

LINEAR LIBRARY
C01 0068 0681



**THEORETICAL AND NUMERICAL
ASPECTS OF PROBLEMS IN
FINITE-STRAIN PLASTICITY**

R. A. Eve

Ph.D. Thesis

1992



The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**THEORETICAL AND NUMERICAL
ASPECTS OF PROBLEMS IN
FINITE-STRAIN PLASTICITY**

by

R. A. Eve

A thesis submitted in fulfilment of the requirements for
the degree of Doctor of Philosophy

Department of Applied Mathematics
University of Cape Town

April 1992

ABSTRACT

A new internal variable theory of plasticity is presented. This theory is developed within a framework of non-smooth convex analysis; a unification of ideas concerning the postulates of plasticity is achieved by using the powerful tools provided by results in this branch of mathematics. A firm mathematical foundation for the study of qualitative aspects of problems involving plastic deformations is provided. Among the features of the theory is the establishment of a clear relationship between conventional formulations, which make use of yield functions, and those formulated in terms of a dissipation function. The role of the principle of maximum plastic work is also made precise.

Attention is focussed on application of the theory to finite-strain plasticity. Quasi-static initial-boundary-value problems involving large plastic deformations are considered. An incremental form of such problems arises from a discretisation in time. A variational form of the incremental boundary-value problem is derived using the new theory. This incremental formulation is based on a generalised midpoint rule, evolution equations for plastic variables are defined in terms of a dissipation function, and an assumption of isochoric plastic deformation is imposed explicitly. A spatially discrete form of the incremental problem is obtained by application of the finite element method. An algorithm for solving this discrete problem, based on the Newton-Raphson procedure and having the typical predictor-corrector structure used in computational plasticity, is proposed and investigated. This algorithm is implemented in NOSTRUM, the in-house finite element code of The FRD/UCT Centre for Research in Computational and Applied Mechanics, at the University of Cape Town. A number of standard example problems are analysed using this code and results are compared with those obtained by others.

It is shown that a corrector algorithm based on use of a dissipation function is a viable alternative to the conventional return mapping algorithms. While this alternative approach is not necessarily better than the conventional one for simple models of plasticity, it may prove valuable when considering more complex models for materials which exhibit dissipative behaviour. The manner in which an assumption of isochoric plastic deformation is incorporated into the incremental form of the problem is shown to play an important role.

DECLARATION

I, Robin Andrew Eve, hereby declare that this thesis is essentially my own work and that no part of it has been submitted for a degree at any other university.

R A EVE

April 1992

To my parents and Helen

ACKNOWLEDGEMENTS

I wish to express my gratitude to the following:

My supervisor Professor B/D Reddy, for his guidance, encouragement and great patience throughout this work.

Professor J B Martin, for his help and encouragement.

Professor R T Rockafellar, for the hours of interesting discussion we had during his visits to Cape Town in January 1988 and April 1991, and in particular for his invaluable assistance in establishing the proof to the main theorem in Chapter III.

The various visitors to UCT with whom I have had the opportunity to discuss this work and whose comments have proved useful, in particular Professors M Ortiz, J C Simo, P Wriggers and Dr J Nagtegaal.

The members of CERECAM for their support, in particular Colin Mercer and Mike Eastman.

The Foundation for Research and Development for their financial assistance.

My parents and friends for their encouragement and continuous support over the past few years.

TABLE OF CONTENTS

Abstract	i
Declaration	ii
Dedication	iii
Acknowledgements	iv
Table of contents	iv

Chapter I INTRODUCTION

1.1 Overview	1
1.2 Background	4
1.3 Objectives	6
1.4 Literature Survey	6
1.5 The plan of this thesis	12

Chapter II TENSORS AND NOTATION

2.1 Introduction	13
2.2 Scalars and vectors	13
2.3 Tensors	16
2.4 Tensor functions	21
2.5 Tensor fields	22

Chapter III RESULTS FROM CONVEX ANALYSIS

3.1 Introduction	27
3.2 Sets	28
3.3 Functions	30
3.4 Special results	36

Chapter IV KINEMATICS OF FINITE-STRAIN PLASTICITY

4.1 Introduction	43
4.2 Deforming bodies	44
4.3 Elastic-plastic deformations	49

4.4	Decomposition into dilational and volume-preserving parts	53
-----	---	----

Chapter V CONTINUUM THERMODYNAMICS

5.1	Introduction	55
5.2	Stress measures	56
5.3	Laws of thermodynamics	57
5.4	Internal variables	58
5.5	Constitutive relations	59
5.6	Remarks	61

Chapter VI AN INTERNAL VARIABLE THEORY OF PLASTICITY

6.1	Introduction	63
6.2	The dissipation inequality	64
6.3	The principle of maximum plastic work	66
6.4	Evolution laws	67
6.5	The internal variable theory	67
6.6	The relationship between yield and dissipation functions	70
6.7	Constitutive equations in Kuhn–Tucker form	74
6.8	Further constitutive assumptions	76
6.8.1	Plastic spin	76
6.8.2	Isochoric plastic deformation	77

Chapter VII APPLICATION OF THE INTERNAL VARIABLE THEORY

7.1	Introduction	79
7.2	The free energy function	79
7.2.1	The elastic strain energy	80
7.2.2	The plastic potentials	82
7.3	Flow laws based on the von Mises yield condition	83

Chapter VIII THE INCREMENTAL PROBLEM

8.1	Introduction	87
8.2	The initial–boundary–value problem	87

8.3	The incremental boundary value problem	89
8.4	Variational form of the incremental problem	94
8.4.1	The equilibrium equation	94
8.4.2	The evolution and constraint equations	95
8.4.3	Statement of the problem	100

Chapter IX FINITE ELEMENT APPROXIMATION

9.1	Introduction	103
9.2	The finite element method	104
9.2.1	The approximation of displacements	106
9.2.2	Approximation of internal variables and the Lagrange multiplier	107
9.3	The discrete form of the incremental boundary problem	110

Chapter X SOLUTION ALGORITHMS

10.1	Introduction	113
10.2	The Newton–Raphson procedure	114
10.3	The predictor–corrector scheme	115
10.3.1	The predictor phase	116
10.3.2	The corrector phase	118
10.4	Remarks	121
10.5	Summary	123

Chapter XI NUMERICAL EXAMPLES

11.1	Introduction	125
11.2	The NOSTRUM finite element code	126
11.3	Imposition of the plastic volume constraint	127
11.4	Reduced integration	128
11.5	Numerical examples	130
11.5.1	<i>EXAMPLE 1</i> : Expansion of a thick walled cylinder	130
11.5.2	<i>EXAMPLE 2</i> : Upsetting of an axisymmetric billet	137
11.6	Discussion of results	143

Chapter XII SUMMARY AND CONCLUSIONS

CHAPTER I

INTRODUCTION

1.1 Overview

In this thesis an internal variable constitutive theory of finite-strain plasticity is presented. This new theory is used as a basis for the formulation of a variational form of incremental boundary-value problems involving large plastic deformations. An approximation of this problem by finite elements and an algorithm for solving the resulting discrete problem are discussed.

The new theory of plasticity is developed within a framework of non-smooth convex analysis. A unification of ideas concerning the postulates of plasticity is achieved by using the powerful tools provided by results in this branch of mathematics. A general theory of continuum thermodynamics with internal variables is used as the foundation for the theory. Although the new theory is applicable to both small and finite-strain plasticity we restrict attention to its application to the finite-strain case. An important feature of the theory is that it clearly relates conventional formulations which make use of yield functions, with those based on the use of a dissipation function. The role of the principle of maximum plastic work is also made precise.

The special case of purely elastic behaviour should be readily recoverable from any plasticity theory. In many of the proposed finite-strain theories of plasticity this feature has not been included in a satisfactory manner. By careful choice of variables we have included it in our finite-strain plasticity theory in a most natural way; with this choice we can use standard hyperelastic constitutive theory to model the elastic component of the material behaviour.

Initial boundary-value problems involving quasi-static rate-independent plasticity form a class of problems which are of considerable practical importance. The solution of such problems consists of determining the history of response of a body to a given program of loading. For convenience the program of loading, be it prescribed displacements

or forces or a combination thereof, is described with the aid of a time-like parameter. Because elastoplastic behavior is history dependent, the response at a given point on the loading path is dependent on the complete record of plastic deformation which has taken place along the path up to that point. We refer to derivatives taken with respect to the time-like parameter as rates. The classical approach to solving problems in plasticity is to express the constitutive laws in rate form, that is, to relate rates of change of stresses to rates of change in deformation or strain. The problem then becomes one of integrating these rates along the loading path; this we refer to as the *rate problem*. For certain programs of loading, for example, proportional loading, a valid solution can be obtained without integrating in this way, and hence, the response can be considered path-independent; behaviour of this type is called *holonomic*. Deformation theories of plasticity are based on an assumption of path independence, see [70]; these theories have played an important role in the development of more general holonomic methods, see [35,36,98].

In general the problems of continuum mechanics cannot be solved explicitly; various methods of approximation are employed to recast these problems in a form for which a solution can be found. For quasi-static problems such as those of rate-independent plasticity a semi-discretisation in time is applied and finite difference methods are used to approximate rate quantities. In this way an incremental form of the problem is derived. The finite element method is usually used for spatial discretisation of this incremental problem.

The incremental formulation we develop is based on a generalised midpoint rule in which rate quantities are approximated by an Euler backward difference. The evolution equations for plastic variables are defined in terms of a dissipation function and an assumption of isochoric plastic deformation is imposed explicitly. Application of the finite element method follows a standard procedure which reduces the problem to one of finding increments in displacement values at node points and internal variable values at element integration points, at each of a series of time steps. The solution procedure proposed and investigated is based on the Newton-Raphson procedure and has the typical predictor-corrector structure of algorithms used in computational plasticity.

In early implementations of this approach it was common practice to use an *elastic predictor*, however, with this approximation convergence of the procedure is slow. It has subsequently been shown that with the use of a so called *consistent predictor*, convergence approaches the quadratic rate of the Newton method. We derive a consistent predictor for our formulation. We have also investigated use of a quasi-Newton method, namely the BFGS method [32,73]; the objective of this method is to improve the computational efficiency of the solution process. It does not, however, enjoy the good convergence characteristics exhibited by consistent methods.

The algorithm developed for the corrector phase differs considerably from conventional return mapping algorithms. The novelty of this algorithm arises from the use of the dissipation function in formulating the evolution equations and the explicit imposition of the plastic volume constraint. A predictor-corrector approach used in conjunction with the type of discretisation indicated above results in an uncoupling of the evolution equations for internal variables associated with different integration points. Corrector algorithms are therefore required to solve a number of small systems of equations, each system associated with a single point. A Lagrange multiplier method is used to impose the volume constraint; by perturbing the resulting saddle-point problem the multiplier may be condensed out of the system of equations associated with each point.

The proposed solution algorithm is implemented in NOSTRUM, the in-house finite element code of the The Centre for Research in Computational and Applied Mechanics, at the University of Cape Town. The capabilities of this code were previously restricted to the solution of small-strain nonlinear problems. A number of example problems are analysed to investigate the effectiveness of the algorithm. These include problems involving both perfect plasticity and nonlinear isotropic hardening. Program development was carried out on a Micro Vax 2000 and example problems were run on either a Vax 3100 workstation or a Vax 3100 multi-user machine; the limited speed of these facilities restricts the size and complexity of examples which can be studied in detail.

1.2 Background

The present investigation essentially continues that reported in [35] in 1986 (see also [36] and [98]). In this work a holonomic theory based on the concept of extremal work paths proposed by Martin and co-workers [11,36,93,98] was further developed, and the incremental holonomic problem and the classical rate problem in small-strain elastoplasticity were formulated as variational inequalities. The initial aim of the work reported in this thesis was to extend some of the ideas used in [35] and to apply them to the formulation of problems involving finite-strain plasticity. A survey of the literature pertaining to finite-strain plasticity revealed some inconsistency, controversy and a lack of clarity in issues such as the decomposition of deformation into elastic and plastic parts, the choice of stress-rate measures and the extension to finite-strains of classical stability postulates. The stability postulate proposed by Drucker [25] and its dual form, proposed by Ilyushin [48], play a central role in the arguments employed in developing the small-strain incremental holonomic formulation used in [35] and in [11] and [98].

With regard to the stability postulates, in order to come to a greater understanding of this problem the application of these postulates in the small-strain problem was studied. It was found that while both Drucker and Ilyushin's postulates provide sufficient conditions for a stable classical theory of plasticity complete with normality laws for convex yield surfaces, neither is a necessary condition. Both approaches result in the inclusion of some excess information, essentially the difference between necessity and sufficiency of the conditions postulated. These extra components lead to problems when the postulates are applied to the case of finite plastic deformations. The principle of maximum plastic work is contained in both postulates; in the case of Ilyushin's postulate it appears in dual form as the principle of maximum plastic dissipation. These principles carry over to the finite-strain case without any ambiguity, provided that appropriate pairs of stress and strain-rate or strain and stress-rate measures are used. The importance of the role of the principle of maximum plastic work, or maximum plastic dissipation, in the development of theories of plasticity has been noted by a number of people, in particular Hill [41,42,43,44] (see also [14,63,70,78,106]). In Hill's work the principle of maximum plastic work is clearly identified as an essential postulate.

A number of other important topics such as the role of potential functions, the use of elements of continuum thermodynamics with internal variables, and the use of convex analysis were identified as important aspects of the development of theories of plasticity; they are reviewed in the literature survey section of this Chapter and we discuss them further in later Chapters.

A study of these topics together with the ideas mentioned above led to the development of a new internal variable theory of small-strain plasticity. This theory was developed using the tools of convex analysis, which made it possible to unify previously loosely connected ideas, resulting in some new features of an otherwise classical theory. The maximum plastic work inequality is the basis of the theory; an important result is the establishment of the equivalence of this postulate to the flow laws expressed in terms of yield or dissipation functions. This work was published in 1990 (see Eve, Reddy and Rockafellar [29]) and is also included in this thesis, although here we develop the theory in the context of finite-strain plasticity.

Having developed an improved small-strain theory, attention returned to consideration of the theory of finite-strain plasticity. The importance of nonlinear hardening laws in the constitutive theory of finite-strain plasticity was recognised and so before attempting an extension of the theory to take into account finite-strains the inclusion of such laws in the new small-strain theory was investigated. This investigation was carried out as joint work with Mr Tekin Gültop who also contributed to the many discussions which lead to the extension of the theory to include finite-strain plasticity. The investigation revealed several interesting aspects of the theory and novel ways of viewing the structure of hardening laws. A full investigation of nonlinear kinematic hardening has yet to be completed, however. An outline of the new internal variable theory of finite-strain plasticity developed was also published in 1990 (see Eve, Gültop and Reddy [28]); it is reviewed in this thesis.

Having established the basis for a new approach to the formulation of problems in finite-strain plasticity we were able to refine and focus the objectives of the thesis.

1.3 Objectives

The primary objective of this thesis is to present the development of a new internal variable theory of finite-strain plasticity and to discuss its application to solving problems involving finite plastic deformations using an incremental approach.

To achieve this objective both theoretical and numerical aspects of problems involving finite-strain plasticity have been studied. The inclusion of the plastic volume constraint in the incremental form of the problem is an aspect to which we have paid some particular attention.

The main objective of the numerical component of this work is to investigate the application of the formulation we have developed in practical situations. A specific goal is to compute solutions to a few typical practical problems.

1.4 Literature Survey

We include here a brief overview of the development of finite-strain plasticity as a branch of computational and applied mechanics. It is not always easy and indeed not always desirable to discuss separately those developments concerned specifically with finite-strain plasticity from those of general plasticity theory or even those specifically limited in scope to the small-strain case.

Interest in plasticity as an area of mechanics has been apparent since the turn of the century and significant progress has been made since the mid 1900's. In most early work attention was restricted to small-strain plasticity. An interest in finite-strain plasticity emerged later, paralleling the emergence of improved engineering materials and the consequent interest in finite-elasticity (see early papers of Green and Naghdi [33], Hill [43,44] and Lee and co-workers [56,58,60]). Since these early contributions there has been a growing research effort in the development of both small and finite-strain theories of plasticity. The revolutionary advances in computer hardware which have taken place during this period have had a marked influence on this work.

Rather than attempting to provide a detailed historical account of developments de-

scribed in the now voluminous body of literature pertaining to the solution of problems involving plastic deformations, we outline some specific issues. We concentrate on issues which have arisen from the consideration of finite plastic deformations and pay particular attention to those which have relevance to the developments described in this thesis. In the literature much of the discussion concerning computational issues is coupled to particular theoretical formulations. Several basic approaches have been used; we single out only a few specific developments.

Over the past thirty years the theory of thermodynamics has played an increasingly important role in the theory of continuum mechanics. It has proved to be particularly powerful in the development of constitutive relations for dissipative processes of which plastic deformation is a prime example [33,125]. The development of finite-strain plasticity theory has been greatly influenced by the early paper (1965) of Green and Naghdi [33] in which the proposed theory is firmly based on the principles of thermodynamics. The modern theory of continuum thermodynamics was being developed at the same time as this early paper appeared (see the work of Coleman and Gurtin [18]). Among the significant features of the modernised theory is the idea of including internal variables to model dissipative processes. The use of such internal variables in the formulation of problems in plasticity dates back to the work of Kestin and Rice [49]. There has been some resistance to this approach, due mainly to arguments over the precise physical meaning of some of the specific variables and associated conjugate force quantities which have been used. In recent years an internal variable approach to modelling plasticity has become more popular (see for examples [68,10,22,68,71,99]).

Various authors, notably Halphen and Nguyen [38], Kim and Oden [51] and [52], Dashner [23] and Hill [40] (see also [10,66]), have made use of potential functions and their complements or dual functions in developing evolution equations for plasticity. Such potential functions may or may not be associated with an underlying thermodynamic framework. The dissipation function is one such potential function; while it has been identified by a number of authors it has not been used in any significant way by anyone with the notable exception of Martin and co-workers [68,72].

The notion of convexity enters into the classical theory of plasticity in several ways, a

simple example of this being the usual assumption of convex yield surfaces; the validity of this assumption for the case of finite plastic deformations is discussed by Naghdi and Trapp [78]. The normality rule for plastic flow is intimately connected with this assumption. Many of the potential functions used in plasticity are nondifferentiable in nature, and hence several authors have made use of the notion of subdifferentials when dealing with this situation. The subdifferential is just one of a number of concepts which has been developed in the general area of convex analysis. This and other tools of convex analysis have however not been fully exploited and their usefulness is underemphasised in the bulk of plasticity literature. Nevertheless, some authors have made significant use of elements of convex analysis in the derivation of various theories of plasticity (see for examples [1,30,38,51,66] and in particular Moreau [75] and [76]). Moreau established some duality relationships between yield and dissipation functions similar to those found in the theory we have developed.

Most theories of plasticity have been formulated with reference to stress-space. In contrast, strain-space formulations have attracted little attention; exceptions being the work of Naghdi and Trapp [79] and Dafalias [20]. The duality of these approaches has been commented on by several authors but none have made much practical use of these ideas. There are many interesting connections between the ideas put forward by these authors and the relationships established in our theory.

An important aspect of any finite-strain plasticity theory is the choice of an appropriate kinematic description of the elastic and plastic components of deformation. In the canonical contribution of Green and Naghdi [33] no specific description of the kinematics of finite elastoplastic deformation was made; it was simply assumed that there exists a unique additive decomposition of total strain into elastic and plastic parts. Subsequently the description of the kinematics of finite plastic deformation proved to be a topic which attracted much attention.

The multiplicative decomposition of deformation gradient into elastic and plastic parts proposed by Lee [56,57] was for many years the centre of much controversy and debate [13,12,34,56,57,62,61,83,84]. This multiplicative decomposition has, however, emerged as the most widely accepted basis for the description of finite elastoplastic deformation.

The decomposition has its roots in the early work of Backman [4] where the intermediate plastic coordinates were introduced and a simple chain rule was used to obtain the decomposition. Lee essentially formalised and generalised this basic concept, adding the notion of a stress-free intermediate configuration. Much of the ensuing debate was concerned with the physical interpretation of this intermediate state. Later additive decompositions of strain and deformation rate were derived using the multiplicative decomposition as a basis for the description of finite elastoplastic deformation [53,114]. The development of these additive decompositions did much for the popularity of the multiplicative decomposition. The use of logarithmic strains has recently attracted attention (see for example [92,27]). We describe the multiplicative decomposition in more detail in Chapter IV since we use it as a basis for the finite-strain theory of plasticity developed in this thesis.

It is worth noting that the multiplicative decomposition of deformation gradient is not the only possibility to be considered. An example of an alternative description is that proposed by Kim and Oden in [51,52]. This alternative approach makes use of a mathematically rigorous procedure to construct an additive decomposition of the deformation gradient into elastic and plastic parts. This decomposition has not attracted much attention, however, it is shown in [51] that under certain conditions it reduces to the more restrictive multiplicative decomposition described above.

As mentioned earlier, construction of rate forms of the equations of elastoplastic materials behavior has been the most popular approach to the formulation of problems in elastoplasticity. This approach often admits the development of a general theory of viscoplastic behaviour from which a model for rate-independent plasticity is obtained as special limiting case. It has been shown by a number of authors, see for example [19,47,80,114,106,121], that care must be taken in evaluating the stress rates which play a central role in such formulations. The *Jaumann rate* (see [67] for a definition) has proved effective in many formulations and is widely used. Algorithms developed using this approach are required to be incrementally objective (see Hughes and Winget [47] or Ortiz and Simo [87] for discussion of the implications of this requirement).

The problem of choosing appropriate rate quantities prompted some interest in the use

of convected coordinates in the description of elastoplastic deformation (see in particular the work of Ortiz and Simo [87,114,106] also [91], and [67] for background and theory). Using this approach the equivalence of a number of possible rate quantities has been shown; this equivalence is established by showing that these belong to a set of rate quantities known as Lie derivatives. Interest in this approach appears to have waned, presumably because it is considered by many to require an unwarranted level of mathematical complexity, the basis for which is not readily accessible to many of those who work in the area of elastoplasticity.

A natural question to arise from the consideration of alternative stress rates concerns the role of plastic spin in the description of elastoplastic behaviour. This question is one which is of relevance in any approach to the formulation of theories of plasticity. It has received considerable attention in recent years (see for examples [21,64,63,74]); however, it remains a largely unresolved issue. It has, however, been proposed that plastic spin only plays a significant role in problems involving considerable kinematic hardening, or in problems involving high strain rates, whereas in simpler rate-independent situations it has been suggested that at least initially plastic spin can be ignored.

Much of the theory of plasticity has been developed specifically for modelling the plastic deformation of metals. For metals it is commonly assumed that there is no change in volume associated with plastic flow, that is, plastic deformation is considered to be isochoric. This feature is often referred to as the *plastic volume constraint*. As a consequence of this assumption it is the deviatoric part of the stress that is used in constructing yield conditions and evolution equations for plastic variables. In the case of small-strain plasticity this constraint is a linear one and is easily taken into account. In the case of finite-strain plasticity treatment of this constraint is less straight-forward, in particular in incremental forms of such problems. The important effect that this assumption has on the overall solution procedure was first recognised and addressed in the classical paper of Nagtegaal, Parks and Rice [81]. The mean dilation approach advocated by Nagtegaal *et al* is related to methods proposed by Belytschko and co-workers [5,31] for the control of hourglass modes which occur in underintegrated finite elements. A comprehensive discussion of methods used to incorporate the plastic volume constraint in the formulation of the problem is given by Simo, Taylor and Pister [113].

One method involves the use of a Lagrange multiplier.

The modelling of elastic behaviour forms an integral part of any model of elastoplasticity. In many formulations of finite-strain plasticity hypoelastic models are used. In the case of metal plasticity elastic strain is assumed to be small and linear elastic stress-strain relations are often used. More recently hyperelastic formulations have been used, [87,113,114].

In the opening of this Chapter we briefly described how problems in plasticity are recast in incremental form. The bulk of developments have been aimed at developing implicit methods for solving these problems. The advantage of this approach lies in the conditional stability of implicit methods, and hence reasonably large increments can be used. Very small increments are required if explicit methods are used; the advantage of using implicit methods is that much larger increments can be used. Recent advances in computer hardware, in particular the parallelisation of processing, has nevertheless resulted in renewed interest in the use of explicit methods. Explicit methods are better suited to the development of algorithms which take advantage of these new tools.

For some time attention was restricted to the use of Euler backward schemes. More recently a generalised midpoint rule has attracted some interest [99,110], and trapezoidal rules have also recently been investigated [100,101]. We use a generalised midpoint rule for the construction of the incremental form of the problem.

Use of these implicit methods results in the need to solve large systems of nonlinear equations. Newton methods are widely used; the basic structure of these iterative methods is a predictor-corrector procedure. In its application to problems of plasticity this structure is exploited in such a way as to overcome the problem of determining which plastic evolution equations must be taken into account, that is, where yielding is taking place. The first step in each iteration, that is the predictor phase, involves constructing a tangent operator. Various methods of calculating this tangent have been proposed; the use of a so-called consistent tangent has been shown to be superior to other approximations, in particular to the use of an elastic tangent. The development of consistent predictors has been the focus of considerable research effort (see [9,87,106,105]). In conventional algorithms the corrector phase of each iteration is often referred to as a return

or return mapping algorithm. The development and analysis of these algorithms has also received considerable attention, see for example [88,116,110,111] They are popular partly because they lend themselves to geometric interpretation in stress-space and because they are known to be stable. We describe the Newton method and its application to the solution of problems in plasticity in detail in Chapter X.

1.5 The plan of this thesis

In this Chapter we have given an overview of the scope of the work which is to be discussed. The background to the thesis has been described and the main objectives stated. We have also included a brief review of relevant literature; further discussion of various contributions is included in other Chapters.

Tensor methods are of fundamental importance in the development of the theory which describes the deformation of continuous bodies; we therefore give a brief account of the tensor algebra and analysis used. This account is given in Chapter II. In Chapter III we collect together results of convex analysis which are used later in the development of the theory of plasticity. A Chapter describing kinematics of finite-strain plasticity follows. Then in Chapter V we give a brief account of the theory of continuum thermodynamics with internal variables. In Chapter VI the new internal variable of plasticity is developed using the results reviewed in Chapter III. The application of the theory is described in Chapter VII. In Chapter VIII a simple initial-boundary-value problem which involves finite plastic deformations is posed; an incremental form of this problem is constructed and a variational formulation of this incremental problem is developed. Finite element approximations are described in Chapter IX. A procedure for solving the discrete problem obtained in Chapter IX is proposed in Chapter X. The implementation and various aspects of the performance of the proposed algorithm are described in Chapter XI. Also presented in Chapter XI are the solutions to a selection of standard example problems obtained with the code developed as part of this thesis. We conclude the thesis in Chapter XII with a brief summary of the work, indicating what we believe to be the main contributions made in respect of both theoretical and numerical aspects of the developments described; directions for further study are also indicated.

CHAPTER II

TENSORS AND NOTATION

2.1 Introduction

In this chapter we give a summary of the tensor algebra and analysis which play a crucial role in the developments discussed in this thesis. We also establish the notation used in later chapters. In the first section we discuss scalars and vectors; we then define tensors at a point. We concentrate on second order cartesian tensors, but include mention of higher order tensors; both coordinate free and component forms of the various tensor quantities are given. We include a brief section on tensor functions in which we give a number of relationships involving derivatives of tensor functions since these play an important role in later developments. Tensor fields are then discussed, particular attention being paid to the calculus of tensor fields. Further discussion of tensor algebra and analysis may be found in, for example, Chadwick [15], Lai Rubin and Krempel [55] or Ogden [86].

2.2 Scalars and vectors

We denote by \mathbf{R} the set of real numbers. The finite-dimensional space \mathbf{R}^n is equivalent to the set of all ordered n -tuples of real numbers, $n < \infty$. Vectors may be identified with elements of \mathbf{R}^n and scalars with elements of \mathbf{R} . In general we shall use lower case bold letters to denote vectors. A vector space V is a set with the operations of vector addition and of multiplication by a scalar defined on it, and which has the properties:

- a) $\mathbf{u} + \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ and $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$, $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$;
- b) V contains an element $\mathbf{0}$, the zero element, such that $\mathbf{u} + \mathbf{0} = \mathbf{u} \quad \forall \mathbf{u} \in V$ and every element $\mathbf{u} \in V$ has an inverse denoted $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$;
- c) $\alpha \mathbf{u} \in V$, $1\mathbf{u} = \mathbf{u}$, $\alpha(\beta \mathbf{u}) = (\alpha\beta)\mathbf{u}$, $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$, $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$,
 $\forall \alpha, \beta \in \mathbf{R}$, $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

A *Euclidean vector space* denoted \mathbf{E} is a real vector space equipped with both a *scalar product* and a *vector product*. Problems in mechanics are cast in a three-dimensional physical space; therefore unless an indication to the contrary is given \mathbf{E} is assumed to be a three-dimensional vector space, the elements of which may be identified with elements of \mathbf{R}^3 . The scalar product is a bilinear mapping from $\mathbf{E} \times \mathbf{E}$ to \mathbf{R} where $\mathbf{E} \times \mathbf{E}$ is the Cartesian product defined as the set of ordered pairs (\mathbf{u}, \mathbf{v}) of vectors $\mathbf{u}, \mathbf{v} \in \mathbf{E}$; the scalar product defines an inner product on \mathbf{E} . The vector product is a mapping from $\mathbf{E} \times \mathbf{E}$ to \mathbf{E} . Hence, for any pair of vectors $\mathbf{u}, \mathbf{v} \in \mathbf{E}$ there is a scalar denoted by $(\mathbf{u} \cdot \mathbf{v})$ and a vector denoted by $\mathbf{u} \wedge \mathbf{v}$ with the properties:

$$\text{a) } (\mathbf{u} \cdot \mathbf{v}) = (\mathbf{v} \cdot \mathbf{u}) \quad \forall \mathbf{u}, \mathbf{v} \in V,$$

$$\text{b) } (\mathbf{u} \cdot \mathbf{u}) \geq 0 \quad \forall \mathbf{u} \in V, \text{ equality iff } \mathbf{u} = \mathbf{0},$$

$$\text{c) } (\alpha \mathbf{u} + \beta \mathbf{v}) \cdot \mathbf{w} = \alpha(\mathbf{u} \cdot \mathbf{w}) + \beta(\mathbf{v} \cdot \mathbf{w}) \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{E}, \alpha, \beta \in \mathbf{R},$$

$$\text{d) } \mathbf{u} \wedge \mathbf{v} = -\mathbf{v} \wedge \mathbf{u} \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{E}, \text{ so that } \mathbf{u} \wedge \mathbf{u} = \mathbf{0}. \quad \forall \mathbf{u} \in \mathbf{E},$$

$$\text{e) } (\alpha \mathbf{u} + \beta \mathbf{v}) \wedge \mathbf{w} = \alpha(\mathbf{u} \wedge \mathbf{w}) + \beta(\mathbf{v} \wedge \mathbf{w}) \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{E}, \alpha, \beta \in \mathbf{R},$$

from which it follows that

$$\text{f) } \mathbf{u} \cdot (\mathbf{u} \wedge \mathbf{v}) = 0 \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{E},$$

$$\text{g) } (\mathbf{u} \wedge \mathbf{v}) \cdot (\mathbf{u} \wedge \mathbf{v}) = (\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}) - (\mathbf{u} \cdot \mathbf{v})^2 \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{E}.$$

The *magnitude* or modulus of a vector $\mathbf{u} \in \mathbf{E}$, denoted $|\mathbf{u}|$, is defined by

$$|\mathbf{u}| = \sqrt{(\mathbf{u} \cdot \mathbf{u})}. \tag{2.1}$$

A vector \mathbf{u} is called a *unit vector* if $|\mathbf{u}| = 1$. A pair of vectors \mathbf{u} and \mathbf{v} are *orthogonal* if $(\mathbf{u} \cdot \mathbf{v}) = 0$. The magnitude of a vector is a norm on \mathbf{E} generated by the scalar (inner) product. In addition to the above the scalar and vector products have the properties

$$\text{a) } (\mathbf{u} \cdot \mathbf{v}) = |\mathbf{u}| |\mathbf{v}| \cos \theta,$$

$$\text{b) } \mathbf{u} \wedge \mathbf{v} = |\mathbf{u}| |\mathbf{v}| \sin \theta \mathbf{k}$$

where θ is the angle between the directions of vectors \mathbf{u} and \mathbf{v} ($0 \leq \theta \leq \pi$) and \mathbf{k} is a unit vector normal to the plane containing \mathbf{u} and \mathbf{v} . The quantity $|\mathbf{u}| |\mathbf{v}| \sin \theta$ is the *area* of a parallelogram with adjacent sides \mathbf{u} and \mathbf{v} .

The *scalar triple product*, also known as a *box product* or *triple product*, and denoted by $[\mathbf{u}, \mathbf{v}, \mathbf{w}]$ is defined by

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w}) . \quad (2.2)$$

Its value is the *volume* of a parallelepiped defined by the triad $\mathbf{u}, \mathbf{v}, \mathbf{w}$. The properties of the box product are easily derived from those of the scalar and vector products.

A fundamental property of finite-dimensional spaces equipped with an inner product is the existence of an *orthonormal basis*. In a Euclidean vector space there exists a set of vectors corresponding to the orthonormal basis of the underlying finite-dimensional space. Hence, for a three-dimensional Euclidean vector space \mathbf{E} there exists such a set $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ denoted collectively $\{\mathbf{e}_i\}$; the *range* of the subscript index is 3, that is, i takes values 1, 2 or 3. This basis has the property

$$(\mathbf{e}_i \cdot \mathbf{e}_j) = \delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j \end{cases} . \quad (2.3)$$

The quantity δ_{ij} defined by the second equation is known as the *Kronecker delta*. In this thesis we shall only make use of proper orthogonal bases, that is, bases which satisfy $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] = 1$.

Corresponding to any vector $\mathbf{u} \in \mathbf{E}$ an ordered triple of scalars (u_1, u_2, u_3) exists such that

$$\mathbf{u} = u_1 \mathbf{e}_1 + u_2 \mathbf{e}_2 + u_3 \mathbf{e}_3 = u_i \mathbf{e}_i \quad (2.4)$$

for a given basis $\{\mathbf{e}_i\}$; these scalars are the *components* of the vector relative to this basis.

The *summation convention* has been used in (2.4); the appearance of an index twice implies the sum of all terms obtained by replacing the letter with each of the values in the range of the index. We adopt this convention throughout unless an indication to the contrary is given.

The orthonormal basis used above is not unique. A second basis $\{\mathbf{e}'_i\}$ can be obtained from a rotation of vectors in the basis $\{\mathbf{e}_i\}$. This is expressed by

$$\mathbf{e}'_i = Q_{ij}\mathbf{e}_j, \quad \text{where } Q_{ij} = (\mathbf{e}'_i \cdot \mathbf{e}_j). \quad (2.5)$$

The coefficients Q_{ij} are the direction cosines of the vectors \mathbf{e}_i relative to the vectors \mathbf{e}_j ; they may be collected into a 3×3 *proper orthogonal matrix* \mathbf{Q} .

2.3 Tensors

We concentrate on second-order Cartesian tensors in this section since we consider only orthonormal bases and make little use of tensors of higher order. In general, we shall use bold-face uppercase roman letters to indicate tensors; however, in some situations we deviate from this convention. We indicate the nature of all quantities in such cases.

