das, B. M. (1983). *Fundamentals of Soil Dynamics*, Elsevier, New York, Amsterdam and Oxford.

Das, B. M. (1997). *Advanced Soil Mechanics*, 2nd edition, Taylor & Francis, Washington, D.C.

de Boer, R. (1996). Highlights in the historical development of the porous media theory: Towards a consistent macroscopic theory, *Applied Mechanics Reviews ASME* **49**: 201–262.

de Borst, R. (1991). Simulation of strain localization: A reappraisal of the Cosserat continuum, *Engineering Computations* **8**: 317–332.

de Borst, R. and Mühlhaus, H. B. (1991). Continuum models for discontinuous media, in *Fracture Processes in Concrete, Rock and Ceramics*, J. G. M. van Mier, J. G. Rots and A. Bakker (eds), Vol. 2, E&FN Spon, London, pp. 601–618.

de Borst, R. and Mühlhaus, H. B. (1992). Gradient-dependent plasticity: Formulation and algorithmic aspects, *International Journal for Numerical Methods in Engineering* **35**: 521–539.

de Groot, S. R. and Mazur, P. (1962). *Nonequilibrium Thermodynamics*, North-Holland, Amsterdam.

de St. Vénant, B. (1870). *Comptes Rendus de l' Académie des Sciences, Paris* **70**: 368, 473.

Dennis, J. E. and Schnabel, R. B. (1983). *Numerical Methods for Unconstrained Minimization and Nonlinear Equations*, Prentice Hall, Englewood Cliffs, New Jersey.

Derski, W., Izbicki, R., Kisiel, I. and Mróz, Z. (1989). *Rock and Soil Mechanics*, Developments in geotechnical engineering, Elsevier, Amsterdam.

Desai, C. S. and Siriwardane, H. J. (1984). *Constitutive Laws for Engineering Materials, with Emphasis on Geological Materials*, Prentice-Hall, Englewood Cliffs, New Jersey.

Dienes, J. K. (1979). On the analysis of rotation and stress rate in deforming bodies, *Archives of Mechanics* **32**: 217–232.

DiMaggio, F. L. and Sandler, I. S. (1971). Material model for granular soils, *Journal of the Engineering Mechanics Division, ASCE* **97**: 935–950.

Dischinger, F. (1937). Untersuchungen über die Kriechsicherheit, die Elastische Verformung und das Kriechen des Betons bei Bogenbrücken, *Bauingenieur* **18**: 487–520, 539–552, 595–621.

Dischinger, F. (1939). Elastische und plastische Verformungen bei Eisenbetontragwerken, *Bauingenieur* **20**: 53–63, 286–294, 426–437, 563–572.

Disque, R. O. (1973). Inelastic K-factor for column design, *Engineering Journal AISC* **10**: 33–35.

Dogui, A. and Sidoroff, F. (1985). Kinematic hardening in large elastoplastic strain, *Engineering Fracture Mechanics* **21**: 685–695.

Dougill, J. W. (1976). On stable progressively fracturing solids, *Zeitschrift für angewandte Mathematik und Physik* **27**: 423–437.

Doyle, T. C. and Ericksen, J. L. (1956). Nonlinear elasticity, *Advances in Applied Mechanics* **4**: 53–115.

Dragon, A. and Mróz, Z. (1979). A continuum model for plastic-brittle behavior of rock and concrete, *International Journal of Engineering Science* **17**: 121–137.

Drucker, D. C. (1950). Some implications of work hardening and ideal plasticity, *Quarterly of Applied Mathematics* **7**: 411–418.

Drucker, D. C. (1952). A more fundamental approach to plastic stress-strain relations, *Proc. 1st U. S. National Congress of Applied Mechanics*, Chicago, pp. 487–491.

Drucker, D. C. (1956). On uniqueness in the theory of plasticity, *Quarterly Journal of Applied Mathematics* **14**: 35–42.

Drucker, D. C. (1959). A definition of stable inelastic material, *Journal of Applied Mechanics, ASME* **26**: 101–106.

Drucker, D. C. and Prager, W. (1952). Soil mechanics and plasticity analysis of limit design, *Quarterly Journal of Applied Mathematics* **10**: 157–162.

Drucker, D. C., Greenberg, H. J. and Prager, W. (1951). The safety factor of an elastic-plastic body in plane strain, *Journal of Applied Mechanics, ASME* **18**: 371–378.

Drucker, D. C., Prager, W. and Greenberg, H. J. (1952). Extended limit design theorems for continuous media, *Quarterly Journal of Applied Mathematics* **9**: 381–389.

Drugan, W. J. and Willis, J. R. (1996). A micromechanics-based nonlocal constitutive equation and estimates of representative volume element size for elastic composites, *Journal of the Mechanics and Physics of Solids* **44**: 497–524.

Dubé, J.-F., Pijaudier-Cabot, G. and La Borderie, C. (1996). Rate dependent damage model for concrete in dynamics, *Journal of Engineering Mechanics, ASCE* **122**: 939–947.

Dugdale, D. S. (1960). Yielding of steel sheets containing slits, *Journal of the Mechanics and Physics of Solids* **8**: 100–108.

Duhem, P. (1903). *Recherches sur l'Hydrodynamique*, Gauthier-Villars, Paris.

Duvaut, G. and Lions, J. L. (1972). *Les inéquations en mécanique et en physique*, Dunod, Paris.

Ekeland, I. and Temam, R. (1976). *Convex Analysis and Variational Problems*, North Holland, Amsterdam.

Engelen, R. A. B., Geers, M. G. D. and Baaijens, F. P. T. (2001). Nonlocal implicit gradient-enhanced elasto-plasticity for the modelling of softening behaviour, *International Journal of Plasticity*, in press.

Engesser, F. (1895). Über die Knickfragen, *Schweizerische Bauzeitung* **26**: 24–26.

England, G. L. and Illston, J. M. (1965). Methods of computing stress in concrete from a history of measured strain, *Civil Engineering and Public Works Review* pp. 513–517, 692–694, 845–847.

Eringen, A. C. (1966a). A unified theory of thermomechanical materials, *International Journal of Engineering Science* **4**: 179–202.

Eringen, A. C. (1966b). Mechanics of micromorphic materials, *Proc. 11th International Congress of Applied Mechanics*, Springer, Berlin, Heidelberg and New York, pp. 131–138.

Eringen, A. C. (1972). Linear theory of nonlocal elasticity and dispersion of plane waves, *International Journal of Engineering Science* **10**: 425–435.

Eringen, A. C. (1981). On nonlocal plasticity, *International Journal of Engineering Science* **19**: 1461–1474.

Eringen, A. C. (1983). Theories of nonlocal plasticity, *International Journal of Engineering Science* **21**: 741–751.

Estrin, Y., Sluys, L. J., Brechet, Y. and Molinari, A. (1998). A dislocation based gradient plasticity model, *Journal de Physique IV* **8**: 135–141.

Eterovic, A. L. and Bathe, K.-J. (1990). A hyperelastic-based large strain elasto-plastic constitutive formulation with combined isotropic-kinematic hardening using the logarithmic stress and strain measures, *International Journal for Numerical Methods in Engineering* **30**: 1099–1114.

Etse, G. and Willam, K. J. (1994). A fracture-energy based constitutive formulation for inelastic behavior of plain concrete, *Journal of Engineering Mechanics, ASCE* **120**: 1983–2011.

Exadaktylos, G. E. and Vardoulakis, I. (1998). Surface instability in gradient elasticity with surface energy, *International Journal of Solids and Structures* **35**: 2251–2277.

Faria, R., Oliver, J. and Cervera, M. (1998). A strain-based plastic viscous-damage model for massive concrete structures, *International Journal of Solids and Structures* **35**: 1533–1558.

Feenstra, P. H. and de Borst, R. (1995a). A plasticity model and algorithm for mode-I cracking in concrete, *International Journal for Numerical Methods in Engineering* **38**: 2509–2529.

Feenstra, P. H. and de Borst, R. (1995b). Constitutive model for reinforced concrete, *Journal of Engineering Mechanics, ASCE* **121**: 587–595.

Feenstra, P. H. and de Borst, R. (1996). A composite plasticity model for concrete, *International Journal of Solids and Structures* **33**: 707–730.

Fleck, N. A. and Hutchinson, J. W. (1993). A phenomenological theory for strain gradient effects in plasticity, *Journal of the Mechanics and Physics of Solids* **41**: 1825–1857.

Fleck, N. A. and Hutchinson, J. W. (1997). Strain gradient plasticity, *Advances in Applied Mechanics* **33**: 295–361.

Fleck, N. A., Muller, G. M., Ashby, M. F. and Hutchinson, J. W. (1994). Strain gradient plasticity: Theory and experiment, *Acta Metallurgica et Materialia* **42**: 475–487.

Flory, T. J. (1961). Thermodynamic relations for highly elastic materials, *Transactions of the Faraday Society* **57**: 829–838.

Fried, I. (1984). Orthogonal trajectory accession to the nonlinear equilibrium curve, *Computer Methods in Applied Mechanics and Engineering* **47**: 283–298.

Gao, H. and Huang, Y. (2001). Taylor-based nonlocal theory of plasticity, *International Journal of Solids and Structures* **38**: 2615–2637.

Gao, H., Huang, Y., Nix, W. D. and Hutchinson, J. W. (1999). Mechanism-based strain gradient plasticity – I. Theory, *Journal of the Mechanics and Physics of Solids* **47**: 1239–1263.

Gass, S. I. (1984). *Linear Programming: Methods and Applications*, 5th edition, McGraw-Hill, New York.

Geers, M. G. D. (1997). *Experimental Analysis and Computational Modeling of Damage and Fracture*, PhD thesis, Eindhoven University of Technology, The Netherlands.