A second-order tensor \mathbf{A} is a linear mapping of \mathbf{E} into itself, $\mathbf{A} : \mathbf{E} \rightarrow \mathbf{E}$. That is, there is a vector denoted $\mathbf{A}\mathbf{a} \in \mathbf{E}$ for any arbitrary vector $\mathbf{a} \in \mathbf{E}$. This we write as

$$\mathbf{b}(\mathbf{a}) = \mathbf{A}\mathbf{a}. \quad (2.6)$$

We denote by $\mathbf{L}(\mathbf{E}, \mathbf{E})$ the set of all linear mappings from \mathbf{E} to \mathbf{E} ; $\mathbf{L}(\mathbf{E}, \mathbf{E})$ is a vector space with element $\alpha\mathbf{A} + \beta\mathbf{B}$ defined by

$$(\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{a} = \alpha\mathbf{A}\mathbf{a} + \beta\mathbf{B}\mathbf{a} \quad \forall \mathbf{A}, \mathbf{B} \in \mathbf{L}(\mathbf{E}, \mathbf{E}), \alpha, \beta \in \mathbf{R}, \mathbf{a} \in \mathbf{E}. \quad (2.7)$$

The *zero tensor* \mathbf{O} maps every vector in \mathbf{E} to the zero vector $\mathbf{0}$, and the *identity tensor* \mathbf{I} maps every vector in \mathbf{E} to itself:

$$\mathbf{0}\mathbf{a} = \mathbf{0} \quad \text{and} \quad \mathbf{I}\mathbf{a} = \mathbf{a} \quad \forall \mathbf{a} \in \mathbf{E} . \quad (2.8)$$

The *tensor product* or *dyadic product* of two vectors \mathbf{u} and \mathbf{v} is a tensor denoted $\mathbf{u} \otimes \mathbf{v}$ and is defined through its action on an arbitrary vector \mathbf{a} by

$$(\mathbf{u} \otimes \mathbf{v})\mathbf{a} = (\mathbf{a} \cdot \mathbf{v})\mathbf{u} \quad \forall \mathbf{a} \in \mathbf{E} . \quad (2.9)$$

It has the properties

$$(\alpha\mathbf{u} + \beta\mathbf{v}) \otimes \mathbf{w} = \alpha(\mathbf{u} \otimes \mathbf{w}) + \beta(\mathbf{v} \otimes \mathbf{w}) \quad \text{and}$$

$$\mathbf{u} \otimes (\alpha\mathbf{v} + \beta\mathbf{w}) = \alpha(\mathbf{u} \otimes \mathbf{v}) + \beta(\mathbf{u} \otimes \mathbf{w}) \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{E}, \alpha, \beta \in \mathbf{R} .$$

Tensors may be identified with members of the set of all bilinear functions over $\mathbf{E} \times \mathbf{E}$ which forms a vector space over \mathbf{R} ; this set is denoted $\mathbf{L}(\mathbf{E} \times \mathbf{E}, \mathbf{R})$ or $\mathbf{L}(\mathbf{E}^2, \mathbf{R})$. Tensors of order n are members of $\mathbf{L}(\mathbf{E}^n, \mathbf{R})$, hence vectors may be regarded as tensors of order one while scalars are of order zero. The tensor product between two vectors results in a second-order tensor, similarly the tensor product of two second-order tensors results in a fourth-order tensor. Note that not all second-order tensors can be generated by the tensor product of two vectors; and similarly, for tensors of higher order.

We can define *components* A_{ij} of a tensor \mathbf{A} relative to an orthonormal basis $\{\mathbf{e}_i\}$ by

$$A_{ij} = \mathbf{e}_i \cdot \mathbf{A}\mathbf{e}_j . \quad (2.10)$$

The representation

$$\mathbf{A} = A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j \quad (2.11)$$

then follows from this definition and (2.9).

The *composition* of two tensors \mathbf{A} and \mathbf{B} , written \mathbf{AB} , is defined by

$$(\mathbf{AB})\mathbf{a} = \mathbf{A}(\mathbf{B}\mathbf{a}) \quad \forall \mathbf{A}, \mathbf{B} \in \mathbf{L}(\mathbf{E}, \mathbf{E}), \mathbf{a} \in \mathbf{E} ; \quad (2.12)$$

in component form \mathbf{AB} is written

$$\mathbf{AB} = A_{ij}B_{jk} \mathbf{e}_i \otimes \mathbf{e}_k. \quad (2.13)$$

The fourth-order tensor generated by the tensor product of \mathbf{A} and \mathbf{B} can be written in component form as

$$\mathbf{A} \otimes \mathbf{B} = A_{ij}B_{kl} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l. \quad (2.14)$$

The scalar product of two tensors \mathbf{A} and \mathbf{B} is defined by

$$\mathbf{A} \cdot \mathbf{B} = A_{ij}B_{ij} \quad (2.15)$$

and the *trace* of a tensor \mathbf{A} , denoted $\text{tr}\mathbf{A}$, is defined by

$$\text{tr}\mathbf{A} = (\mathbf{A} \cdot \mathbf{I}) = A_{ii}. \quad (2.16)$$

The composition \mathbf{AA} is written \mathbf{A}^2 , similarly $\mathbf{A}^2\mathbf{A}$ is denoted by \mathbf{A}^3 . The scalars defined by $\text{tr}\mathbf{A}$, $\text{tr}\mathbf{A}^2$ and $\text{tr}\mathbf{A}^3$ are invariants of the tensor \mathbf{A} known as the *moments* of \mathbf{A} .

Associated with every tensor \mathbf{A} there is another set of three scalars, I_A , II_A and III_A called the *principal invariants* of \mathbf{A} . These values are defined by

$$[\mathbf{Aa}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \mathbf{Ab}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \mathbf{Ac}] = I_A[\mathbf{a}, \mathbf{b}, \mathbf{c}], \quad (2.17)$$

$$[\mathbf{a}, \mathbf{Ab}, \mathbf{Ac}] + [\mathbf{Aa}, \mathbf{b}, \mathbf{Ac}] + [\mathbf{Aa}, \mathbf{Ab}, \mathbf{c}] = II_A[\mathbf{a}, \mathbf{b}, \mathbf{c}], \quad (2.18)$$

$$[\mathbf{Aa}, \mathbf{Ab}, \mathbf{Ac}] = III_A[\mathbf{a}, \mathbf{b}, \mathbf{c}] \quad (2.19)$$

for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{E}$. The first invariant I_A is the trace of \mathbf{A} . The trace is a linear function mapping \mathbf{L} to \mathbf{R} , the other invariants are not linear. The third invariant III_A is known as the determinant of \mathbf{A} , written $\det \mathbf{A}$; it has the properties

$$\det \alpha \mathbf{A} = \alpha^3 \det \mathbf{A}, \quad (2.20)$$

$$\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}, \quad (2.21)$$

$$\det \mathbf{I} = 1. \quad (2.22)$$

These invariants are expressed in terms of the moments of \mathbf{A} by

$$I_A = \text{tr } \mathbf{A} , \quad (2.23)$$

$$II_A = \frac{1}{2}((\text{tr } \mathbf{A})^2 - (\text{tr } \mathbf{A}^2)) , \quad (2.24)$$

$$III_A = \frac{1}{6}\{(\text{tr } \mathbf{A})^3 - 3(\text{tr } \mathbf{A})(\text{tr } \mathbf{A}^2) + 2(\text{tr } \mathbf{A}^3)\} . \quad (2.25)$$

If $\det \mathbf{A} \neq 0$ then \mathbf{A} is *invertible*, that is, there exists a unique tensor denoted \mathbf{A}^{-1} called the *inverse* of \mathbf{A} such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} . \quad (2.26)$$

We note the following properties of inverses:

$$\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1} , \quad (2.27)$$

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} . \quad (2.28)$$

Associated with every tensor \mathbf{A} there is a unique tensor \mathbf{A}^T called the *transpose* of \mathbf{A} defined by

$$\mathbf{a} \cdot \mathbf{A}^T \mathbf{b} = \mathbf{b} \cdot \mathbf{A} \mathbf{a} \quad \forall \mathbf{a}, \mathbf{b} \in \mathbf{E} , \quad (2.29)$$

with the properties

$$\text{a) } (\mathbf{A}^T)^T = \mathbf{A} ,$$

$$\text{b) } (\alpha \mathbf{A} + \beta \mathbf{B})^T = \alpha \mathbf{A}^T + \beta \mathbf{B}^T \quad \forall \mathbf{A}, \mathbf{B} \in \mathbf{L}, \alpha, \beta \in \mathbf{R} ,$$

$$\text{c) } (\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T \quad \forall \mathbf{A}, \mathbf{B} \in \mathbf{L} .$$

In component form, if $\mathbf{A} = A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$ then $\mathbf{A}^T = A_{ij} \mathbf{e}_j \otimes \mathbf{e}_i$.

Also associated with every tensor \mathbf{A} in \mathbf{L} there is a unique tensor called the *adjugate* of \mathbf{A} and denoted $\text{adj } \mathbf{A}$, defined by

$$\text{adj } \mathbf{A}(\mathbf{a} \wedge \mathbf{b}) = (\mathbf{A}\mathbf{a}) \wedge (\mathbf{A}\mathbf{b}) \quad \forall \mathbf{a}, \mathbf{b} \in \mathbf{E} . \quad (2.30)$$

We note that in a matrix representation of second-order tensors the adjugate tensor corresponds to the cofactor matrix, written $[\text{cof } \mathbf{A}]$. We note the following relationship between the adjugate and the inverse of an invertible tensor \mathbf{A} :

$$\text{adj } \mathbf{A}^T = (\det \mathbf{A}) \mathbf{A}^{-1} . \quad (2.31)$$

An *orthogonal tensor* is one which preserves scalar products: a tensor \mathbf{Q} is orthogonal if

$$(\mathbf{Q}\mathbf{a}) \cdot (\mathbf{Q}\mathbf{b}) = \mathbf{a} \cdot \mathbf{b} \quad \forall \mathbf{a}, \mathbf{b} \in \mathbf{E} . \quad (2.32)$$

A consequence of this is that $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$, and so $\det \mathbf{Q} = \pm 1$, \mathbf{Q} is invertible and $\mathbf{Q}^{-1} = \mathbf{Q}^T$. We restrict attention to proper orthogonal tensors, that is those for which $\det \mathbf{Q} = 1$. The coefficients of the matrix \mathbf{Q} defined by (2.32) are the components of an orthogonal matrix.

A *symmetric tensor* \mathbf{S} is defined as one for which $\mathbf{S} = \mathbf{S}^T$, while a skew-symmetric tensor \mathbf{W} has the property $\mathbf{W}^T = -\mathbf{W}$. Any tensor $\mathbf{A} \in \mathbf{L}$ can be expressed as the sum of symmetric and skew-symmetric parts as

$$\begin{aligned} \mathbf{A} &= \text{sym} \mathbf{A} + \text{skew} \mathbf{A} , \\ &= \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^T) . \end{aligned} \quad (2.33)$$

A *spherical tensor* \mathbf{P} is one which can be written in the form

$$\mathbf{P} = \alpha \mathbf{I} \quad (2.34)$$

where α is a scalar. Any tensor $\mathbf{A} \in \mathbf{L}$ can also be decomposed into *deviatoric* and *spherical* parts. The deviatoric part, denoted $\text{dev } \mathbf{A}$, is defined by

$$\text{dev } \mathbf{A} = \mathbf{A} - \frac{1}{3} \text{tr}(\mathbf{A}) \mathbf{I} \quad (2.35)$$

and the spherical part is given by

$$\frac{1}{3}\text{tr}(\mathbf{A})\mathbf{I} . \quad (2.36)$$

In addition to the additive decompositions described above any arbitrary tensor $\mathbf{A} \in \mathcal{L}$ has a pair of unique multiplicative decompositions. These multiplicative decompositions are called the right and left *polar decompositions*, and are defined respectively by

$$\mathbf{A} = \mathbf{Q}\mathbf{U} \quad \text{and} \quad \mathbf{A} = \mathbf{V}\mathbf{R} \quad (2.37)$$

where \mathbf{U} and \mathbf{V} are positive definite symmetric tensors and \mathbf{Q} and \mathbf{R} are orthogonal tensors.

For any tensor \mathbf{A} having nonzero determinant another unique multiplicative decomposition is defined by

$$\mathbf{A} = \hat{\mathbf{A}}\bar{\mathbf{A}} \quad (2.38)$$

where $\bar{\mathbf{A}} = (\det \mathbf{A})^{-\frac{1}{3}}\mathbf{A}$ and $\hat{\mathbf{A}} = (\det \mathbf{A})^{\frac{1}{3}}\mathbf{I}$. As a consequence of these definitions $\det \bar{\mathbf{A}} = 1$ and $\det \hat{\mathbf{A}} = \det \mathbf{A}$. This type of decomposition is used in the context of the description of deformation; we discuss it further in Chapter IV.

2.4 Tensor functions

Scalar functions of tensor quantities play a major role in later developments, in particular isotropic scalar functions which are used in the constitutive theory. The invariants defined in (2.23), (2.24) and (2.25) are examples of scalar functions of second-order tensors. We shall need to differentiate these functions, so for reference purposes we record the relationships

$$\frac{\partial I_A}{\partial \mathbf{A}} = \frac{\partial(\text{tr} \mathbf{A})}{\partial \mathbf{A}} = \mathbf{I} , \quad (2.39)$$

$$\frac{\partial II_A}{\partial \mathbf{A}} = (\text{tr} \mathbf{A})\mathbf{I} - \mathbf{A} , \quad (2.40)$$

$$\frac{\partial III_A}{\partial \mathbf{A}} = \frac{\partial(\det \mathbf{A})}{\partial \mathbf{A}} = (\det \mathbf{A})\mathbf{A}^{-T} . \quad (2.41)$$

The inverse of a tensor and the composition of a tensor with itself, or, for example, with its transpose, are simple examples of tensor-valued functions of second order tensors. Derivatives of these with respect to the relevant tensors are given for \mathbf{A}^{-1} by

$$\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} = -\mathbf{A}^{-1} \otimes \mathbf{A}^{-1}, \quad (2.42)$$

and for $\mathbf{B} = \mathbf{A}\mathbf{A}^T$ by

$$\frac{\partial B_{ij}}{\partial A_{kl}} = A_{il}\delta_{jk} + \delta_{ik}A_{jl} \quad (2.43)$$

in component form.

2.5 Tensor fields

Many of the tensor quantities (including scalars and vectors) used in continuum mechanics represent physical quantities, the values of which depend on position and time. The tensors and related properties discussed so far have been defined at a point. We consider tensor fields defined over a domain $\mathbf{Q} = \Omega \times [0, T] \subset \mathcal{E} \times \mathbf{R}$, where Ω is the spatial domain, \mathcal{E} is the three dimensional Euclidean point space in which physical phenomena are observed and $[0, T]$ is the time interval over which the quantities are defined.

A *scalar field* is a mapping of the form

$$\phi : \mathbf{Q} \rightarrow \mathbf{R},$$

a *vector field* is a mapping of the form

$$\mathbf{v} : \mathbf{Q} \rightarrow \mathbf{E}$$

and a *tensor field* is a mapping of the form

$$\mathbf{T} : \mathbf{Q} \rightarrow \mathbf{L}(\mathbf{E}, \mathbf{E}).$$

Given an origin point $o \in \mathcal{E}$, corresponding to all points $x \in \mathcal{E}$ there is a vector \mathbf{x} . These *position vectors* form a vector field on Ω , hence we replace $\Omega \in \mathcal{E}$ by $\Omega_o \in \mathbf{E}$ where the subscript indicates the choice of origin. We consider scalar, vector and tensor fields as functions of the form $\phi = \phi_o(\mathbf{x}, t)$, $\mathbf{v} = \mathbf{v}_o(\mathbf{x}, t)$ and $\mathbf{T} = \mathbf{T}_o(\mathbf{x}, t)$ respectively, for all ordered pairs (\mathbf{x}, t) in the domain $\Omega_o \times [0, T]$. For the moment we shall drop the subscript o and assume that the choice of origin is apparent from the context. Later, when considering the deformation of bodies we reintroduce a subscript which has further meaning.

Our main concern is with the calculus of continuous tensor fields defined on open connected sets or domains. For the purpose of defining spatial derivatives we consider fields which do not depend on time.

A scalar field ϕ on Ω is *continuous* if

$$\lim_{\alpha \rightarrow 0} |\phi(\mathbf{x} + \alpha \mathbf{a}) - \phi(\mathbf{x})| = 0 \quad \forall \mathbf{x} \in \Omega, \mathbf{a} \in \mathbf{E}; \alpha \in \mathbf{R} \quad (2.44)$$

and *differentiable* if there exists a vector field \mathbf{w} such that

$$\lim_{\alpha \rightarrow 0} |\mathbf{w}(\mathbf{x}) \cdot \mathbf{a} - \alpha^{-1} \{\phi(\mathbf{x} + \alpha \mathbf{a}) - \phi(\mathbf{x})\}| = 0 \quad \forall \mathbf{x} \in \Omega, \mathbf{a} \in \mathbf{E}; \alpha \in \mathbf{R} \quad (2.45)$$

The vector field \mathbf{w} is unique for differentiable ϕ ; it is called the *gradient* of ϕ and is denoted $\text{grad}\phi$. We write

$$\mathbf{w} = \text{grad}\phi = \frac{\partial \phi}{\partial \mathbf{x}}. \quad (2.46)$$

Vector and tensor fields \mathbf{u} and \mathbf{T} on Ω are continuous and differentiable if (2.44) and (2.45) hold for the scalar fields defined by $(\mathbf{u} \cdot \mathbf{a})$ and $(\mathbf{a} \cdot \mathbf{T}\mathbf{b})$ for all vectors \mathbf{a} and \mathbf{b} .

The *gradient of a vector field* \mathbf{u} is a tensor field, denoted $\text{grad}\mathbf{u}$, which is defined by

$$\{\text{grad}\mathbf{u}(\mathbf{x})\}^T \mathbf{a} = \text{grad}\{\mathbf{u}(\mathbf{x}) \cdot \mathbf{a}\} \quad \forall \mathbf{x} \in \Omega, \mathbf{a} \in \mathbf{E}. \quad (2.47)$$

We shall not require gradients of tensor fields.

The *divergence* of a vector field \mathbf{u} is a scalar field, denoted $\text{div } \mathbf{u}$, and the divergence of a tensor field is a vector field, denoted $\text{div } \mathbf{T}$; these are defined by

$$\text{div } \mathbf{u} = \text{tr}(\text{grad } \mathbf{u}) \quad (2.48)$$

and

$$\{\text{div } \mathbf{T}\} \cdot \mathbf{a} = \text{div } \{\mathbf{T}\mathbf{a}\} . \quad (2.49)$$

In later developments we make use of component forms of field quantities and their derivatives. To express position vectors in component form we require a basis. The choice of basis and origin defines a coordinate system on \mathcal{E} . In most of this thesis we use a rectangular Cartesian coordinate system $\{o, \mathbf{e}_i\}$, where $\{\mathbf{e}_i\}$ is the orthogonal basis introduced earlier. With this choice of coordinates we can write $\mathbf{x} = x_i \mathbf{e}_i$ ($i = 1, 2, 3$) and obtain the representations

$$\text{grad } \phi = \frac{\partial \phi}{\partial x_i} \mathbf{e}_i = \phi_{,i} \mathbf{e}_i , \quad (2.50)$$

$$\text{grad } \mathbf{u} = \frac{\partial u_i}{\partial x_j} \mathbf{e}_i \otimes \mathbf{e}_j = u_{i,j} \mathbf{e}_i \otimes \mathbf{e}_j , \quad (2.51)$$

$$\text{div } \mathbf{u} = \frac{\partial u_i}{\partial x_i} = u_{i,i} \quad (2.52)$$

and

$$\text{div } \mathbf{T} = \frac{\partial T_{ij}}{\partial x_i} \mathbf{e}_j = T_{i,j,i} \mathbf{e}_j \quad (2.53)$$

where a comma subscript indicates differentiation with respect to the coordinates corresponding to the indices which follow.

We refer to derivatives with respect to time as *rates* and denoted them by a superposed dot : scalar, vector and tensor rates are then written as

$$\dot{\phi} = \frac{\partial \phi}{\partial t}, \quad \dot{\mathbf{v}} = \frac{\partial \mathbf{v}}{\partial t}, \quad \text{and} \quad \dot{\mathbf{T}} = \frac{\partial \mathbf{T}}{\partial t}. \quad (2.54)$$

When the choice of basis is independent of time we have

$$\dot{\mathbf{v}} = \dot{v}_i \mathbf{e}_i \quad \text{and} \quad \dot{\mathbf{T}} = \dot{T}_{ij} \mathbf{e}_i \otimes \mathbf{e}_j. \quad (2.55)$$

We shall also need to consider the integration of scalar, vector and tensor fields. The *divergence theorem* is employed to convert integrals over the boundary $\partial\Omega$ of Ω , to integrals over Ω . One form of the theorem states that

$$\int_{\Omega} (\text{div } \mathbf{T}^T \cdot \mathbf{v}) dv + \int_{\Omega} (\mathbf{T}^T \cdot \text{grad } \mathbf{v}) dv = \int_{\partial\Omega} (\mathbf{T} \mathbf{n} \cdot \mathbf{v}) ds \quad \forall \mathbf{T}, \mathbf{v} \quad (2.56)$$

where dv and ds are elements of volume and surface area respectively, and \mathbf{n} is the unit normal vector to the boundary $\partial\Omega$. We assume the boundary to be smooth so that \mathbf{n} is unique at every boundary point.

CHAPTER III

RESULTS FROM CONVEX ANALYSIS

3.1 Introduction

The theory of plasticity which we present in Chapter VI is developed within the framework of convex analysis. In this Chapter we collect together a number of results from convex analysis which are of importance in this theory. No familiarity with the concepts of convex analysis is assumed. We provide an adequate introduction for the reader who is unfamiliar with these ideas, while at the same time providing an easily accessible record for later discussion. Alternative and more detailed presentations of the standard results reviewed here can be found in Rockafellar [103], Aubin [3] or Oden [85], or in other reviews such as those found in Moreau [75,76] and Kim and Oden [51]. We include the proof of a new theorem, presented for the first time in [29], which plays a crucial role in the development of the theory of plasticity.

We are concerned exclusively with finite-dimensional spaces, for example, spaces of vectors or tensors at a point and functions acting on them (these are defined and discussed in Chapter II). We denote such a finite-dimensional space by E ; E is isomorphic to (and may be identified with) \mathbf{R}^n for appropriate n . The dual space of E is denoted by E^* . For the case of second-order tensors with an underlying three dimensional space $E = \mathbf{L}(E, E)$, $E^* = \mathbf{L}(E^2, \mathbf{R})$ and $n = 9$. Later we shall use a combination of tensors and scalars so that n will in general be greater than 9. In this chapter examples where $n = 1$ or 2 are used to illustrate the various concepts since simple two- and three-dimensional diagrams can be drawn for these cases. We denote elements of E and E^* by \mathbf{x} and \mathbf{x}^* respectively, and the action of \mathbf{x}^* on \mathbf{x} is denoted by $\langle \mathbf{x}^*, \mathbf{x} \rangle$. The Riesz representation theorem states that given $\mathbf{x}^* \in E^*$, there exists a unique element $\mathbf{y} \in E$ such that $\mathbf{x} \cdot \mathbf{y} = \langle \mathbf{x}^*, \mathbf{x} \rangle$ for all $\mathbf{x} \in E$. Thus E^* may be identified with E , but it will be convenient to maintain the distinction.

3.2 Sets

The *neighbourhood* of a point $\bar{\mathbf{x}}$ in E is the set \mathcal{N} defined by

$$\mathcal{N}(\bar{\mathbf{x}}, \epsilon) = \{\mathbf{x} : \mathbf{x} \in E, |\mathbf{x} - \bar{\mathbf{x}}| < \epsilon, 0 < \epsilon \in \mathbf{R}\} \quad (3.1)$$

where ϵ is the radius of the neighbourhood and $|\cdot|$ is the Euclidean length.

Let S be a subset of E . The *interior*, *closure* and *boundary* of S are denoted by $\text{int } S$, $\text{cl } S$ and $\text{bd } S$, respectively. A point $\bar{\mathbf{x}} \in S$ is an *interior point* of S if there exists a neighbourhood of $\bar{\mathbf{x}}$ with finite radius which contains only members of S . If every point in S is an interior point the set S is an *open set*. A point $\bar{\mathbf{x}}$ not necessarily in S is a *limit point* of S if every neighbourhood of $\bar{\mathbf{x}}$ contains at least one point distinct from $\bar{\mathbf{x}}$ which is in S . The *closure* of S is the union of S and all its limit points. A *closed set* contains all its limit points. If T is a subset of E , then S is *dense* in T if $\text{cl } S = T$. The *boundary* of S is the set of points in S which are not interior points. S is *bounded* if it has finite diameter, that is, if

$$|\mathbf{x} - \bar{\mathbf{x}}| < \infty \quad \forall \mathbf{x}, \bar{\mathbf{x}} \in S \quad \text{or} \quad \sup_{\mathbf{x}, \mathbf{y} \in S} |\mathbf{x} - \mathbf{y}| < \infty .$$

The set S is *convex* if, for any \mathbf{x}, \mathbf{y} in S

$$\theta \mathbf{x} + (1 - \theta) \mathbf{y} \in S, \quad 0 < \theta < 1 . \quad (3.2)$$

Intervals on the real line \mathbf{R} are examples of convex sets; Figure 3.1 shows examples of convex and nonconvex subsets of \mathbf{R}^2 .

A *cone* is a set containing the origin which is closed under positive scalar multiplication; a cone is therefore the union of half lines emanating from the origin.

The *normal cone* to a convex set S at \mathbf{x} , denoted by $N_S(\mathbf{x})$, is the set defined by

$$N_S(\mathbf{x}) = \{\mathbf{x}^* \in E^* : \langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle \leq 0 \quad \forall \mathbf{y} \in S\} . \quad (3.3)$$

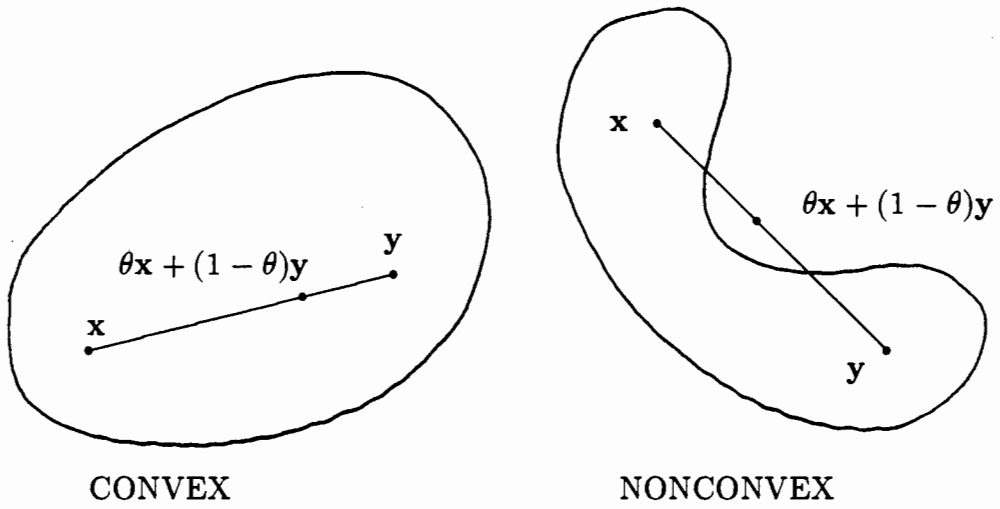


Figure 3.1: Convex and nonconvex sets

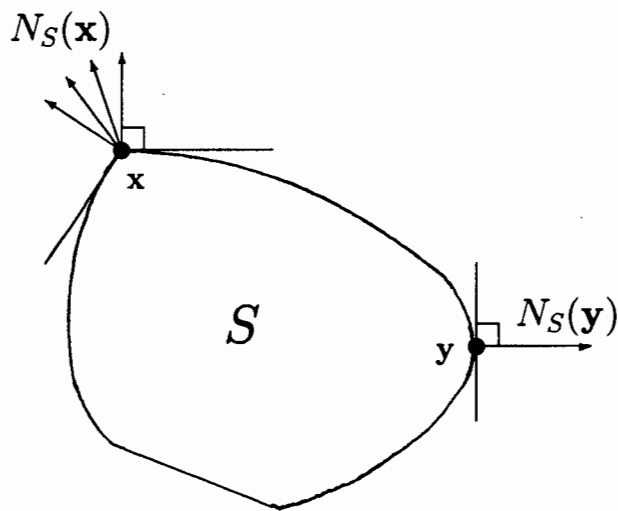


Figure 3.2: Normal cones to a convex set S at nonsmooth and smooth points x and y

When $\mathbf{x} \in \text{int}S$ we clearly have $N_S(\mathbf{x}) = \{\mathbf{0}\}$, whereas for $\mathbf{x} \in \text{bd}S$, $N_S(\mathbf{x})$ consists of the cone of normals at \mathbf{x} ; this is illustrated in Figure 3.2 for the case $S \subset \mathbb{R}^2$.

The *polar* set of S is denoted by S° , and is defined by

$$S^\circ = \{\mathbf{x}^* : \forall \mathbf{x} \in S, \langle \mathbf{x}^*, \mathbf{x} \rangle \leq 1\}. \quad (3.4)$$

If S is a closed convex set containing the origin, then S° is bounded if and only if $\mathbf{0} \in \text{int}S$, and dually S is bounded if and only if $\mathbf{0} \in \text{int}S^\circ$.

3.3 Functions

Let f be a function whose domain is E and whose values are real or $\pm\infty$. The *effective domain* of f , denoted $\text{dom}f$, is defined by

$$\text{dom}f = \{\mathbf{x} \in E : f(\mathbf{x}) < \infty\}. \quad (3.5)$$

The *graph* of f is the set of ordered pairs $(\mathbf{x}, f(\mathbf{x}))$ in $E \times \mathbb{R}$, and the *epigraph* of f , denoted $\text{epi} f$, is the set of ordered pairs

$$\text{epi} f = \{(\mathbf{x}, \alpha) \in E \times \mathbb{R} : f(\mathbf{x}) \leq \alpha\} \quad (3.6)$$

(see Figure 3.3).

The function f is *convex* if

$$f(\theta\mathbf{x} + (1 - \theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1 - \theta)f(\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in E, 0 < \theta < 1 \quad (3.7)$$

and is *strictly convex* if \leq is replaced by $<$ in this condition. Examples of convex, strictly convex and non-convex functions on \mathbb{R} are given in Figure 3.4.

The function f is *positively homogeneous* if

$$f(\alpha\mathbf{x}) = \alpha f(\mathbf{x}) \quad \forall \mathbf{x} \in E, 0 < \alpha \in \mathbb{R}, \quad (3.8)$$

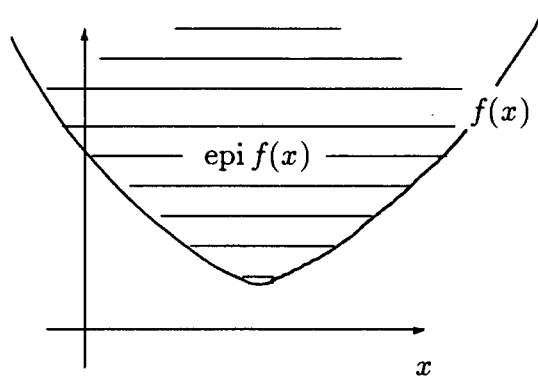


Figure 3.3: The graph and epigraph of a function

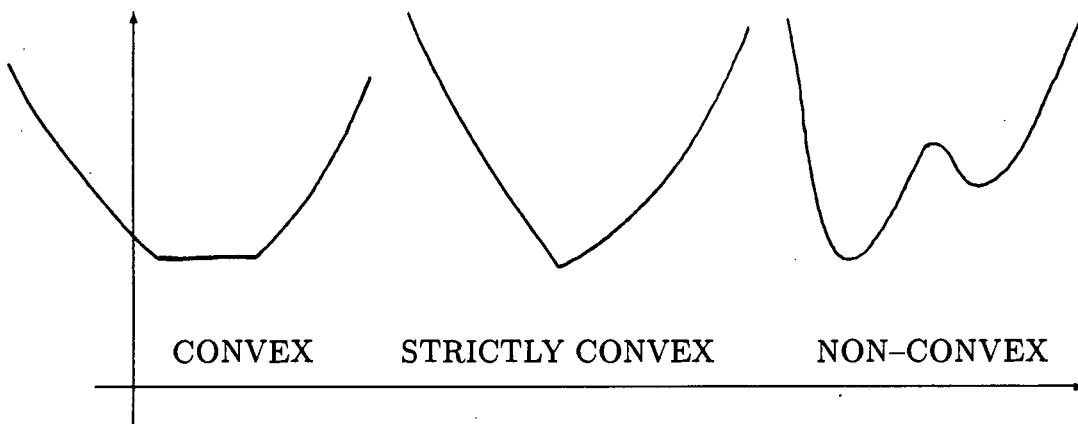


Figure 3.4: Convex, strictly convex and non-convex functions

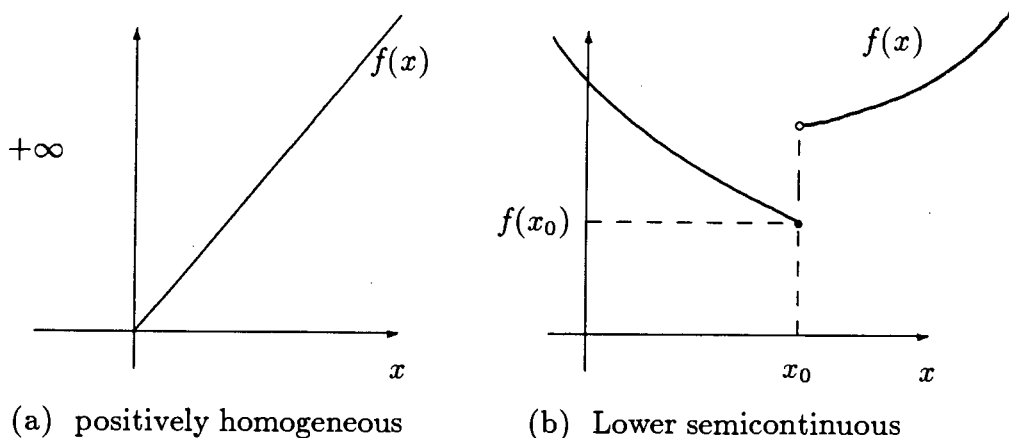


Figure 3.5: Positively homogeneous and lower semicontinuous functions

and *lower semicontinuous*, written *lsc*, if

$$\liminf_{n \rightarrow \infty} f(\mathbf{x}_n) \geq f(\mathbf{x}) \quad (3.9)$$

for any sequence $\{\mathbf{x}_n\}$ converging to \mathbf{x} . Examples of these types of functions are given in Figure 3.5.

It is well known that f is *lsc* if and only if $\text{epi} f$ is closed, and that a convex function is continuous on any open subset of E where its values are finite.

A convex function f is *proper* if $f(\mathbf{x}) \leq +\infty$ for at least one \mathbf{x} and $f(\mathbf{x}) > -\infty$ for every \mathbf{x} .

Let f be a proper *lsc* convex function on E . The *conjugate function* f^* of f is defined by

$$f^*(\mathbf{x}^*) = \sup\{\langle \mathbf{x}^*, \mathbf{x} \rangle - f(\mathbf{x}) : \forall \mathbf{x} \in E\}. \quad (3.10)$$

We have the further result that if f is proper, convex and *lsc*, then so is f^* , and furthermore

$$f^{**} \equiv (f^*)^* = f. \quad (3.11)$$

For any set $S \subset E$, the *indicator function* I_S of S is defined on E by

$$I_S(\mathbf{x}) = \begin{cases} 0 & , \mathbf{x} \in S \\ +\infty & , \mathbf{x} \notin S \end{cases} , \quad (3.12)$$

and the *support function* σ_S of S is defined on E^* by

$$\sigma_S(\mathbf{x}^*) = \sup\{\langle \mathbf{x}^*, \mathbf{x} \rangle : \mathbf{x} \in S\} . \quad (3.13)$$

From the definition (3.13) and (3.10) we see that the support function σ_S is *conjugate* to the indicator function I_S :

$$I_S^* = \sigma_S . \quad (3.14)$$

In particular, I_S is proper, convex and *lsc* whenever S is closed and convex, so for such a set

$$I_S = \sigma^*_S = I^{**}_S . \quad (3.15)$$

Differentiation of non-smooth convex functions is very important in the theory we develop later. Given a convex function f on E , for any $\mathbf{x} \in E$ the *subdifferential* $\partial f(\mathbf{x})$ of f at \mathbf{x} is the subset (possibly empty) of E^* defined by

$$\partial f : E \rightarrow 2^{E^*} , \quad \partial f(\mathbf{x}) = \{\mathbf{x}^* \in E^* : f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle \quad \forall \mathbf{y} \in E\} \quad (3.16)$$

where 2^{E^*} denotes the class of subsets of E^* . The members of $\partial f(\mathbf{x})$ are called *subgradients*. In Figure 3.6 we show the subdifferential of a function on \mathbb{R} . According to the definition, $\partial f(\mathbf{x}) = \emptyset$ if $\mathbf{x} \notin \text{dom } f$. If f is differentiable at \mathbf{x} then clearly

$$\partial f(\mathbf{x}) = \{\nabla f(\mathbf{x})\} \quad (3.17)$$

(this is also illustrated in Figure 3.6).

We also define the *partial subdifferential* of a function $f(\mathbf{x}, \mathbf{z})$, defined on $E \times E$, at point (\mathbf{x}, \mathbf{z}) with respect to \mathbf{x} as the set

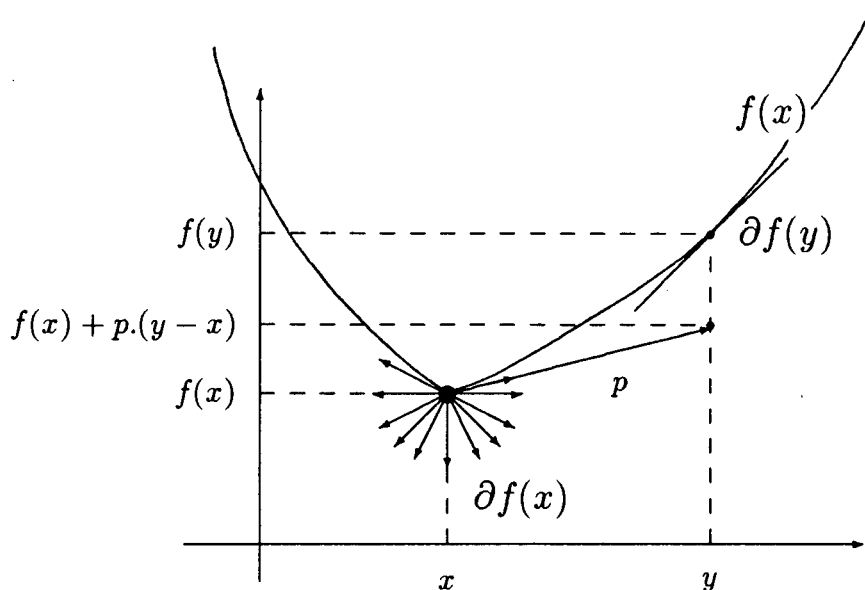


Figure 3.6: The subdifferential of a nonsmooth convex function

$$\partial_{\mathbf{x}} f(\mathbf{x}, \mathbf{z}) = \{\mathbf{x}^* \in E^* : f(\mathbf{y}, \mathbf{z}) \geq f(\mathbf{x}, \mathbf{z}) + \langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle \quad \forall \mathbf{y} \in E\}. \quad (3.18)$$

We have the important result that

$$\mathbf{x}^* \in \partial f(\mathbf{x}) \quad \text{iff} \quad \mathbf{x} \in \partial f^*(\mathbf{x}^*). \quad (3.19)$$

For the special case of the indicator function, it is evident from (3.3) and (3.16) that

$$\partial I_S(\mathbf{x}) = N_S(\mathbf{x}). \quad (3.20)$$

A function $g : E \rightarrow [0, \infty]$ is called a *gauge* if g is convex, positively homogeneous and $g(\mathbf{0}) = 0$; this implies that $\text{epi } g$ is a convex cone in $E \times \mathbf{R}$ which contains the origin but does not contain any vectors (\mathbf{x}, μ) such that $\mu < 0$. These properties imply that g is *lsc*. A gauge function satisfies

$$g(\mathbf{x}) = \inf \{\mu > 0 : \mathbf{x} \in \mu C\} \quad (3.21)$$

where C is a non-empty convex set. In general C is not unique, but we are interested in the case for which

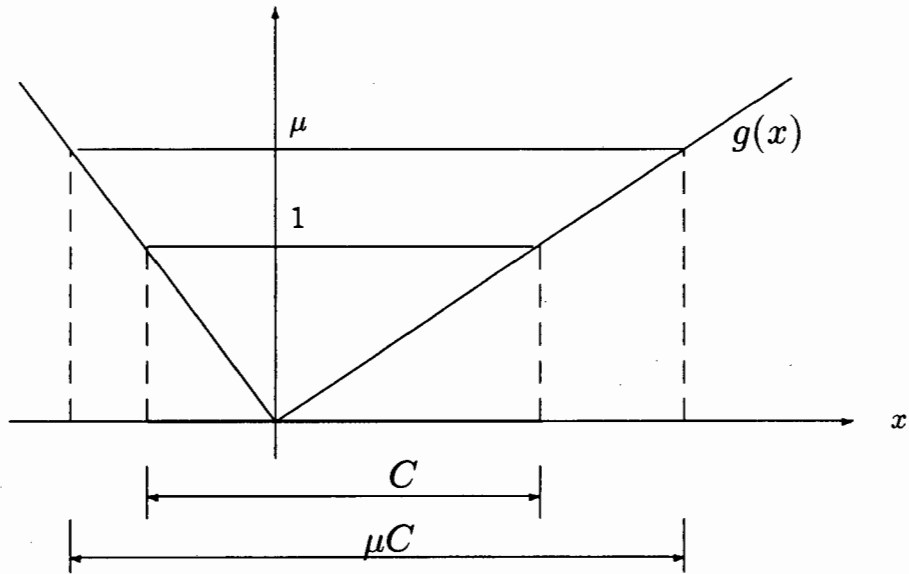


Figure 3.7: The gauge function of a convex set C

$$C = \{\mathbf{x} : g(\mathbf{x}) \leq 1\}. \quad (3.22)$$

In this case if g is proper, convex and *lsc* then C is a unique convex set containing the origin and g is called *the gauge function of C* ; this is illustrated in Figure 3.7.

The *polar* of a function f is denoted f° and is defined by

$$f^\circ(\mathbf{x}^*) = \inf\{\mu \geq 0 : \langle \mathbf{x}^*, \mathbf{x} \rangle \leq \mu f(\mathbf{x}), \forall \mathbf{x}\}. \quad (3.23)$$

If in addition to the properties given above f is finite everywhere, then f° can be written as

$$f^\circ(\mathbf{x}^*) = \sup_{\mathbf{x} \neq 0} \frac{\langle \mathbf{x}^*, \mathbf{x} \rangle}{f(\mathbf{x})}. \quad (3.24)$$

The polar g° of the gauge function g of a set C can equivalently be defined by

$$g^\circ(\mathbf{x}^*) = \inf\{\mu > 0 : \mathbf{x}^* \in \mu C^\circ\} \quad (3.25)$$

where C° is the set polar to C ; this result is given as Theorem 15.1 in [103]. We also have the result that $g^{\circ\circ} = g$. A corollary to the theorem is that if C is a closed convex set containing the origin, the gauge function of C and the support function of C are gauges which are polar to each other.

Gauges that are polar to each other have the property that

$$\langle \mathbf{x}^*, \mathbf{x} \rangle \leq g(\mathbf{x})g^\circ(\mathbf{x}^*), \quad \forall \mathbf{x} \in \text{dom } g, \quad \forall \mathbf{x}^* \in \text{dom } g^\circ. \quad (3.26)$$

3.4 Special results

In this Section we present the main theorem used in the development of the new theory of plasticity presented in Chapter VI, but before doing so we give some results which are of particular importance in the proof of this theorem.

We set

$$\text{dom}(\partial f) = \{\mathbf{x} \in E : \partial f(\mathbf{x}) \neq \emptyset\} \quad (3.27)$$

and note the following results:

Lemma 3.1

- a) Let f be a proper, convex, *lsc* function on E . Then $\text{dom}(\partial f) \neq \emptyset$ and $\text{dom}(\partial f)$ is dense in $\text{dom } f$.
- b) Let f and g be proper, convex, *lsc* functions on E . Then $\partial f(\mathbf{x}) = \partial g(\mathbf{x}) \quad \forall \mathbf{x} \in E$ if and only if $f = g + \text{const}$.

Lemma 3.2

Let g be non-negative and convex, with $g(\mathbf{0}) = 0$ and \mathbf{x} a point in the interior of $\text{dom } g$ such that $g(\mathbf{x}) > 0$. Set $C = \{\mathbf{z} : g(\mathbf{z}) \leq g(\mathbf{x})\}$. Then $\mathbf{x}^* \in N_C(\mathbf{x})$ if and only if there exists $\lambda \geq 0$ such that $\mathbf{x}^* \in \lambda \partial g(\mathbf{x})$.

Remark. Lemma 3.2 appears in [103] as a Corollary (see Cor. 23.7.1). We give a simple proof here.

Proof. Assume first that $\mathbf{x}^* \in \lambda \partial g(\mathbf{x})$ for some $\lambda \geq 0$. Then $\mathbf{x}^* = \lambda \mathbf{y}^*$ for $\mathbf{y}^* \in \partial g(\mathbf{x})$, and

$$g(\mathbf{z}) \geq g(\mathbf{x}) + \langle \mathbf{y}^*, (\mathbf{z} - \mathbf{x}) \rangle.$$

For $\mathbf{z} \in C$, $g(\mathbf{z}) \leq g(\mathbf{x})$, so that

$$\langle \mathbf{y}^*, (\mathbf{z} - \mathbf{x}) \rangle \leq 0 \Rightarrow \langle \mathbf{x}^*, (\mathbf{z} - \mathbf{x}) \rangle \leq 0 \quad \forall \mathbf{z} \in C,$$

hence $\mathbf{x}^* \in N_C(\mathbf{x})$.

Now assume that $\mathbf{x}^* \in N_C(\mathbf{x})$. We want to show that there exists $\lambda \geq 0$ such that $\mathbf{x}^* = \lambda \mathbf{z}^*$ and

$$g(\mathbf{y}) \geq g(\mathbf{x}) + \langle \mathbf{z}^*, (\mathbf{y} - \mathbf{x}) \rangle \quad \forall \mathbf{y} \quad (\text{i}).$$

If $\mathbf{x}^* \neq 0$ and $\langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle = 0$ then $\mathbf{x} \in \text{bd } C$ and since C is convex, $\mathbf{y} \in E \setminus \text{int } C$, the complement of $\text{int } C$. Hence $g(\mathbf{y}) \geq g(\mathbf{x})$, so that (i) holds. Assume next that $\langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle < 0$. Then for $\mathbf{y} \in E \setminus C$ we have $g(\mathbf{y}) - g(\mathbf{x}) \geq 0$ so that (i) holds for any $\lambda \geq 0$. Finally suppose that $\mathbf{y} \in C$. Then for some $\mu > 0$

$$g(\mathbf{y}) - g(\mathbf{x}) - \mu \langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle = \left\{ \frac{g(\mathbf{y}) - g(\mathbf{x})}{\langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle} - \mu \right\} \langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle \quad (\text{ii})$$

which is non-negative provided that $\mu > \max\{(g(\mathbf{y}) - g(\mathbf{x})) / \langle \mathbf{x}^*, (\mathbf{y} - \mathbf{x}) \rangle : \mathbf{y} \in C\}$. The result follows with $\lambda = \mu^{-1}$. \square

We now define a special class of mappings which we call *maximal responsive relations* and develop a theorem involving these. This theorem is an integral component of the theory of plasticity described in Chapter V. In preparation for the later application of the results recorded here we adjust the notation and replace \mathbf{x} by \mathbf{p} and \mathbf{x}^* by \mathbf{Z} .

The relationship between plastic flow and conjugate forces is usually given in the form

$$\mathbf{p} \in F(\mathbf{Z}) \quad (3.28)$$

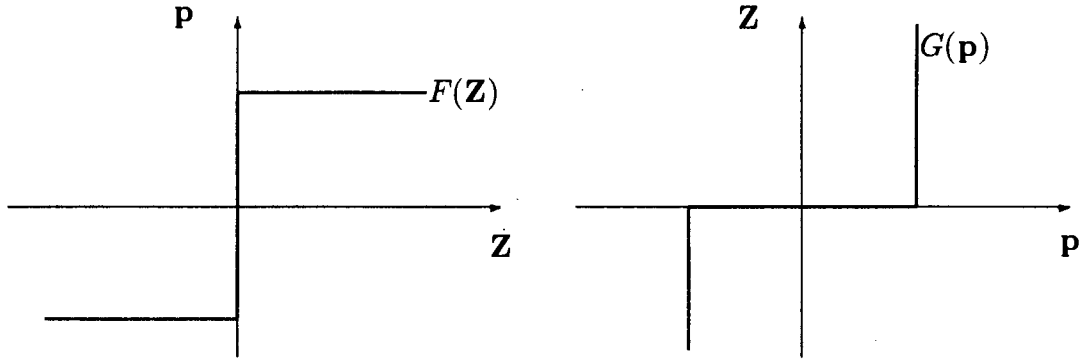


Figure 3.8: Multivalued mappings $F(\mathbf{Z})$ and $G(\mathbf{p})$

where \mathbf{p} and \mathbf{Z} represent values of the plastic flow and conjugate stresses respectively and F is *multivalued*. A multivalued map $F : E^* \rightarrow 2^E$ is one under which the image of a single element $\mathbf{Z} \in E^*$ is a set in E ; this set may contain more than one element. Such maps are also referred to as *set-valued maps*. We shall only consider maps for which the image set of any point is a connected set.

Consider a correspondence $G : \mathbf{p} \rightarrow G(\mathbf{p})$ which associates with each $\mathbf{p} \in E$ a set $G(\mathbf{p})$ (possibly empty for some \mathbf{p}) in E^* . The relation inverse to (3.28) is thus of the form

$$\mathbf{Z} \in G(\mathbf{p}) . \quad (3.29)$$

This equation reflects the multivaluedness of G . In Figure 3.8 we give a simple one-dimensional example of mappings, F and G , of the type which occur in perfect plasticity.

Definition 3.1

The map G is said to be *responsive* if

$$\mathbf{0} \in G(\mathbf{0}) \quad (3.30)$$

and if for any $\mathbf{p}_0, \mathbf{p}_1 \in E$,

$$\langle \mathbf{Z}_0 - \mathbf{Z}_1, \mathbf{p}_0 \rangle \geq 0 \quad \text{and} \quad \langle \mathbf{Z}_1 - \mathbf{Z}_0, \mathbf{p}_1 \rangle \geq 0 \quad (3.31)$$

whenever $\mathbf{Z}_0 \in G(\mathbf{p}_0)$ and $\mathbf{Z}_1 \in G(\mathbf{p}_1)$.

Definition 3.2

Let $G : E \rightarrow 2^{E^*}$ be a responsive map. G is said to be *maximal responsive* if there is no other responsive map whose graph properly includes the graph of G .

Remark. The notion of maximality occurs in various branches of nonlinear analysis, for example, the theory of maximal monotone operators (see, for example, [85,104,124]). In the present context maximal responsiveness is equivalent to the property that for any $\mathbf{p}_0 \in E$ and $\mathbf{Z}_0 \in E^*$, the condition

$$\langle \mathbf{Z}_0 - \mathbf{Z}_1, \mathbf{p}_0 \rangle \geq 0 \quad \text{and} \quad \langle \mathbf{Z}_1 - \mathbf{Z}_0, \mathbf{p}_1 \rangle \geq 0 \quad \forall \mathbf{p}_1, \mathbf{Z}_1 \in G(\mathbf{p}_1) \quad (3.32)$$

which *implies* that $\mathbf{Z}_0 \in G(\mathbf{p}_0)$.

Lemma 3.3

Let \mathcal{D} be a lower semicontinuous gauge on E and define the closed convex set K by

$$K = \{ \mathbf{Z} \in E^* : \langle \mathbf{Z}, \mathbf{p} \rangle \leq \mathcal{D}(\mathbf{p}) \quad \forall \mathbf{p} \in E \} . \quad (3.33)$$

Then

a) \mathcal{D} is the support function of K :

$$\mathcal{D}(\mathbf{p}) = \sup_{\mathbf{Z} \in K} \langle \mathbf{Z}, \mathbf{p} \rangle ;$$

b) the function \mathcal{D}^* conjugate to \mathcal{D} is the *indicator function* of K :

$$\mathcal{D}^*(\mathbf{Z}) = \begin{cases} 0, & \mathbf{Z} \in K, \\ +\infty, & \mathbf{Z} \notin K; \end{cases}$$

c) $K = \partial \mathcal{D}(\mathbf{0})$;

d) $\mathbf{Z} \in \partial \mathcal{D}(\mathbf{p}) \Leftrightarrow \mathbf{p} \in \partial \mathcal{D}^*(\mathbf{Z}) = N_K(\mathbf{Z})$.

(The notation \mathcal{D} is chosen with the later application in mind).

We now come to the theorem which provides the main result of this Section.

Theorem 3.1

Let $G : E \rightarrow 2^{E^*}$ be a multivalued mapping. Then the following are equivalent:

- a) G is a maximal responsive.
- b) There exists a gauge \mathcal{D} on E with the property

$$G(\mathbf{p}) = \partial\mathcal{D}(\mathbf{p}) \quad \forall \mathbf{p} \in E .$$

Furthermore, when G is maximal responsive it determines \mathcal{D} *uniquely* and in terms of the set $\text{dom } G = \{\mathbf{p} \in E : G(\mathbf{p}) \neq \emptyset\}$ satisfies

$$\mathcal{D}(\mathbf{p}) = \begin{cases} \mathbf{Z} \cdot \mathbf{p} & \forall \mathbf{Z} \in G(\mathbf{p}), \mathbf{p} \in \text{dom}G, \\ +\infty & \text{if } \mathbf{p} \notin \text{cl dom}G. \end{cases} \quad (3.34)$$

Remark

The above theorem is similar in nature to results which connect maximal cyclic monotone maps with the subgradients of convex *lsc* functions (see Rockafellar [103], Zeidler [124]).

Proof of Theorem 3.1

We first show that (b) implies (a) and then that (a) implies (b).

(b) \Rightarrow (a):

Assume that the conditions of (b) hold. It follows from Lemma 3.3 that

$$G(\mathbf{p}) = \{\mathbf{Z} \in K : \langle \mathbf{Z}, \mathbf{p} \rangle \geq \langle \mathbf{Y}, \mathbf{p} \rangle \quad \forall \mathbf{Y} \in \mathbf{K}\} \quad (3.35)$$

since $G = \partial\mathcal{D}$, where K is defined by (3.33). In particular, (3.31) holds. Property (3.30) follows from the observation that $\mathbf{0} \in K = \partial\mathcal{D}(\mathbf{0})$ (Lemma 3.3(c)). To show that G is maximal, consider any pair $(\bar{\mathbf{p}}, \bar{\mathbf{Z}})$ such that

$$\langle \mathbf{Z} - \bar{\mathbf{Z}}, \mathbf{p} \rangle \geq 0 \quad \text{and} \quad \langle \bar{\mathbf{Z}} - \mathbf{Z}, \bar{\mathbf{p}} \rangle \geq 0 \quad (3.36)$$

holds for all $\mathbf{p} \in E$ and $\mathbf{Z} \in G(\mathbf{p})$. We must verify that $\bar{\mathbf{Z}} \in G(\bar{\mathbf{p}})$. We have, from Lemma 3.2(d),

$$\langle \bar{\mathbf{Z}}, \mathbf{p} \rangle \leq \mathcal{D}(\mathbf{p}) \quad \forall \mathbf{p} \in E \quad (3.37)$$

so that $\bar{\mathbf{Z}} \in K$. The second part of (3.36) then implies by (3.35) that $\bar{\mathbf{Z}} \in G(\bar{\mathbf{p}})$, as required.

(a) \Rightarrow (b)

To see that

$$G(\mathbf{0}) \supset G(\mathbf{p}) \quad \forall \mathbf{p} \in E, \quad (3.38)$$

consider any $\bar{\mathbf{p}}$ and $\bar{\mathbf{Z}} \in G(\bar{\mathbf{p}})$. From (3.31),

$$\langle \mathbf{Z} - \bar{\mathbf{Z}}, \mathbf{p} \rangle \geq 0 \quad \text{and} \quad \langle \bar{\mathbf{Z}} - \mathbf{Z}, \bar{\mathbf{p}} \rangle \geq 0 \quad (3.39)$$

whenever $\mathbf{Z} \in G(\mathbf{p})$. Hence the pair $(\mathbf{0}, \bar{\mathbf{Z}})$ has the property that

$$\langle \mathbf{Z} - \bar{\mathbf{Z}}, \mathbf{p} \rangle \geq 0 \quad \text{and} \quad \langle \bar{\mathbf{Z}} - \mathbf{Z}, \mathbf{0} \rangle \geq 0 \quad (3.40)$$

whenever $\mathbf{Z} \in G(\mathbf{p})$, and so $(\mathbf{0}, \bar{\mathbf{Z}})$ could be added to the graph of G without violating (3.31). Since G is maximal responsive, we must have $\bar{\mathbf{Z}} \in G(\mathbf{0})$, whence (3.38).

The above argument actually establishes that $G(\mathbf{0})$ coincides with the set

$$K = \{ \bar{\mathbf{Z}} \in E^* : \langle \mathbf{Z} - \bar{\mathbf{Z}}, \mathbf{p} \rangle \geq 0 \quad \forall \mathbf{p} \in E, \mathbf{Z} \in G(\mathbf{p}) \}. \quad (3.41)$$

This set is closed and convex, and contains $\mathbf{0}$, by property (3.30). From (3.38) and (3.41),

$$\bar{\mathbf{Z}} \in G(\bar{\mathbf{p}}) \Rightarrow \bar{\mathbf{Z}} \in G(\mathbf{0}) = K \Rightarrow \bar{\mathbf{p}} \in N_K(\bar{\mathbf{Z}}). \quad (3.42)$$

Let \mathcal{D} be the support function of K . Since \mathcal{D} is the support function of a closed convex set containing $\mathbf{0}$ it is a gauge (see [103] Ch.15) and

$$\bar{\mathbf{Z}} \in \partial \mathcal{D}(\bar{\mathbf{p}}) \Leftrightarrow \bar{\mathbf{p}} \in N_K(\bar{\mathbf{Z}}). \quad (3.43)$$

Moreover $\partial\mathcal{D}$ is a responsive map. Furthermore, (3.42) implies that the graph of G is included in the graph of $\partial\mathcal{D}$. Inasmuch as G is maximal responsive, we may conclude that $G = \partial\mathcal{D}$, whence part (a).

To establish the uniqueness of \mathcal{D} , we recall (Lemma 3.1) that two *lsc* proper convex functions have the same subdifferential if and only if they differ by an additive constant. We fix this constant by the requirement that $\mathcal{D}(\mathbf{0}) = 0$, thereby defining \mathcal{D} uniquely.

To establish (3.34), we note that \mathcal{D} is the support function of K , defined by (3.33), so that from (3.35) $\mathcal{D}(\mathbf{p}) = \mathbf{Z} \cdot \mathbf{p}$ when $\mathbf{Z} \in G(\mathbf{p})$. Since $\text{cl dom}G = \text{cl dom}\partial\mathcal{D} = \text{cl dom}\mathcal{D}$ (Lemma 2.1) we also have $\mathcal{D}(\mathbf{p}) = +\infty$ when $\mathbf{p} \in \text{cl dom}G$, whence (3.34). \square

CHAPTER IV

KINEMATICS OF FINITE-STRAIN PLASTICITY

4.1 Introduction

In our introductory review of developments in finite-strain plasticity, we identified the issue of the kinematic description of finite plastic deformations as one which has in the past been particularly contentious. Since its introduction the multiplicative decomposition of deformation gradients, proposed by Lee [56], has been used by many authors. It is now widely accepted as a suitable basis for describing the kinematics of problems involving elastoplastic deformations. There remains some debate as to the precise physical interpretation of the decomposition, but this does not detract from its usefulness as a mathematical tool. Other decompositions which are commonly assumed, for example, additive decompositions of Lagrangian strain and of deformation rate have been shown to be compatible with the multiplicative decomposition of deformation gradient (see for example Kleiber [53] and Simo and Ortiz [114]).

We use the multiplicative decomposition of deformation gradient as a basis for the kinematics used in the theory of finite-strain plasticity which we develop in later Chapters. In this Chapter we describe the multiplicative decomposition and define and derive all the kinematic quantities we shall need. For the purposes of later discussion we include definitions of some of the other quantities which are commonly used and describe the interrelationships between these. While some standardisation of the notation used in finite-strain plasticity literature has emerged, some variation in the definition of certain quantities remains. Here we establish the notation which we shall use, where possible using standard symbols and conventions. We give a concise account of the main concepts used in describing the deformation of bodies and define the relevant kinematic quantities before beginning the discussion of plastic deformations.

4.2 Deforming bodies

We consider the mathematical description of the deformation of bodies made up of continuously distributed material.

A body B is defined as a set of particles, a typical particle being denoted by X . The body B occupies a region Ω in three-dimensional Euclidean point space \mathcal{E} , hence the particles X are referred to as material points. As indicated in Chapter II, each point x in \mathcal{E} may be put in one-to-one correspondence with a place denoted \boldsymbol{x} , and referred to as the position vector of the point in \mathcal{E} . The position vectors are defined relative to a fixed but arbitrarily chosen origin o in \mathcal{E} . The region Ω is defined by a mapping $\chi : B \rightarrow \Omega \subset \mathcal{E}$; χ is assumed to be invertible and twice continuously differentiable, and it defines a configuration of the body:

$$\boldsymbol{x} = \chi(X) \quad \text{hence} \quad X = \chi^{-1}(\boldsymbol{x}). \quad (4.1)$$

The region $\Omega \subset \mathcal{E}$ is that region occupied by the body B in *configuration* χ ; we shall somewhat loosely refer to Ω as the configuration of the body.

Physical observations of bodies are made in specific configurations. It is therefore convenient to identify a *reference configuration* $\tilde{\Omega}$ defined by $\tilde{\chi} : B \rightarrow \tilde{\Omega} \subset \mathcal{E}$. This configuration is a fixed, but arbitrary configuration. To define the position vector field \boldsymbol{X} an origin $\mathbf{0}$ is chosen so that

$$\boldsymbol{X} = \tilde{\chi}(X). \quad (4.2)$$

To express vectors and tensors defined on $\tilde{\Omega}$ in component form a basis is also needed. We denote the orthonormal basis chosen for the reference configuration by $\{\mathbf{E}_A\}$ ($A=1,2,3$). Material points in the body are identified by their places \boldsymbol{X} in $\tilde{\Omega}$. Components $X_A = \boldsymbol{X} \cdot \mathbf{E}_A$ are referred to as the material coordinates of the point \boldsymbol{X} . Note that when defining referential quantities we have deviated from the convention of using lower case letters for vectors.

A *motion* of a body B is a one-parameter family of configurations $\chi_t : B \rightarrow \Omega_t \subset \mathcal{E}$. The *path* traced out by a particle \boldsymbol{X} is thus given by

$$\mathbf{x}(\mathbf{X}, t) = \chi_t(\mathbf{X}) \quad \text{and} \quad \mathbf{X} = \chi_o(\mathbf{X}). \quad (4.3)$$

We consider quasi-static problems so that the parameter t need only be *time-like*; we shall however refer to it as *time*. A configuration Ω_t is called the *current configuration* at time t . Note that the reference configuration need not belong to the family Ω_t and is not dependent on *time*. It is however often convenient to choose the initial configuration at time $t = 0$ as the reference; we shall always do this and therefore use Ω_o rather than $\tilde{\Omega}$ for the reference configuration.

We use \mathbf{o} , \mathbf{x} and $\{\mathbf{e}_i\}$ to denote the origin, position vectors and basis used in the current configuration.

The notion of an observer and associated frame of reference is fundamental to the theory of physical phenomena. We do not attempt to describe this abstract notion here; detailed discussion of observers is given by, among others, Ogden in [86] and Truesdell in [120]. For practical purposes an observer is equivalent to a particular choice of frame of reference, that is a choice of origin and set of basis vectors. We note that an observer transformation or change of observer is characterised by

$$\mathbf{x}^* = \mathbf{c}(t) + \mathbf{Q}(t)\mathbf{x} \quad (4.4)$$

where \mathbf{x} and \mathbf{x}^* are the positions of a point observed by observers O and O^* using origins \mathbf{o} and \mathbf{o}^* respectively. The vector \mathbf{c} is the shift of the origin and the proper orthogonal tensor \mathbf{Q} represents the rotation of basis vectors.

A *deformation* of a body is a change in the relative positions of particles which occurs when there is a change in the configuration of the body. The *deformation gradient*, denoted \mathbf{F} is the gradient of the vector field describing the position of material points in the current (deformed) configuration with respect to some referential coordinates. It is defined by

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \text{Grad } \mathbf{x}; \quad (4.5)$$

note that $\text{Grad}(\cdot)$ is used to indicate differentiation with respect to referential coordinates \mathbf{X} and $\text{grad}(\cdot)$ will be used for differentiation with respect to current coordinates

\mathbf{x} . Later we shall use only the former, which we then denote by $\nabla(\cdot)$. The deformation gradient has the following representation relative to the bases $\{\mathbf{e}_i\}$ and $\{\mathbf{E}_A\}$:

$$\mathbf{F} = \frac{\partial x_i}{\partial X_A} \mathbf{e}_i \otimes \mathbf{E}_A = F_{iA} \mathbf{e}_i \otimes \mathbf{E}_A . \quad (4.6)$$

Quantities associated with the reference configuration are referred to as *Lagrangian* or *referential* quantities while quantities associated with the deformed or current configuration are called *Eulerian*, *spatial* or *current* quantities. We differentiate between these by using upper and lower case letters to denote Lagrangian and Eulerian quantities respectively.

The deformation of bodies is often expressed in terms of the *displacement* from one configuration to another. The displacement of a body forms a vector field \mathbf{u} over the body; this field is defined by

$$\mathbf{u}(\mathbf{X}, t) = \mathbf{x}(\mathbf{X}, t) - \mathbf{X} , \quad (4.7)$$

from which it follows that

$$\mathbf{F} = \mathbf{I} + \text{Grad } \mathbf{u} . \quad (4.8)$$

Material line elements at points $\mathbf{x} \in \Omega_t$ and $\mathbf{X} \in \Omega_0$ are denoted by $d\mathbf{x}$ and $d\mathbf{X}$ respectively. The material making up $d\mathbf{X}$ moves under a deformation with gradient \mathbf{F} so as to lie along $d\mathbf{x}$, so that

$$d\mathbf{x} = \mathbf{F}d\mathbf{X} . \quad (4.9)$$

According to the polar decomposition theorem, we may write \mathbf{F} in the form

$$\mathbf{F} = \mathbf{R}\mathbf{U} \quad \text{or} \quad \mathbf{F} = \mathbf{V}\mathbf{R} \quad (4.10)$$

where \mathbf{R} is a proper orthogonal tensor which characterises the rotation of line elements, and \mathbf{U} and \mathbf{V} are symmetric positive-definite tensors which provide a measure of the

stretch of line elements. Although \mathbf{U} and \mathbf{V} appear to be the simplest measures of the local deformation, in practice it is more convenient to work with either the *right* or the *left Cauchy–Green deformation tensors*. The right Cauchy–Green deformation tensor \mathbf{C} is the Lagrangian quantity defined by

$$\mathbf{C} = \mathbf{F}^T \mathbf{F} = \mathbf{U}^2 \quad (4.11)$$

while the left Cauchy–Green deformation tensor \mathbf{b} is the Eulerian quantity defined by

$$\mathbf{b} = \mathbf{F} \mathbf{F}^T = \mathbf{V}^2. \quad (4.12)$$

Strain measures are defined in terms of either \mathbf{U} or \mathbf{V} ; the most commonly used strain tensors are the Lagrangian strain \mathbf{E} and the Eulerian strain \mathbf{e} defined respectively by

$$\mathbf{E} = \frac{1}{2}(\mathbf{U}^2 - \mathbf{I}) = \frac{1}{2}(\mathbf{C} - \mathbf{I}) \quad (4.13)$$

and

$$\mathbf{e} = \frac{1}{2}(\mathbf{I} - \mathbf{V}^{-2}) = \frac{1}{2}(\mathbf{I} - \mathbf{b}^{-1}). \quad (4.14)$$

Note that $\mathbf{E} = \mathbf{F}^T \mathbf{e} \mathbf{F}$ and $\mathbf{e} = \mathbf{F}^{-T} \mathbf{E} \mathbf{F}^{-1}$. We use a form of Cauchy–Green tensor rather than strain tensors in our description of deformation.

A *natural* configuration of a body is usually defined as one which is *unstressed*; in this thesis we assume that stress is dependent only on the elastic deformation of the body. We consider a natural configuration to be one which is *unstressed* and *undeformed*; this definition takes into account the possibility of deformations which are only plastic. Some simplification of the theory is achieved by always choosing the reference configuration to be a natural configuration.

An element of volume in the reference configuration, denoted dV , may be defined using the box product of a triad of line elements $d\mathbf{X}_i$. Under a deformation characterised by \mathbf{F} these line elements transform to the line elements $d\mathbf{x}_i$, the box product of which is the volume element dv in the deformed configuration. We introduce the notation $J = \det \mathbf{F}$ and record the relationship

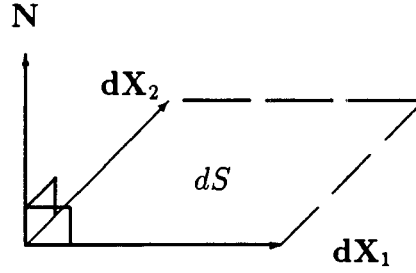


Figure 4.1: Element of area in the reference configuration

$$J = \frac{dv}{dV} \quad (4.15)$$

which is obtained directly from (2.19).