Geers, M. G. D., Engelen, R. A. B. and Ubachs, R. J. M. (2001). On the numerical modelling of ductile damage with an implicit gradient-enhanced formulation, *Revue européenne des éléments finis* **10**: 173–191.

Gens, A., Jouanna, P. and Schrefler, B. A. (1995). *Modern Issues in Non-Saturated Soils*, Courses and lectures, International Center for Mechanical Sciences, Springer, Wien.

Gerard, G. and Becker, H. (1951). Column behavior under conditions of impact, *Journal of Aeronautical Sciences* **19**: 58–65.

Gerard, G. and Becker, H. (1957). Handbook of structural stability: Part i, buckling of flat plates, *Technical Note* 3781, National Advisory Committee for Aeronautics.

Germain, P. (1973). *Mécanique des milieux continus*, Masson, Paris.

Germain, P., Nguyen, Q. S. and Suquet, P. (1983). Continuum thermodynamics, *Journal of Applied Mechanics, ASME* **50**: 1010–1020.

Glanville, W. H. (1933). Creep of concrete under load, *The Structural Engineer* **11**: 54–73.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, Massachusetts.

Goldfarb, D. and Todd, M. J. (1989). Linear programming, in *Optimization*, G. L. Nemhauser, A. H. G. R. Kan and M. J. Todd (eds), North-Holland, Amsterdam, chapter 2, pp. 73–170.

Green, A. E. and Rivlin, R. S. (1964). Multipolar continuum mechanics, *Archive for Rational Mechanics and Analysis* **17**: 113–147.

Green, G. (1839). On the laws of reflection and refraction of light at the common surface of two non-crystallised media, *Trans. Cambridge Philosophical Society* **7**: 1–24.

Green, G. (1841). On the propagation of light in crystallised media, *Trans. Cambridge Philosophical Society* **7**: 121–140.

Green, S. J. and Swanson, S. R. (1973). Static constitutive relations for concrete, *Technical Report AFWL-TR-72-2*, Air Force Weapons Laboratory, Kirtland Air Force Base.

Greenberg, H. J. and Prager, W. (1951). Limit design of beams and frames, *Proceedings ASCE* **77**(59): 1–12.

Grüning, M. (1926). *Die Tragfähigkeit statisch unbestimmter Tragwerke aus Stahl bei beliebig häufig widerholter Belastung*, Julius Springer, Berlin.

Gudehus, G. (1973). Elastoplastische Stoffgleichungen für trockenen Sand, *Ingenieur-Archiv* **42**: 151–169.

Guggenheim, E. A. (1959). Thermodynamics, classical and statistical, in *Encyclopedia of Physics*, S. Flügge (ed), Vol. III/2, Springer-Verlag, Berlin, pp. 1–118.

Gvozdev, A. A. (1938). Determination of the value of failure load for statically indeterminate systems subject to plastic deformations (in Russian), in *Conference on Plastic Deformations 1936*, B. G. Galerkin (ed), Akademia Nauk SSSR, Moscow and Leningrad, pp. 19–38.

Gvozdev, A. A. (1960). The determination of the value of the collapse load for statically indeterminate systems undergoing plastic deformation, *International Journal of Mechanical Sciences* **1**: 322–335.

Halphen, B. and Nguyen, Q. S. (1975). Sur les matériaux standards généralisés, *Journal de Mécanique* **14**: 39–63.

Hamilton, R., Boyle, J. T., Shi, J. and Mackenzie, D. (1996). A simple upper-bound method for calculating approximate shakedown loads, *Journal of Presure Vessel Technology ASME* **120**: 195–199.

Hanson, J. A. (1953). A ten-year study of creep properties of concrete, *Concrete Lab. Report SP-38*, US Department of the Interior, Bureau of Reclamation, Denver, Colorado.

Harboe, E. M. *et al.* (1958). A comparison of the instantaneous and the sustained modulus of elasticity of concrete, *Concrete Lab. Report C-854*, Division of Engineering Laboratories, US Department of the Interior, Bureau of Reclamation, Denver, Colorado.

Hardy, G. M. and Riesz, M. (1915). The general theory of Dirichlet series, *Cambridge Tracts in Mathematics and Mathematical Physics*, No. 18, Cambridge University Press, Cambridge, MA.

Hartmann, L. (1896). *Distribution des déformations dans les métaux soumis à des efforts*, Berger-Levrault, Paris.

Hartmann, S., Lührs, G. and Haupt, P. (1997). An efficient stress algorithm with applications in viscoplasticity and plasticity, *International Journal for Numerical Methods in Engineering* **40**: 991–1013.

Haupt, P. and Tsamakis, C. (1986). On kinematic hardening and large plastic deformations, *International Journal of Plasticity* **2**, 279–293.

Hencky, H. (1923). Über einige statisch bestimmte Fälle des Gleichgewichts in plastischen Körpern, *Zeitschrift für angewandte Mathematik und Mechanik* **3**: 241–351.

Hencky, H. (1924). Zur Theorie plastischer Deformationen und der hierdurch im Material hervorgerufenen Nachspannungen, *Zeitschrift für angewandte Mathematik und Mechanik* **4**: 323–334.

Herren, S., Lowe, T. C., Asaro, R. J. and Needleman, A. (1989). Analysis of large-strain shear in rate-dependent face-centered cubic polycrystals: correlations of micro- and macromechanics, *Philosophical Transactions of the Royal Society of London A* **328**: 443–500.

Hibbler, R. (1994). *Structural Analysis*, 3rd edition, Prentice Hall, Englewood Cliffs, New Jersey.

Hill, R. (1947). A theory of the yielding and plastic flow of anisotropic materials, *Proceedings of the Royal Society A* **193**: 281–297.

Hill, R. (1949). The plastic yielding of notched bars under tension, *Quarterly Journal of Applied Mathematics* **2**: 40–52.

Hill, R. (1950). *The Mathematical Theory of Plasticity*, reprinted 1986, Oxford University Press, Oxford, UK.

Hill, R. (1951). On the state of stress in a plastic-rigid body at the yield point, *Philosophical Magazine* **42**: 868–875.

Hill, R. (1952). A note on estimating the yield-point loads in a plastic-rigid body, *Philosophical Magazine* **43**: 353–355.

Hill, R. (1966a). Continuum micromechanics of elastoplastic polycrystals, *Journal of the Mechanics and Physics of Solids* **13**: 89–101.

Hill, R. (1966b). Generalized constitutive relations for incremental deformation of metal crystals by multislip, *Journal of the Mechanics and Physics of Solids* **14**: 95–102.

Hill, R. (1972). On constitutive macro-variables for heterogeneous solids at finite strain, *Proceedings of the Royal Society A* **326**: 131–147.

Hill, R. and Rice, J. R. (1972). Constitutive analysis of elastic-plastic crystal at arbitrary strain, *Journal of the Mechanics and Physics of Solids* **20**: 401–413.

Hillerborg, A. (1990). Fracture mechanics concepts applied to rotational capacity of reinforced concrete beams, *Engineering Fracture Mechanics* **53**: 233–240.

Hillerborg, A., Modéer, M. and Peterson, P. E. (1976). Analysis of crack propagation and crack growth in concrete by means of fracture mechanics and finite elements, *Cement and Concrete Research* **6**: 773–782.

Hodge, P. G. (1957). Interaction curves for shear and bending of plastic beams, *Journal of Applied Mechanics, ASME* **24**: 453–456.

Hodge, P. G. (1959). *Plastic Analysis of Structures*, McGraw-Hill, New York.

Hoek, E. and Brown, E. T. (1980). Empirical strength criterion for rock masses, *Journal of the Geotechnical Engineering Division, ASCE* **106**: 1013–1035.

Hoffman, O. (1967). The brittle strength of orthotropic materials, *Journal of Composite Materials* **1**: 200–206.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor.

Holnicki-Szulc, J. and Gierlinski, J. T. (1989). Structural modifications simulated by virtual distortions, *International Journal for Numerical Methods in Engineering* **28**: 645–666.

Holzapfel, G. A. (2000). *Nonlinear Solid Mechanics: A Continuum Approach for Engineering*, Wiley, Chichester.

Honneff, H. and Mecking, H. (1978). A method for the determination of the active slip systems and orientation changes during single crystal deformation, in *Proc. 5th Int. Conf. on Texture of Materials (IOCOTOM 5)*, G. Gottstein and K. Lucke (eds), pp. 265–275.

Horne, M. R. (1949). Fundamental propositions in the plastic theory of structures, *Journal of the Institute Civil Engineers (London)* **34**: 174–177.

Huang, Y., Gao, H., Nix, W. D. and Hutchinson, J. W. (2000). Mechanism-based strain gradient plasticity – II. Analysis, *Journal of the Mechanics and Physics of Solids* **48**: 99–128.

Hughes, D. A. (1995). The evolution of deformation microstructure and local orientation, in *Proc. 16th Riso Int. Symp. on Materials Science: Microstructural and Crystallographic Aspects of Recrystallization*, N. Hansen, Y.-L. Liu, D. J. Jensen and B. Ralph (eds), Riso National Laboratory, Roskilde, Denmark, pp. 63–85.

Hughes, D. A. and Hansen, N. (1993). High angle boundaries and orientation distribution at large strains, *Scripta Metallurgica et Materialia* **33**: 315–321.

Hughes, T. J. R. (1980). Generalization of selective integration procedures to anisotropic and nonlinear media, *International Journal for Numerical Methods in Engineering* **15**: 1413–1418.

Hughes, T. J. R. (1987). *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Prentice Hall, Englewood Cliffs, New Jersey.

Hughes, T. J. R. and Pister, K. S. (1978). Consistent linearization in mechanics of solids and structures, *Computers and Structures* **9**: 391–397.

Hughes, T. J. R. and Shakib, F. (1986). Psedo-corner theory: A simple enhancement of J_2-flow theory for applications involving non-proportional loading, *Engineering Computations* **3**: 116–120.