An element of surface area denoted \mathbf{dS} with magnitude dS and unit normal \mathbf{N} in the reference configuration, as shown in Figure 4.1, is defined by

$$\mathbf{dS} = \mathbf{dX}_1 \wedge \mathbf{dX}_2 = \mathbf{N}dS .$$

Under a deformation characterised by \mathbf{F} this surface element transforms to the surface element \mathbf{ds} defined by

$$\mathbf{ds} = \mathbf{dx}_1 \wedge \mathbf{dx}_2 = \mathbf{n}ds$$

where \mathbf{n} is the unit normal to the surface in the deformed configuration. An arbitrary volume defined by $\mathbf{dS} \cdot \mathbf{dX}$ deforms to $\mathbf{ds} \cdot \mathbf{dx}$, that is, $\mathbf{ds} \cdot \mathbf{dx} = J\mathbf{dS} \cdot \mathbf{dX}$ and hence

$$\mathbf{n}ds = J\mathbf{F}^{-T}\mathbf{N}dS . \quad (4.16)$$

This relationship is known as *Nanson's formula*; it plays an important role in the definition of stress measures.

4.3 Elastic-plastic deformations

We use a multiplicative decomposition of deformation gradient \mathbf{F} as a basis for the description of the kinematics of elastic-plastic behaviour. The deformation gradient is written as

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p \quad (4.17)$$

where \mathbf{F}^e and \mathbf{F}^p are respectively its elastic and plastic parts.

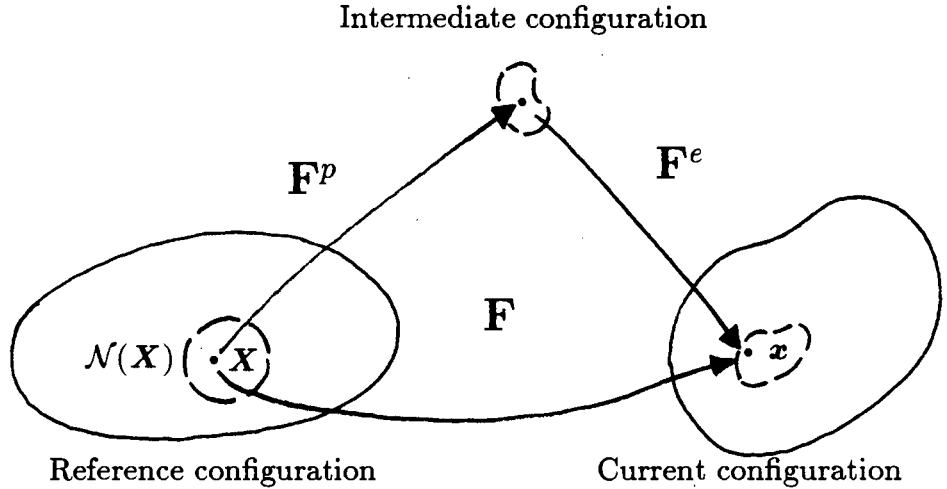
Physical interpretation of this multiplicative decomposition has in the past been the subject of considerable controversy [57,62,59,61,83,84]. Essentially the concept of an intermediate configuration is added to that of a natural reference and current configuration. This interpretation can however only be applied locally. We emphasise that \mathbf{F}^p and \mathbf{F}^e cannot be interpreted as the gradients of motion from reference to intermediate configuration and from intermediate to current configurations respectively. Here the intermediate configuration represents a state in which the neighbourhood of each material point in the body is only plastically deformed and is therefore unstressed. This concept is illustrated in Figure 4.2.

The multiplicative decomposition (4.17) is nonunique since an arbitrary rigid body motion of the intermediate configuration leaves the total deformation unchanged. We therefore have for given \mathbf{F}^e and \mathbf{F}^p , and an arbitrary proper orthogonal tensor \mathbf{Q} say, that

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p = \mathbf{F}^e \mathbf{Q} \mathbf{Q}^T \mathbf{F}^p = \tilde{\mathbf{F}}^e \tilde{\mathbf{F}}^p \quad (4.18)$$

where $\tilde{\mathbf{F}}^e = \mathbf{F}^e \mathbf{Q}$ and $\tilde{\mathbf{F}}^p = \mathbf{Q}^T \mathbf{F}^p$. This lack of uniqueness may be avoided by using the polar decomposition theorem and making an assumption about the rotational components of \mathbf{F}^e and \mathbf{F}^p . One may assume that either \mathbf{F}^e or \mathbf{F}^p is rotation free, hence symmetric and positive-definite, and write

$$\mathbf{F}^e = \mathbf{V}^e \quad \text{and} \quad \mathbf{F}^p = \mathbf{V}^p \mathbf{R} \quad (4.19)$$



The reference and current configurations of a material body and reference, intermediate and current configurations of a neighbourhood $\mathcal{N}(X)$ of a material point X

Figure 4.2: Multiplicative decomposition of deformation gradient.

or, alternatively,

$$\mathbf{F}^e = \mathbf{R}\mathbf{U}^e \quad \text{and} \quad \mathbf{F}^p = \mathbf{U}^p \quad (4.20)$$

so that either $\mathbf{F} = \mathbf{V}^e\mathbf{V}^p\mathbf{R}$, that is $\mathbf{V} = \mathbf{V}^e\mathbf{V}^p$, or $\mathbf{F} = \mathbf{R}\mathbf{U}^e\mathbf{U}^p$, that is $\mathbf{U} = \mathbf{U}^e\mathbf{U}^p$. The first of these two alternatives has been used by a number of authors [62,82]. Some benefit is derived in situations where elastic deformations are small; there are, however, some problems associated with this choice. The second alternative is less commonly explicitly chosen, though it is sometimes implicitly imposed by a combination of other assumptions; this is the case in our formulation. Later we choose a natural configuration as the fixed reference configuration for our total Lagrangian formulation. In this natural configuration we choose the reference value of \mathbf{F}^p to be the identity \mathbf{I} at all points in the body. We also use an evolution law for \mathbf{F}^p in which the plastic spin \mathbf{W}^p (defined below) is zero. It is found that these choices always render \mathbf{F}^p symmetric, that is $\mathbf{F}^p = \mathbf{U}^p$. It is not necessary to specify which of these choices is made in the derivation of constitutive laws and in the subsequent formulation of the variational form of the boundary-value

problem, but for these purposes we need only assume that the decomposition has been defined in a unique way.

We consider the displacement \mathbf{u} and the plastic part of the deformation gradient \mathbf{F}^p to be primary variables and rewrite (4.17) so as to define \mathbf{F}^e as a function of \mathbf{u} and \mathbf{F}^p :

$$\mathbf{F}^e = \mathbf{F}\mathbf{F}^{p-1} \quad \text{so} \quad \mathbf{F}^e = (\mathbf{I} + \nabla\mathbf{u})\mathbf{F}^{p-1} = \mathbf{F}^e(\mathbf{u}, \mathbf{F}^p). \quad (4.21)$$

When no plastic deformation has taken place, we have

$$\mathbf{F}^p = \mathbf{I} \quad \text{and} \quad \mathbf{F}^e = \mathbf{F}\mathbf{I}^{-1} = \mathbf{F}.$$

We define elastic and plastic counterparts of the right and left Cauchy–Green deformation tensors by

$$\mathbf{C}^e = \mathbf{F}^{eT}\mathbf{F}^e, \quad (4.22)$$

$$\mathbf{C}^p = \mathbf{F}^{pT}\mathbf{F}^p, \quad (4.23)$$

$$\mathbf{b}^e = \mathbf{F}^e\mathbf{F}^{eT}, \quad (4.24)$$

$$\mathbf{b}^p = \mathbf{F}^p\mathbf{F}^{pT}. \quad (4.25)$$

The tensor \mathbf{C}^p is a referential quantity while the tensors \mathbf{C}^e and \mathbf{b}^p are intermediate quantities and \mathbf{b}^e is a spatial quantity. Various strain measures may be defined using these quantities.

In the early paper of Green and Naghdi [33] there is little discussion of the kinematics of elastic–plastic deformation. It is however postulated that it is always possible to define an additive decomposition of strain into elastic and plastic parts. Many authors have followed Green and Naghdi and have used such strain measures in developing theories of finite–strain plasticity. We shall not use strain measures in our theory; we shall consider deformations measured relative to the intermediate configuration using \mathbf{C}^e as a measure of elastic deformation. It has been shown that an additive decomposition of strain can be made compatible with the multiplicative decomposition of deformation gradient [114] and [53].

The velocity of the particles X of a body during a motion are described by a vector field \mathbf{v} defined over the current configuration of the body:

$$\mathbf{v} = \frac{\partial \chi(X, t)}{\partial t} = \mathbf{v}(\mathbf{x}, t) . \quad (4.26)$$

The *velocity gradient* is denoted \mathbf{L} ; we have the elementary identity

$$\mathbf{L} = \text{grad } \mathbf{v} = \dot{\mathbf{F}}\mathbf{F}^{-1} ; \quad (4.27)$$

note \mathbf{L} is an Eulerian quantity.

The velocity gradient may be decomposed into symmetric and anti-symmetric parts, that is,

$$\mathbf{L} = \mathbf{D} + \mathbf{W} \quad (4.28)$$

where $\mathbf{D} = \frac{1}{2}(\mathbf{L} + \mathbf{L}^T)$ and $\mathbf{W} = \frac{1}{2}(\mathbf{L} - \mathbf{L}^T)$. We refer to the symmetric part \mathbf{D} as the *deformation rate* and the skew-symmetric part \mathbf{W} as the *spin*.

As with the multiplicative decomposition of deformation gradient into elastic and plastic parts, the velocity gradient can be decomposed in a natural way into elastic and plastic parts. This decomposition is obtained from

$$\dot{\mathbf{F}} = \dot{\mathbf{F}}^e \mathbf{F}^p + \mathbf{F}^e \dot{\mathbf{F}}^p . \quad (4.29)$$

Using this in (4.27) we get

$$\mathbf{L} = \mathbf{L}^e + \mathbf{L}^p \quad (4.30)$$

where

$$\mathbf{L}^e = \dot{\mathbf{F}}^e \mathbf{F}^p \mathbf{F}^{-1} \quad \text{and} \quad \mathbf{L}^p = \mathbf{F}^e \dot{\mathbf{F}}^p \mathbf{F}^{-1} . \quad (4.31)$$

Each of these tensors may be decomposed into symmetric and skew-symmetric parts:

$$\mathbf{L}^e = \mathbf{D}^e + \mathbf{W}^e \quad \text{and} \quad \mathbf{L}^p = \mathbf{D}^p + \mathbf{W}^p. \quad (4.32)$$

The symmetric parts \mathbf{D}^e and \mathbf{D}^p are called the elastic and plastic deformation rates while the skew-symmetric parts \mathbf{W}^e and \mathbf{W}^p are elastic and plastic spins respectively.

It is easily shown that

$$\mathbf{D} = \mathbf{D}^e + \mathbf{D}^p \quad \text{and} \quad \mathbf{W} = \mathbf{W}^e + \mathbf{W}^p. \quad (4.33)$$

It can also easily be shown that

$$\dot{\mathbf{C}}^e = 2\mathbf{F}^{eT} \mathbf{D}^e \mathbf{F}^e; \quad (4.34)$$

this relationship is particularly useful later.

4.4 Decomposition into dilational and volume-preserving parts

It is sometimes convenient to consider separately that part of a deformation which involves a change in volume. To do this the decomposition

$$\mathbf{F} = \hat{\mathbf{F}} \bar{\mathbf{F}} \quad (4.35)$$

of the deformation gradient may be used; here $\hat{\mathbf{F}} = J^{\frac{1}{3}} \mathbf{I}$, which is purely a change in volume (recall $J = \det \mathbf{F}$), and $\bar{\mathbf{F}} = J^{-\frac{1}{3}} \mathbf{F}$. We shall refer to $\hat{\mathbf{F}}$ and $\bar{\mathbf{F}}$ as the *dilational* and *volume-preserving* parts of \mathbf{F} respectively. Note $\det \hat{\mathbf{F}} = J$ and $\det \bar{\mathbf{F}} = 1$.

This decomposition results in a natural decomposition of \mathbf{L} into deviatoric and spherical parts. We find that

$$(\det \bar{\mathbf{F}})^{\circ} = \text{tr } \bar{\mathbf{D}} = 0 \quad (4.36)$$

so that

$$\mathbf{L} = \text{dev } \mathbf{L} + \frac{1}{3} J^{-1} \dot{J} \mathbf{I} \quad (4.37)$$

where $\text{dev } \mathbf{L} = \dot{\bar{\mathbf{F}}}\bar{\mathbf{F}}^{-1}$ and the spherical part is equal to $\dot{\hat{\mathbf{F}}}\hat{\mathbf{F}}^{-1}$. Of course such a decomposition of \mathbf{L} could have been obtained directly. The relationship between the above decomposition of deformation gradient and a deviatoric-spherical decomposition of velocity gradient is nonetheless interesting.

The decomposition (4.35) can be applied directly to \mathbf{F}^e and \mathbf{F}^p using the fact that

$$J = J^e J^p \quad (4.38)$$

where $J^e = \det \mathbf{F}^e$ and $J^p = \det \mathbf{F}^p$. We can write

$$\mathbf{L} = \text{dev } \mathbf{L}^e + \text{dev } \mathbf{L}^p + \frac{1}{3}J^{e-1}\dot{J}^e\mathbf{I} + \frac{1}{3}J^{p-1}\dot{J}^p\mathbf{I} \quad (4.39)$$

where

$$\text{dev } \mathbf{L}^e = \dot{\bar{\mathbf{F}}}\bar{\mathbf{F}}^p\bar{\mathbf{F}}^{-1} \quad \text{and} \quad \text{dev } \mathbf{L}^p = \bar{\mathbf{F}}^e\dot{\bar{\mathbf{F}}}^p\bar{\mathbf{F}}^{-1}. \quad (4.40)$$

The relationships described above are of particular interest in situations where constraints involving the volumetric parts of elastic and/or plastic deformations are considered. This approach has been exploited by Simo and co-workers [106,112] in the description of isochoric plastic behaviour. For this behaviour a constraint is placed on the dilational part of plastic deformation, that is, $J^p = 1$ so that $J^e = J$. In the formulations proposed by Simo J^p is eliminated from the outset.

CHAPTER V

CONTINUUM THERMODYNAMICS

5.1 Introduction

The development of theories of plasticity has been greatly influenced by the establishment of general theories of continuum thermodynamics. This is particularly true of the finite-strain theory. In the well-known early paper of Green and Naghdi 1965 [33] the theory proposed is firmly based on the axioms of thermodynamics, and many authors have followed this early example [10,23,50,102]. The central role of thermodynamics is not restricted to problems involving thermal effects, and it is widely used as a basis for describing restrictions on the constitutive equations of isothermal processes. We describe the thermodynamic laws which apply to thermomechanical behaviour and then consider the special case of isothermal dissipative processes.

The use of internal variables to model dissipative internal rearrangements of material structure provides a flexible framework for the development of a general form for the constitutive relations for plasticity. Internal variables were first included in a general theory of thermodynamics by Coleman and Gurtin [18,37]. The account of this theory given by Gurtin [37] is particularly clear and has been referred to extensively in the preparation of this Chapter.

Kestin and Rice [49] introduced the use of internal variables in modelling plastic phenomena. This approach has become progressively more popular in recent years; see for example the work of Martin and co-workers [69,68,72], Carroll [10], Halphen and Nguyen [38] and the recently published text [65]. In spite of the advantages gained by making use of internal variables in the construction of constitutive relations for plasticity there is still some resistance to the idea. This resistance stems from uncertainty concerning the precise meaning of these variables.

We begin by defining the stress measures which will be used in this and subsequent Chapters. We then apply the laws of thermodynamics and obtain the resulting restric-

tions on constitutive relations for plastically deforming media. A brief description is included of the internal variables used in the general theory developed here. We make use of these results in Chapter VI where we discuss further the form taken by evolution equations introduced in this Chapter.

5.2 Stress measures

The Cauchy stress, denoted by $\bar{\mathbf{T}}$, is the fundamental stress measure. It is sometimes referred to as the *true stress* or *engineering stress*. It has a direct relation to the tractive force \mathbf{t} acting on an element of area $d\mathbf{s} = \mathbf{n}ds$ within a body; that is

$$\mathbf{t} = \bar{\mathbf{T}}\mathbf{n} . \quad (5.1)$$

A detailed discussion of Cauchy stress can be found in many introductory texts on continuum mechanics for example [15,86] or [2]. We note that the Cauchy stress is a symmetric Eulerian quantity.

Nanson's formula is used to relate the surface traction acting on an element of area in the current configuration to that which acts on the corresponding area in the reference configuration. This gives rise to a stress measure known as the *first Piola-Kirchhoff* or *non-symmetric Piola-Kirchhoff stress*, denoted by \mathbf{P} and related to $\bar{\mathbf{T}}$ by

$$\mathbf{P} = J\bar{\mathbf{T}}\mathbf{F}^{-T} . \quad (5.2)$$

The transpose of \mathbf{P} is sometimes referred to as the nominal stress.

For convenience we also introduce the *Kirchhoff stress* \mathbf{T} , related to Cauchy stress $\bar{\mathbf{T}}$ by

$$\mathbf{T} = J\bar{\mathbf{T}} . \quad (5.3)$$

We recall that $J = \det\mathbf{F} = \rho_0/\rho$, ρ_0 being the referential mass density.

To motivate the introduction of further stress measures we consider the stress power per unit volume in the reference configuration. The stress power is defined as the scalar

product of stress and deformation rate; the following are equivalent expressions for stress power:

$$(\mathbf{P} \cdot \dot{\mathbf{F}}), \quad (\bar{\mathbf{S}} \cdot \dot{\mathbf{E}}), \quad \frac{1}{2}(\bar{\mathbf{S}} \cdot \dot{\mathbf{C}}) \quad \text{and} \quad (\mathbf{T} \cdot \mathbf{D}). \quad (5.4)$$

The stress measure $\bar{\mathbf{S}}$, conjugate to the Lagrangian strain rate $\dot{\mathbf{E}}$, is known as the *second* or *symmetric Piola–Kirchhoff stress*. The definition of conjugate stress measures is based on the requirement that the stress power be unaltered by the choice of rate quantity used; hence $\bar{\mathbf{S}}$ is obtained from the third and fourth expressions in (5.4), and is given by

$$\bar{\mathbf{S}} = \mathbf{F}^{-1} \mathbf{T} \mathbf{F}^{-T}. \quad (5.5)$$

Conjugate stress analysis is discussed in some detail in [86].

We define a symmetric "Piola–Kirchhoff" stress \mathbf{S} relative to the intermediate configuration by

$$\mathbf{S} = (\mathbf{F}^e)^{-1} \mathbf{T} (\mathbf{F}^e)^{-T}. \quad (5.6)$$

The standard symmetric Piola–Kirchhoff stress is recovered in the absence of plastic deformation, that is, when $\mathbf{F}^p = \mathbf{I}$ and $\mathbf{F}^e = \mathbf{F}$.

We record the following relationships between \mathbf{S} , \mathbf{P} and \mathbf{T} for later use:

$$\mathbf{P} = \mathbf{F}^e \mathbf{S} \mathbf{F}^{p-T}; \quad \mathbf{T} = \mathbf{F}^e \mathbf{S} \mathbf{F}^{eT}. \quad (5.7)$$

5.3 Laws of thermodynamics

The thermodynamics of deformable media has as its basis the equations of balance of linear momentum, balance of angular momentum, balance of energy and an inequality which governs the rate of entropy production. The equation of balance of energy is essentially the first law of thermodynamics and the inequality expresses the second law.

In local form the balance equations are:

$$\text{balance of momentum} \quad \operatorname{div} \mathbf{P} + \rho_0 \mathbf{b} = \rho_0 \mathbf{a} \quad (5.8)$$

$$\text{balance of angular momentum} \quad \mathbf{T} = \mathbf{T}^T \quad (5.9)$$

and

$$\text{balance of energy} \quad \rho_0 \dot{e} = \mathbf{P} \cdot \dot{\mathbf{F}} + \operatorname{div} \mathbf{q} + \rho_0 r \quad (5.10)$$

where ρ_0 is the mass density in the reference configuration, \mathbf{b} is the specific body force, \mathbf{a} is the acceleration vector, e is the specific internal energy density, \mathbf{q} is the heat flux vector and r is the specific external heat supply; the superposed dot denotes material differentiation with respect to time as defined in Chapter IV.

The local form of the second law (the Clausius–Duhem inequality) is

$$\rho_0 \dot{s} \geq -\operatorname{div} \left(\frac{\mathbf{q}}{\theta} \right) + \frac{\rho_0 r}{\theta} \quad (5.11)$$

where s is the specific entropy density and θ is the absolute temperature.

By introducing the free energy density ψ , defined by

$$\psi = e - \theta s \quad (5.12)$$

and using the first definition in (5.4) for the stress power, the first and second laws of thermodynamics may be combined to give the *reduced dissipation inequality*

$$\rho_0 (\dot{\psi} + s\dot{\theta}) - \mathbf{T} \cdot \mathbf{D} + \theta^{-1} \mathbf{q} \cdot \nabla \theta \leq 0. \quad (5.13)$$

5.4 Internal variables

We introduce internal variables which describe effects associated with internal rearrangements of the material such as plastic deformation and hardening behaviour. For the purpose of formulating our internal variable theory of plasticity we consider the

internal variables \mathbf{F}^p , γ and \mathbf{A} . The variable γ is a scalar while \mathbf{A} is a positive definite symmetric second-order tensor with respect to the reference configuration. The variable \mathbf{F}^p has been described in Chapter IV; it is fundamental to the description of plastic deformation. We exclude from consideration internal variables which are vectors since it is known that these would lead to a violation of the principle of material frame indifference, (see [37]). In general there will be an array of both scalar and tensor internal variables; however, for simplicity we include in addition to \mathbf{F}^p only one of each kind in the development of the theory. The extension to arrays of internal variables may be accomplished in a straightforward manner.

We remark that \mathbf{F}^p enters into the description of the material behaviour implicitly through the definition of the elastic part of the deformation of the material. The other internal variables may be functions of \mathbf{F}^p , for example $\mathbf{A} = \mathbf{C}^p$. The choice $\mathbf{A} = \mathbf{E}^p$ is sometimes used where \mathbf{A} characterises kinematic hardening, this choice arises from a direct use of the form which linear kinematic hardening takes in the small-strain case.

Later we restrict attention to the cases of perfect plasticity and isotropic hardening for which we need only consider the internal variables \mathbf{F}^p and γ .

5.5 Constitutive relations

We consider a general class of thermomechanical behaviour which is governed by a set of constitutive equations of the form in which $\{\psi, \mathbf{S}, s, \mathbf{q}\}$ are each functions of $\{\mathbf{F}^e, \mathbf{A}, \gamma, \theta, \nabla\theta\}$, and a set of evolution equations for the internal variables which also depend on these quantities. The evolution equations reflect the dissipative nature of plastic deformation; detailed discussion of evolution laws is postponed to Chapter VI.

This set of constitutive equations must obey the principle of material frame indifference. A change of frame manifests itself as the transformation $\mathbf{F} \rightarrow \mathbf{Q}\mathbf{F}$, where \mathbf{Q} is a time-dependent proper orthogonal transformation as defined in Chapter IV. From the decomposition (4.17) it is clear that this amounts to the transformation $\mathbf{F}^e \rightarrow \mathbf{Q}\mathbf{F}^e$, the component \mathbf{F}^p evidently being left unchanged by the change of frame. Likewise, both γ and \mathbf{A} are left unchanged by the transformation, the first because it is a scalar and

the second because it is a referential quantity. Under these circumstances the principle of material frame indifference requires that for example

$$\psi(\mathbf{F}^e, \gamma, \mathbf{A}, \theta) = \psi(\mathbf{QF}^e, \gamma, \mathbf{A}, \theta) . \quad (5.14)$$

By arguing in the usual way we obtain

$$\psi = \psi(\mathbf{C}^e, \gamma, \mathbf{A}, \theta) , \quad (5.15)$$

and similarly for the remaining functions.

Substitution of these constitutive equations in the reduced dissipation inequality (5.13) yields the result that ψ , \mathbf{S} and s are independent of $\nabla\theta$, and that

$$\mathbf{S} = \rho_0 \frac{\partial \psi}{\partial \mathbf{C}^e} \quad \text{and hence} \quad \mathbf{T} = 2\rho_0 \mathbf{F}^e \frac{\partial \psi}{\partial \mathbf{C}^e} \mathbf{F}^{eT} , \quad (5.16)$$

$$s = -\rho_0 \frac{\partial \psi}{\partial \theta} \quad (5.17)$$

and

$$\rho_0 \frac{\partial \psi}{\partial \gamma} \cdot \dot{\gamma} + \rho_0 \frac{\partial \psi}{\partial \mathbf{A}} \cdot \dot{\mathbf{A}} - \mathbf{T} \cdot \mathbf{D}^p + \frac{1}{\theta} \mathbf{q} \cdot \nabla \theta \leq 0. \quad (5.18)$$

We are concerned with constitutive theories for inelastic materials, and in order to focus on issues which are central to materials of this type we confine attention to *isothermal* behaviour, so that $\nabla\theta = \mathbf{0}$, $\mathbf{q} = \mathbf{0}$, and $r = 0$, and θ is omitted as a variable.

We define thermodynamic forces g and \mathbf{G} conjugate to γ and \mathbf{A} , respectively, by

$$g = -\rho_0 \frac{\partial \psi}{\partial \gamma}, \quad \text{and} \quad \mathbf{G} = -\rho_0 \frac{\partial \psi}{\partial \mathbf{A}} . \quad (5.19)$$

Using these definitions we can rewrite (5.18) as

$$\mathbf{T} \cdot \mathbf{D}^p + g\dot{\gamma} + \mathbf{G} \cdot \dot{\mathbf{A}} \geq 0. \quad (5.20)$$

This form of the dissipation inequality plays a pivotal role in the construction of the evolution equations.

Instead of setting out explicitly the form which these equations take, we remark simply that we will seek a set of equations in which *rates* of change of $\{\mathbf{F}^p, \gamma, \mathbf{A}\}$ are each sought as functions of $\{\mathbf{F}^e, \gamma, \mathbf{A}\}$. The reason for this vagueness is that, rather than impose *a priori* a particular structure, we will deduce an appropriate form for the evolution equation from the inequality (5.20) together with assumptions which embody notions of convexity of particular sets and functions (see Chapter VI).

5.6 Remarks

It is worth noting that in the classical small-strain theory of plasticity the free energy density and hence stress are taken to be functions of the *linearised total strain* $\boldsymbol{\varepsilon}(\mathbf{u})$ and internal variables. It is clear that the plastic strain $\boldsymbol{\varepsilon}^p$ must be included as one of the internal variables and that the total strain only enters through terms involving the elastic strain $\boldsymbol{\varepsilon}^e(\mathbf{u}, \boldsymbol{\varepsilon}^p) = \boldsymbol{\varepsilon}(\mathbf{u}) - \boldsymbol{\varepsilon}^p$. The free energy can therefore be written as a function of $\boldsymbol{\varepsilon}^e$ and internal variables, in which case $\boldsymbol{\varepsilon}^p$ need not appear explicitly. The small-strain theory with total strain as a variable is viable precisely because of the linearity which pervades this theory. In the finite-strain counterpart to this theory, however, use of total strain (through \mathbf{F}) is unduly restrictive in its range of applicability. In the case of finite-strains it proves to be judicious to assume, as we have done, that the free energy density and hence the stress and other thermodynamic conjugate forces depend on the *elastic* part of the deformation (that is, on \mathbf{F}^e) and on the internal variables.

CHAPTER VI

AN INTERNAL VARIABLE THEORY OF PLASTICITY

6.1 Introduction

In this Chapter we develop an internal variable theory of plasticity. We use the thermodynamic framework described in Chapter V and the tools of nonsmooth convex analysis reviewed in Chapter III. Amongst other things this theory makes specific the form taken by the evolution equations introduced in Chapter V. It is constructed in such a way as to include the essential features of classical theories, such as the existence of a convex set of admissible stresses, the principle of maximum plastic work and the normality law. We assume from the outset that these basic assumptions are valid in the finite-strain case. A further assumption, concerning the plastic spin, is required to complete the characterisation of finite-strain plasticity. An assumption of isochoric plastic deformation is also included in the theory.

The theory we discuss here was initially developed for the small-strain case [29]. This small-strain theory is similar in many respects to that proposed by Moreau [75]; however, it includes some important additional features. In this Chapter we essentially extend the theory described in [29] to the finite-strain case. The main task in extending the theory in this way is that of identifying the relevant quantities to use as the components of the arrays of internal variables, internal variable rates and their conjugate forces. The basis for the theory has already been established in the form of the main theorem presented in Chapter III. To complete the theory it remains to assign physical significance to the sets and functions which appear in this theorem and in the other results from convex analysis which we use.

We first identify the quantities which are used in the finite-strain theory; this is done by reconsidering the dissipation inequality. We then state the principle of maximum plastic work and identify the class of evolution laws which we consider. The internal variable theory is then developed. We include discussion of the relationship between

yield and dissipation functions, and show how the classical Kuhn–Tucker form of the constitutive equations can be recovered from the new theory. These Sections are followed by discussion of additional assumptions which we have made. We discuss the application of the theory in Chapter VII.

6.2 The dissipation inequality

It is necessary first of all to decide on the configuration in which to postulate the existence of the appropriate convex sets or functions. The most important set to be considered is the elastic region. The boundary of this region is traditionally described by a yield function. We are guided by the important observation [42,44] that, for metal plasticity at any rate, it is the yield function *as a function of Cauchy stress* which describes a convex elastic region. It is this notion of convexity which we will incorporate into the structure of the evolution equation, though we will postulate convexity with respect to Kirchhoff stress: since the Kirchhoff and Cauchy stresses differ by a scalar, any set or function which is convex with Cauchy stress as the variable, remains convex with Kirchhoff stress as the variable.

In order to ensure consistency in the arguments which follow it is necessary that all quantities appearing in the evolution law be *spatial* in character. Now the relevant quantities are those appearing in the dissipation inequality (5.18). Of these, g and γ are scalars, and hence insensitive to the particular description, while \mathbf{T} and \mathbf{D}^p are spatial quantities. This leaves the tensorial internal variable \mathbf{A} and its conjugate force \mathbf{G} , both of which are referential quantities. The manner in which these two quantities are transformed into spatial quantities will depend on the specific interpretation given them; for example, \mathbf{A} might represent a stress-like or a deformation-like quantity, and its spatial counterpart will differ accordingly. Rather than be unduly prescriptive, we assume that spatial counterparts Λ and Γ of \mathbf{A} and \mathbf{G} , respectively, can be defined.

In the case where \mathbf{A} is a deformation-like quantity we consider the relationships (4.13) and (4.14) from which we conclude that the transformation between spatial and referential quantities would take the form

$$\mathbf{A} = \mathbf{F}^T \boldsymbol{\Lambda} \mathbf{F} \quad \text{and} \quad \mathbf{G} = \mathbf{F}^{-1} \boldsymbol{\Gamma} \mathbf{F}^{-T} . \quad (6.1)$$

This transformation has to be taken into account when relating the rates of \mathbf{A} and $\boldsymbol{\Lambda}$; it is found that

$$\dot{\mathbf{A}} = \mathbf{F}^T \dot{\boldsymbol{\Lambda}} \mathbf{F} + \mathbf{F}^T \mathbf{L}^T \boldsymbol{\Lambda} \mathbf{F} + \mathbf{F}^T \boldsymbol{\Lambda} \mathbf{L} \mathbf{F} \quad (6.2)$$

and hence in this case we define an objective rate of change $\check{\boldsymbol{\Lambda}}$ of $\boldsymbol{\Lambda}$ by

$$\check{\boldsymbol{\Lambda}} = \dot{\boldsymbol{\Lambda}} + \mathbf{L}^T \boldsymbol{\Lambda} + \boldsymbol{\Lambda} \mathbf{L} , \quad (6.3)$$

so that $\dot{\mathbf{A}} = \mathbf{F}^T \check{\boldsymbol{\Lambda}} \mathbf{F}$.

Similarly, in the case where \mathbf{A} is a stress-like quantity we assume that

$$\mathbf{A} = \mathbf{F}^{-1} \boldsymbol{\Lambda} \mathbf{F}^{-T} \quad \text{and} \quad \mathbf{G} = \mathbf{F}^T \boldsymbol{\Gamma} \mathbf{F} , \quad (6.4)$$

so that

$$\dot{\mathbf{A}} = \mathbf{F}^{-1} \dot{\boldsymbol{\Lambda}} \mathbf{F}^{-T} - \mathbf{F}^{-1} \mathbf{L} \boldsymbol{\Lambda} \mathbf{F}^{-T} - \mathbf{F}^{-1} \boldsymbol{\Lambda} \mathbf{L}^T \mathbf{F}^{-T} \quad (6.5)$$

and hence the appropriate objective rate is

$$\check{\boldsymbol{\Lambda}} = \dot{\boldsymbol{\Lambda}} - \mathbf{L} \boldsymbol{\Lambda} - \boldsymbol{\Lambda} \mathbf{L}^T . \quad (6.6)$$

In both of these cases $\check{\boldsymbol{\Lambda}}$ has been defined in such a way that

$$\mathbf{G} \cdot \dot{\mathbf{A}} = \boldsymbol{\Gamma} \cdot \check{\boldsymbol{\Lambda}} ; \quad (6.7)$$

for further discussion of objective rates see Marsden and Hughes [67] or Ogden [86].

With all the relevant spatial quantities at our disposal we can rewrite the dissipation inequality (5.18) in the form

$$\mathbf{T} \cdot \mathbf{D}^p + g\dot{\gamma} + \Gamma \cdot \dot{\check{\Lambda}} \geq 0. \quad (6.8)$$

We now define an array \mathbf{p} , of rate quantities, by

$$\mathbf{p} = (\mathbf{D}^p, \dot{\gamma}, \dot{\check{\Lambda}}) \quad (6.9)$$

and an array \mathbf{Z} , of conjugate forces, by

$$\mathbf{Z} = (\mathbf{T}, g, \Gamma) \quad (6.10)$$

so that the dissipation inequality (6.8) can be expressed in the compact form

$$\mathbf{Z} \cdot \mathbf{p} \geq 0. \quad (6.11)$$

An interesting alternative is discussed in [74], where the evolution law is constructed using quantities defined with respect to the *intermediate* configuration, and where in particular a convex yield function is defined with respect to such quantities. In the case of metals, for which elastic strains are small, there would presumably be little difference between evolution laws which differ only in their use of intermediate rather than spatial quantities as primary variables. We have chosen to make use of spatial variables, but should there be compelling physical evidence for the suitability of intermediate variables, the theory presented here can be modified with little effort, by replacing conjugate pairs of spatial variables with corresponding conjugate intermediate variables.

6.3 The principle of maximum plastic work

In its most general form the principle of maximum plastic work states that plastic deformation takes place in such a manner that the work done by forces conjugate to the variables which describe the plastic deformation is a maximum. This principle may be characterised by the condition

$$(\mathbf{Z} - \bar{\mathbf{Z}}) \cdot \mathbf{p} \geq 0 \quad \forall \text{ admissible } \bar{\mathbf{Z}}. \quad (6.12)$$

The actual conjugate force \mathbf{Z} is the one which maximises the lefthand side of this inequality.

The principle of maximum plastic work is also referred to as the maximum dissipation postulate.