Hughes, T. J. R. and Taylor, R. L. (1978). Unconditionally stable algorithms for quasi-static elasto/viscoplastic finite element analysis, *Computers and Structures* **8**: 169–173.

Hughes, T. J. R. and Winget, J. (1980). Finite rotation effects in numerical integration of rate constitutive equations arising in large-deformation analysis, *International Journal for Numerical Methods in Engineering* **15**: 1862–1867.

Hutchinson, J. W. (1965). Bounds and self-consistent estimates for creep of polycrystalline materials, *Proceedings of the Royal Society of London A* **348**: 101–127.

Hutchinson, J. W. (1970). Elastic plastic behavior of polycrystalline metals and composites, *Proceedings of the Royal Society of London A* **319**: 247–272.

Hutchinson, J. W. (1974). Plastic buckling, *Advances in Applied Mechanics* **14**: 67–144.

Hutchinson, J. W. (1997). Linking scales in mechanics, in *Advances in Fracture Research*, B. L. Karihaloo, Y.-W. Mai, M. I. Ripley and R. O. Ritchie (eds), Pergamon Press, New York, pp. 1–14.

Huterer, J., Brown, D. G., Yanchula, S. and Cheng, P. C. (1985). Darlington generating station vacuum building, *SMiRT8*, No. H6/2, Brussels.

Ioffe, A. D. and Tikhomirov, V. M. (1979). *Theory of Extremal Problems*, North Holland, Amsterdam, New York.

Ivey, H. J. (1961). Plastic stress-strain relations and yield surfaces for aluminum alloys, *Journal of Mechanical Engineering Science* **3**: 15–31.

Iwan, W. D. (1967). On a class of models for the yielding behavior of continuous and composite systems, *Journal of Applied Mechanics, ASME* **34**: 612–617.

Jaumann, G. (1911). Geschlossenes System physikalischer und chemischer Differentialgesetze, *Sitzungsberichte, Akademie der Wissenschaften in Wien* **120**: 385–530.

Jeremić, B. and Sture, S. (1998). Tensor objects in finite element programming, *International Journal for Numerical Methods in Engineering* **41**: 113–126.

Jirásek, M. (1997). Analytical and numerical solutions for frames with softening hinges, *Journal of Engineering Mechanics, ASCE* **123**: 8–14.

Jirásek, M. (1998). Nonlocal models for damage and fracture: Comparison of approaches, *International Journal of Soils and Structures*, **35**: 4133–4145.

Jirásek, M. and Bažant, Z. P. (1995). Macroscopic fracture characteristics of random particle systems, *International Journal of Fracture* **69**: 201–228.

Jirásek, M. and Zimmermann, T. (1998). Rotating crack model with transition to scalar damage, *Journal of Engineering Mechanics, ASCE* **124**: 277–284.

Johansen, K. W. (1943). *Brudlinieteorier (Yield Line Theory)*, Jul. Gjellerups Forlag, Copenhagen, translated by Cement and Concrete Association, London, 1962.

Johansson, M. (1996). *Generalized Viscoplasticity Coupled to Damage for Modeling Creep-Related Phenomena*, Master's thesis, Chalmers University of Technology, Göteborg, Sweden.

Joule, J. P. (1845). On the existence of an equivalent relation between heat and the ordinary forms of mechanical power, *Philosophical Magazine* **27**: 205–207.

Kamenjarzh, J. A. (1996). *Limit Analysis of Solids and Structures*, CRC Press, Boca Raton, Florida.

Kang, H. D. (1997). *Triaxial Constitutive Model for Plain and Reinforced Concrete Behavior*, PhD thesis, University of Colorado, Boulder, Colorado.

Kang, H. D. and Willam, K. J. (1999). Localization characteristics of triaxial concrete model, *Journal of Engineering Mechanics, ASCE* **125**: 941–950.

Kang, H. D. and Willam, K. J. (2000). Performance evaluation of elastoviscoplastic concrete model, *Journal of Engineering Mechanics, ASCE* **126**: 995–1000.

Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming, *Combinatorica* **4**: 373–396.

Karush, W. (1939). *Minima of Functions of Several Variables with Inequalities as Side Conditions*, Master's thesis, Department of Mathematics, University of Chicago.

Kazinczy, G. (1914). Experiments with clamped beams (in Hungarian), *Betonszemle* **2**: 68, 83, 101.

Key, S. W. and Krieg, R. D. (1982). On the numerical implementation of inelastic time dependent and time independent, finite strain constitutive equations in structural mechanics, *Computer Methods in Applied Mechanics and Engineering* **33**: 439–452.

Kirchhoff, G. (1852). Über die Gleichungen des Gleichgewichts eines elastischen Körpers bei nicht unendlich kleinen Verschiebungen seiner Teile, *Sitzungsberichte, Akademie der Wissenschaften in Wien* **9**: 762–773.

Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983). Optimization by simulated annealing, *Science* **220**: 671–680.

Kist, N. C. (1917). Does a stress analysis based on the proportionality of force and deformation lead to a good design of steel bridges and buildings? (in Dutch), *De Ingenieur* **4**: 743.

Kist, N. C. (1920). Die Zähigkeit des Materials als Grundlage für die Berechnung von Brücken, Hochbauten, und ähnlichen Konstruktionen aus Flußeisen, *Der Eisenbau* **11**: 425.

Koiter, W. T. (1953a). On the stress-strain relations and the general theorems of plasticity, *Technical report*, Labor. voor Toegepaste Mechanica, Techn. Hogeschool Delft.

Koiter, W. T. (1953b). Stress-strain relations, uniqueness and variational theorems for elastic-plastic materials with a singular yield surface, *Quarterly Journal of Applied Mathematics* **11**: 350–354.

Koiter, W. T. (1955). Some remarks on plastic shakedown theorems, *Proc. 8th International Congress of Applied Mechanics*, Istanbul, p. 220.

Koiter, W. T. (1956). A new general theorem on shake-down of elastic-plastic structures, *Proceedings Koninklijke Nederlandse Akademie van Wetenshappen B* **59**: 24–34.

Koiter, W. T. (1960). General theorems for elastic-plastic solids, *Progress in Solid Mechanics*, Vol. 1, North-Holland, Amsterdam, pp. 165–221.

Kolymbas, D. (1987). A novel constitutive law for soils, in *Constitutive Laws for Engineering Materials*, C. S. Desai (ed), Elsevier, New York, pp. 319–326.

Kratochvíl, J. (1988). Dislocation pattern formation in metals, *Revue de physique appliquée* **23**: 419–429.

Krieg, R. D. and Key, S. W. (1976). Implementation of a time dependent plasticity theory into structural computer programs, in *Constitutive Equations in Viscoplasticity: Computational and Engineering Aspects*, J. A. Stricklin and K. J. Saczlski (eds), ASME, New York, p. AMD-20.

Kröner, E. (1961). Zur plastischen Verformung des Vielkristalls, *Acta Metallurgica* **9**: 155–161.

Kröner, E. (1967). Elasticity theory of materials with long-range cohesive forces, *International Journal of Solids and Structures* **3**: 731–742.

Krumhansl, J. A. (1965). Generalized continuum field representation for lattice vibrations, in *Lattice Dynamics*, R. F. Wallis (ed), Pergamon Press, London and Oxford, pp. 627–634.

Kuhn, H. W. and Tucker, A. W. (1951). Nonlinear programming, in *Proc. of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman (ed), University of California Press, Berkeley and Los Angeles, pp. 481–492.

Kupfer, H. B. and Gerstle, K. H. (1973). Behavior of concrete under biaxial stresses, *Journal of the Engineering Mechanics Division, ASCE* **99**: 853–866.

Kupfer, H., Hilsdorf, H. K. and Rüsch, H. (1969). Behavior of concrete under biaxial stresses, *Journal of the American Concrete Institute* **66**: 656–666.

Křístek, V. and Bažant, Z. P. (1987). Shear lag effect and uncertainty in concrete box girder creep, *Journal of Structural Engineering, ASCE* **113**: 557–574.

Lacy, T. E., McDowell, D. L. and Talreja, R. (1999). Gradient concepts for evolution of damage, *Mechanics of Materials* **31**: 831–860.

Lade, P. V. (1977). Elastic-plastic stress-strain theory for cohesionless soils with curved yield surfaces, *International Journal of Solids and Structures* **13**: 1019–1035.

Lade, P. V. and Duncan, J. M. (1975). Elasto-plastic stress-strain theory for cohesionless soil, *Journal of the Geotechnical Engineering Division, ASCE* **101**: 1037–1053.

Lade, P. V. and Kim, M. K. (1995). Single hardening constitutive model for soil, rock and concrete, *International Journal of Solids and Structures* **32**: 1963–1978.

Ladevèze, P. (1992). A damage computational method for composite structures, *Computers and Structures* **44**: 79–87.

Lai, W. M., Rubin, D. and Krempl, E. (1991). *Introduction to Continuum Mechanics*, 3rd edition, Butterworth-Heinemann, Woburn, MA.

Launay, P. and Gachon, H. (1971). Strain and ultimate strength of concrete under triaxial stress, in *Proc. 1st Int. Conf. on Struct. Mech. in Reactor Technol. (SMiRT1)*, T. Jaeger (ed.), Commission of European Communities, Brussels, Vol. 3, pp. 1–12.

Lazić, J. D. and Lazić, V. B. (1984). Generalized age-adjusted effective modulus method for creep in composite beam structures. Part I–theory, *Cement and Concrete Research* **14**: 819–932.

Lazić, J. D. and Lazić, V. B. (1985). Generalized age-adjusted effective modulus method for creep in composite beam structures. Part II–applications, *Cement and Concrete Research* **15**: 1–12.

Leblond, J. B., Perrin, G. and Devaux, J. (1994). Bifurcation effects in ductile metals with non-local damage, *Journal of Applied Mechanics, ASME* **61**: 236–242.

Lee, E. H. (1969). Elastic-plastic deformation at finite strains, *Journal of Applied Mechanics, ASME* **36**: 1–6.

Lee, E. H. and Liu, D. T. (1967). Finite-strain elastic-plastic theory with applications to plane wave analysis, *Journal of Applied Physics* **38**: 19–27.

Lehmann, T. (1972). Einige Bemerkungen zu einer allgemeinen Klasse von Stoffgesetzen für große elasto-plastische Formänderungen, *Ingenieur-Archiv* **41**: 297–310.

Lemaitre, J. and Chaboche, J.-L. (1990). *Mechanics of Solid Materials*, Cambridge University Press, Cambridge, UK.

Lemke, C. E. (1954). The dual method of solving the linear programming problem, *Naval Research Logistics Quarterly* **1**: 36–47.

Leon, A. (1935). Über die Scherfestigkeit des Betons, *Beton und Eisen* **34**.

Levine, H. S. (1982). A two-surface plastic and microcracking model for plain concrete, *Proc. of the ASME Winter Annual Meeting*, ASME, New York, pp. 27–47.

Lévy, M. (1870). *Comptes Rendus de l'Académie des Sciences, Paris* **70**: 1323.

Lewis, R. W. and Schrefler, B. A. (1987). *The Finite Element Method in the Deformation and Consolidation of Porous Media*, Wiley, Chichester.

Lewis, R. W. and Schrefler, B. A. (1998). *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media*, 2nd edition, Wiley, Chichester.

Lin, F.-B. and Bažant, Z. P. (1986). Convexity of smooth yield surface of frictional material, *Journal of Engineering Mechanics, ASCE* **112**: 1259–1262.

Lin, F.-B., Bažant, Z. P., Chern, J.-C. and Marchertas, A. H. (1987). Concrete model with normality and sequential identification, *Computers and Structures* **26**: 1011–1025.

Lin, T. H. and Ito, M. (1965). Theoretical plastic distortion of a polycrystalline aggregate under combined and reversed stresses, *Journal of the Mechanics and Physics of Solids* **13**: 103–115.

Lin, T. H. and Ito, M. (1966). Theoretical plastic stress-strain relationship of a polycrystal, *International Journal of Engineering Science* **4**: 543–561.

Lorentz, E. (1999). *Lois de comportement à gradients de variables internes: Construction, formulation variationelle et mise en oeuvre numérique*, PhD thesis, LMT, Ecole Normale Supérieure de Cachan, Cachan, France.

Loret, B. and Prevost, J. H. (1990). Dynamic strain localization in elasto-(visco-) plastic solids, Part 1. General formulation and one-dimensional examples, *Computer Methods in Applied Mechanics and Engineering* **83**: 247–273.

Lubarda, V. and Krajcinovic, D. (1993). Damage tensors and the crack density distribution, *International Journal of Solids and Structures* **30**: 2859–2877.

Lubliner, J. (1986). Normality rules in large-deformation plasticity, *Mechanics of Materials* **5**: 29–34.

Lubliner, J. (1990). *Plasticity Theory*, Macmillan, New York.

Lubliner, J. (1991). A simple model of generalized plasticity, *International Journal of Solids and Structures* **28**: 769–778.

Lubliner, J., Oliver, J., Oller, S. and Oñate, E. (1989). A plastic-damage model for concrete, *International Journal of Solids and Structures* **25**: 299–326.

Lüders (1854). *Dinglers Polytechnisches Journal.*

Luenberger, D. G. (1973). *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Massachusetts.

Maier, G. (1969). Shakedown theory in perfect elastoplasticity with associated and non-associated flow laws: A finite element linear programming approach, *Meccanica, AIMETA* **4**: 250–260.

Maier, G. (1973). A shakedown matrix theory allowing for work hardening and second order geometric effects, in A. Sawczuk (ed.), *Foundations of Plasticity*, Vol. **1**, Noordoff, Leyden, pp. 417–433.

Maier, G. and Zavelani, A. (1970). On the behavior of axially compressed metallic beams (in Italian), *Costruzioni Metalliche* **22**: 282–297.

Maier, G., Garvelli, V. and Cocchetti, G. (2000). On direct methods for shakedown and limit analysis, *European Journal of Mechanics/A: Solids* **19**: S79-S100.

Maier, G., Giacomini, S. and Paterlini, F. (1979). Combined elastoplastic and limit analysis via restricted basis linear programming, *Computer Methods in Applied Mechanics and Engineering* **19**: 21–48.

Maier, G., Srinivasan, R. and Save, M. A. (1976). On limit design of frames using linear programming, *Journal of Structural Mechanics* **4**: 349–378.

Majid, K. I. and Celik, T. (1985). The elastic-plastic analysis of frames by the theorems of structural variation, *International Journal for Numerical Methods in Engineering* **21**: 671–681.

Malkus, D. and Hughes, T. (1978). Mixed finite element methods–reduced and selective integration techniques: A unification of concepts, *Computer Methods in Applied Mechanics and Engineering* **15**: 63–81.

Malvern, L. E. (1969). *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, New Jersey.

Mandel, J. (1964a). Conditions de stabilité et postulat de Drucker, in *Rheology and Soil Mechanics*, J. Kravtchenko and P. M. Sirieyes (eds), Springer, Berlin, pp. 56–68.

Mandel, J. (1964b). Contribution théorique à l'étude de l'écrouissage et des lois de l'écoulement plastique, *Proc. of the 11th International Congress on Applied Mechanics*, pp. 502–509.

Mandel, J. (1972). *Plasticité classique et viscoplasticité*, Springer-Verlag, Wien.

Mandel, J. (1974). Thermodynamics and plasticity, in *Foundations of Continuum Thermodynamics*, J. J. D. Domingos, M. N. R. Nina and J. H. Whitelaw (eds), Macmillan.

Marsden, J. E. and Hughes, T. J. R. (1983). *Mathematical Foundation of Elasticity*, Prentice-Hall, Englewood Cliffs, New Jersey.

Marti, P. (1979). Plastic analysis of reinforced concrete shear walls, *Colloquium on Plasticity in Reinforced Concrete, Copenhagen*, Vol. 28, International Association for Bridge and Structural Engineering, Zurich, Switzerland, pp. 51–69.

Marti, P. and Thürlimann, B. (1977). Fließbedingung für Stahlbeton mit Berücksichtigung der Betonzugfestigkeit, *Beton- und Stahlbetonbau* **72**: 7–12.

Maslov, G. N. (1941). Thermal stress states in concrete masses, with consideration of concrete creep (in Russian), *Izvestia Nauchno-Issledovatelskogo Instituta Gidrotekhniki, Gosenergoizdat* **28**: 175–188.

Matthies, H. and Strang, G. (1979). The solution of nonlinear finite element equations, *International Journal for Numerical Methods in Engineering* **14**: 1613–1626.

Maugin, G. A. (1992). *The Thermodynamics of Plasticity and Fracture*, Cambridge University Press, Cambridge, UK.

Mayer, J. R. (1845). *Die organische Bewegung in ihrem Zusammenhange mit dem Stoffwechsel*, Dechsler, Heilbronn.

McHenry, D. (1943). A new aspect of creep in concrete and its application to design, *Proceedings of the American Society for Testing and Materials* **43**: 1069–1086.

McMeeking, R. M. and Rice, J. R. (1975). Finite-element formulations for problems of large elastoplastic deformations, *International Journal of Solids and Structures* **11**: 601–616.

McMillan, F. R. (1916). Method of designing reinforced concrete slabs, discussion by A. C. Janni, *Transactions ASCE* **80**: 1738.

Melan, E. (1936). Theorie statisch unbestimmter Systeme aus idealplastischem Baustoff, *Sitzungsberichte, Akademie der Wissenschaften in Wien, Abt. II/a* **145**: 195–218.

Melan, E. (1938a). Der Spannungszustand eines Miseschen-Henckyschen Kontinuums bei veränderlicher Belastung, *Sitzungsberichte, Akademie der Wissenschaften in Wien, Abt. II/a* **147**: 73–87.

Melan, E. (1938b). Zur Plastizität des räumlichen Kontinuums, *Ingenieur-Archiv* **9**: 116–126.

Menzel, A. and Steinmann, P. (2000). On the continuum formulation of higher gradient plasticity for single and polycrystals, *Journal of the Mechanics and Physics of Solids* **48**: 1777–1796.

Menétrey, P. and Willam, K. J. (1995). A triaxial failure criterion for concrete and its generalization, *ACI Structural Journal* **92**: 311–318.

Miehe, C. and Schröder, J. (2001). A comparative study of stress update algorithms for rate-independent and rate-dependent crystal plasticity, *International Journal for Numerical Methods in Engineering* **50**: 273–298.

Mindlin, R. D. (1964). Microstructure in linear elasticity, *Archive for Rational Mechanics and Analysis* **16**: 51–78.

Mindlin, R. D. (1965). Second gradient of strain and surface tension in linear elasticity, *International Journal of Solids and Structures* **1**: 417–438.

Mohr, O. (1900). Welche Umstände bedingen die Elastizitätsgrenze und den Bruch eines Materials? *Civilingenieur: Zeitschrift des Vereins deutscher Ingenieure, VDI* **44–45**: 1524–1530, 1572–1577.

Moncarz, P. D. and Gerstle, K. H. (1981). Steel frames with nonlinear connections, *Journal of Structural Engineering, ASCE* **107**: 1427–1441.

Moreau, J. J. (1963). *Comptes Rendus de l'Académie des Sciences, Paris* **257**: 4117.

Moreau, J. J. (1966). *Fonctionnelles convexes*, Séminaire sur les équations aux dérivées partielles, Collège de France, Paris.

Moreau, J. J. (1970). Sur les lois de frottement, de viscosité et de plasticité, *Comptes Rendus de l'Académie des Sciences, Paris* **271**: 608–611.