6.4 Evolution laws

In the thermodynamic theory, reviewed in Chapter V, the general form of the evolution laws for the internal variables \mathbf{F}^p , γ and \mathbf{A} is

$$\{\dot{\mathbf{F}}^p, \dot{\gamma}, \dot{\mathbf{A}}\} = F(\mathbf{F}, \mathbf{F}^p, \gamma, \mathbf{A}). \quad (6.13)$$

We restrict attention to a special class of such evolution laws where the evolution of the rate quantities appearing in the array \mathbf{p} depend on $(\mathbf{F}, \mathbf{F}^p, \gamma, \mathbf{A})$ through the conjugate forces \mathbf{Z} . Evolution laws of this type take the simple form

$$\mathbf{p} = F(\mathbf{Z}) \quad (6.14)$$

where $\mathbf{Z} = \tilde{\mathbf{Z}}(\mathbf{F}^e, \gamma, \mathbf{A})$. We prefer, however, to work with the dual form of (6.14), which is

$$\mathbf{Z} \in G(\mathbf{p}). \quad (6.15)$$

6.5 The internal variable theory

We now develop the internal variable theory using the results from convex analysis reviewed in Chapter III. We use as the general setting in which to present these results, an unspecified finite dimensional space E and its dual E^* with elements \mathbf{x} and \mathbf{x}^* respectively. In order to apply the results we choose either E and E^* or E^* and E to be spaces of conjugate forces and rate quantities respectively. The arrays \mathbf{Z} and \mathbf{p} then replace elements \mathbf{x} and \mathbf{x}^* or \mathbf{x}^* and \mathbf{x} . To apply Theorem 3.1 we make the second of these two choices; a change of notation which reflects this choice has already been made in the presentation of the theorem in Chapter III.

Much of the information which is vital to the development of the new theory of plasticity is contained in Lemma 3.3; recall that this lemma provides the basis for the proof of Theorem 3.1. We then recognise that the function \mathcal{D} , introduced in the lemma, is what is known in classical plasticity as *the dissipation function*. This connection is confirmed by the the last part of the theorem where \mathcal{D} is shown to be equal to $\mathbf{Z} \cdot \mathbf{p}$. The assumption of non-negativity in the definition of a gauge such as \mathcal{D} reflects the second law of thermodynamics as expressed by the dissipation inequality (6.11). Thus $\mathcal{D}(\mathbf{p})$ gives the rate of plastic (dissipative) work.

The set K introduced in Lemma 3.3 is easily recognised as the convex elastic region used in classical plasticity. This connection is made by noting that it is only those \mathbf{Z} which lie on the boundary of K which have associated with them non-zero values of \mathbf{p} . The boundary $\text{bd } K$ of K is known as the *yield surface*. We discuss further relationships between $\text{bd } K$ and classical yield functions in the next section of this Chapter.

The fact that \mathcal{D} is the support function of K (part (a) of Lemma 3.3.) is used to establish the dual relationship between the dissipation and yield functions which have been found by other authors such as Halphen and Nguyen [38], Hill [40], Martin [72] and Moreau [75]. Of interest is the consequent relationship between the subdifferential of the dissipation function and the normal cone to the elastic region K (part (d) of Lemma 3.3.). We have

$$\mathbf{Z} \in \text{bd } K \Rightarrow \mathbf{p} \in N_K(\mathbf{Z}) \neq \{\mathbf{0}\} \quad \text{and} \quad \mathbf{Z} \in \text{int } K \Rightarrow N_K(\mathbf{Z}) = \{\mathbf{0}\} \quad (6.16)$$

so that

$$\mathbf{p} = \mathbf{0} \quad \text{for} \quad \mathbf{Z} \in \text{int } K, \quad (6.17)$$

which is the normality rule.

We now look at the application of Theorem 3.1. We begin by considering the connection with the classical maximum plastic work inequality (6.12). This connection is made by considering the notion of a *maximal responsive map* as defined in Chapter III. The

principle of maximum plastic work corresponds precisely to the existence of such a map. The theorem provides the result that a maximal responsive map G may be written as

$$\mathbf{Z} \in \partial\mathcal{D}(\mathbf{p}), \quad (6.18)$$

so that $G = \partial\mathcal{D}$, which identifies \mathcal{D} as a *pseudopotential* for \mathbf{Z} (see Moreau [75]). What is interesting here, is the equivalence of properties (a) and (b) in the theorem: maximal responsiveness of G is *necessary* and *sufficient* for the existence of \mathcal{D} with the requisite properties.

Since \mathcal{D}^{**} , the dual of the indicator function \mathcal{D}^* of K , is equal to \mathcal{D} , we have embodied in Theorem 3.1 and Lemma 3.3 *three equivalent formulations* of the evolution equation, summarised below:

$$\begin{array}{l} G \text{ maximal responsive} \\ \mathbf{Z} \in G(\mathbf{p}) \end{array} \quad (I)$$

$$\Updownarrow$$

$$\begin{array}{l} \mathcal{D} \text{ convex, pos.hom., lsc, } \mathcal{D}(\mathbf{p}) \geq 0, \mathcal{D}(\mathbf{0}) = 0 \\ \mathbf{Z} \in \partial\mathcal{D}(\mathbf{p}) \end{array} \quad (II)$$

$$\Updownarrow$$

$$\begin{array}{l} K \text{ closed, convex, contains zero} \\ \mathcal{D}^* = \text{indicator function of } K \\ \mathbf{p} \in \partial\mathcal{D}^*(\mathbf{Z}). \end{array} \quad (III)$$

The small-strain equivalent of formulation (III) is well known and goes back to Moreau; in the finite-strain case equivalent forms have been proposed by Halphen and Nguyen [38] and Kim and Oden [51]. Formulation (II) is sometimes mentioned as a consequence of (III) but is seldom given prominence in its own right. An exception is the work of Martin [68,72] (in small-strain plasticity), in which the evolution equation is assumed to be of the form (II). Formulation (I), is new, though there is some connection with the work of Rice [102] and of Hill [40,41,42], who regard the maximum plastic work inequality as a fundamental property of a plastic material.

These three formulations show clearly the *minimal* assumptions which need to be made if an acceptable classical theory of plasticity is to emerge. In particular, we see that (I) and (II) do not require the assumption of an elastic region and yield surface: these are *consequences*. Practical considerations would dictate which of these formulations would be most appropriate for the problem at hand. For example, (III) is most often used, in one guise or another; (II) has been used in [72]; (I) may have limitations in that it is not simple or natural to formulate equations of this form, except perhaps for one-dimensional problems. The major benefit of (I), though, is that it resolves the issue of how much information needs to be added to the assumption of the principle of maximum plastic work, in order that it form the basis of an internal variable theory of plasticity. The results of experimental work designed to determine the physical properties of plastic materials may also be better interpreted using (I) rather than (II) or (III); the latter could of course then be retrieved using the correspondence described above.

6.6 The relationship between yield and dissipation functions

It is common in engineering practice to identify a region of admissible or elastic stresses $K_{\mathbf{T}}$ in stress space (or deviatoric stress space). This is usually defined to be the zero level set of a given function $\varphi(\mathbf{T}, \mathbf{A}, \gamma)$ say, where the particular form of φ depends on spatial forms of the internal variables \mathbf{A} and γ and is called the *yield function*, thus

$$K_{\mathbf{T}} = \{\mathbf{T} : \varphi(\mathbf{T}; \mathbf{A}, \gamma) \leq 0\}. \quad (6.19)$$

With an appropriate choice of free energy function ψ and transformation of the type described in (6.1), we can redefine φ as a function of the conjugate forces \mathbf{Z} of the form

$$\tilde{\varphi}(\mathbf{Z}) = f(\mathbf{Z}) - 1 \quad (6.20)$$

and hence define the set K in the space of conjugate forces by

$$K = \{\mathbf{Z} : f(\mathbf{Z}) \leq 1\}, \quad (6.21)$$

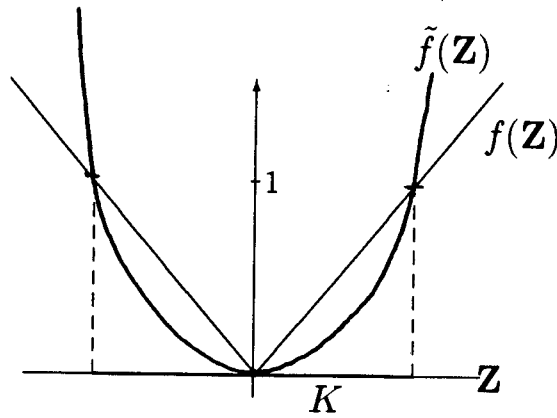


Figure 6.1: Yield functions in one-dimension with canonical yield function f

as in Lemma 3.3. The set $K_{\mathcal{T}}$ is simply the projection of K in stress space. In Chapter VII the relationship between φ and f is illustrated by considering the example of the von Mises yield condition.

By a slight abuse of terminology we refer to f as the yield function. The function f is not in general unique, as can be seen in Figure 6.1 where for a simple one-dimensional example two possible functions, f and \tilde{f} are shown. We make a particular choice of f which is convenient from the point of view of interpreting this yield function as the polar conjugate of the dissipation function \mathcal{D} . We call this choice of f the *canonical yield function*.

Using the results of convex analysis reviewed in Chapter III we show that given an elastic region $K \subset E^*$ which is closed, convex and non-empty it is always possible to construct a yield function f which is a *gauge*. Thus, $\text{epi } f$ is a closed convex cone containing the origin. We now establish the dual relationships between f and \mathcal{D} .

So far we know that \mathcal{D} and the indicator function \mathcal{D}^* of K are *conjugate*, in the sense that

$$\mathbf{Y} \cdot \mathbf{q} \leq \mathcal{D}^*(\mathbf{Y}) + \mathcal{D}(\mathbf{q}) \quad \forall \mathbf{Y} \in E^*, \mathbf{q} \in E \quad (6.22)$$

with equality when \mathbf{Y} is related to \mathbf{q} through

$$\mathcal{D}^*(\mathbf{Y}) = \sup_{\mathbf{r} \in E} \{\mathbf{Y} \cdot \mathbf{r} - \mathcal{D}(\mathbf{r})\} = \mathbf{Y} \cdot \mathbf{q} - \mathcal{D}(\mathbf{q}) \quad (6.23)$$

where \mathbf{q} is the maximising member.

Given the closed convex set K defined by (6.21), we choose the function f to be the *canonical yield function* defined by

$$f(\mathbf{Z}) = \inf \{\mu > 0 : \mathbf{Z} \in \mu K\} \quad (6.24)$$

where $\mu K = \{\mu \mathbf{Y} : \mathbf{Y} \in K\}$. From (3.21) it is evident that f is a *non-negative, positively homogeneous convex* function whose level set at 1 is equal to K , that is, f is a *gauge*. For the simple example shown in Figure 6.1 f clearly has these properties (compare with Figures 3.5 and 3.7). Note that $f(\mathbf{Z})$ can take the value $+\infty$.

With this choice we find from Theorem 15.1 in [103] that f and \mathcal{D} are also dual in the sense that they are *polars* of each other.

From the definitions (3.33), (6.24) of K and of f , an alternative form for f is

$$f(\mathbf{Z}) = \inf \{\mu > 0 : \mathbf{Z} \cdot \mathbf{q} \leq \mu \mathcal{D}(\mathbf{q}) \quad \forall \mathbf{q}\}. \quad (6.25)$$

We have from the definition of a gauge such as \mathcal{D} that $\mathcal{D}(\mathbf{q}) = 0$ if $\mathbf{q} = \mathbf{0}$, we assume that if $\mathcal{D}(\mathbf{q}) = 0$ then $\mathbf{q} = \mathbf{0}$, and so f and \mathcal{D} are related by

$$f(\mathbf{Z}) = \sup_{\substack{\mathbf{q} \in \text{dom } \mathcal{D} \\ \mathbf{q} \neq \mathbf{0}}} \frac{\mathbf{Z} \cdot \mathbf{q}}{\mathcal{D}(\mathbf{q})}; \quad (6.26)$$

in other words, we have the inequality

$$\mathbf{Z} \cdot \mathbf{q} \leq f(\mathbf{Z}) \mathcal{D}(\mathbf{q}) \quad \forall \mathbf{Z} \in E, \mathbf{q} \in \text{dom } \mathcal{D} \quad (6.27)$$

(we adopt the convention $0 \cdot \infty = \infty \cdot 0 = 0$, thus avoiding ambiguities in the event that the right-hand side of (6.27) takes these forms). Consider $\mathbf{Z} \in \text{bd } K$: then

$$\sup_{\substack{\mathbf{q} \in \text{dom } \mathcal{D} \\ \mathbf{q} \neq \mathbf{0}}} \frac{\mathbf{Z} \cdot \mathbf{q}}{\mathcal{D}(\mathbf{q})} \leq 1 \quad (6.28)$$

and the supremum is achieved when $\mathbf{q} = \mathbf{p}$, say, and \mathbf{p} is conjugate to \mathbf{Z} in the sense of Lemma 3.3(d). Thus, for $\mathbf{Z} \in K$ and $\mathbf{p} \in \partial f(\mathbf{Z})$, $\mathbf{p} \neq \mathbf{0}$,

$$\mathbf{Z} \cdot \mathbf{p} = f(\mathbf{Z})\mathcal{D}(\mathbf{p}). \quad (6.29)$$

Thus, whereas \mathcal{D} and \mathcal{D}^* are *conjugate* in the sense of (6.22), (6.23), \mathcal{D} and f are *polar* in the sense of (6.28), (6.29). Furthermore, just as $\mathcal{D}^{**} = \mathcal{D}$, it can be shown [103] that if $f = \mathcal{D}^\circ$, then $\mathcal{D}^{\circ\circ} = \mathcal{D}$ (for this we need the lower semicontinuity of \mathcal{D}).

The introduction of the canonical yield function f also allows us to express the normality rule (6.16) in another form. For this we need Lemma 3.2. Various results follow from this lemma. First, by setting the function $g = f$ and $\mathbf{x}^* = \mathbf{X}$ where $f(\mathbf{Z}) = 1$, so that the set C is identified with K , it is found that

$$\mathbf{p} \in \lambda \partial f(\mathbf{Z}). \quad (6.30)$$

If f is differentiable at \mathbf{Z} we recover the normality rule in its classical form

$$\mathbf{p} = \lambda \frac{\partial f}{\partial \mathbf{Z}}, \quad \lambda \geq 0. \quad (6.31)$$

It is also possible to characterise λ by using the properties of f . Indeed, from (6.30) we have

$$\mathbf{p} = \lambda \mathbf{p}' \quad \text{where } \mathbf{p}' \in \partial f(\mathbf{Z}), \quad (6.32)$$

that is,

$$f(\mathbf{Y}) - f(\mathbf{Z}) - \mathbf{p}' \cdot (\mathbf{Y} - \mathbf{Z}) \geq 0 \quad \forall \mathbf{Y} \in E^*. \quad (6.33)$$

Since $f(\mathbf{Z}) = 1$ and $f(\mathbf{0}) = 0$, we have $\mathbf{Z} \cdot \mathbf{p}' \geq 1$. Furthermore, setting $\mathbf{Y} = 2\mathbf{Z}$, since $f(2\mathbf{Z}) = 2f(\mathbf{Z})$ we obtain $\mathbf{Z} \cdot \mathbf{p}' \leq 1$. In other words,

$$\mathbf{Z} \cdot \mathbf{p}' = 1 \quad (6.34)$$

or

$$\lambda = \lambda \mathbf{Z} \cdot \mathbf{p}' = \mathbf{Z} \cdot \mathbf{p} = \mathcal{D}(\mathbf{p}) . \quad (6.35)$$

Thus the scalar multiplier λ associated with \mathbf{p} has the same value as $\mathcal{D}(\mathbf{p})$.

Lemma 3.2 may also be applied to the dissipation function; setting $g = \mathcal{D}$ and defining

$$C = \{\mathbf{q} : \mathcal{D}(\mathbf{q}) \leq \mathcal{D}(\mathbf{p})\} \quad (6.36)$$

for given $\mathbf{p} \neq \mathbf{0}$ we have immediately, for \mathbf{Z} related to \mathbf{p} through (6.18),

$$\mathbf{Z} \in \partial \mathcal{D}(\mathbf{p}) \Rightarrow \mathbf{Z} \in N_C(\mathbf{p}) . \quad (6.37)$$

There is also the result

$$\mathbf{Z} \in N_C(\mathbf{p}) \Rightarrow \mathbf{Z} \in \lambda \partial \mathcal{D}(\mathbf{p}) \quad \text{for some } \lambda > 0 \quad (6.38)$$

(we exclude the possibility $\lambda = 0$ since $\mathbf{Z} \neq \mathbf{0}$) or, using the positive homogeneity of \mathcal{D} ,

$$\mathbf{Y} \in N_C(\mathbf{p}) \Rightarrow \mathbf{Y} \in \partial \mathcal{D}(\lambda \mathbf{p}) . \quad (6.39)$$

This situation is illustrated in Figure 6.2: in E^* the conjugate pair (\mathbf{Z}, \mathbf{p}) is such that \mathbf{p} lies in the normal cone to K (the level set $f(\mathbf{Y}) \leq 1$) at \mathbf{Z} , while in E we find that \mathbf{Z} lies in the normal cone to C (the level set $\mathcal{D}(\mathbf{q}) \leq \mathcal{D}(\mathbf{p})$) at \mathbf{p} .

6.7 Constitutive equations in Kuhn–Tucker form

We conclude this section by recovering the plastic constitutive equations in their conventional classical form : to do this it is necessary to add a further axiom, the *consistency condition*, which states that at $\mathbf{Z} \in \text{bd } K$,

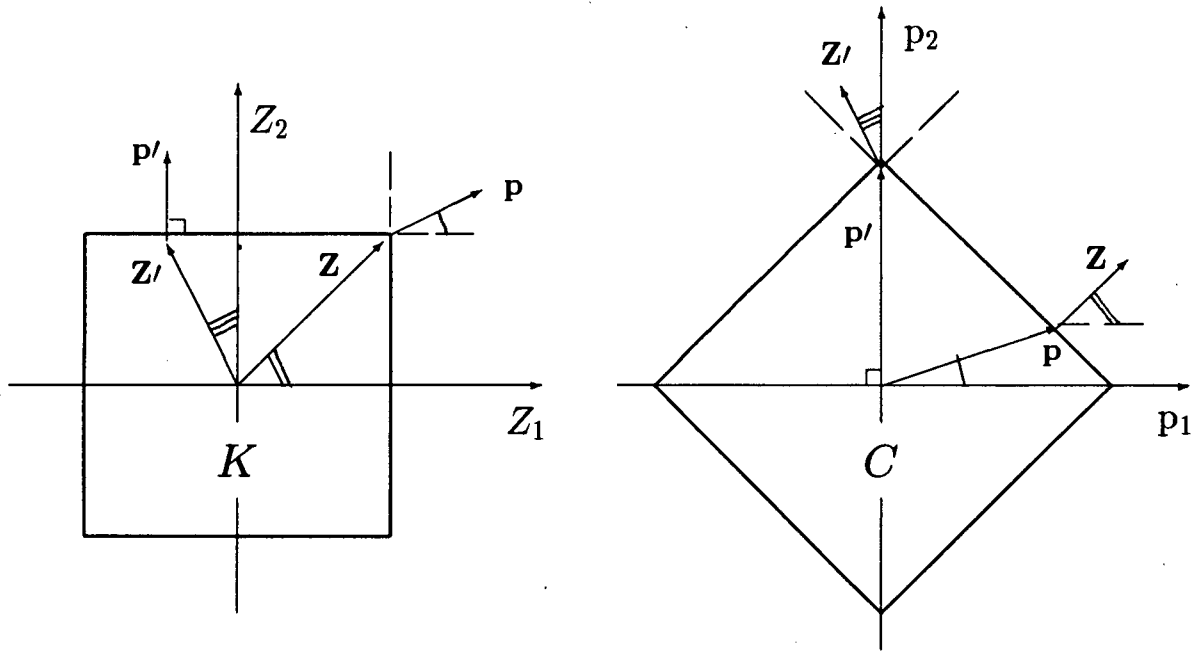


Figure 6.2: Relationship between K and C , a level set of \mathcal{D}

$$\lambda \dot{f} = 0 \tag{6.40}$$

where \dot{f} is the change in f associated with the change in \mathbf{Z} accompanying \mathbf{p} . Thus either

$$\lambda > 0, \quad \dot{f} = 0 \quad \text{or} \quad \dot{f} < 0, \quad \lambda = 0 \tag{6.41}$$

($\dot{f} > 0$ is not permitted since we require $\mathbf{Z} \in K$ always). Then one form of the complete set of equations prescribing the evolution of the internal variables is

$$\mathbf{p} \in \lambda \partial f(\mathbf{Z}) \tag{6.42}$$

where

$$\lambda \geq 0 \quad \text{if} \quad f(\mathbf{Z}) = 1, \quad \dot{f} = 0 \tag{6.43}$$

and

$$\lambda = 0 \quad \left\{ \begin{array}{l} \text{if } f(\mathbf{Z}) = 1 \text{ and } \dot{f} < 0 \\ \text{or } f(\mathbf{Z}) < 1 . \end{array} \right. \quad (6.44)$$

Furthermore $\lambda = \mathcal{D}(\mathbf{p})$.

6.8 Further constitutive assumptions

6.8.1 Plastic spin

The evolution laws which we have defined above are incomplete; we must add information regarding the *skew* part of \mathbf{L}^p in order to have complete information about the evolution of \mathbf{F}^p . Here various options are available, the choice depending to some extent on the particular kind of elastic-plastic material being modelled. We describe two possibilities.

One appropriate form of the evolution law for the skew part of \mathbf{L}^p is obtained simply by assuming that there is no plastic spin:

$$\mathbf{W}^p = \mathbf{0} . \quad (6.45)$$

The physical appropriateness and implications of this assumption are discussed by Moran, Ortiz and Shih [74], and Lubliner [64], for example. In this case

$$\mathbf{D}^p = \mathbf{L}^p = \mathbf{F}^e \dot{\mathbf{F}}^p \mathbf{F}^{-1} , \quad (6.46)$$

whence one may find $\dot{\mathbf{F}}^p$, given \mathbf{D}^p , \mathbf{F} and \mathbf{F}^p . When written out in component form, relative to any basis, this expression represents a set of nine equations for the nine components of $\dot{\mathbf{F}}^p$.

A second option is to assume that $\mathbf{F}^p = \mathbf{U}^p$ where \mathbf{U}^p is the positive-definite symmetric part of the polar decomposition of \mathbf{F}^p , so that

$$\mathbf{D}^p = \text{sym}(\mathbf{F} \mathbf{U}^{p-1} \dot{\mathbf{U}}^p \mathbf{U}^{p-2} \mathbf{F}^{-1}) \quad (6.47)$$

and once again it is possible to determine $\dot{\mathbf{F}}^p = \dot{\mathbf{U}}^p$ from a knowledge of \mathbf{D}^p , \mathbf{F} and \mathbf{F}^p . When written out in component form, this expression represents a set of six equations for the six components of $\dot{\mathbf{U}}^p$.

6.8.2 Isochoric plastic deformation

Another assumption which is commonly made is that there is no volume change associated with plastic deformation, although there seems to be experimental evidence to the contrary for many metals [6,7]. This assumption may be expressed in the form

$$J^p = 1 \quad (6.48)$$

or alternatively as

$$\dot{J}^p = J^p \text{tr} \mathbf{D}^p = 0, \quad (6.49)$$

which, since $J^p > 0$, implies that

$$\text{tr} \mathbf{D}^p = 0. \quad (6.50)$$

This assumption is often included as part of the kinematic description of elastoplastic behaviour; this is unduly restrictive in view of the possible need to accommodate plastic volume changes in a general theory. Thus it is more correctly incorporated as a constitutive assumption.

As has been shown in Chapter IV, \mathbf{L}^p and hence \mathbf{D}^p can be decomposed into deviatoric and spherical parts that are related to a decomposition of \mathbf{F}^p into dilational and volume-preserving parts (see [106] and [112]). When substituted into the dissipation inequality (6.8), this gives

$$\text{dev} \mathbf{T} \cdot \text{dev} \mathbf{D}^p + (\text{tr} \mathbf{T})(\text{tr} \mathbf{D}^p) + g\dot{\gamma} + \Gamma \cdot \dot{\mathbf{\Lambda}} \geq 0. \quad (6.51)$$

Suitable adjustments to the definition of \mathbf{Z} and \mathbf{p} can then be made and the general theory which we develop can be used to construct the required evolution laws for $\bar{\mathbf{F}}^p$,

J^p and the other internal variables. We note that in many classical treatments \mathbf{T} is simply replaced by $\text{dev } \mathbf{T}$ in (5.13). Since little is known about the form which such an evolution for J^p law should take we use our original definition of \mathbf{Z} and \mathbf{p} and impose the plastic volume constraint explicitly; this we discuss in more detail in Chapter VIII.

CHAPTER VII

APPLICATION OF THE INTERNAL VARIABLE THEORY

7.1 Introduction

The internal variable theory described in Chapters V and VI has been applied to the modelling of a variety of phenomena other than classical plasticity in [29]. In this thesis our main concern is the modelling of finite-strain plasticity so here we restrict discussion to this application. The objective of this Chapter is to introduce a simple but typical model of elastoplastic behaviour which is based on the use of internal variables; we use this example extensively in later Chapters. We consider a hyperelastic constitutive model for the elastic behaviour, and plastic behaviour which is governed by the von Mises yield criterion. We discuss the form taken by both the elastic and plastic components of the free energy function. In this Chapter we consider both kinematic and nonlinear isotropic hardening; however, later we shall restrict attention to perfect plasticity and isotropic hardening.

7.2 The free energy function

We recall from Chapter V that the free energy function is a function of \mathbf{F}^e or of \mathbf{C}^e , and of the internal variables. The plastic deformation gradient \mathbf{F}^p does not appear explicitly as one of these variables, but enters into the problem through the definition of \mathbf{F}^e . Kinematic hardening is accounted for by the introduction of a tensorial internal variable \mathbf{A} which is assumed symmetric, and isotropic hardening is characterised by a scalar variable γ .

We consider a free energy function in the separable form

$$\psi(\mathbf{F}^e, \mathbf{A}, \gamma) = \psi_0(\mathbf{F}^e) + \psi_1(\mathbf{A}) + \psi_2(\gamma) \quad (7.1)$$

where the function ψ_0 represents the contribution of elastic strain energy and the functions ψ_1 and ψ_2 are plastic potentials for kinematic and isotropic hardening respectively.

Separation of the free energy into elastic and plastic parts in this way has been employed by many authors (see for example [10,23] and [114]). The validity of decomposing the plastic potential into parts which describe separately the kinematic and isotropic hardening is perhaps questionable in the case of fully nonlinear combined kinematic and isotropic hardening (see for example [77]); we shall assume, however, that is valid for our purposes and proceed to make suitable choices for ψ_0 , ψ_1 and ψ_2 .

7.2.1 The elastic strain energy

In seeking an appropriate function ψ_0 we are guided by the notion that when the material is elastic, that is when $\mathbf{F}^e = \mathbf{F}$, the corresponding boundary value problem should be well-posed, in that at least one solution exists in a suitable Sobolev space. While there is as yet no existence theory for elastic-plastic problems with finite strains it is reasonable to assume that such a theory, based on a free energy of the form (7.1), would require the function ψ_0 to have at least the properties which the elastic theory demands, for existence of solutions. Realistic sufficient conditions for this to be the case in elasticity were first given by Ball [5] in respect of the strain energy function (see Ciarlet [16] for an account of this theory). In the present context these conditions amount to the following: it is required that

- (a) ψ_0 be polyconvex, that is, ψ_0 be expressible as a convex function of \mathbf{F}^e , $\text{cof } \mathbf{F}^e$ and $\det \mathbf{F}^e$;
- (b) $\psi_0(\mathbf{F}^e) \rightarrow +\infty$ as $\det \mathbf{F}^e \rightarrow 0^+$;
- (c) ψ_0 be coercive, in the sense that there exist constants α, β, p, q, r such that

$$\psi_0(\mathbf{F}^e) \geq \alpha (\|\mathbf{F}^e\|^p + \|\text{cof } \mathbf{F}^e\|^q + (\det \mathbf{F}^e)^r) + \beta.$$

for all invertible \mathbf{F}^e , where $\alpha > 0$, $p \geq 2$, $q \geq \frac{p}{p-1}$, $r > 1$.

We use an elastic strain energy function of the type proposed by Ciarlet and Geymonat [17] (see also [16]) which possesses all of the above properties and which has the further merit that it is easily adjustable to experimental results; this function is of the form

$$\psi_0(\mathbf{F}^e) = a\|\mathbf{F}^e\|^2 + b\|\text{cof } \mathbf{F}^e\|^2 + l(\det \mathbf{F}^e) + e \quad (7.2)$$

where $a > 0$, $b > 0$, and $l(\delta) = c\delta^2 - d\log \delta$, with $c > 0$, $d > 0$, and $e \in \mathbb{R}$. Further restrictions on these material constants can be formulated in terms of the Lamé constants μ and λ evaluated for small deformations; these restrictions are described in detail in [16].

A compressible Neo-Hookean model of the elastic material behaviour is obtained by choosing $b \approx 0$ so that the second term may be dropped. With this assumption the material constants a , c , and d can be fixed in terms of μ and λ as

$$a = \frac{\mu}{2}, \quad c = \frac{\lambda}{4} \quad \text{and} \quad d = \frac{\lambda}{2} - \mu; \quad (7.3)$$

e is then chosen so that $\psi_0 = 0$ when $\mathbf{F}^e = \mathbf{I}$.

Simo and co-workers [106,112,115] propose that for a compressible Neo-Hookean material model an appropriate form of ψ_0 is

$$\psi_0(\bar{\mathbf{g}}, \mathbf{b}^e, \mathbf{F}) = \frac{1}{2}\mu(\bar{I}_{\mathbf{b}^e} - 3) + U(J) \quad (7.4)$$

where \mathbf{g} is a spatial metric $\bar{\mathbf{g}} = J^{-\frac{2}{3}}\mathbf{g}$ is the volume preserving part of this metric, $\mathbf{b}^e = \mathbf{F}^e\mathbf{F}^{eT}$, $\bar{I}_{\mathbf{b}^e} = \mathbf{b}^e \cdot \bar{\mathbf{g}}$ and $U(J) = \frac{1}{2}K(\log J)^2$ (see [106]), or $U(J) = \frac{1}{2}K(J^2 - 1)$ (see [112]), μ being the initial shear modulus and K the initial bulk modulus. This function satisfies the conditions for the existence of a solution described above. We note that in constructing this function a dilational/volume-preserving decomposition of deformations is used; this allows J^p to be eliminated as a variable when an assumption of isochoric plastic behaviour is made. In our treatment of the problem we impose the assumption of no plastic volume change explicitly so that, where we use this form of the elastic free energy potential, J is replaced by J^e and we rewrite ψ_0 as

$$\psi_0(\mathbf{F}^e) = \frac{1}{2}\mu J^{e-\frac{2}{3}}\|\mathbf{F}^e\|^2 + U(J^e) + e \quad (7.5)$$

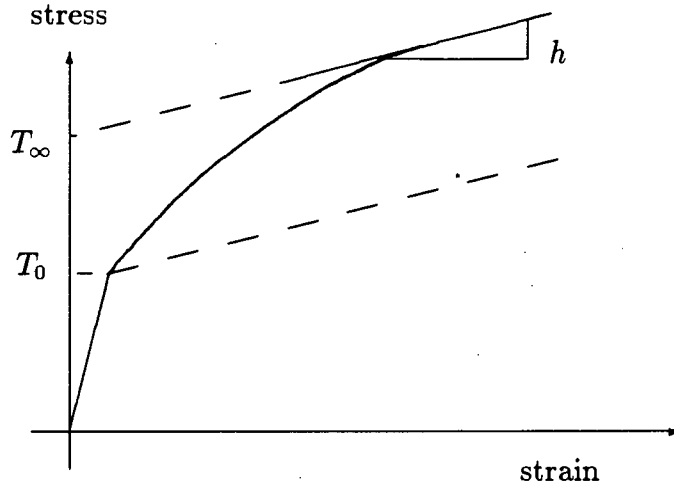


Figure 7.1: Nonlinear isotropic hardening

7.2.2 The plastic potentials

We first describe the isotropic hardening potential. We let the internal variable γ represent the accumulated plastic work or some scalar measure of accumulated plastic strain, and to accommodate nonlinear isotropic hardening consider an expression for ψ_2 of the form

$$\psi_2(\gamma) = \frac{1}{2}c_1\gamma^2 + (c_2 - c_0)(\gamma + \beta^{-1}e^{-\beta\gamma}) \quad (7.6)$$

where c_0, c_1, c_2 and β are material constants with $c_2 > c_0$. The force g conjugate to γ is then given by

$$g(\gamma) = -\psi_2'(\gamma) = -c_1\gamma + (c_0 - c_2)(1 - e^{-\beta\gamma}). \quad (7.7)$$

The physical meaning of the various material constants can be interpreted from the results of a uniaxial tension test, as shown in Figure 7.1. The constant c_1 is $\sqrt{2/3}$ times the initial slope of the hardening law while c_0 and c_2 are $\sqrt{2/3}$ times the initial and limiting yield stress, T_0 and T_∞ , respectively; β is the saturation exponent. Similar nonlinear isotropic hardening laws have been used in [50,108,114] and [106].

We shall not describe the kinematic hardening potential $\psi_1(\mathbf{A})$ in any detail. The conjugate force $\mathbf{\Gamma}$ is assumed to be a stress-like quantity (*the back stress*). We recall from Chapter V that

$$\mathbf{G} = -\rho_0 \frac{\partial \psi_1}{\partial \mathbf{A}}$$

and that spatial counterparts $\mathbf{\Lambda}$ and $\mathbf{\Gamma}$ of \mathbf{G} and \mathbf{A} are defined by

$$\mathbf{\Gamma} = \mathbf{F}\mathbf{G}\mathbf{F}^T \quad \text{and} \quad \mathbf{\Lambda} = \mathbf{F}^{-T}\mathbf{A}\mathbf{F}^{-1}.$$

For consistency with the form that the flow law takes it is necessary to assume that the objective time rate $\dot{\mathbf{\Lambda}}$ of $\mathbf{\Lambda}$, like \mathbf{D}^p , is trace-free. Then the dissipation inequality (5.13) can be written in the form

$$\text{dev } \mathbf{T} \cdot \mathbf{D}^p + \text{dev } \mathbf{\Gamma} \cdot \dot{\mathbf{\Lambda}} + g\dot{\gamma} \geq 0, \quad (7.8)$$

with the result that the ordered triples \mathbf{p} and \mathbf{Z} take the form

$$\mathbf{p} = (\mathbf{D}^p, \dot{\mathbf{\Lambda}}, \dot{\gamma}) \quad \text{and} \quad \mathbf{Z} = (\text{dev } \mathbf{T}, \text{dev } \mathbf{\Gamma}, g). \quad (7.9)$$

7.3 Flow laws based on the von Mises yield condition

According to our theory of plasticity we need only specify one of the properties H , \mathcal{D} or K of the plastic behaviour in order to construct a flow law for the internal variables. We construct our flow law in terms of the dissipation function \mathcal{D} ; however, we begin with a classical description of plastic behaviour where an elastic region K is specified by way of a yield condition φ . We first find the equivalent canonical yield function of the form $f(\mathbf{Z})$, as described in Chapter VI. We require this form of the yield function in order to apply our theory.