Mróz, Z. (1967). On the description of anisotropic work-hardening, *Journal of the Mechanics and Physics of Solids* **15**: 163.

Mróz, Z. (1969). An attempt to describe the behavior of metals under cyclic loads using a more general work hardening model, *Acta Mechanica* **7**: 199.

Mróz, Z. (1980). On hypoelasticity and plasticity approaches to constitutive modeling of inelastic behavior of soils, *International Journal of Numerical and Analytical Methods in Geomechanics* **4**: 45–55.

Mróz, Z., Norris, V. A. and Zienkiewicz, O. C. (1979). Application of an anisotropic hardening model in the analysis of elasto-plastic deformation of soils, *Géotechnique* **29**: 1–34.

Mörsch, E. (1920). *Der Eisenbetonbau – Seine Theorie und Anwendung*, Part 1, 5th edition, Wittwer, Stuttgart.

Mörsch, E. (1922). *Der Eisenbetonbau – Seine Theorie und Anwendung*, Part 2, 5th edition, Wittwer, Stuttgart.

Mühlhaus, H. B. and Aifantis, E. C. (1991). A variational principle for gradient plasticity, *International Journal of Solids and Structures* **28**: 845–858.

Mühlhaus, H. B. and Vardoulakis, I. (1987). The thickness of shear band in granular materials, *Géotechnique* **37**: 271–283.

Müller, I. (1975). Thermodynamics of mixture of fluids, *Journal de Mécanique* **14**: 267–303.

Nádai, A. (1923). Der Beginn des Fließvorganges in einem tordierten Stab, *Zeitschrift für angewandte Mathematik und Mechanik* **3**: 442–454.

Nádai, A. (1950). *Theory of Flow and Fracture of Solids*, McGraw-Hill, New York, Toronto and London.

Naghdi, P. M., Essenburg, F. and Koff, W. (1958). An experimental study of initial and subsequent yield surfaces in plasticity, *Journal of Applied Mechanics, ASME* **25**: 201–209.

Nagtegaal, J. C. (1982). On the implementation of inelastic constitutive equations with special reference to large deformation problems, *Computer Methods in Applied Mechanics and Engineering* **33**: 469–484.

Nagtegaal, J. C. and Jong, J. E. D. (1981). Some computational aspects of elastic-plastic large strain analysis, *International Journal for Numerical Methods in Engineering* **17**: 15–41.

Nagtegaal, J. C., Parks, D. M. and Rice, J. (1974). On numerically accurate finite element solutions in the fully plastic range, *Computer Methods in Applied Mechanics and Engineering* **4**: 153–177.

Neal, B. G. (1956). *The Plastic Methods of Structural Analysis*, Chapman & Hall, London.

Needleman, A. (1987a). A continuum model for void nucleation by inclusion debonding, *Journal of Applied Mechanics, ASME* **54**: 525–531.

Needleman, A. (1987b). Material rate dependence and mesh sensitivity in localization problems, *Computer Methods in Applied Mechanics and Engineering* **67**: 68–85.

Needleman, A. and Tvergaard, V. (1998). Dynamic crack growth in a nonlocal progressively cavitating solid, *European Journal of Mechanics/A: Solids* **17**: 421–438.

Nielsen, M. P. and Braestrup, M. W. (1978). Shear strength of reinforced concrete beams without web reinforcement, *Magazine of Concrete Research* **30**, No. 104.

Nilsson, C. (1994). On nonlocal plasticity, strain softening and localization, *Report TVSM-1007*, Division of Structural Mechanics, Lund Institute of Technology, Lund, Sweden.

Nilsson, C. (1997). Nonlocal strain softening bar revisited, *International Journal of Solids and Structures* **34**: 4399–4419.

Noll, W. (1967). Material uniform inhomogeneous material bodies, *Archive for Rational Mechanics and Analysis* **27**: 1–32.

Noll, W. (1972). A new mathematical theory of simple materials, *Archive for Rational Mechanics and Analysis* **48**: 1–50.

Norris, C. H., Wilbur, J. B. and Utku, S. (1991). *Elementary Structural Analysis*, 4th edition, McGraw Hill, New York.

Norton, F. H. (1929). *The Creep of Steel at High Temperatures*, McGraw-Hill, New York.

Nova, R. and Wood, D. M. (1979). A constitutive model for sand in triaxial compression, *International Journal of Numerical and Analytical Methods in Geomechanics* **3**: 255–278.

Nye, J. F. (1953). Some geometrical relations in dislocated crystal, *Acta Metallurgica* **1**: 153–162.

Odqvist, F. K. G. (1933a). Die Verfestigung von flußeisenähnlichen Körpern, *Zeitschrift für angewandte Mathematik und Mechanik* **13**: 360–363.

Odqvist, F. K. G. (1933b). *Plasticity Theory with Applications (in Swedish)*, Royal Swedish Academy of Engineering Sciences, Stockholm.

Oldroyd, J. G. (1950). On the formulation of rheological equations of state, *Proceedings of the Royal Society A* **200**: 523–541.

Oliver, J. (1989). A consistent characteristic length for smeared cracking models, *International Journal for Numerical Methods in Engineering* **28**: 461–474.

Oller, S., Oñate, E., Miquel, J. and Botello, S. (1996). A plastic damage constitutive model for composite materials, *International Journal of Solids and Structures* **33**: 2501–2518.

Oller, S., Oñate, E., Oliver, J. and Lubliner, J. (1990). Finite element nonlinear analysis of concrete structures using a 'plastic-damage model', *Engineering Fracture Mechanics* **35**: 219–231.

Onsager, L. (1931). Reciprocal relations in irreversible processes, *Physical Review* **37**: 405–426.

Ortiz, M. and Pinsky, P. P. (1981). Global analysis method for the solution of elastoplastic and viscoplastic dynamic problems, *Report UCB/SESM 81/08*, Department of Civil Engineering, University of California, Berkeley.

Ortiz, M. and Popov, E. P. (1985). Accuracy and stability of integration algorithms for elastoplastic constitutive relations, *International Journal for Numerical Methods in Engineering* **21**: 1561–1576.

Ortiz, M. and Simo, J. (1986). An analysis of a new class of integration algorithms for elastoplastic constitutive relations, *International Journal for Numerical Methods in Engineering* **29**: 353–366.

Ottosen, N. S. (1977). A failure criterion for concrete, *Journal of Engineering Mechanics, ASCE* **103**: 527–535.

Ožbolt, J. and Bažant, Z. P. (1996). Numerical smeared fracture analysis: Nonlocal microcrack interaction approach, *International Journal for Numerical Methods in Engineering* **39**: 635–661.

Pamin, J. (1994). *Gradient-Dependent Plasticity in Numerical Simulation of Localization Phenomena*, PhD thesis, Delft University of Technology, Delft, The Netherlands.

Pan, J. and Rice, J. R. (1983). Rate sensitivity of plastic flow and implications for yield surface vertices, *International Journal of Solids and Structures* **19**: 973–987.

Pande, G. N. and Sharma, K. G. (1981). Implementation of computer procedures and stress-strain laws in geotechnical engineering, in *Implementation of Computer Procedures and Stress-Strain Laws in Geotechnical Engineering*, C. S. Desai and S. K. Saxena (eds), Acorn Press, Durham, N.C., pp. 575–590.

Pande, G. N. and Sharma, K. G. (1982). Multi-laminate model of clays – A numerical evaluation of the influence of rotation of the principal stress axis, *Report*, Deptartment of Civil Engineering, University College of Swansea.

Pande, G. N. and Xiong, W. (1982). An improved multi-laminate model of jointed rock masses, in *Proceedings of the International Symposium on Numerical Models in Geomechanics*, R. Dungar, G. N. Pande and G. Studder (eds), Balkema, Rotterdam, pp. 218–226.

Papalambros, P. Y. and Wilde, D. J. (2000). *Principles of Optimal Design*, 2nd edition, Cambridge University Press, Cambridge.

Park, R. and Gamble, W. L. (1980). *Reinforced Concrete Slabs*, Wiley, New York.

Pauw, A. (1960). Static modulus of elasticity of concrete as affected by density, *Journal of the American Concrete Institute* **32**: 679–687.

Peerlings, R. H. J., de Borst, R., Brekelmans, W. A. M. and de Vree, J. H. P. (1996). Gradient-enhanced damage for quasi-brittle materials, *International Journal for Numerical Methods in Engineering* **39**: 3391–3403.

Peerlings, R. H. J., Geers, M. G. D., de Borst, R. and Brekelmans, W. A. M. (2001). A critical comparison of nonlocal and gradient-enhanced softening continua, *International Journal of Solids and Structures*, in press.

Pekau, O. A. and Zhang, Z. X. (1994). Strain-space cracking model for concrete and its application, *Computers and Structures* **51**: 151–162.

Perić, D. (1993). On a class of constitutive equations in viscoplasticity: Formulation and computational issues, *International Journal for Numerical Methods in Engineering* **36**: 1365–1393.

Perić, D., Owen, D. R. J. and Honnor, M. E. (1992). A model for finite strain elastoplasticity based on logarithmic strains: Computational issues, *Computer Methods in Applied Mechanics and Engineering* **94**: 35–61.

Perzyna, P. (1963). The constitutive equations for rate sensitive plastic materials, *Quarterly of Applied Mathematics* **20**: 321–332.

Perzyna, P. (1966). Fundamental problems in viscoplasticity, *Advances in Applied Mechanics* **9**: 243–377.

Perzyna, P. (1971). Thermodynamic theory of viscoplasticity, *Advances in Applied Mechanics* **11**: 313–354.

Phillips, A. and Gray, G. A. (1961). Experimental investigation of corners in the yield surface, *Journal of Basic Engineering* **83**: 275–289.