We choose for simplicity a von Mises yield condition, which has been altered so as to accommodate kinematic and isotropic hardening. For this condition the yield function φ is given by

$$\varphi(\mathbf{T}, \mathbf{\Gamma}, \gamma) = |\text{dev } \mathbf{T} + \text{dev } \mathbf{\Gamma}| - (c_0 - g(\gamma)) \leq 0 . \quad (7.10)$$

We rewrite this in the form

$$\tilde{\varphi}(\mathbf{T}, \mathbf{\Gamma}, \gamma) = \frac{|\text{dev } \mathbf{T} + \text{dev } \mathbf{\Gamma}| + g(\gamma)}{c_0} - 1 \leq 0 \quad (7.11)$$

and then set

$$f(\mathbf{Z}) = \frac{|\text{dev } \mathbf{T} + \text{dev } \mathbf{\Gamma}| + g(\gamma)}{c_0} \quad (7.12)$$

to get the required relationship

$$\tilde{\varphi}(\mathbf{T}, \mathbf{\Gamma}, \gamma) = f(\mathbf{X}) - 1 \leq 0 . \quad (7.13)$$

It is easily established that the function f is the canonical yield function, that is, $f(\mathbf{Z})$ is a gauge: it clearly satisfies the conditions $f(\mathbf{Z}) \geq 0$, $f(\mathbf{0}) = 0$, and it is positively homogeneous; f is in fact continuous, hence lower semicontinuous, and convexity follows from the linearity of $\text{dev}(\cdot)$ and repeated use of the triangle inequality using the norm $|\cdot|$ on E .

The flow law in its classical form now gives

$$\mathbf{p} \equiv \begin{pmatrix} \mathbf{D}^p \\ \check{\mathbf{\Lambda}} \\ \dot{\gamma} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{N} \\ \mathbf{N} \\ 1 \end{pmatrix} \quad (7.14)$$

where the constant c_0 has been absorbed into λ , and

$$\mathbf{N} = \frac{\text{dev}(\mathbf{T} + \mathbf{\Gamma})}{|\text{dev}(\mathbf{T} + \mathbf{\Gamma})|} . \quad (7.15)$$

Thus we find that

$$\dot{\gamma} = \lambda \quad \text{and} \quad \mathbf{D}^p = \check{\mathbf{\Lambda}} . \quad (7.16)$$

In view of (7.14) and (7.16) we may restrict attention to the subspace consisting of those vectors \mathbf{p} of the form $\mathbf{p} = (\mathbf{K}, \mathbf{K}, \alpha)$ with $\text{tr } \mathbf{K} = 0$, and for vectors in this subspace the dissipation function \mathcal{D} is

$$\begin{aligned}
 \mathcal{D}(\mathbf{p}) &= \sup_{f(\mathbf{X}) \leq 1} \mathbf{X} \cdot \mathbf{p} \\
 &= \sup_{f(\mathbf{X}) = 1} \mathbf{X} \cdot \mathbf{p} \\
 &= \sup_{f(\mathbf{X}) = 1} \{ \text{dev}(\mathbf{T} + \mathbf{\Gamma}) \cdot \mathbf{D}^p + g\dot{\gamma} \} \\
 &= \sup_{f(\mathbf{X}) = 1} \{ g(\dot{\gamma} - |\mathbf{D}^p| \cos \theta) + c_0 |\mathbf{D}^p| \cos \theta \}
 \end{aligned} \tag{7.17}$$

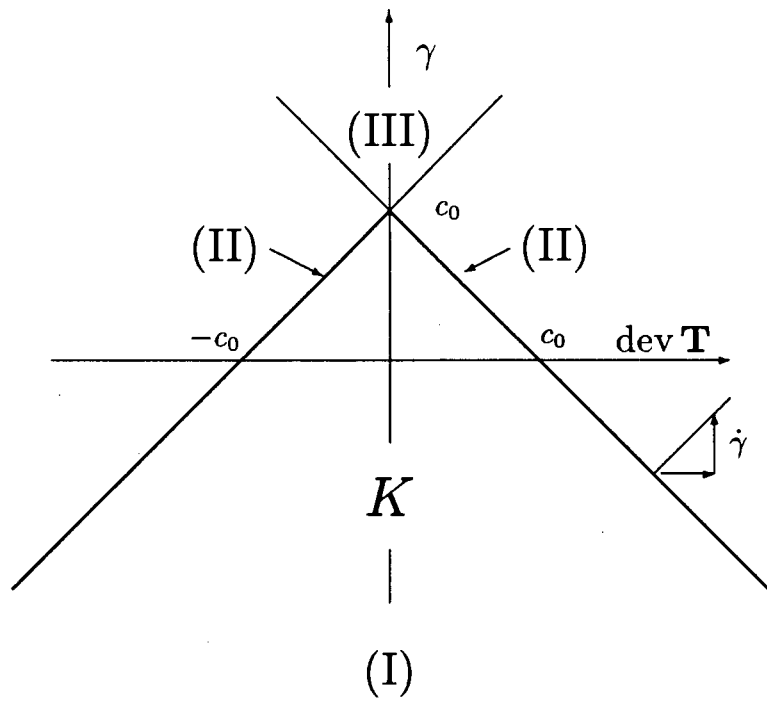
where

$$\cos \theta = \frac{(\text{dev}(\mathbf{T} + \mathbf{\Gamma})) \cdot \mathbf{D}^p}{|\text{dev}(\mathbf{T} + \mathbf{\Gamma})| |\mathbf{D}^p|} . \tag{7.18}$$

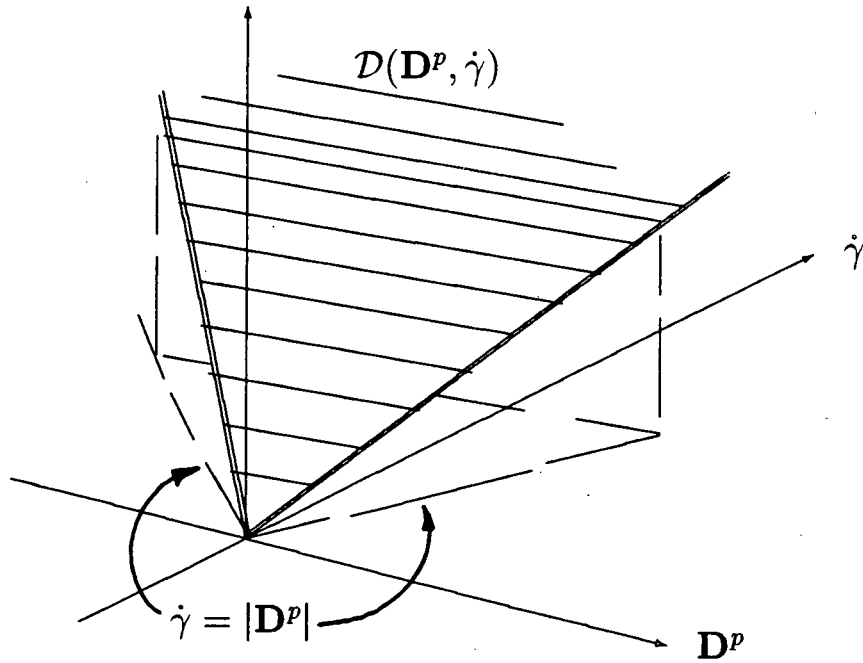
Three situations are possible: $\dot{\gamma}$ is either less than, equal to, or greater than, $|\mathbf{D}^p|$. For each of these situations we evaluate the supremum on the righthand side of (7.17) to obtain

$$\begin{aligned}
 (I) \quad \dot{\gamma} < |\mathbf{D}^p| : \quad \mathcal{D}(\mathbf{p}) &\rightarrow \infty \\
 (II) \quad \dot{\gamma} = |\mathbf{D}^p| : \quad \mathcal{D}(\mathbf{p}) &= \dot{\gamma} \\
 (III) \quad \dot{\gamma} > |\mathbf{D}^p| : \quad \mathcal{D}(\mathbf{p}) &= \dot{\gamma} .
 \end{aligned} \tag{7.19}$$

The situation is summarised pictorially in Figure 7.2, which shows for a one-dimensional situation the canonical yield function and the associated dissipation function. We note that while the theory gives rise to the three possibilities summarised above, in practice it is only (II) which will be possible: the region $g > 0$ in the set K is not accessible since we have at all times $\dot{\gamma} \geq 0 \iff g \leq 0$, so that case (III) will not arise. It is clear from the normality law and the shape of K that case (I) will not arise either. This observation has important consequences which affect the formulation and solution of the incremental boundary value problem discussed in Chapters VIII and IX.



(a) K the set of admissible conjugate forces $\mathbf{Z} = (\text{dev } \mathbf{T}, g)$



(b) dissipation function $\mathcal{D}(\mathbf{p})$, $\mathbf{p} = (\mathbf{D}^p, \dot{\gamma})$

Figure 7.2: K and \mathcal{D} for a one-dimensional example involving isotropic hardening

CHAPTER VIII

THE INCREMENTAL PROBLEM

8.1 Introduction

In this Chapter we consider the application of the internal variable theory of finite-strain plasticity developed in Chapters V and VI to the solution of a simple class of initial-boundary-value problems which involve finite plastic deformations. We wish to model the deformation of a body consisting of an isotropic elastic-plastic isotropically hardening material, subjected to a slowly applied program of loading. Under these conditions any rate dependence of the plastic behaviour can be ignored, as can inertial terms in the equation of motion for the body. The resulting problem is one of quasistatic equilibrium. The case of elastic-perfectly plastic materials is obtained simply by choosing the appropriate hardening constants to be zero.

We begin by stating the general form of an initial-boundary-value problem of the type described above. We then approximate this problem by discretising in time to obtain an incremental problem. This approximation is reasonably straightforward; however, some care needs to be taken to ensure that the plastic volume constraint is accommodated correctly in the incremental problem. To solve the incremental problem a further discretisation, in space, is required; we shall use the finite element method for this. To apply the finite element method we need to express the problem in variational form. The main purpose of this Chapter is to describe how we derive a variational form of the incremental boundary-value problem. We do this for the example which we introduced in Chapter VII and which we consider in later Chapters. The finite element approximation is discussed in Chapter IX.

8.2 The initial-boundary-value problem

We consider the quasistatic behaviour of an elastic-plastic body under the influence of prescribed boundary tractions, body forces and boundary displacements given as

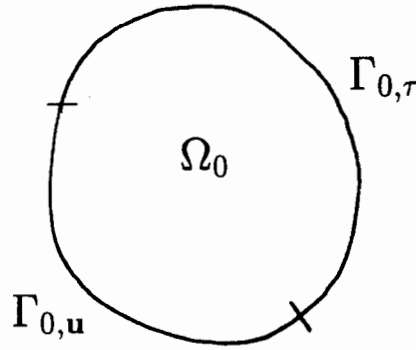


Figure 8.1: The reference configuration with prescribed loading

functions of time t ($0 \leq t \leq T$). This is illustrated in Figure 8.1.

The initial configuration of the body, Ω_0 , is assumed to be a natural configuration. We follow a *Total Lagrangian approach* in formulating the problem, in that this natural configuration is used as the fixed reference configuration. (Recall that in Chapter IV we described a natural configuration as one which is unstressed and undeformed.)

The body force is described relative to the reference configuration, that is, $\mathbf{b} = \mathbf{b}(\mathbf{X}, t)$ on Ω_0 . Surface tractions $\boldsymbol{\tau}(\mathbf{X}, t)$ are described on a part $\Gamma_{0,\tau}$ of the boundary in the initial configuration and displacements are prescribed on the remainder of the boundary $\Gamma_{0,u}$. The problem we wish to solve is:

Find the displacement field $\mathbf{u}(\mathbf{X}, t)$, the plastic part of the deformation gradient $\mathbf{F}^p(\mathbf{X}, t)$ and the hardening parameter $\gamma(\mathbf{X}, t)$ which satisfy:

$$\text{the equation of equilibrium} \quad \text{Div } \mathbf{P}(\mathbf{u}, \mathbf{F}^p) + \mathbf{b} = \mathbf{0} , \quad (8.1)$$

$$\text{the constitutive equations} \quad \mathbf{T} = 2\rho_0 \mathbf{F}^e \frac{\partial \psi_0(\mathbf{C}^e)}{\partial \mathbf{C}^e} \mathbf{F}^{eT} \quad (8.2)$$

$$\text{and} \quad g = -\rho_0 \frac{\partial \psi_2(\gamma)}{\partial \gamma} , \quad (8.3)$$

$$\text{the evolution equations} \quad (\text{dev } \mathbf{T}, g) \in \partial \mathcal{D}(\mathbf{D}^p, \dot{\gamma}), \quad (8.4)$$

$$\text{and the constraints} \quad \mathbf{W}^p = \mathbf{0} \quad (8.5)$$

$$\text{and} \quad J^p = 1 , \quad (8.6)$$

in the domain Ω_0 for all t in $[0, T]$;

$$\left. \begin{array}{l} \text{the boundary conditions} \\ \text{and} \end{array} \right\} \begin{array}{l} \mathbf{u} = \bar{\mathbf{u}}(t) \quad \text{on } \Gamma_{0,u} \\ \mathbf{PN} = \boldsymbol{\tau}(t) \quad \text{on } \Gamma_{0,\tau} \end{array} \quad (8.7)$$

where \mathbf{N} is the unit outward normal to the boundary in the reference configuration Ω_0 ;

$$\left. \begin{array}{l} \text{the initial conditions} \\ \text{and} \end{array} \right\} \begin{array}{l} \mathbf{u}(\mathbf{X}, 0) = \mathbf{0} \\ \mathbf{F}^p(\mathbf{X}, 0) = \mathbf{I} \\ \gamma(\mathbf{X}, 0) = 0 \end{array} \quad (8.8)$$

which imply that $\mathbf{F}(\mathbf{X}, 0) = \mathbf{I}$, $\mathbf{F}^e(\mathbf{X}, 0) = \mathbf{F}(\mathbf{X}, 0) = \mathbf{I}$, $J^p(\mathbf{X}, 0) = 1$ and that $g(\mathbf{X}, 0) = 0$.

8.3 The incremental boundary value problem

In general it is not possible to find an exact solution to an initial-boundary-value problem of the type described above. We therefore discretise the problem and look for an approximate solution. Following Martin and Reddy [72,99] we first semidiscretise with respect to time t : we partition the interval $[0, T]$ so that

$$0 = t_0 \leq t_1 \leq \dots \leq t_N = T,$$

and write

$$A_n = A(t_n) \quad \text{and} \quad \Delta A = A_{n+1} - A_n \quad (8.9)$$

for any function A . We then construct a procedure for finding increments $\Delta \mathbf{u}$, $\Delta \mathbf{F}^p$ and $\Delta \gamma$ which when added to the solution $\mathbf{w}_n = \{\mathbf{u}_n, \mathbf{F}^p_n, \gamma_n\}$, obtained at t_n give the solution $\mathbf{w}_{n+1} = \{\mathbf{u}_{n+1}, \mathbf{F}^p_{n+1}, \gamma_{n+1}\}$ at t_{n+1} ($n = 0, \dots, N-1$).

A finite difference method is used to approximate rate quantities so that the problem can be written in terms of the known solution at t_n and the unknown increments identified above. In a recent paper by Simo and Govindjee [110] the use of a *generalised midpoint rule* has been investigated; see also Simo [109], Reddy and Martin [99] and Ortiz and Popov [89]. Generalised trapezoidal rules have also been considered, but these have been shown to be less effective than the midpoint rules (see [89]). In algorithms based

on the generalised midpoint rule the equations governing the problem are assumed to hold at the intermediate time $t_{n+\alpha} = \alpha t_{n+1} + (1 - \alpha)t_n$ where the parameter α lies in the interval $[0, 1]$. The rate of any function $A(t)$ is approximated by an Euler backward difference:

$$\dot{A} \approx \frac{A_{n+1} - A_n}{t_{n+1} - t_n} \equiv \frac{\Delta A}{\Delta t}, \quad (8.10)$$

so that

$$A_{n+\alpha} \approx A_n + \alpha \Delta A. \quad (8.11)$$

In the case of small-strain plasticity, stability of the method is ensured for $\alpha \in [\frac{1}{2}, 1]$; we expect this result to carry over to the finite-strain case. The choice $\alpha = 0$ results in a fully explicit form of the problem, while the commonly used Euler backward difference scheme is recovered with the choice $\alpha = 1$. In many situations an Euler backward scheme results in an overestimation of plastic dissipation; it has been noted, however, (see [110]) that this choice is found to be the most practical in many applications and is widely used. The choice $\alpha = \frac{1}{2}$ has been used to alleviate problems arising in the evaluation of the *covariant stress rates* required when rate forms of the constitutive equations are used [95]. We do not use this approach though it has been suggested that this method could also be used to overcome the problem of imposing the volume constraint of isochoric plastic flow [81]. We investigate the use of $\alpha = \frac{1}{2}$ in Chapter XI.

We write

$$\dot{\mathbf{F}}^p \approx \frac{\mathbf{F}^p_{n+1} - \mathbf{F}^p_n}{t_{n+1} - t_n} \equiv \frac{\Delta \mathbf{F}^p}{\Delta t} \quad (8.12)$$

and similarly

$$\dot{\gamma} \approx \frac{\gamma_{n+1} - \gamma_n}{t_{n+1} - t_n} \equiv \frac{\Delta \gamma}{\Delta t}. \quad (8.13)$$

The displacement \mathbf{u} enters into the problem only through \mathbf{F} (recall (4.8)), so that for given displacements \mathbf{u}_n the $\mathbf{F}_{n+\alpha}$ is obtained from

$$\mathbf{F}_{n+\alpha} = \mathbf{I} + \nabla(\mathbf{u}_n) + \alpha \nabla(\Delta \mathbf{u}) = \mathbf{F}_n + \alpha \nabla(\Delta \mathbf{u}) . \quad (8.14)$$

In the incremental problem we may therefore consider \mathbf{F} to be simply a function of $\Delta \mathbf{u}$. We also write $\mathbf{F}_{n+\alpha}^p$ in terms of $\Delta \mathbf{F}^p$, using (8.11), and hence express $\mathbf{F}_{n+\alpha}^e$ as a function of $\Delta \mathbf{u}$ and $\Delta \mathbf{F}^p$ for given α (recall (4.21)).

Using the above we can approximate the plastic part of the velocity gradient and, using (8.12), write it in terms of the unknown increments as

$$\mathbf{L}_{n+\alpha}^p \approx \mathbf{F}_{n+\alpha}^e(\Delta \mathbf{u}, \Delta \mathbf{F}^p) \frac{\Delta \mathbf{F}^p}{\Delta t} \mathbf{F}_{n+\alpha}^{-1}(\Delta \mathbf{u}) . \quad (8.15)$$

We recall that the dissipation function is positively homogeneous of degree one so that its subdifferential is positively homogeneous of degree zero:

$$\partial \mathcal{D}(\theta \mathbf{p}) = \partial \mathcal{D}(\mathbf{p}) , \quad 0 < \theta \in \mathbf{R} . \quad (8.16)$$

Choosing $\theta = \frac{1}{\Delta t}$ in (8.16) and using the approximation (8.15), we may write the evolution equations for the incremental problem as

$$\text{dev } \mathbf{T}_{n+\alpha}(\Delta \mathbf{u}, \Delta \mathbf{F}^p) \in \partial_{\mathbf{D}^p} \mathcal{D}(\mathbf{D}^p(\Delta \mathbf{u}, \Delta \mathbf{F}^p), \Delta \gamma) \quad (8.17)$$

and

$$g_{n+\alpha} \in \partial_{\dot{\gamma}} \mathcal{D}(\mathbf{D}^p(\Delta \mathbf{u}, \Delta \mathbf{F}^p), \Delta \gamma) . \quad (8.18)$$

Before stating the incremental problem we must consider how to impose the plastic volume constraint in this discrete case. We recall from Chapter VI that the dissipation inequality can be written in the form

$$\text{dev } \mathbf{T} \cdot \text{dev } \mathbf{D}^p + \frac{1}{3}(\text{tr } \mathbf{T})(\text{tr } \mathbf{D}^p) + g \dot{\gamma} \geq 0 . \quad (8.19)$$

The term involving the spherical part of the stress was eliminated by requiring that $\dot{J}^p = 0$ or equivalently that $\text{tr } \mathbf{D}^p = 0$; this resulted in the choice of $\text{dev } \mathbf{T}$ as the force

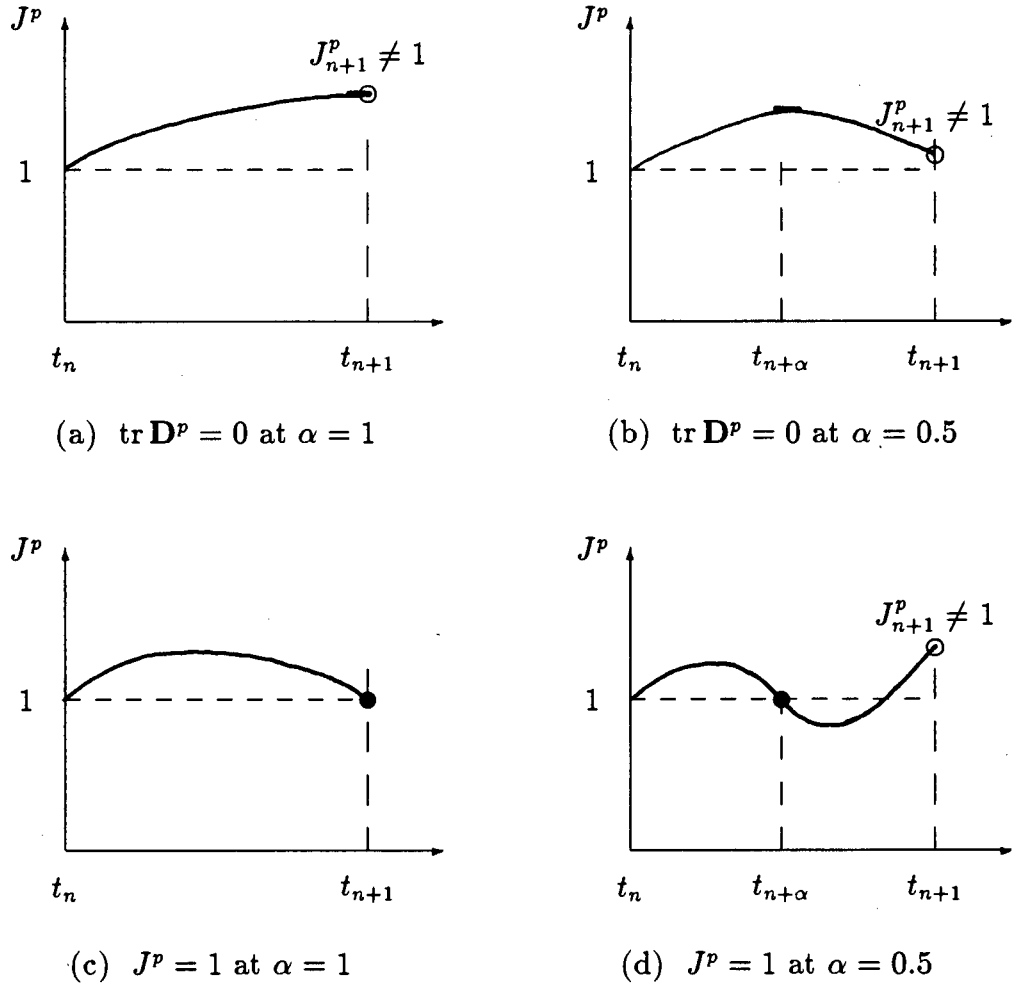


Figure 8.2: Discrete forms of the assumption of isochoric plastic flow

conjugate to $\mathbf{D}^p = \text{dev } \mathbf{D}^p$. On its own $\dot{J}^p = 0$ implies only that J^p is constant. In the continuous case the choice $J^p = 1$ at $t = 0$ together with \dot{J}^p fixes $J^p = 1$ for all time; since we use this initial condition the assumption of isochoric plastic deformation may be included in the formulation of the problem simply by setting $\text{tr } \mathbf{D}^p = 0$.

In the discrete case the constraint $J^p = 1$ cannot simply be replaced by the condition $\dot{J}^p = 0$ as in the continuous case. If it were it would effectively only be applied at discrete points in time, $t = t_{n+\alpha}$; because of the discrete nature of the problem the connection to the initial condition would be lost and hence the value of J^p would not be fixed. We illustrate this observation in Figure 8.2. In the discrete case it is therefore

necessary to add the constraint $J^p = 1$ to ensure that isochoric plastic deformation is maintained; since this implies $\dot{J}^p = 0$ we eliminate the middle term from (8.19). We then proceed as before but replace \mathbf{D}^p with $\text{dev } \mathbf{D}^p$.

Using the generalised midpoint rule the constraint becomes

$$J_{n+\alpha}^p - 1 = 0. \quad (8.20)$$

We remark that an approximation of \dot{J}^p by

$$j^p \approx \frac{J_{n+1}^p - J_n^p}{\Delta t} = 0, \quad (8.21)$$

with the assumption that only isochoric plastic deformation takes place up to time t_n so that $J_n^p = 1$, amounts to the constraint being applied in the form

$$J_{n+1}^p - 1 = 0, \quad (8.22)$$

which is independent of α ; of course in the case $\alpha = 1$ (8.20) and (8.22) are identical.

We can now state the general form of the incremental problem:

Given the solution $\mathbf{w}_n = \{\mathbf{u}_n, \mathbf{F}_n^p, \gamma_n\}$ at t_n and the data \mathbf{b}_{n+1} and $\boldsymbol{\tau}_{n+1}$ find $\Delta \mathbf{w}$ such that $\mathbf{w}_{n+\alpha} = \mathbf{w}_n + \alpha \Delta \mathbf{w}$ satisfies:

$$\text{the equilibrium equation} \quad \text{Div } \mathbf{P}(\mathbf{w}_{n+\alpha}) + \mathbf{b}_{n+\alpha} = 0, \quad (8.23)$$

$$\text{the constitutive equations} \quad \mathbf{P}_{n+\alpha} = 2\rho_0 \mathbf{F}_{n+\alpha} \mathbf{F}_{n+\alpha}^{p-1} \left. \frac{\partial \psi_0(\mathbf{C}^e)}{\partial \mathbf{C}^e} \right|_{t_{n+\alpha}} \mathbf{F}_{n+\alpha}^{p-T} \quad (8.24)$$

$$\text{and} \quad g_{n+\alpha} = -\rho_0 \left. \frac{\partial \psi_2(\gamma)}{\partial \gamma} \right|_{t_{n+\alpha}}, \quad (8.25)$$

$$\text{the evolution equations} \quad \text{dev } \mathbf{T}_{n+\alpha}(\Delta \bar{\mathbf{w}}) \in \partial_{\mathbf{D}^p} \mathcal{D}(\text{dev } \mathbf{D}^p(\Delta \bar{\mathbf{w}}), \Delta \gamma) \quad (8.26)$$

$$\text{and} \quad g_{n+\alpha}(\Delta \gamma) \in \partial_\gamma \mathcal{D}(\text{dev } \mathbf{D}^p(\Delta \bar{\mathbf{w}}), \Delta \gamma) \quad (8.27)$$

$$\text{subject to constraints} \quad \mathbf{W}_{n+\alpha}^p = \mathbf{0} \quad (8.28)$$

$$\text{and} \quad J_{n+\alpha}^p - 1 = 0 \quad (8.29)$$

in the domain Ω_0 ;

$$\left. \begin{array}{l} \text{the boundary conditions} \\ \text{and} \end{array} \right\} \begin{array}{ll} \mathbf{u}_{n+\alpha} = \bar{\mathbf{u}}_{n+\alpha} & \text{on } \Gamma_{0,u} \\ \mathbf{P}_{n+\alpha} \mathbf{N} = \boldsymbol{\tau}_{n+\alpha} & \text{on } \Gamma_{0,\tau} \end{array} \quad (8.30)$$

where $\Delta \bar{\mathbf{w}} = \{\Delta \mathbf{u}, \Delta \mathbf{F}^p\}$.

We conclude this Section by considering the specific model of plastic behaviour discussed later; that is, plastic behaviour which conforms to the von Mises yield criterion (7.10). In Chapter VII we showed that for this type of yield function $\dot{\gamma} = |\mathbf{D}^p|$; the subdifferential of \mathcal{D} with respect to $\dot{\gamma}$ was used to obtain this relationship. With this result we can express $\Delta \gamma$ in terms of $\Delta \mathbf{u}$ and $\Delta \mathbf{F}^p$ as

$$\Delta \gamma(\Delta \bar{\mathbf{w}}) = |\text{sym}(\mathbf{F}^e(\Delta \bar{\mathbf{w}}) \Delta \mathbf{F}^p \mathbf{F}^{-1}(\Delta \mathbf{u}))|. \quad (8.31)$$

This equation describes fully the evolution of γ for the increment and can be used in place of equation (8.27) in the incremental problem; $\Delta \gamma$ is therefore not required as a primary variable.

8.4 Variational form of the incremental problem

To construct a variational or weak form of the incremental problem we assume that $\Delta \mathbf{u}$, $\Delta \mathbf{F}^p$ and $\Delta \gamma$ belong to suitable spaces of functions H , Q and R .

8.4.1 The equilibrium equation

We choose the space H of displacement variations to be defined by

$$H = \{\mathbf{v} = (v_i) : v_i \in H^1(\Omega_0), v_i = 0 \text{ on } \Gamma_{0,u_i}\}. \quad (8.32)$$

We remark that the space H^1 to which the components of \mathbf{v} belong is too restrictive for a general theory (see Ball [5]), or for perfect plasticity (see Teman and Strang [119] or Reddy [96]) but it suffices for our purposes. The variational form of the equilibrium equation (8.23) is then obtained in the usual way, by first multiplying by arbitrary displacements $\mathbf{v} \in H$ and integrating over Ω_0 to get

$$\int_{\Omega_0} \mathbf{v} \cdot \operatorname{div} \mathbf{P} \, d\mathbf{X} - \int_{\Omega_0} \mathbf{v} \cdot \mathbf{b} \, d\mathbf{X} = 0 \quad \forall \mathbf{v} \in H \quad (8.33)$$

and then applying the divergence theorem (2.56) and the boundary conditions implied by (8.32) to obtain

$$\int_{\Omega_0} \nabla \mathbf{v} \cdot \mathbf{P}(\mathbf{w}_n) \, d\mathbf{X} - \int_{\Omega_0} \mathbf{b} \cdot \mathbf{v} \, d\mathbf{X} - \int_{\Gamma_{0,r}} \mathbf{v} \cdot \boldsymbol{\tau}_n \, d\mathbf{S} = 0 \quad \forall \mathbf{v} \in H . \quad (8.34)$$

Of course by writing $\mathbf{P} = \mathbf{P}(\mathbf{w})$ we have included the constitutive equation (8.24).

8.4.2 The evolution and constraint equations

It is less straightforward to cast the evolution equations (8.26) in variational form because of the nondifferentiability of the dissipation function. We shall consider only dissipation functions which are differentiable everywhere except at the origin. The dissipation function which corresponds to the von Mises yield condition is of this type.

At any given time we divide the reference domain of the problem into two nonoverlapping subdomains Ω^E and Ω^P , so that

$$\bar{\Omega}_0 = \bar{\Omega}^E \cup \bar{\Omega}^P, \quad \Omega^E \cap \Omega^P = \emptyset . \quad (8.35)$$

On Ω^E the material behaves elastically so that $\mathbf{L}^P = \mathbf{0}$, which implies that $\mathbf{D}^P = \mathbf{0}$, $\mathbf{W}^P = \mathbf{0}$ and $\dot{\mathbf{J}}^P = 0$; hence no constraint equations are required. In our model the internal variable γ does not depend on elastic deformation and hence $\dot{\gamma} = 0$ also. Having $\mathbf{D}^P = \mathbf{0}$ in turn implies that \mathcal{D} is nondifferentiable; however, since the behaviour is elastic we do not need any evolution equations.

On Ω^P yielding takes place and the material behaves plastically, that is $\mathbf{D}^P \neq \mathbf{0}$; within the class of dissipation functions being considered this implies that \mathcal{D} is differentiable. The crucial point is that Ω^P is unknown *a priori*; this is overcome by the manner in which algorithms for solving the problem are constructed. For the moment we shall assume that Ω^P has been found.

By definition \mathbf{F} and \mathbf{F}^p are never zero, hence $\mathbf{D}^p = \mathbf{0}$ implies that $\dot{\mathbf{F}}^p = \mathbf{0}$ which in our approximation implies that $\Delta\mathbf{F}^p = \mathbf{0}$. Thus on Ω^E , $\Delta\mathbf{F}^p$ is known and we need only consider variations on the subdomain Ω^P where \mathcal{D} is differentiable.

The form of the evolution equations at any point on Ω^P depends on the particular dissipation function being used. For our example (von Mises plasticity with isotropic hardening) we have

$$\operatorname{dev} \mathbf{T} = k \frac{\operatorname{dev} \mathbf{D}^p}{|\operatorname{dev} \mathbf{D}^p|} \quad (8.36)$$

where $k = c_0 - g(\gamma)$. Since both sides of this equation involve symmetric tensors it amounts to a system of six independent equations in the nine components of $\Delta\mathbf{F}^p$. The assumption that $\mathbf{W}^p = \mathbf{0}$ provides a further three equations; adding these and rearranging, (8.36) becomes

$$-\frac{\operatorname{dev} \mathbf{T}}{k} + \frac{\operatorname{dev} \mathbf{L}^{pT}}{|\operatorname{dev} \mathbf{D}^p|} = \mathbf{0}. \quad (8.37)$$

Using the expression (8.18), the function $g(\gamma)$ can be rewritten in terms of $\Delta\mathbf{F}^p$ and $\Delta\mathbf{u}$; hence, given $\Delta\mathbf{u}$, (8.37) contains all the information required to determine $\Delta\mathbf{F}^p$ on Ω^P . From this solution $\Delta\gamma$ can also be found.

We define a space Q of tensor fields by

$$Q = \{\mathbf{q} = (q_{\beta A}) : q_{\beta A} \in L^2(\Omega_0)\} \quad (8.38)$$

where the components $q_{\beta A}$ are with respect to any pair of bases ($\{\mathbf{E}_A\}$ in the reference configuration and $\{\bar{\mathbf{E}}_\beta\}$ in the intermediate configuration, (see Chapter II) and the space $L^2(\Omega_0)$ is the space of functions which are square-integrable on Ω_0 . For hardening materials the requirement that the components $q_{\alpha A}$ be in $L^2(\Omega_0)$ is consistent with the physical interpretation of \mathbf{F}^p ; in the case of perfect plasticity, however, localisation of plastic deformation can occur which results in the formation of *sliplines*; hence in this case the components $q_{\alpha A}$ are in general bounded measures (see [119] for a detailed discussion of this point). We shall seek solutions in $L^2(\Omega_0)$ in both the case of hardening and perfect plasticity, that is, we shall ignore the possible formation of sliplines.

We must still add the constraint $J^p = 1$. In the continuous case it is sufficient to replace $\text{dev } \mathbf{D}^p$ by \mathbf{D}^p in (8.36); the orthogonality of deviatoric and spherical tensors ensures that a solution of (8.37) satisfies this condition since the righthand side is deviatoric.