Pickett, G. (1956). Effect of aggregates on shrinkage of concrete and a hypothesis concerning shrinkage, *ACI Journal* **52**: 581–590.

Pierce, D., Asaro, R. J. and Needleman, A. (1983). Material rate dependence and localized deformation in crystalline solids, *Acta Metallurgica* **31**: 1951–1976.

Pietruszczak, S. and Mróz, Z. (1981). Finite element analysis of deformation of strain-softening materials, *International Journal for Numerical Methods in Engineering* **17**: 327–334.

Pijaudier-Cabot, G. and Bažant, Z. P. (1987). Nonlocal damage theory, *Journal of Engineering Mechanics, ASCE* **113**: 1512–1533.

Pinsky, P., Ortiz, M. and Pister, K. S. (1983). Numerical integration of rate constitutive equations in finite deformation analysis, *Computer Methods in Applied Mechanics and Engineering* **40**: 137–158.

Planas, J. and Elices, M. (1993). Drying shrinkage effect on the modulus of rupture, in *Creep and Shrinkage of Concrete*, Z. P. Bažant and I. Carol (eds), E&FN Spon (Chapman & Hall), London, pp. 357–368.

Planas, J., Elices, M. and Guinea, G. V. (1993). Cohesive cracks versus nonlocal models: Closing the gap, *International Journal of Fracture* **63**: 173–187.

Planas, J., Guinea, G. V. and Elices, M. (1996). Basic issues on nonlocal models: Uniaxial modeling, *Technical Report 96-jp03*, Departamento de Ciencia de Materiales, ETS de Ingenieros de Caminos, Universidad Politécnica de Madrid, Spain.

Polizzotto, C. and Borino, G. (1996). Shakedown and steady-state responses of elastic-plastic solids in large displacements, *International Journal of Solids and Structures* **33**: 3415–3437.

Polizzotto, C., Borino, G. and Fuschi, P. (1997). Gradient plasticity – a thermodynamic formulation, in *Computational Plasticity*, D. R. J. Owen, E. Oñate and E. Hinton (eds), CIMNE, Barcelona, pp. 481–488.

Polizzotto, C., Borino, G. and Fuschi, P. (1998). A thermodynamically consistent formulation of nonlocal and gradient plasticity, *Mechanics Research Communications* **25**: 75–82.

Ponter, A. R. S. and Carter, K. F. (1997). Shakedown state simulation techniques based on linear elastic solutions, *Computer Methods in Applied Mechanics and Engineering* **140**: 259–279.

Potts, D. M. and Zdravković, L. (1999). *Finite Element Analysis in Geotechnical Engineering*, Thomas Telford, London.

Powrie, W. (1997). *Soil Mechanics: Concepts and Applications*, E & FN Spon, London.

Prager, W. (1948). Problem types in the theory of perfectly plastic materials, *Journal of Aeronautical Science* **15**: 337–341.

Prager, W. (1955). *Probleme der Plastizitätstheorie*, Birkhäuser, Basel, Stuttgart.

Prager, W. (1956). A new method of analyzing stresses and strains in work hardening plastic solids, *Journal of Applied Mechanics, ASME* **23**: 493–496.

Prager, W. (1959). *Introduction to Plasticity*, Addison-Wesley, Reading, MA.

Prager, W. and Hodge, P. G. (1951). *Theory of Perfectly Plastic Solids*, Wiley, New York.

Pramono, E. and Willam, K. (1989). Fracture energy-based plasticity formulation of plain concrete, *Journal of Engineering Mechanics, ASCE* **115**: 1183–1203.

Prandtl, L. (1903). Zur Torsion von prismatischen Stäben, *Zeitschrift für Physik* **4**: 758–759.

Prandtl, L. (1923). Anwendungsbeispiele zu einem Henckyschen Satz über das plastische Gleichgewicht, *Zeitschrift für angewandte Mathematik und Mechanik* **3**: 402–406.

Prandtl, L. (1924). Spannungsverteilung in plastischen Körpern, *Proc. 1st International Congress of Applied Mechanics*, Delft, pp. 43–54.

Prat, P. C. and Bažant, Z. P. (1991). Microplane model for triaxial deformation of saturated cohesive soils, *Journal of Geotechnical Engineering, ASCE* **117**: 891–912.

Prevost, J. H. (1977). Mathematical modeling of monotonic and cyclic undrained clay behavior, *International Journal of Numerical and Analytical Methods in Geomechanics* **1**: 195–216.

Prevost, J. H. (1978a). Anisotropic undrained stress-strain behavior of clays, *International Journal of Numerical and Analytical Methods in Geomechanics* **2**: 1075–1090.

Prevost, J. H. (1978b). Plasticity theory for soil stress-strain behavior, *Journal of the Engineering Mechanics Division, ASCE* **104**: 1177–1194.

Radenković, D. (1961). Théorèmes limites pour un matériaux de Coulomb à dilatation non standartisée, *Comptes Rendus de l'Académie des Sciences, Paris* **252**: 4103–4104.

Ramberg, W. and Osgood, W. (1943). Description of stress-strain curves by three parameters, *Technical Note 902*, National Advisory Committee for Aeronautics.

Ramm, E. (1981). Strategies for tracing the nonlinear response near limit point, in *Nonlinear Finite Element Analysis in Structural Mechanics*, W. Wunderlich, E. Stein and K.-J. Bathe (eds), Springer-Verlag, Berlin, pp. 63–89.

Rankine, W. J. M. (1858). *Manual of Applied Mechanics*, Glasgow.

Resende, L. (1987). A damage mechanics constitutive theory for the inelastic behavior of concrete, *Computer Methods in Applied Mechanics and Engineering* **60**: 57–93.

Reuß, A. (1930). Berücksichtigung der elastischen Formänderungen in der Plastizitätstheorie, *Zeitschrift für angewandte Mathematik und Mechanik* **10**: 266–274.

Reuß, A. (1933). Vereinfachte Berechnung der plastischen Formänderungs geschwindigkeiten bei Voraussetzung der Schubspannungs-fließbedingung, *Zeitschrift für angewandte Mathematik and Mechanik* **13**: 356–360.

Rice, J. R. (1970). On the structure of stress-strain relations for time-dependent plastic deformation of metals, *Journal of Applied Mechanics, ASME* **37**: 728–737.

Rice, J. R. (1971). Inelastic constitutive relations for solids: An internal variable theory and its application to metal plasticity, *Journal of the Mechanics and Physics of Solids* **19**: 433–455.

Rice, J. R. (1975). On the stability of dilatant hardening of saturated rock masses, *Journal of Geophysical Research* **80**: 1531–1536.

Rice, J. R. (1976). The localization of plastic deformation, in *Preprints, 14th Congress of International Union of Theoretical and Applied Mechanics*, W. Koiter (ed), North Holland, Amsterdam, pp. 207–220.

Richart, F. E., Brandtzaeg, A. and Brown, R. L. (1928). A study of the failure of concrete under combined compressive stresses, *Bulletin No. 185*, University of Illinois, Engineering Experiment Station, Urbana, Illinois.

Riks, E. (1979). An incremental approach to the solution of snapping and buckling problems, *International Journal of Solids and Structures* **15**: 529–551.

Riks, E. (1984). Some computational aspects of the stability analysis of nonlinear structures, *Computer Methods in Applied Mechanics and Engineering* **47**: 219–259.

Ristinmaa, M. and Ottosen, N. S. (1998). Viscoplasticity based on an additive split of the conjugated force, *European Journal of Mechanics/A: Solids* **17**: 207–235.

Ritter, W. (1899). Die Bauweise Hennebique, *Schweizerische Bauzeitung Zürich* **33**: 59–61.

Rivlin, R. S. (1948). Large elastic deformations of isotropic materials, I. Fundamental concepts, *Philosophical Transactions of the Royal Society of London A* **240**: 459–490.

Rockafellar, R. T. (1970). *Convex Analysis*, Princeton University Press, Princeton, New Jersey.

Rogula, D. (1965). Influence of spatial acoustic dispersion on dynamical properties of dislocations, *Bulletin de l'Académie Polonaise des Sciences, Série des sciences techniques* **13**: 337–343.

Rogula, D. (1982). Introduction to nonlocal theory of material media, in *Nonlocal Theory of Material Media*, D. Rogula (ed.), number 268 in *CISM Courses and Lectures*, Springer Verlag, Wien and New York, pp. 125–222.

Rollet, A. D., Stout, M. G. and Kocks, U. F. (1989). Polycrystal plasticity as applied to the problem of in-plane anisotropy in rolled cubic metals, in *Plasticity (2nd International Symposium on Plasticity and its Current Applications)*, A. S. Kahn and M. Tokuda (eds), Mie, Japan, pp. 69–72.

Roscoe, K. H. and Burland, J. B. (1968). On the generalised stress-strain behavior of "wet clay", in *Engineering Plasticity*, J. Heyman and F. A. Leckie (eds), Cambridge University Press, Cambridge, pp. 535–609.

Roscoe, K. H. and Schofield, A. N. (1963). Mechanical behavior of an idealised "wet clay", in *Proceedings of the 2nd European Conference on Soil Mechanics and Foundation Engineering*, Wiesbaden, pp. 47–54.

Rots, J. G. (1988). *Computational Modeling of Concrete Fracture*, PhD thesis, Delft University of Technology, Delft, The Netherlands.

Rubinstein, R. and Atluri, S. N. (1983). Objectivity of incremental constitutive relations over finite time steps in computational finite deformation analyses, *Computer Methods in Applied Mechanics and Engineering* **36**: 277–290.

Rudnicky, J. W. and Rice, J. R. (1975). Conditions for the localization of deformation in pressure-sensitive dilatant materials, *Journal of the Mechanics and Physics of Solids* **23**: 371–394.

Runesson, K., Ristinmaa, M. and Mähler, L. (1999). A comparison of viscoplasticity formats and algorithms, *Mechanics of Cohesive-Frictional Materials* **4**: 75–98.

Rüsch, H., Jungwirth, D. and Hilsdorf, H. (1973). Kritische Sichtung der Einflüsse von Kriechen und Schwinden des Betons auf das Verhalten der Tragwerke, *Beton- und Stahlbetonbau* **68**: 49–60, 76–86, 152–158.

Rzhanitsyn, A. R. (1949). *Structural Analysis with Plastic Properties of Materials Taken Into Account (in Russian)*, Stroyizdat, Moscow.

Sakarovitch, M. (1983). *Linear Programming*, Springer-Verlag, New York.

Salençon, J. (1974). Un example de non validité de la théorie classique des charges limites pour un système non standard, *Problems of Plasticity*, Noordhoff, Leyden, pp. 432–439.

Salençon, J. (1977). *Applications of the Theory of Plasticity in Soil Mechanics*, Wiley, Chichester, New York.

Sandler, I. S., DiMaggio, F. L. and Baladi, G. Y. (1976). Generalized cap model for geologic materials, *Journal of the Geotechnical Engineering Division, ASCE* **102**: 683–699.

Saouridis, C. (1988). *Identification et numérisation objectives des comportements adoucissants: Une approche multiéchelle de l'endommagement du béton*, PhD thesis, Université Paris VI.

Saouridis, C. and Mazars, J. (1992). Prediction of the failure and size effect in concrete via a bi-scale damage approach, *Engineering Computations* **9**: 329–344.

Save, M. and Prager, W. (eds) (1985). *Structural Optimization*, Vol. 1: Optimality Criteria, Plenum Press, New York and London.

Save, M. and Prager, W. (eds) (1990). *Structural Optimization*, Vol. 2: Mathematical Programming, Plenum Press, New York and London.

Saxcé, G. D. and Bousshine, L. (1998). Limit analysis theorems for implicit standard materials: Application to the unilateral contact with dry friction and non-associated flow rules in soils and rocks, *International Journal of Mechanical Sciences* **40**: 387–398.

Schapery, R. A. (1962). Approximate methods of transform inversion for viscoelastic stress analysis, *Proc. 4th U.S. National Congress on Applied Mechanics*, Vol. 2, ASME, Berkeley, CA, pp. 1075–1085.

Schlaich, J., Schafer, K. and Jannewein, M. (1987). Toward a consistent design for structural concrete, *PCI Journal* **32**: 75–150.

Schleicher, F. (1926). Der Spannungszustand an der Fließgrenze (Plastizitätsbedingung), *Zeitschrift für angewandte Mathematik und Mechanik* **6**: 199–216.

Schofield, A. N. and Wroth, C. P. (1968). *Critical State Soil Mechanics*, McGraw-Hill, New York.

Schreyer, H. and Chen, Z. (1986). One-dimensional softening with localization, *Journal of Applied Mechanics, ASME* **53**: 791–797.

Shanley, F. R. (1947). Inelastic column theory, *Journal of the Aeronantical Sciences* **14**: 261–268.

Simo, J. C. (1992). Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory, *Computer Methods in Applied Mechanics and Engineering* **99**: 62–112.

Simo, J. C. and Govindjee, S. (1991). Non-linear B-stability and symmetry preserving return mapping algorithms for plasticity and viscoplasticity, *International Journal for Numerical Methods in Engineering* **31**: 151–176.

Simo, J. C. and Hughes, T. J. R. (1998). *Computational Inelasticity*, Springer-Verlag, New York.

Simo, J. C. and Ortiz, M. (1985). A unified approach to finite deformation elastoplasticity based on the use of hyperelastic constitutive equations, *Computer Methods in Applied Mechanics and Engineering* **49**: 221–245.

Simo, J. C. and Pister, K. S. (1984). Remarks on rate constitutive equations for finite deformation problems: Computational implications, *Computer Methods in Applied Mechanics and Engineering* **46**: 201–215.

Simo, J. C. and Rifai, M. S. (1990). A class of mixed assumed strain methods and the method of incompatible modes, *International Journal for Numerical Methods in Engineering* **29**: 1595–1638.

Simo, J. C. and Taylor, R. L. (1985). Consistent tangent operators for rate independent elasto-plasticity, *Computer Methods in Applied Mechanics and Engineering* **48**: 101–118.

Simo, J. C. and Taylor, R. L. (1986). A return mapping algorithm for plane stress elastoplasticity, *International Journal for Numerical Methods in Engineering* **29**: 649–670.

Simo, J. C., Kennedy, J. G. and Govindjee, S. (1988). Non-smooth multisurface plasticity and viscoplasticity, *International Journal for Numerical Methods in Engineering* **26**: 2161–2186.

Simo, J. C., Taylor, R. L. and Pister, K. (1985). Variational and projection methods for the volume constraint in finite deformation elastoplasticity, *Computer Methods in Applied Mechanics and Engineering* **51**: 177–208.

Sluys, L. J. (1992). *Wave Propagation, Localisation and Dispersion in Softening Solids*, PhD thesis, Delft University of Technology, Delft, The Netherlands.

Smith, D. L. (ed.) (1990). *Mathematical Programming Methods in Structural Plasticity*, Springer-Verlag, New York.

Smith, S. (1987). *On Fundamental Aspects of Concrete Behavior*, Master's thesis, University of Colorado, Boulder, Colorado.

Stein, E. and Huang, Y. (1994). An analytical method for shakedown problems with linear kinematic hardening materials, *International Journal of Solids and Structures* **31**: 2433–2444.

Steinmann, P. and Willam, K. (1992). Localization within the framework of micropolar elasto-plasticity, in *Advances in Continuum Mechanics*, O. Brüller, V. Mannl and J. Najar (eds), Springer-Verlag, Berlin, pp. 296–313.

Strang, G. (1986). *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, Massachusetts.

Strayer, J. K. (1989). *Linear Programming and its Applications*, Undergraduate texts in mathematics, Springer-Verlag, New York.

Stroud, A. H. (1971). *Approximate Calculation of Multiple Integrals*, Prentice-Hall, Englewood Cliffs, New Jersey.

Strömberg, L. and Ristinmaa, M. (1996). FE-formulation of a nonlocal plasticity theory, *Computer Methods in Applied Mechanics and Engineering* **136**: 127–144.

Stumpf, H. (1993). Theoretical and computational aspects in the shake-down analysis of finite elastoplasticity, *International Journal of Plasticity* **9**: 583–602.

Stüssi, F. and Kollbrunner, C. F. (1935). Beitrag zum Traglastverhalten, *Die Bautechnik* Vol. 13, No. 21.

Svendsen, B., Arndt, S., Klingbeil, D. and Sievert, R. (1998). Hyperelastic models for elastoplasticity with non-linear isotropic and kinematic hardening at large deformation, *International Journal of Solids and Structures* **35**: 3363–3389.

Symonds, P. S. (1951). Shakedown in continuous media, *Journal of Applied Mechanics, ASME* **18**: 85–89.

Symonds, P. S. and Neal, B. G. (1951). Recent progress in the plastic methods of structural analysis, *Journal of the Franklin Institute* **252**: 383.

Tamagnini, C., Viggiani, G. and Chambon, R. (2000). A review of two different approaches to hypoplasticity, *in* D. Kolymbas (ed), *Constitutive Modeling of Granular Materials*, Springer-Verlag, Berlin pp. 107–145.

Taylor, G. I. (1934). The mechanics of plastic deformation of crystals. Part I Theoretical, *Proceedings of the Royal Society of London* **145**: 362–387.

Taylor, G. I. (1938). Plastic strain in metals, *Journal of the Institute of Metals* **62**: 307–324.

Taylor, G. I. and Elam, C. F. (1923). The distortion of an aluminium crystal during a tensile test, *Proceedings of the Royal Society of London A* **102**: 643–667.

Taylor, G. I. and Elam, C. F. (1925). The plastic extension and fracture of aluminium crystals, *Proceedings of the Royal Society of London A* **108**: 28–51.

Taylor, G. I. and Elam, C. F. (1926). The distortion of iron crystals, *Proceedings of the Royal Society of London A* **112**: 337–361.

Taylor, G. I. and Quinney, H. (1931). The plastic distortion of metals, *Philosophical Transactions of the Royal Society of London A* **230**: 323–362.

Terzaghi, K. (1936). The shearing resistance of a saturated soil and the angle between the planes of shear, *Proceedings of the 1st International Conference of Soil Mechanics and Foundation Engineering*, Harvard, Massachusetts, pp. 54–56.

Terzaghi, K. (1946). *Theoretical Soil Mechanics*, Chapman & Hall, London.

Terzaghi, K. and Peck, R. B. (1948). *Soil Mechanics in Engineering Practice*, Wiley, New York.

Terzaghi, K., Peck, R. B. and Mesri, G. (1996). *Soil Mechanics in Engineering Practice*, 3rd edition, Wiley, New York.

Thomson, W. (1851). On the dynamical theory of heat, with numerical results deduced from Mr. Joule's equivalent of a thermal unit, and M. Regnault's observations on steam, *Transactions of the Royal Society of Edinburgh* **20**: 261–268, 289–298.

Toupin, R. A. (1962). Elastic materials with couple-stress, *Archive for Rational Mechanics and Analysis* **11**: 385–414.

Tresca (1868). Mémoire sur l'écoulement des corps solides, *Mémoires présentés par divers savants* **18**: 733–799.

Triantafyllidis, N. and Aifantis, E. C. (1986). A gradient approach to localization of deformation: I. Hyperelastic materials, *Journal of Elasticity* **16**: 225–237.

Trost, H. (1967). Auswirkungen des Superpositionsprinzip auf Kriech- und Relaxations-Probleme bei Beton und Spannbeton, *Beton- und Stahlbetonbau* **62**: 230–238, 261–269.