In the discrete case we have argued that the appropriate form of the constraint is $J_{n+\alpha}^p - 1 = 0$; we therefore require that \mathbf{q} and the solution $\Delta \mathbf{F}^p$ lie in the subset $\hat{Q}_{n+\alpha}$ of Q defined by

$$\hat{Q}_{n+\alpha} = \{\mathbf{q} \in Q \mid \det(\mathbf{F}_{n+\alpha}^p + \alpha \mathbf{q}) = 1\}. \quad (8.39)$$

We remark that the elements of $\hat{Q}_{n+\alpha}$ are not required to satisfy the condition

$$\text{tr}(\mathbf{F}_{n+\alpha}(\mathbf{F}_{n+\alpha}^p + \alpha \mathbf{q})^{-1})\mathbf{q}(\mathbf{F}_{n+\alpha})^{-1} = 0,$$

which is essentially the condition $\text{tr } \mathbf{D}^p = 0$ at $t_{n+\alpha}$.

In the remainder of this Chapter we shall drop the subscript $n + \alpha$, unless otherwise stated, since all quantities are evaluated at time $t_{n+\alpha}$.

A weak form of (8.37) is obtained by taking the dot product of both sides with $\mathbf{F}^e \mathbf{q} \mathbf{F}^{-1}$ and integrating. Assuming for the moment that $\Delta \mathbf{u}$ is known, the problem becomes:

Find the solution $\Delta \mathbf{F}^p \in \hat{Q}_{n+\alpha}$ that satisfies $\Delta \mathbf{F}^p = \mathbf{0}$ on Ω^E , and

$$\int_{\Omega^p} \mathbf{F}^{eT} \left(-\frac{\text{dev } \mathbf{T}}{k} + \frac{\text{dev } \mathbf{L}^{pT}}{|\text{dev } \mathbf{D}^p|} \right) \mathbf{F}^{-1} \cdot \mathbf{q} \, d\mathbf{X} = 0 \quad \forall \mathbf{q} \in \hat{Q}_{n+\alpha}. \quad (8.40)$$

We remark that in this form the constraint $J^p = 1$ is included through the definition of the space $\hat{Q}_{n+\alpha}$. It is necessary to use $\mathbf{F}^e \mathbf{q} \mathbf{F}^{-1}$ rather than simply \mathbf{q} to make the product meaningful, that is, we *push forward* the intermediate configuration quantity \mathbf{q} to the current configuration (see [67,114]). We also note that, when \mathbf{q} is equal to the solution $\Delta \mathbf{F}^p$, equation (8.40) becomes

$$\int_{\Omega} (k|\mathbf{D}^p| - (\mathbf{T} \cdot \mathbf{D}^p)) \, d\mathbf{X} = 0, \quad (8.41)$$

which implies that the dissipation $(\mathbf{T} \cdot \mathbf{D}^p)$ is equal to $\mathcal{D}(\mathbf{D}^p) = k|\mathbf{D}^p|$.

In the algorithm used to solve the incremental problem we solve the evolution equations for $\Delta \mathbf{F}^p$ holding the displacements fixed. If we do not restrict the solution to $\hat{Q}_{n+\alpha}$ but proceed as we did for the continuous case by setting $\mathbf{D}^p = \text{dev } \mathbf{D}^p$ and ignoring the constraint $J^p = 1$, thus allowing the solution to be in Q , we find that in the discrete case, for fixed \mathbf{u} , equation (8.37) has a solution $\Delta \mathbf{F}^{p*}$ for which $J^{p*} = J$, in addition to the required solution $\Delta \mathbf{F}^p$ for which $J^p = 1$. The reason for this is that if $J^p = J$ then $J^e = 1$ and the constitutive equation gives $\mathbf{T} = \text{dev } \mathbf{T}$. As it turns out, when a Newton-Raphson method is employed to solve (8.37) at a point, the procedure converges only to the solution $\Delta \mathbf{F}^{p*}$.

In practice it is very awkward to work with the space $\hat{Q}_{n+\alpha}$, so to avoid this difficulty we make use of a Lagrange multiplier. We set

$$\lambda = \frac{J^{p-1}}{3k} \text{tr } \mathbf{T} \quad (8.42)$$

and define an appropriate space L of Lagrange multipliers; the space L could be chosen to be $L^2(\Omega)$, for example. Then using the fact that $\text{dev } \mathbf{T} = \mathbf{T} - \frac{1}{3}(\text{tr } \mathbf{T})\mathbf{I}$ we rewrite (8.40) and add a weak form of the constraint equation so that the problem becomes:

find $\Delta \mathbf{F}^p \in Q$ and $\lambda \in L$ that satisfy

$$\int_{\Omega^p} \left(-\frac{\mathbf{F}^e T \mathbf{T} \mathbf{F}^{-T}}{k} + \lambda J^p \mathbf{F}^{p-T} + k \mathbf{F}^e T \frac{\text{dev } \mathbf{L}^{pT}}{|\text{dev } \mathbf{D}^p|} \mathbf{F}^{-T} \right) \cdot \mathbf{q} \, d\mathbf{X} = 0 \quad \forall \mathbf{q} \in Q \quad (8.43)$$

and

$$\int_{\Omega^p} \mu (J^p - 1) \, d\mathbf{X} = 0 \quad \forall \mu \in L. \quad (8.44)$$

We now consider an interesting alternative justification for recasting the problem in this way, and in particular for the form the λ term takes in (8.43). We multiply equation (8.40) by α and note that this equation is of the form

$$\int_{\Omega^p} G(\Delta \mathbf{F}^p) \cdot \mathbf{q} \, d\mathbf{X} = 0, \quad (8.45)$$

for given \mathbf{u} and α . Now suppose that there is a minimisation problem

$$\min_{\mathbf{q} \in \hat{Q}_{n+\alpha}} W(\mathbf{q}) \quad (8.46)$$

whose solution is $\Delta \mathbf{F}^p$. That is (8.45) is equivalent to the requirement

$$\langle DW(\Delta \mathbf{F}^p), \mathbf{q} \rangle = 0 \quad (8.47)$$

where D denotes the Gateaux derivative, so that $G = DW$ for some functional $W(\cdot)$. Then, rather than working with the constrained minimisation problem (8.46), we could consider the *equivalent* saddlepoint problem:

$$\min_{\mathbf{q} \in Q} \max_{\mu \in L} K(\mathbf{q}, \mu) \quad (8.48)$$

where

$$K(\mathbf{q}, \mu) = W(\mathbf{q}) + \int_{\Omega^p} \mu (\det(\mathbf{F}^p_n + \alpha \mathbf{q}) - 1) d\mathbf{X}$$

(recall that $J_{n+\alpha}^p = \det(\mathbf{F}^p_n + \alpha \mathbf{q})$ for $\mathbf{q} = \Delta \mathbf{F}^p$.) Necessary conditions for $\{\Delta \mathbf{F}^p, \lambda\}$ to be a solution of this saddlepoint problem are

$$\langle D_{\mathbf{q}} K(\Delta \mathbf{F}^p, \lambda), \mathbf{q} \rangle = 0 \quad (I),$$

$$\langle D_{\mu} K(\Delta \mathbf{F}^p, \lambda), \mu \rangle = 0 \quad (II),$$

where $D_{\mathbf{q}}$ and D_{μ} denote partial Gateaux derivatives with respect to \mathbf{q} and μ respectively. (I) is equivalent to

$$\left. \frac{d}{d\epsilon} K(\Delta \mathbf{F}^p + \epsilon \mathbf{q}, \lambda) \right|_{\epsilon=0} = 0 \quad (8.49)$$

or

$$\left. \frac{d}{d\epsilon} W(\Delta \mathbf{F}^p + \epsilon \mathbf{q}) \right|_{\epsilon=0} + \left. \frac{d}{d\epsilon} \langle \lambda, (\det(\mathbf{F}^p_n + \alpha(\Delta \mathbf{F}^p + \epsilon \mathbf{q})) - 1) \rangle \right|_{\epsilon=0} = 0. \quad (8.50)$$

Using the the expression for the derivative of a determinant, (8.50) becomes

$$\langle DW(\Delta \mathbf{F}^p), \mathbf{q} \rangle + \alpha \int_{\Omega^P} (\lambda J_{n+\alpha}^p \mathbf{F}_{n+\alpha}^{p-T} \cdot \mathbf{q}) \, d\mathbf{X} = 0; \quad (8.51)$$

dividing by α we get (8.43). Similarly from (II) we get (8.44).

Later we shall find it convenient to be able to eliminate λ from the system of equations; hence, with this objective in mind we introduce the perturbed form

$$J^p - 1 + \epsilon \lambda = 0, \quad 0 < \epsilon \ll 1 \quad (8.52)$$

of the constraint equation, and replace (8.44) by

$$\int_{\Omega^P} \mu (J^p - 1) \, d\mathbf{X} + \epsilon \int_{\Omega^P} \mu \lambda \, d\mathbf{X} = 0 \quad \forall \mu \in L. \quad (8.53)$$

The constraint is recovered in the limit as ϵ goes to zero.

8.4.3 Statement of the problem

In summary of this Section we state the variational form of the full incremental boundary-value problem for our example. We extend the prescribed boundary displacements $\bar{\mathbf{u}}_{n+1}$ to a function $\bar{\mathbf{u}}_{n+1}$ in $H^1(\Omega_0)$ so that the problem becomes

Given $\mathbf{w}_n = \{\mathbf{u}_n, \mathbf{F}^p_n, \gamma_n\}$ find $\Delta \mathbf{u}$ such that $(\Delta \mathbf{u} - \Delta \bar{\mathbf{u}}) \in H$, $\Delta \mathbf{F}^p \in Q$, $\lambda \in L$ and $\Omega_{n+\alpha}^P$ such that

$$\int_{\Omega_0} \nabla \mathbf{v} \cdot \mathbf{P} \, d\mathbf{X} - \int_{\Omega_0} \mathbf{v} \cdot \mathbf{b}_{n+\alpha} \, d\mathbf{X} = 0 \quad \forall \mathbf{v} \in H, \quad (8.54)$$

$$\int_{\Omega_{n+\alpha}^P} \left(-\frac{\mathbf{F}^{eT} \mathbf{T} \mathbf{F}^{-T}}{k} + \lambda J^p \mathbf{F}^{p-T} + \frac{\mathbf{F}^{eT} \operatorname{dev} \mathbf{L}^p \mathbf{T} \mathbf{F}^{-T}}{|\operatorname{dev} \mathbf{D}^p|} \right) \cdot \mathbf{q} \, d\mathbf{X} = 0 \quad \forall \mathbf{q} \in Q \quad (8.55)$$

and

$$\int_{\Omega_{n+\alpha}^P} \mu (J^p - 1 + \epsilon \lambda) \, d\mathbf{X} = 0 \quad \forall \mu \in L, \quad 0 < \epsilon \ll 1 \quad (8.56)$$

where the quantities \mathbf{P} , \mathbf{T} , \mathbf{F} , \mathbf{F}^p , \mathbf{F}^e etc. are evaluated at $t_{n+\alpha}$.

It is important to keep in mind that the definition of Ω^P is dependent on time and is unknown *a priori*.

It is interesting to note that the Lagrange multiplier λ which we have introduced is essentially the force which would have been chosen as conjugate to \dot{J}^p , had J^p been used as an independent internal variable from the outset; this is easily verified by inspection of (8.19). Use of J^p as a variable has been investigated in conjunction with a mixed method and use of a dilational/volume-preserving decomposition of deformation, by Simo and co-workers [106,112,113].

We also note that in the final form of the problem which we solve a finite element discretisation and numerical integration scheme are used. The particular manner in which these approximations are employed makes this problem equivalent to one in which the evolution laws are imposed pointwise on Ω^P ; we discuss this further in Chapter IX.

CHAPTER IX

FINITE ELEMENT APPROXIMATION

9.1 Introduction

In Chapter VIII we recognised that in general it is not possible to obtain an exact solution to the initial boundary-value problem described in Section 8.2. As a first step towards obtaining an approximate solution we semidiscretised and derived an incremental form of the problem which we then recast in variational form. A further discretisation, in space, is required before we get a form of the problem for which a solution can be found.

We obtain a discrete form of the incremental problem by replacing the functions $\Delta \mathbf{u}, \mathbf{v} \in H$, $\Delta \mathbf{F}^p, \mathbf{q} \in Q$, and $\lambda, \mu \in L$ which appear in (8.54), (8.55), and (8.56) by approximations $\Delta \mathbf{u}_h, \mathbf{v}_h \in H_h$, $\Delta \mathbf{F}^p_h, \mathbf{q}_h \in Q_h$ and $\lambda_h, \mu_h \in L_h$ where H_h , Q_h and L_h are finite-dimensional subspaces of H , Q and L respectively. The primary task in establishing a discrete form of the problem is the construction of these finite dimensional spaces. The finite element method provides us with a systematic procedure for doing this.

In this Chapter we describe the application of the finite element method. We begin by giving a brief and very general description of the finite element method; it is described in detail in many texts (see for example [24,45,46,97]) which have been referred to while preparing this Chapter. We then discuss separately the construction of the finite element spaces H_h , and Q_h and L_h . We follow a standard procedure to construct H_h and hence an approximation of displacements; the approach described by Griffin, Reddy and Martin [36] is used to construct Q_h and approximate the internal variables. The same basis functions as those used in constructing Q_h are used to construct L_h . This particular choice of displacement, internal variable and Lagrange multiplier finite element approximations leads to an uncoupling of the equations to be solved in the corrector phase of the algorithm described in Chapter X.

9.2 The finite element method

The finite element method is a method used to obtaining a discrete form of a boundary-value problem; it is a particular form of the Galerkin or Ritz method. These methods make use of the basic notion of Fourier analysis, that any function f belonging to a Hilbert space F can be represented by a sum of the form

$$f(\mathbf{X}) = \sum_{I=1}^{\infty} \Psi^I(\mathbf{X}) a^I \quad (9.1)$$

where the set of functions $\{\Psi^I\}_{I=1}^{\infty}$ forms a basis for the space F . A finite number of these basis functions forms a basis for a finite dimensional subspace F_h of F ; an approximation $f_h \in F_h$ to $f \in F$ of the form

$$f_h(\mathbf{X}) = \sum_{I=1}^N \Psi^I(\mathbf{X}) a^I \quad (N < \infty) \quad (9.2)$$

is obtained for a given choice of the N basis functions. In both the Galerkin and Ritz methods an approximation of this type is used in a variational form of the boundary-value problem. The finite element method amounts to a procedure for choosing a suitable set of basis functions Ψ^I in such an approximation.

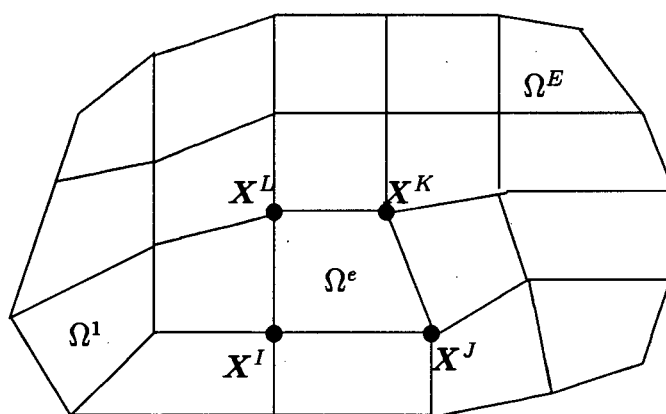


Figure 9.1: Division of a domain into finite elements Ω^e

The domain Ω is first divided into E subdomains or *elements* Ω^e as shown in Figure 9.1. The finite element basis functions are then chosen in a way that enables the fundamental property of integration that

$$\int_{\Omega} (\dots) d\mathbf{X} = \sum_{e=1}^E \int_{\Omega^e} (\dots) d\mathbf{X} , \quad (9.3)$$

to be exploited in the construction of the discrete form of the problem.

For the particular class of finite element approximations we consider the basis functions Ψ^I have the following properties:

1. $\Psi^I \in F$. This property ensures that $F_h \subset F$; in the class of approximations we are considering F_h is a subset of the space of continuous functions on Ω , that is $C^0(\Omega)$, or at least a subset of the space of piecewise continuous functions.
2. For a specific set of *node* points \mathbf{X}^J on Ω ,

$$\Psi^I(\mathbf{X}^J) = \begin{cases} 1 & \text{if } I = J \\ 0 & \text{if } I \neq J \end{cases} \quad (I, \dots, J = 1, N) . \quad (9.4)$$

This property implies that $a^I = f(\mathbf{X}^I)$; these scalars are referred to as *nodal degrees of freedom* of the function f_h .

3. The functions $\psi^{(e)I} = \Psi^I|_{\Omega^e}$ are polynomials. This property implies that the global basis functions are piecewise polynomials. The functions $\psi^{(e)I}$ are referred to as *local or element basis functions*; since they are obtained from the restriction of a global basis functions Ψ to a given element they inherit the property that

$$\psi^I(\mathbf{X}^J) = \begin{cases} 1 & \text{if } I = J \\ 0 & \text{if } I \neq J \end{cases} \quad (9.5)$$

for nodes \mathbf{X}^I on Ω^e . The number and position of nodes required on each element is governed by the generic shape of Ω^e , the continuity requirements of the functions Ψ^I and the degree, r , of the polynomial functions $\psi^{(e)I}$. In particular a number of nodes are required to be on the boundaries of each element if the function f is required to be continuous over Ω .

The combination of elements and nodes defines the *finite element mesh* used to approximate the function f on the domain Ω .

With a finite element approximation of this type, functions to be integrated on a particular element are given in terms of the simple polynomial local basis functions $\psi^{(e)I}$ corresponding to the nodes appearing on this element. A numerical integration scheme such as Gaussian quadrature can be used to evaluate the element integrals since all the functions appearing in the integral are smooth, by definition polynomials. A systematic procedure results, whereby element contributions are calculated in a routine manner and are accumulated to obtain the required integrals over Ω .

9.2.1 The approximation of displacements

The displacement quantities we use are $\Delta \mathbf{u}$ and \mathbf{v} ; these belong to the space of vector functions H defined in Chapter VIII. It is the n components of these vector functions relative to a given basis that we approximate, where each of these components belongs to the space $H^1(\Omega_0)$. It is common practice to use a local numbering system for the nodes and nodal degrees of freedom on each element. Since the same finite-dimensional subspace of $H^1(\Omega_0)$ is used for the approximation of each of the components of the displacements there are n nodal displacement degrees of freedom associated with each node in the mesh. There are thus $N^e = n \times n^e$ degrees of freedom associated with each element where n^e is the number of nodes on the element. We collect the nodal displacement degrees of freedom on each element into vectors; for $n = 2$

$$\mathbf{a}^{(e)} = \left\{ \begin{array}{c} \Delta u_1(\mathbf{X}^1) \\ \Delta u_2(\mathbf{X}^1) \\ \Delta u_1(\mathbf{X}^2) \\ \Delta u_2(\mathbf{X}^2) \\ \dots \\ \dots \\ \Delta u_1(\mathbf{X}^{n^e}) \\ \Delta u_2(\mathbf{X}^{n^e}) \end{array} \right\} = \left\{ \begin{array}{c} a^1 \\ a^2 \\ \dots \\ \dots \\ \dots \\ \dots \\ a^{N^e-1} \\ a^{N^e} \end{array} \right\}. \quad (9.6)$$

On each element the components of \mathbf{v} and $\Delta \mathbf{u}$ are related to the displacement degrees of freedom by a mapping $\mathbf{N}^{(e)}$;

$$v_i = \mathbf{N}_i^{(e)I}(\mathbf{X}) v^I \quad \text{and} \quad \Delta u_i = \mathbf{N}_i^{(e)I}(\mathbf{X}) a^I \quad (9.7)$$

where summation is implied on $I = 1, \dots, N^e$ and $i = 1, \dots, n$. The element mapping $\mathbf{N}^{(e)}$ can be written in matrix form; for the case $n = 2$

$$[\mathbf{N}_i^{(e)I}] = \left[\begin{array}{cc|cc|ccc|cc} \psi^{(e)1} & 0 & \psi^{(e)2} & 0 & \dots & \dots & \psi^{(e)n^e} & 0 \\ 0 & \psi^{(e)1} & 0 & \psi^{(e)2} & \dots & \dots & 0 & \psi^{(e)n^e} \end{array} \right]. \quad (9.8)$$

Similarly we relate the components of the gradient of displacement on an element to displacement degrees of freedom with a mapping $\mathbf{B}^{(e)}$; for $\text{Grad } \mathbf{v}$ on Ω^e we write

$$\frac{\partial v_i}{\partial X_A} = \mathbf{B}_{iA}^{(e)I}(\mathbf{X}) v^I. \quad (9.9)$$

The total number of displacement degrees of freedom N is equal to n times the total number of nodes in the mesh. We assemble a vector \mathbf{a} of nodal displacement degrees of freedom from the element vectors $\mathbf{a}^{(e)}$; $\mathbf{a} \in \mathbf{R}^N$. We omit technical details of this routine assembly process, and use the notation a^I for the components of both \mathbf{a} and $\mathbf{a}^{(e)}$ assuming that the range of the superposed index is obvious from the context in which they are used; $I = 1, \dots, N^e$ for elements and $I = 1, \dots, N$ globally.

The important point to make is that the functions H^1 are regarded to be continuous on Ω_0 and therefore nodes are required on the boundaries of elements.

9.2.2 Approximation of internal variables and the Lagrange multiplier

The internal variables which enter into the problem are \mathbf{F}^p and γ . In the incremental form of the problem we have shown that $\Delta \gamma$ can be written in terms of $\Delta \mathbf{F}^p$ so we need only construct a spatial discretisation for $\Delta \mathbf{F}^p$. This approximation implies a spatial discretisation of both \mathbf{F}^p and γ since these are obtained by adding the spatially approximated increments at each time step. The plastic volume constraint is imposed on each increment, so no connection need be made between the multipliers used on successive increments.

We shall use the same elements for the discretisation of internal variables and Lagrange multipliers as those used for displacements. We recall that each of the components of

$\Delta \mathbf{F}^p$ are assumed to belong to the space of square-integrable functions, $L^2(\Omega_0)$, and that λ belongs to a similar space. As a consequence these need not be continuous over any particular part of Ω_0 . We use piecewise polynomial finite element basis functions ϕ of the type described in Section 9.2 to approximate both the components of $\Delta \mathbf{F}^p$ and λ . These are continuous over the interior of each element; however, since no continuity is required on inter-element boundaries, they can be constructed using node points which all lie on the interior of the elements.

We choose the numerical integration points \mathbf{X}^g as the node points for the approximation of the internal variables and Lagrange multipliers on each element; $g = 1, \dots, G^e$ where G^e is the number of Gauss points used on each element. We remark that it is not necessary to make this particular choice, although as we explain later there is considerable benefit to be derived from doing so.

For this finite element approximation the local and global basis functions coincide and have the property that

$$\phi^g(\mathbf{X}^k) = \begin{cases} 1 & \text{if } g = k \\ 0 & \text{if } g \neq k \end{cases} \quad (g, k = 1, \dots, G^e). \quad (9.10)$$

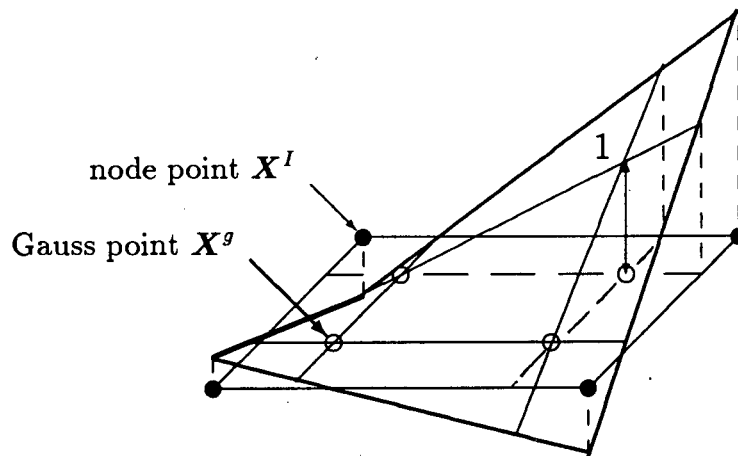


Figure 9.2: Basis function for the approximation of internal variables

In Figure 9.2 we show an example of how the basis functions $\phi^{(g)}$ are constructed for a two-dimensional bilinear element.

We collect the internal variable degrees of freedom at each Gauss point into a vector $\mathbf{p}^{(g)}$. There are nine internal variable degrees of freedom at each point:

$$\mathbf{p}^{(g)} = \begin{pmatrix} \Delta F_{11}^p \\ \Delta F_{12}^p \\ \Delta F_{13}^p \\ \Delta F_{21}^p \\ \Delta F_{22}^p \\ \Delta F_{23}^p \\ \Delta F_{31}^p \\ \Delta F_{32}^p \\ \Delta F_{33}^p \end{pmatrix} = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \\ p^4 \\ p^5 \\ p^6 \\ p^7 \\ p^8 \\ p^9 \end{pmatrix}. \quad (9.11)$$

On each element we approximate the components of \mathbf{q} and $\Delta \mathbf{F}^p$ by

$$q_{\alpha A} = M_{\alpha A}^{(g)R}(\mathbf{X}) q^{(g)R} \quad \text{and} \quad \Delta F_{\alpha A}^p = M_{\alpha A}^{(g)R}(\mathbf{X}) p^{(g)R} \quad (9.12)$$

where summation on g and R is implied, $g = 1, \dots, G^e$ and $R = 1, \dots, 9$. The nine nonzero elements of $M^{(g)}$ take the value $\phi^g(\mathbf{X})$; $M^{(g)}$ does not have a simple matrix representation.

Similarly, on each element we approximate μ and λ by

$$\mu = \phi^g(\mathbf{X}) \mu^g \quad \text{and} \quad \lambda = \phi^g(\mathbf{X}) \lambda^g. \quad (9.13)$$

We recall from Section 9.2 that the quantities μ^g and λ^g are the values of μ and λ at the points \mathbf{X}^g ; because of the simplicity of this approximation we have not used new symbols for these quantities and shall refer to them as Lagrange multipliers when discussing the discrete problem.

The total number of integration points on Ω_0 is $G = E \times G^e$. As we did for the element displacement degrees of freedom, we collect nodal degrees of freedom into a global vector. Unlike the assembly of element displacement quantities, the collection of quantities associated with internal variables into a global vector is reasonably straightforward. In

a manner similar to that used for the displacement degrees of freedom, we shall use the notation p^R for the components of both \mathbf{p} and $\mathbf{p}^{(g)}$, again assuming that the range of the superposed indices is clear from the context; $R = 1, \dots, 9$ for Gauss points and $R = 1, \dots, M$ globally; M has a maximum value of $9G$.

We note that in the discrete case Ω^P is defined as the set of all points \mathbf{X}^g at which yielding takes place and we therefore put $M = 9G^P$, where G^P is the number of Gauss points in Ω^P . At points \mathbf{X}_g in Ω^E associated components of \mathbf{p} are zero so M represents the number of unknown internal variable values. We remark that the division of Ω_0 into elastic and plastic regions, Ω^E and Ω^P , is independent of its division into elements Ω^e .

For the purposes of the following discussion we also collect the Lagrange multipliers λ^g into a vector $\boldsymbol{\lambda} \in \mathbb{R}^{G^P}$. We remark, however, that because of the way we have constructed the finite element approximation in practice we shall not need more than one component of this vector at any given time.

9.3 The discrete form of the incremental boundary problem

We now substitute the finite element approximations into the variational form of the incremental problem. We use the relationships

$$\mathbf{F}_{n+\alpha}(\mathbf{a}) = \mathbf{F}_n + \alpha(\mathbf{B}_{iA}^I a^I) \mathbf{e}_i \otimes \mathbf{E}_A, \quad (9.14)$$

$$\mathbf{F}_{n+\alpha}^p(\mathbf{p}) = \mathbf{F}_n^p + \alpha(\mathbf{M}_{\alpha A}^R p^R) \mathbf{E}_\alpha \otimes \mathbf{E}_A, \quad (9.15)$$

and can therefore write

$$\mathbf{F}_{n+\alpha}^e(\mathbf{a}, \mathbf{p}) = \mathbf{F}_{n+\alpha}(\mathbf{a}) \mathbf{F}_{n+\alpha}^{p-1}(\mathbf{p}) \quad (9.16)$$

and

$$J_{n+\alpha}^p(\mathbf{p}) = \det(\mathbf{F}_{n+\alpha}^p(\mathbf{p})). \quad (9.17)$$

In the approximation of the variational form of the problem the incremental equilibrium equation (8.54) becomes

$$\sum_{e=1}^E b^I \int_{\Omega^e} (\mathbf{B}_{iA}^{(e)I} P_{iA}(\mathbf{a}, \mathbf{p}) + \mathbf{N}_i^{(e)I} b_i) d\mathbf{X} = 0 \quad (9.18)$$

where b_i are components of the body force, and assuming $\Gamma_{0,\tau} = \emptyset$; on Ω_P the evolution equation (8.55) becomes

$$\sum_{g=1}^{G^P} q^{(g)R} \int_{\Omega^P} \mathbf{M}_{\alpha A}^{(g)R} \left(-\frac{F_{i\alpha}^e P_{iA}}{k} - \lambda J^p F_{\alpha A}^{p-1} + \frac{F_{\alpha C}^{p-1} \Delta F_{\beta C}^p F_{\beta A}^{p-1}}{|\mathbf{D}^p|} \right) d\mathbf{X} = 0 \quad (9.19)$$

and the constraint equation (8.56) becomes

$$\sum_{g=1}^{G^P} \mu^g \int_{\Omega^P} \phi^g (J_{n+\alpha}^p - 1 + \epsilon \lambda) b f dX = 0. \quad (9.20)$$

These must hold for any values of b^I , q^R and μ^g ; hence for given Ω^P , and therefore M , the discrete form of the problem becomes:

Find vectors $\mathbf{a} \in \mathbf{R}^N$, $\mathbf{p} \in \mathbf{R}^M$, $\lambda \in \mathbf{R}^{G^P}$ which satisfy

$$\mathcal{F}^I(\mathbf{a}, \mathbf{p}) = 0 \quad I = 1, \dots, N, \quad (9.21)$$

$$\mathcal{G}^R(\mathbf{a}, \mathbf{p}, \lambda) = 0 \quad R = 1, \dots, M, \quad (9.22)$$

$$\mathcal{C}_\epsilon^g(\mathbf{p}, \lambda) = 0 \quad g = 1, \dots, G^P, \quad (9.23)$$

where \mathcal{F}^I , \mathcal{G}^R and \mathcal{C}_ϵ^g are coupled systems of equations obtained by using (9.14), (9.15), (9.16), and (9.17) in the above approximations of the variational equations derived in Chapter VIII.

On inspection it is found that for fixed \mathbf{a} the evolution equations associated with a particular Gauss point are independent of the plastic degrees of freedom and Lagrange multipliers associated with other Gauss points. This is a consequence of simultaneously using Gaussian integration and choosing the Gauss points as the *node points* in the finite element approximation of the internal variables and Lagrange multipliers. In the corrector phase of the solution algorithm described in Chapter X we fix the displacement degrees of freedom and can therefore exploit this feature of the finite element approximation.

It can easily be shown that the discrete incremental problem derived here is equivalent to one obtained from the finite element approximation of an incremental problem in which the constitutive relations are imposed pointwise, at integration points. In this case only a finite element approximation of displacements need be considered. A discretisation of plastic deformations, identical to that used here, is implied since the plastic deformations enter the equilibrium equation only through the stress at Gauss points. The evaluation of stresses involves solving the evolution equation, with fixed displacements, for the increments in plastic deformation at each Gauss point; since the plastic volume constraint must be imposed, a Lagrange multiplier would have to be introduced at each of these points.

We remark that a pointwise imposition of the constitutive relations is the approach used by most authors. This approach seems more natural when attention is restricted to a stress-space formulation of the elastoplastic constitutive relations where the evolution equations are expressed in terms of a yield function. Casting the evolution equations in variational form, as we have done, follows naturally from the use of evolution equations based on the dissipation function. While there is no quantitative difference in the two approaches there may be some benefit to be derived in using the approach which we have followed when considering qualitative aspects of the problem; see Chapter XII for further discussion of this possibility.

In discussing the application of the finite element method we have maintained a level of generality which does not restrict the type of problems for which the formulation of the discrete problem holds. In Chapter XI we have tackled only two-dimensional axisymmetric problems. We have used 4-noded quadrilateral elements for which bilinear basis functions are used; a full description of these elements may be found, for example in [24,45] or [46]. We use reduced integration on these elements for which only a single Gauss point is required. We discuss the use of reduced integration in Chapter XI.

CHAPTER X

SOLUTION ALGORITHMS

10.1 Introduction

The development of algorithms to solve finite element approximations of problems involving plastic deformations has been the subject of much research effort in recent years; [9,36,39,81,80,87,99,107,111,118] are just a few of many contributions in this area. The Newton–Raphson iterative procedure is widely used as a basis for constructing these algorithms. The determination of elastic and plastic regions within the body at time $t_{n+\alpha}$ is an important part of the solution to the incremental problem. Given a division of the domain into such regions, the incremental boundary–value problem consists of finding displacements and plastic internal variables which satisfy the variational equations listed in the summary of the last section of Chapter VIII. In the finite element approximation of this problem these equations are replaced by large coupled systems of nonlinear equations involving a finite number of displacement and internal variable values; Ω^P is defined as a set of Gauss points.

We follow an approach which is widely used and construct an algorithm for solving the discrete incremental problem by incorporating a *predictor–corrector scheme* into each iteration of the standard Newton–Raphson procedure. The purpose of this predictor–corrector scheme is to find the division of the domain into elastic and plastic regions and hence to identify the internal variable increments and evolution equations to be considered in calculating improvements to an estimate of the solution in each iteration.

We begin our discussion of solution algorithms by giving a brief description of the Newton–Raphson method. We then outline the predictor–corrector scheme used in each iteration. We discuss both the predictor and corrector phases in some detail and include the derivation of tangent matrices used in the solution of the particular discrete problem we have formulated. As part of this discussion we show how the Lagrange multipliers introduced in Chapter VIII are eliminated in the construction of a consistent tangent

matrix.

In summary we present an algorithm which can be used to solve the discrete form of the variational incremental boundary-value problem we have formulated. We use this algorithm to solve a number of examples. The implementation and effectiveness of the algorithm are discussed along with the presentation of the results of numerical examples in Chapter XI.

10.2 The Newton–Raphson procedure

The Newton–Raphson iterative procedure is a well-established method for solving systems of nonlinear equations. We outline the general procedure in order to illuminate the essential features of the method and to establish some of the terminology we use later. The method, its derivation and characteristics are fully described in many texts dealing with numerical methods, for example [117,122,94].

We consider the general problem of finding the solution $\mathbf{a} \in \mathbf{R}^N$ to a given system of nonlinear equations $\mathcal{S}^I(\mathbf{a}) = \mathbf{0}$, ($I = 1, \dots, N$). The iterations of the Newton–Raphson procedure produce a sequence of estimates $\{\mathbf{a}^{(i)}\}$ of the solution \mathbf{a} . The procedure is started by making a guess $\mathbf{a}^{(0)}$ at the solution; provided this initial estimate is a suitable one, successive estimates converge towards the solution \mathbf{a} as i increases. In many situations it is convenient to use zero as this *starting value*. A good approximation to \mathbf{a} is therefore found by following the steps outlined below:

Step 1

Find the components of a vector $\hat{\mathbf{a}}$ by solving a system of N equations of the form

$$\mathbf{K}\hat{\mathbf{a}} = \mathbf{r} \tag{10.1}$$

where the *tangent matrix*, \mathbf{K} , and *residual vector*, \mathbf{r} have components

$$K^{IJ} = \left. \frac{\partial \mathcal{S}^I}{\partial a^J} \right|_{\mathbf{a}^{(i-1)}} \quad \text{and} \quad r^I = -\mathcal{S}^I(\mathbf{a}^{(i-1)}) . \tag{10.2}$$

The approximate solution $\mathbf{a}^{(i-1)}$ is that found in the last iteration.