Truesdell, C. (1955a). Hypo-elasticity, *Journal of Rational Mechanics and Analysis* **4**: 83–133.

Truesdell, C. (1955b). The simplest rate theory of pure elasticity, *Communications in Pure and Applied Mathematics* **8**: 123–132.

Truesdell, C. (1969). *Rational Thermodynamics*, McGraw-Hill, New York.

Truesdell, C. and Noll, W. (1960). *The Nonlinear Field Theories*, Handbuch der Physik III/1, Springer-Verlag, Berlin.

Truesdell, C. and Toupin, R. A. (1960). *Classical Field Theories of Mechanics*, Handbuch der Physik III/1, Springer-Verlag, Berlin.

Tsai, S. W. and Wu, E. M. (1971). A general theory of strength for anisotropic materials, *Journal of Composite Materials* **5**: 58–80.

Tsamakis, C. (1996). Kinematic hardening rules at large deformations, Part I: A constitutive approach, *Continuum Mechanics and Thermodynamics* **8**: 215–231.

Tvergaard, V. and Hutchinson, J. W. (1992). The relaion between crack growth resistance and fracture process parameters in elastic-plastic solids, *Journal of the Mechanics and Physics of Solids* **40**: 1377–1397.

Tvergaard, V. and Needleman, A. (1995). Effects of nonlocal damage in porous plastic solids, *International Journal of Solids and Structures* **32**: 1063–1077.

van Mier, J. G. M. (1984). *Strain-Softening of Concrete Under Multiaxial Loading Conditions*, PhD thesis, Technical University of Eindhoven, Eindhoven, The Netherlands.

Vardoulakis, I. and Aifantis, E. C. (1991). A gradient flow theory of plasticity for granular materials, *Archives of Mechanics* **87**: 197–217.

Vecchio, F. and Collins, M. P. (1986). The modified compression field theory for reinforced concrete elements subjected to shear, *ACI Journal* **83**: 219–231.

Vermeer, P. A. and Brinkgreve, R. B. J. (1994). A new effective non-local strain measure for softening plasticity, in *Localization and Bifurcation Theory for Soils and Rocks*, R. Chambon, J. Desrues and I. Vardoulakis (eds), Balkema, Rotterdam, pp. 89–100.

Volterra, V. (1909). Sulle equazioni integrodifferenziali della teoria dell'elasticita, *Atti Reale Accad. Lincei* **18**: 295.

von Kármán, T. (1909). Untersuchungen über die Knickfestigkeit, *Mitteilungen über Forschungsarbeiten auf dem Gebiete des Ingenieurwesens 81*, Springer, Berlin.

von Mises, R. (1913). Mechanik der festen Körper im plastisch-deformablem Zustand, *Göttinger Nachrichten Math.-Phys. Klasse* **1**: 582–592.

von Mises, R. (1928). Mechanik der plastischen Formänderung in Kristallen, *Zeitschrift für angewandte Mathematik und Mechanik* **8**: 161–185.

Wagner, H. (1929). Ebene Blechwandträger mit sehr dünnem Stegblech, *Zeitschrift für Flugtechnik und Motorluftschiffahrt* **20**: 8–12.

Walsh, G. R. (1985). *An Introduction to Linear Programming*, 2nd edition, Wiley, Chichester.

Wang, W. M., Sluys, L. J. and de Borst, R. (1997). Viscoplasticity for instabilities due to strain softening and strain-rate softening, *International Journal for Numerical Methods in Engineering* **40**: 3839–3864.

Weber, G. and Anand, L. (1990). Finite deformation constitutive equations and a time integration procedure for isotropic, hyperelastic-viscoplastic solids, *Computer Methods in Applied Mechanics and Engineering* **79**: 173–202.

Weertman, J. and Weertman, J. R. (1964). *Elementary Dislocation Theory*, Macmillan, New York.

Weichert, D. (1990). Unified formulation of statical shakedown criteria for geometrically nonlinear problems, *International Journal of Plasticity* **6**: 433–447.

Weichert, D. and Maier, G. (eds) (2000). *Inelastic Analysis of Structures under Variable Loads*, Kluwer Academic Publishers, Dordrecht, Boston and London.

Weihe, S. (1989). *Implicit Integration Schemes for Multi-surface Yield Criteria Subjected to Hardening/Softening Behavior*, Master's thesis, University of Colorado at Boulder.

West, H. (1993). *Fundamentals of Structural Analysis*, Wiley, New York.
Wilkins, M. L. (1964). Calculation of elastic-plastic flow, in *Methods of Computational Physics*, B. Adler (ed), Vol. 3, Academic Press, New York, pp. 211–262.
Willam, K. J. and Warnke, E. P. (1974). Constitutive model for the triaxial behavior of concrete, *Concrete Structures Subjected to Triaxial Stresses*, Vol. 19 of *IABSE Report*, International Association of Bridge and Structural Engineers, Zurich, pp. 1–30.
Wittmann, F. H. (1974). Bestimmung physikalischer Eigenschaften des Zementsteins, *Deutscher Ausschuss Stahlbeton*, No. 232.
Wittmann, F. H. (1982). Creep and shrinkage mechanisms, in *Creep and Shrinkage in Concrete Structures*, Z. P. Bažant and F. H. Wittmann (eds), Wiley, London, chapter 6, pp. 129–162.
Wittmann, F. H., Bažant, Z. P., Alou, F. and Kim, J. K. (1987). Statistics of shrinkage test data, *Cement, Concrete and Aggregates* **9**: 129–153.
Wood, D. M. (1990). *Soil Behavior and Critical State Soil Mechanics*, Cambridge University Press, Cambridge.
Woods, L. C. (1975). *The Thermodynamics of Fluid Systems*, Clarendon Press, Oxford.
Yang, B. L., Dafalias, Y. F. and Herrmann, L. P. (1985). A bounding surface plasticity model for concrete, *Journal of Engineering Mechanics, ASCE* **111**: 359–380.
Yazdani, S. and Karnawat, S. (1996). A constitutive theory for brittle solids with application to concrete, *International Journal of Damage Mechanics* **5**: 93–110.
Yazdani, S. and Schreyer, H. L. (1990). Combined plasticity and damage mechanics model for plain concrete, *Journal of Engineering Mechanics, ASCE* **116**: 1435–1450.
Zarka, J., Frelat, J., Inglebert, G. and Kasmai-Navidi, P. (1988). *A New Approach to Inelastic Analysis of Structures*, Martinus Nijhoff Publishers, Dordrecht.
Zbib, H. M. and Aifantis, E. C. (1988). On the localization and postlocalization behavior of plastic deformation, *Res Mechanica* **23**: 261–305.
Ziegler, H. (1959). A modification of Prager's hardening rule, *Quarterly Journal of Applied Mathematics* **17**: 55–65.
Ziegler, H. (1963). Some extremum principles in irreversible thermodynamics with application to continuum mechanics, *Progress in Solid Mechanics*, Vol. IV, North Holland, Amsterdam, pp. 93–193.
Ziegler, H. (1983). *An Introduction to Thermomechanics*, North-Holland, Amsterdam.
Zienkiewicz, O. C. and Pande, G. N. (1977). Time-dependent multi-laminate model of rocks – A numerical study of deformation and failure of rock masses, *International Journal of Numerical and Analytical Methods in Geomechanics* **1**: 219–247.
Zienkiewicz, O. C. and Taylor, R. L. (2000). *The Finite Element Method*, 5th edition, Butterworth-Heinemann, Oxford.
Zienkiewicz, O. C., Watson, M. and King, I. P. (1968). A numerical method of viscoelastic stress analysis, *International Journal of Mechanical Sciences* **10**: 807–827.

Author Index

Subject Index

About the authors

Milan Jirásek received his first degree in Structural Engineering from the Czech Technical University in Prague in 1985, and his Ph.D. in Theoretical and Applied Mechanics from Northwestern University, USA, in 1993. His dissertation, supervised by Professor Zdeněk Bažant, treated various aspects of quasibrittle fracture in concrete, rock, and ice. During the years 1986–91 and 1993–94 he was an Assistant Professor in the Department of Structural Mechanics of the Czech Technical University in Prague, where he conducted research into plastic analysis of structures and nonlinear dynamics. Since 1994 he has worked as a Research Associate in the Laboratory of Structural and Continuum Mechanics of the Swiss Federal Institute of Technology at Lausanne (EPFL). His research activities are centered around theoretical and numerical modeling of inelastic materials and structures, in particular of fracture and damage in quasibrittle materials such as concrete. Milan Jirásek is a member of the Board of Directors of the International Association of Fracture Mechanics of Concrete and Concrete Structures (IA-FraMCoS). He can be reached at Milan.Jirasek@epfl.ch; his web site is located at http://lscpc4.epfl.ch/milan.

Born and educated in Prague, **Zdeněk Bažant** has taught at Northwestern University since 1969 where he was the Director of the Center for Geomaterials during the years 1981–87. He became Professor in 1973 and Murphy Chair Professor in 1990. He is a member of the US National Academy of Engineering and Foreign Member of the Austrian Academy of Sciences, and was the Editor of the American Society of Civil Engineers' (ASCE) Journal of Engineering Mechanics during 1988–94. Bažant was president of two scientific societies (SES and IA–FraMCoS) and the recipient of Honorary Doctorates from the universities of Prague, Karlsruhe, Boulder and Milan as well as many awards, including SES Prager Medal, ASME Warner Medal, and ASCE Newmark Medal and Croes Medal. He is a Registered Structural Engineer of Illinois and has been a consultant to Argonne National Laboratory and many firms. He has authored over 400 refereed journal articles, edited 14 books and authored five other books: *Stability of Structures* (with L. Cedolin), *Fracture and Size Effect* (with J. Planas), *Concrete at High Temperatures* (with M. Kaplan), *Creep of Concrete*, and *Scaling of Structural Strength*.