Step 2

Obtain an improved value for the solution as:

$$\mathbf{a}^{(i)} = \mathbf{a}^{(i-1)} + \hat{\mathbf{a}} . \quad (10.3)$$

This step is often referred to as *correcting* the solution; we remark that the calculations made in the *corrector phase* of the algorithms described below differ considerably from the simple improvement indicated here.

Step 3

Check whether the improved solution satisfies the original equations, at least to within some specified tolerance *tol*. A typical *convergence criterion* is

$$\max |S^I(\mathbf{a}^{(i)})| < tol . \quad (10.4)$$

If this criterion is not satisfied the iteration counter i is incremented ($i = i + 1$) and the procedure is repeated from Step 1. If however the criterion is satisfied, the solution \mathbf{a} is set equal to $\mathbf{a}^{(i)}$ and the iterative procedure is stopped.

10.3 The predictor–corrector scheme

In the discrete form of the incremental problem considered here the full system of equations to be solved is divided into three distinct parts: the equilibrium equations (9.21), the evolution equations (9.22) and the constraint equations (9.23). In each iteration of the Newton–Raphson iterative solution procedure which is used to solve these equations we use a predictor–corrector scheme to ensure that the appropriate evolution equations are considered. We require an estimate or *prediction* of Ω^P before any attempt at finding a solution can be made. The solution algorithm which we describe is constructed in such a way that this prediction can be improved, or *corrected*, at every iteration.

In practice, in each iteration the displacements $\mathbf{a}^{(i)}$ are calculated in the predictor phase of the algorithm, while increments in plastic internal variables, $\mathbf{p}^{(i)}$, are calculated in

the corrector phase. Only the evolution equations at Gauss points which are predicted to be yielding are considered in the corrector phase.

We discuss separately the methods we have used in the predictor and corrector phases of each Newton iteration. In our discussion of the predictor phase we concentrate on the derivation of a *consistent predictor*. We describe in some detail the corrector algorithm which we use. We remark that the corrector phase is not affected by the choice of predictor. The algorithm is summarised at the end of this Chapter in Figures 10.1 and 10.2.

10.3.1 The predictor phase

We begin each iteration of the global Newton–Raphson solution procedure with a prediction of Ω^P ; this makes it possible to identify the evolution equations which are to be used to find an improvement to the estimate of the solution obtained in previous iterations. Initially in each increment we assume that $\Omega^P = \emptyset$. In subsequent iterations the components of $\mathbf{p}^{(i-1)}$ at each Gauss point are checked, and if they are nonzero the point is included in Ω^P . Setting $\Omega^P = \emptyset$ initially implies that $\mathbf{p}^{(0)} = \mathbf{0}$ and therefore that the deformation taking place is purely elastic. At Gauss points which are not in the predicted plastic region we discard the corresponding evolution equations and fix the components of \mathbf{p} equal to zero; constraint equations are not required at these points and therefore the corresponding Lagrange multipliers can also be eliminated. At Gauss points in Ω^P we combine the evolution and constraint equations in the way described later, in Section 10.3.2, when discussing the corrector algorithm. By doing this we eliminate the remaining Lagrange multipliers. We can then write the equations to be solved in the first step of the global Newton–Raphson procedure as

$$\left[\begin{array}{c|c} \mathbf{K} & \mathbf{L} \\ \hline \mathbf{R} & \mathbf{H}^* \end{array} \right] \begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g}^* \end{pmatrix} \quad (10.5)$$

where the components of the matrices \mathbf{K} , \mathbf{L} and \mathbf{R} and the vector \mathbf{f} are given by

$$K^{IJ} = \left. \frac{\partial \mathcal{F}^I}{\partial a^J} \right|_{\mathbf{a}=\mathbf{a}^{(i-1)}, \mathbf{p}=\mathbf{p}^{(i-1)}}, \quad (10.6)$$

$$L^{IS} = \left. \frac{\partial \mathcal{F}^I}{\partial p^S} \right|_{\mathbf{a}=\mathbf{a}^{(i-1)}, \mathbf{p}=\mathbf{p}^{(i-1)}}, \quad (10.7)$$

$$R^{RJ} = \left. \frac{\partial \mathcal{G}^R}{\partial \alpha^J} \right|_{\mathbf{a}=\mathbf{a}^{(i-1)}, \mathbf{p}=\mathbf{p}^{(i-1)}}, \quad (10.8)$$

$$(10.9)$$

and

$$f^I = -\mathcal{F}^I(\mathbf{a}^{(i-1)}, \mathbf{p}^{(i-1)}), \quad (10.10)$$

with $I, J = 1, \dots, N$ and $R, S = 1, \dots, M$; the components of \mathbf{H}^* and \mathbf{g}^* are defined in Section 10.3.2.

The objective of the predictor phase of the algorithm is to obtain only the displacement increments $\mathbf{a}^{(i)}$ and it is unnecessary to find the unknown components of $\hat{\mathbf{p}}$ to achieve this. We therefore eliminate the unknowns $\hat{\mathbf{p}}$ from the system of global equations. This condensation process results in a system of equations which have the form

$$\mathbf{K}^* \hat{\mathbf{a}} = \mathbf{f}^* \quad (10.11)$$

where

$$\mathbf{K}^* = \mathbf{K} - \mathbf{L} \mathbf{H}^{*-1} \mathbf{R} \quad (10.12)$$

and

$$\mathbf{f}^* = \mathbf{f} + \mathbf{L} \mathbf{g}^*. \quad (10.13)$$

The matrix \mathbf{K}^* is known in the computational plasticity literature as the *consistent tangent matrix*. The second term in the residual can be dropped since at convergence in the corrector phase of the previous iteration we require the components of \mathbf{g}^* to be near zero. We solve this reduced system of equations for the components of $\hat{\mathbf{a}}$ and hence obtain $\mathbf{a}^{(i)}$; this concludes the predictor phase of the algorithm.

In theory we could calculate $\hat{\mathbf{p}}$ from the solution of (10.11) using

$$\hat{\mathbf{p}} = -\mathbf{H}^{*-1} \mathbf{R} \hat{\mathbf{a}} ; \quad (10.14)$$

however, except under some special circumstances, the result would not be compatible with the assumed division of the domain into elastic and plastic regions used in obtaining the matrices in this expression. If (10.14) were used, nonzero values of $\mathbf{p}^{(i)}$ may be obtained at points which, in an improved prediction, should be in Ω^E , and yielding might be indicated at Gauss points where the components of \mathbf{p} have been set to zero. To avoid these problems it is clear that, before calculating new estimates for \mathbf{p} , the division of the domain into elastic and plastic regions must be re-examined. This process forms part of the corrector procedure which is described below.

Algorithms based on an *elastic* predictor may be viewed as a simplification of the consistent predictor approach. In the predictor phase of this type of algorithm \mathbf{p} is fixed equal to $\mathbf{p}^{(i-1)}$ and therefore only the equilibrium equations need be considered. Of course $\mathbf{p}^{(i-1)}$ reflects a definition of Ω^P , hence the same prediction of plastic region as that used in the consistent predictor is implied. The tangent matrix involves only the elastic part of the constitutive law and hence $\mathbf{a}^{(i)}$ is called an *elastic prediction* of the displacements. We remark that using an elastic predictor amounts to approximating the consistent tangent matrix by setting $\mathbf{K}^* = \mathbf{K}$.

10.3.2 The corrector phase

As indicated, the objective of the corrector phase is to find values for the components of $\mathbf{p}^{(i)}$ which satisfy the evolution and constraint equations (9.22) and (9.23) for given displacements. Since the evolution equations apply only on Ω^P our first task is to make the best possible prediction of Ω^P . To do this we evaluate the conjugate forces, \mathbf{Z} , using the predicted displacements $\mathbf{a}^{(i)}$ and $\mathbf{p} = \mathbf{0}$; we then check the yield function, $f(\mathbf{Z})$, to obtain a prediction of Ω^P . In this way each prediction of the displacement solution provides a new prediction of Ω^P . In practice these calculations are done at each Gauss point and where yielding is indicated the point is taken to be in Ω^P . As we have indicated in Section 10.3.1 the prediction of displacements depends on $\mathbf{p}^{(i-1)}$. Other than through this influence on the determination of the displacements the corrected solution $\mathbf{p}^{(i)}$ is

obtained independently of estimates of \mathbf{p} and Ω^P found in previous iterations.

In discrete form the evolution equations associated with different Gauss points are coupled only because of the presence of unknown displacements, while the constraint equation associated with a given Gauss point is coupled only to the evolution equations associated with that point. Since in the corrector phase we fix the displacement degrees of freedom we can exploit this property of our finite element discretisation and solve for the components of $\mathbf{p}^{(i)}$ Gauss point by Gauss point.

Thus far the corrector procedure we propose is essentially the same as many others that are described in the literature. It is in solving for $\mathbf{p}^{(i)}$ at each Gauss point that the algorithm differs from the typical return mapping procedures commonly used.

With \mathbf{a} fixed equal to $\mathbf{a}^{(i)}$ we have associated with each Gauss point \mathbf{X}^g in Ω^P nine evolution equations and one constraint equation. We use a standard Newton–Raphson procedure to find the values of λ and the nine components of $\mathbf{p}^{(i)}$ which satisfy each of these small systems of equations. So as to distinguish between the iterations (i) of the global Newton procedure we use (j) as the iteration counter within the corrector algorithm. We drop the global iteration counter and write in place of $\mathbf{p}^{(i,j)}$, $\mathbf{p}^{(j)}$ where $\mathbf{p}^{(i,j)}$ is the value of \mathbf{p} at the j^{th} iteration of corrector calculation in the i^{th} iteration of the global algorithm.

In place of (10.1) in the first step of the Newton–Raphson procedure we have to solve the system of equations

$$\left[\begin{array}{c|c} \mathbf{H} & \mathbf{P} \\ \hline -\mathbf{Q} & \epsilon \end{array} \right] \begin{pmatrix} \hat{\mathbf{p}} \\ \hat{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{g} \\ c \end{pmatrix} \quad (10.15)$$

where

$$\mathbf{H}^{RS} = \left. \frac{\partial \mathcal{G}^R}{\partial p^S} \right|_{\mathbf{a}=\mathbf{a}^{(i)}, \mathbf{p}=\mathbf{p}^{(j-1)}, \lambda=\lambda^{(j-1)}} \quad (10.16)$$

$$\mathbf{P}^R = \left. \frac{\partial \mathcal{G}^R}{\partial \lambda} \right|_{\mathbf{p}=\mathbf{p}^{(j-1)}}, \quad (10.17)$$

$$\mathbf{Q}^S = \left. \frac{\partial \mathcal{C}_\epsilon}{\partial p^S} \right|_{\mathbf{p}=\mathbf{p}^{(j-1)}}, \quad (10.18)$$

$$g^R = -\mathcal{G}^R(\mathbf{a}^{(i)}, \mathbf{p}^{(j-1)}, \lambda^{(j-1)}), \quad (10.19)$$

$$c = -\mathcal{C}_\epsilon(\mathbf{p}^{(j-1)}, \lambda^{j-1}) \quad (10.20)$$

and $R, S = 1, \dots, 9$.

We remark that in this system of equations the *tangent matrix* is in general full, non-symmetric and in the case $\epsilon = 0$ has a zero on the diagonal; these features must be taken into account when choosing a method to solve the equations (see Chapter XI).

If we use the perturbed constraint, that is if $\epsilon \neq 0$, we can condense this system of equations by eliminating $\hat{\lambda}$. To do this we use the last equation to express the unknown $\hat{\lambda}$ in terms of the unknown components of $\hat{\mathbf{p}}$ as

$$\hat{\lambda} = -\frac{1}{\epsilon} \mathbf{Q}^S \hat{\mathbf{p}}^S + \frac{1}{\epsilon} c. \quad (10.21)$$

By substituting for $\hat{\lambda}$ using this expression we reduce the system of equations to the form

$$\mathbf{H}^* \hat{\mathbf{p}} = \mathbf{g}^* \quad (10.22)$$

where

$$\mathbf{H}^* = \mathbf{H} - \mathbf{P} \frac{1}{\epsilon} \mathbf{Q} \quad \text{and} \quad \mathbf{g}^* = \mathbf{g} - \frac{1}{\epsilon} \mathbf{P} c \quad (10.23)$$

Here we remark that since the constraint is a nonlinear one it is important to retain the second contribution to the residual vector of the reduced system; this term is dropped in many formulations where a similar condensation is carried out (see for example [46]).

Our prime motivation for condensing the system is to make it possible to construct the consistent tangent used in the predictor phase of the algorithm. (Recall that we made use of \mathbf{H}^* in the construction of the consistent tangent matrix).

Some care is required in choosing starting values for \mathbf{p} in this procedure. We cannot make what might seem to be the natural choice, $\mathbf{p}^{(0)} = \mathbf{0}$, since the tangent matrix

defined above is then singular. We have found that a reasonably robust method for starting the procedure is to use $\mathbf{p}^{(0)} = \mathbf{0}$ but at the same time, so as to avoid having a singular tangent, set $|\text{dev } \mathbf{D}^p| = \delta$ where δ is a small number (not less than *tolcor*).

An alternative is to use $\mathbf{p}^{(i-1)}$ as the starting value; this can only be used when $i > 1$ and when $\mathbf{p}^{(i-1)} \neq \mathbf{0}$. In general this alternative provides an advantage in that fewer and fewer iterations of the corrector algorithm are required each time the corrector is applied. There are unfortunately circumstances where convergence problems arise, for example when

$$f(\mathbf{Z}(\mathbf{a}^{(i)}, \mathbf{p}^{(i-1)})) \ll 1$$

or when

$$f(\mathbf{Z}(\mathbf{a}^{(i)}, \mathbf{p}^{(i-1)})) \gg f(\mathbf{Z}(\mathbf{a}^{(i)}, \mathbf{0})) > 1 ,$$

essentially when the value of \mathbf{p} obtained in the last iteration is not close to the solution. In the implementation of our algorithm we use both methods; when using the second method we avoid convergence problems by reverting to the first if either of the above conditions arises.

10.4 Remarks

An elastic tangent matrix is computationally much cheaper to construct than the full consistent tangent matrix, hence the early interest in using this approximation. The major disadvantage in using an elastic tangent is that it adversely effects the rate of convergence of the overall algorithm; this is well known and is discussed by a number of authors, see for example [9,89,111]. The improved rate of convergence achieved with a consistent predictor means that far fewer iterations are required to achieve convergence; this compensates for the extra cost of each iteration.

In general the consistent tangent matrix \mathbf{K}^* is nonsymmetric. In the case of small-strain plasticity based on a von Mises yield criterion the equilibrium equation involves

the full stress tensor while the evolution equation involves only the deviatoric part of the stress. Using the simple approach outlined above to construct K^* from the discrete forms of these equations a nonsymmetric form of the matrix is obtained. The lack of symmetry is simply a result of the assumption of isochoric plastic flow since it is this assumption which leads to the deviatoric stress appearing in the evolution equations. A symmetric consistent tangent matrix is, however, easily constructed for this case; the *B-bar* method [46] is commonly used to achieve this. This method makes use of a decomposition of strain (and due to the linearity of the problem also of stress) into spherical and deviatoric parts. This decomposition is used in the equilibrium equation and since by definition the plastic strain is deviatoric it is found that $R = L^T$; the symmetry of K^* then follows from the fact that H is symmetric in this case. A *B-bar* method has also been used in the finite-strain case, but in this case lack of symmetry in the consistent tangent matrix arises for reasons other than simply the assumption of isochoric plasticity. The equilibrium equation involves tensor quantities defined with respect to current and reference configurations, the gradient of \mathbf{v} and the first Piola-Kirchhoff stress, while the evolution equation involves either only current configuration quantities or, in the form we have proposed, tensor quantities defined with respect to the intermediate and reference configurations. The inclusion of the constraint on the plastic spin also adds to the lack of symmetry. The cost of solving a large nonsymmetric system of equations is usually considerably greater than that of solving a symmetric system of the same size; hence, it is important that a symmetric tangent be used. Symmetry can be restored, albeit on a somewhat *ad hoc* basis, by simply dropping the skew part of K^* ; this approach has been used by others (see for example Simo [107]). We have used this approximation and have found, through numerical experiment, that it has no detrimental effect on the solution procedure.

10.5 Summary

We conclude by presenting the complete algorithm for the solution of the discrete incremental problem that we have derived; in Figure 10.1 we show the main steps in procedure and in Figure 10.2 we show the corrector algorithm.

Increment loop: $n = 0$

$n = n + 1$

Newton–Raphson Iteration Loop:

- Initialise: set $i = 0$, $\mathbf{a}^{(0)} = \mathbf{0}$, and $\mathbf{p}^{(0)} = \mathbf{0}$
- Residual: $\mathbf{f}^{*(0)} = \{-F^I(\mathbf{a}^{(0)}, \mathbf{p}^{(0)})\}$

$i = i + 1$

PREDICTOR PHASE:

- Consistent Tangent Stiffness: $\mathbf{K}^*(\mathbf{a}^{(i-1)}, \mathbf{F}_n^{p(i-1)}, \mathbf{p}^{(i-1)})$
- Solve: $\hat{\mathbf{a}} = \mathbf{K}^{*-1} \mathbf{f}^{*(i-1)}$
- Improve displacement solution: $\mathbf{a}^{(i)} = \mathbf{a}^{(i-1)} + \hat{\mathbf{a}}$

CORRECTOR PHASE:

- Find internal variable increments $\mathbf{p}^{(i)}$ (see Figure 10.2)
- Residual: $\mathbf{f}^{*(i)} = \{-\mathcal{F}^I(\mathbf{a}^{(i)}, \mathbf{p}^{(i)})\}$
- Convergence Check: *maximum residual*
 If $\max |f^{*I(i)}| > tol$ then go to next iteration
 else set $\mathbf{a} = \mathbf{a}^{(i)}$ and $\mathbf{p} = \mathbf{p}^{(i)}$

end Newton–Raphson loop

Increment Update:

$$\mathbf{u}_n = \mathbf{u}_{(n-1)} + \Delta \mathbf{u}(\mathbf{a}), \quad \mathbf{F}_n^p = \mathbf{F}_{n-1}^p + \Delta \mathbf{F}^p(\mathbf{p}) \quad \text{and} \quad \gamma_n = \gamma_{n-1} + \Delta \gamma(\mathbf{p})$$

end increment loop

Figure 10.1: The solution algorithm for the discrete incremental problem

At each Gauss point in turn:

- Evaluate $\mathbf{F}_{n+\alpha}^{(i)}$, set $\mathbf{F}_{n+\alpha}^p = \mathbf{F}_n^p$ and $\gamma_{n+\alpha} = \gamma_n$
and hence calculate the predicted conjugate forces $\mathbf{Z}_{n+\alpha}$
- Check for yielding:

If $F(\mathbf{Z}) < 1$ then

set $\Delta\mathbf{F}^p = \mathbf{0}$ (stored as components of $\mathbf{p}_n^{(i)}$) and go to next Gauss point
else:

Newton-Raphson Iteration Loop

- Initialise: set $j = 0$
Set starting value $\Delta\mathbf{F}^{p(0)}$ with components obtained from $\mathbf{p}_n^{(i-1)}$
- Calculate quantities:
 $\mathbf{F}_{n+\alpha}^{p(j)} = \mathbf{F}_n^p + \alpha\Delta\mathbf{F}^{p(j)}$, $\mathbf{F}_{n+\alpha}^{e(j)} = \mathbf{F}_{n+\alpha}^{(i)}\mathbf{F}_{n+\alpha}^{p(j)-1}$
and their inverses, determinants etc. hence also $\mathbf{D}_{n+\alpha}^p$, $|\text{dev}\mathbf{D}^p|$. (*)
- Set $\lambda^{(0)}$ and $\Delta\gamma = |\text{dev}\mathbf{D}^p|$. If $|\text{dev}\mathbf{D}^p|$ is zero set it to a small number, δ .
- Residual: $\mathbf{g}^{*(0)} = \{-\mathcal{G}^R(\mathbf{a}^{(i)}, \mathbf{p}^{(0)}, \lambda^{(0)})\}$

$j = j + 1$

- Form the Tangent Matrix: $\mathbf{H}^*(\mathbf{F}_{n+\alpha}^{e(j-1)}, \mathbf{F}_{n+\alpha}^{p(j-1)}, \dots)$
- Solve: $\hat{\mathbf{p}} = \mathbf{H}^{*-1}\mathbf{g}^{*(j-1)}$
- Correct: set $\hat{\lambda} = -\frac{1}{\epsilon}\mathbf{Q}\hat{\mathbf{p}}$ then $\mathbf{p}^{(j)} = \mathbf{p}^{(j-1)} + \hat{\mathbf{p}}$ and $\lambda^{(j)} = \lambda^{(j-1)} + \hat{\lambda}$
- Recalculate all quantities (*)
- Residual: $\mathbf{g}^{*(j)} = \{-\mathcal{G}^R(\mathbf{a}^{(i)}, \mathbf{p}^{(j)}, \lambda^{(j)})\}$
- Convergence Check: *maximum residual*
If $\max |\mathbf{g}^{*R(j)}| > \text{tolcor}$ then next iteration
else set $\mathbf{p} = \mathbf{p}^{(j)}$ and $\Delta\gamma = |\text{dev}\mathbf{D}^p|$

end Newton-Raphson loop

- Update internal variables increments: $\mathbf{p}_n^{(i)} = \mathbf{p}$ next Gauss point

Figure 10.2: The corrector algorithm

CHAPTER XI

NUMERICAL EXAMPLES

11.1 Introduction

The quest for accurate and efficient methods for numerically modelling the elastoplastic behaviour for both small and large deformations has been the focus of considerable research effort in recent times. A number of different material models have been proposed and several approaches have been used to formulate discrete approximations to the problems of interest. Many important developments to both theory and solution algorithms have resulted from experience gained through attempts to solve practical problems.

In this thesis we have proposed a new internal variable theory of finite-strain plasticity which, from a theoretical point of view, offers some pleasing features. This theory would, however, be of little value if it could not be used to solve practical problems. We have shown how the theory is used, in conjunction with the generalised midpoint rule and the finite element method, to obtain a discrete incremental form of a class of initial-boundary-value problems. We have proposed an algorithm for the solution of these discrete problems. It remains to show how the algorithm performs in practice.

In the literature concerning the solution of problems in plasticity a variety of example problems have been used to test theories and solution procedures which have been proposed. A reasonably widely accepted set of test examples involving finite plastic deformations has evolved. We have selected two of these, namely, *expansion of a thick walled cylinder* and *upsetting of an axisymmetric billet*. Solutions to both these problems obtained by other researchers are reported in the literature (see for example [81,107] and [118]).

We have implemented the algorithm described in Chapter X in a new version of NOSTRUM, the in-house finite element code of The FRD/UCT Centre for Research in Computational and Applied Mechanics at The University of Cape Town. We begin this

Chapter by briefly describing the NOSTRUM finite element code; we indicate the extent of development which has been undertaken in order to incorporate the new algorithm. We do not describe the implementation of the algorithm in detail; we do, however, include discussion of some particular aspects which have a bearing on the discussion of results. We then describe in some detail the two example problems we have selected. The solutions we have obtained are presented and compared with those obtained by others. Using the thick-walled cylinder example we show how the use of the midpoint rule affects the solution procedure. Using the upsetting example we show that the algorithm can give good results with large time steps. An indication of the performance of the algorithm is included. In the final section of this Chapter we summarise the conclusions made in respect of numerical aspects of the thesis. We make some general remarks about the practical application of our theory, and comment on some of the specific problems encountered. We also indicate ways in which we believe the algorithm might be improved and suggest further investigations which would be useful.

11.2 The NOSTRUM finite element code

NOSTRUM, the in-house finite element code of The FRD/UCT Center for Research in Computational and Applied Mechanics, Cerecam (formerly The Applied Mechanics Research Unit, AMRU) was initiated in the early 1980s. The first versions of the code were based on the code described by Owen and Hinton in their well known book [90]. Since these early developments the code has undergone many changes and a variety of special purpose versions have been developed. Recent developments have included the implementation of a reasonably sophisticated data management system and revision of the input phase of the program. The code is severely limited with respect to pre- and post-processing, no plotting procedures are provided and an interface to available pre- and post-processing software has not yet been established.

In the most recent revision the central core of the program has been completely reorganised. A set of user facilities has also been developed; these include user solution procedures, user elements and user materials. A prototype of this major revision has been used as the basis for the implementation of finite-strain algorithms studied in this

thesis. The development of this prototype has been carried out simultaneously with this work.

The NOSTRUM code was previously restricted to the analysis of problems involving small-strains. The basic structure of the code is designed around an incremental iterative solution strategy; this is typical of most nonlinear finite element codes.

The code is written in FORTRAN and development is currently carried out on a Micro Vax 2000. Example problems have been run on a Vax 3100 workstation.

11.3 Imposition of the plastic volume constraint

In preceding Chapters we have described how the plastic volume constraint is applied using a Lagrange multiplier. In the corrector phase of the algorithm the system of equations to be solved at each Gauss point is small; Gaussian elimination is used to solve the equations in each iteration of the local Newton procedure. We remark that it is not necessary to condense out the Lagrange multiplier at this stage in the solution procedure. If the exact form of the constraint (that is with $\epsilon = 0$) is used, to solve the full set of equations given by (10.15) at each Gauss point, a total or partial pivoting Gaussian elimination algorithm is required since a zero appears on the diagonal of the system. Computationally these procedures are expensive. An alternative is to use the perturbed constraint and solve the reduced system of equations (10.22) obtained by condensing out the Lagrange multiplier λ in the manner described in Chapter VIII. Only the unknown components of $\hat{\mathbf{p}}$ are calculated directly; by substituting these in (10.21) the value of $\hat{\lambda}$ is retrieved, and hence, both $\mathbf{p}^{(j)}$ and $\lambda^{(j)}$ can be evaluated in each iteration of the local Newton procedure. In practice, having to retrieve $\hat{\lambda}$ to a large extent cancels the benefit of reducing the size of the already small system of equations to be solved and there is therefore little gain in efficiency to be derived from using this approach. Another alternative is to add the perturbation but to retain the full set of equations. By doing this the zero on the diagonal is eliminated and hence the relatively expensive pivoting procedures can be replaced by straightforward Gauss elimination routines and the values of both $\hat{\mathbf{p}}$ and $\hat{\lambda}$ are obtained directly.

For the construction of the consistent tangent used in the predictor phase of the algorithm it is essential that the perturbed form of the constraint, (8.56), be used. It is assumed that the value of ϵ is small; in practice, if it is chosen to be too small rounding errors cause deterioration in the convergence of the solution procedure. We have found that the predictor phase of the solution process is more sensitive to the value of ϵ than is the corrector phase. There is no reason that the same value of ϵ should be used in both phases of the algorithm; it is in the corrector phase that plastic deformations are evaluated so that it is in this phase of the procedure that the constraint should be applied most accurately. In practice we have used $\epsilon = 10^{-8}$ in the predictor phase and $\epsilon = 10^{-16}$ in the corrector phase. Practical experience indicates that with these values convergence of the solution procedure is only affected when components of the residual are reduced to values of the order of 10^{-8} and below.

In imposing the plastic volume constraint we have not required that $\text{tr } \mathbf{D}^p = 0$. In practice the values calculated for $\text{tr } \mathbf{D}^p$ are small. It has been observed that these values tend to become smaller with a decrease in step size. This observation is made for the case of $\alpha = 1$ in the midpoint rule.

11.4 Reduced integration

We have used only 4-noded bilinear elements. It is essential to use reduced integration schemes for these elements. By full integration we mean a 4 point (2×2) rule, while for reduced integration only a single Gauss point is required. The great advantage of reduced integration is a considerable saving in computational effort. Most of the time-consuming calculations are associated with Gauss points, in particular the calculation of contributions to consistent tangent matrices and the calculation of internal variable increments in the corrector phase of the algorithm; the saving is made simply because fewer Gauss points are required. The reason that it is important to use a reduced integration scheme is, however, that a better approximation of *volume strain* is obtained at the centre of the element than at other Gauss point positions; in our formulation of the discrete problem the volume strain enters the problem as $J = \det \mathbf{F}$, calculated at Gauss points. The overly stiff behaviour of fully integrated quadrilateral elements has

been investigated by Nagtegaal and co-workers [80,81].

The well-known disadvantage of reduced integration is the existence of zero energy modes, also known as *hourglass modes*; a detailed description of this phenomenon may be found, for example, in Hughes [46]. An artificial stiffness can be introduced to counter the development of hourglassing. The basic philosophy used is to allow a small amount of energy to develop due to the addition of the hourglass stiffness. This spurious energy is restricted in size so that it is very small when compared with the total strain energy, and hence, inclusion of hourglass control has negligible effect on the the solution of the problem at hand. A simple method for calculating hourglass stiffnesses for linear problems is described in [46]. The determination of hourglass stiffness for nonlinear problems is less straightforward; it is discussed at length by Flanagan and Belytschko [31]. The method described by Hughes may be adapted to provide hourglass control for nonlinear problems in the way described in [31]. We have implemented this method of hourglass control. We note the recent publication of Belytschko and Bindeman [8] in which the control of hourglass modes for nonlinear problems is discussed and approximation of dilational deformations is given particular attention.

An alternative to simply using reduced integration to overcome problems associated with the calculation of the volumetric component of deformation is the use of enhanced strain elements; this approach is currently under investigation. The use of mixed methods has also received some attention (see [110,112]).

11.5 Numerical examples

11.5.1 EXAMPLE 1: Expansion of a thick walled cylinder

This relatively simple example is widely used to test models for plasticity as well as in testing other material models. An infinitely long thick walled cylinder is subjected to internal pressure. We consider a cylinder with an inner radius of 10mm and outer radius of 20mm. Rather than apply pressure loading we prescribe an increase in the internal radius, from 10 to 85mm; the internal pressure is obtained as the reaction on the inner surface. It is assumed that no extension of the cylinder takes place. The

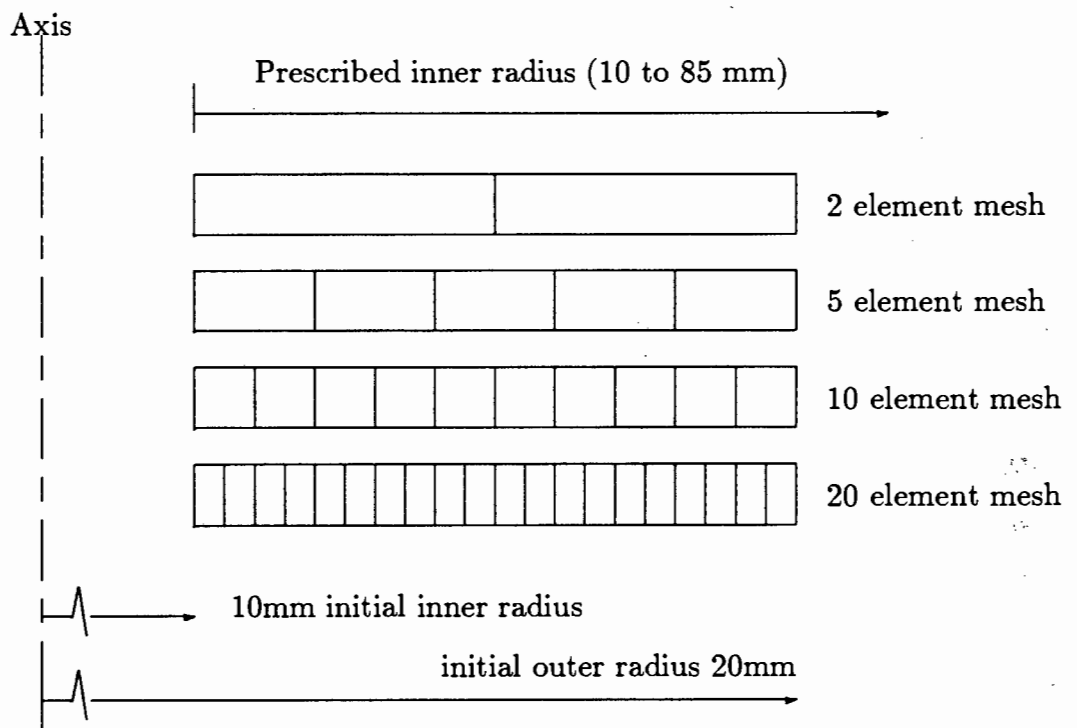


Figure 11.1: Geometry and meshes for Example 1

problem is essentially one-dimensional since all the displacements are radial. However, deformation of the cylinder walls takes place in both radial and tangential directions. These deformations are large: in the radial direction the thickness of the cylinder wall reduces by more than 83% (from 10 to approximately 1.7mm) while in the tangential direction strains of between approximately 77% (at outer surface) and 88% (at the

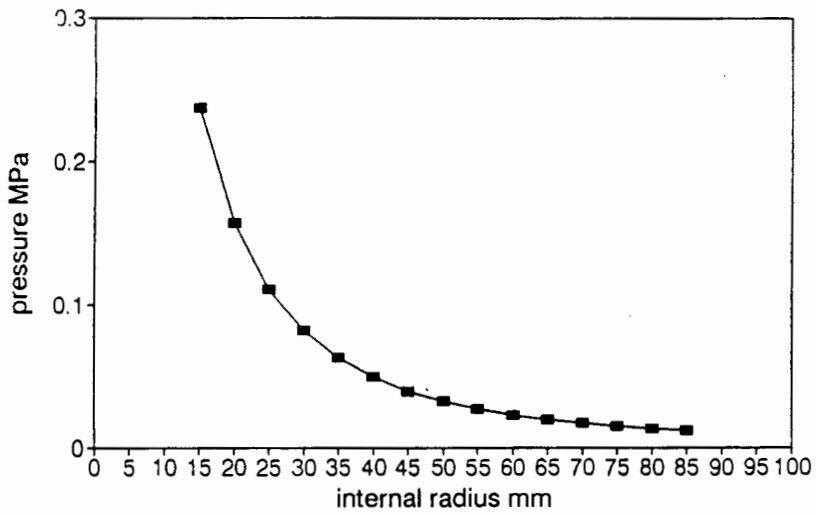
inner surface) occur. We consider the problem as a two-dimensional axisymmetric one and model the cylinder wall using 4-noded bilinear isoparametric elements; reduced integration is used on these elements. We have used meshes of 2, 5, 10 and 20 elements. The geometry of the problem and these finite element meshes are shown in Figure 11.1. The displacements computed using all four meshes are the same; the stress results are not as good for the coarser meshes; this is particularly true for the internal pressure, which is obtained by an extrapolation of the Gauss point radial stress.

Elastic properties:	
Shear modulus	$\mu = 3.8 \text{ GPa}$
Bulk modulus	$K = 40.0 \text{ GPa}$
Perfect plasticity:	
Flow stress	$T_0 = 0.5 \text{ Mpa}$
$(c_0 = \sqrt{\frac{2}{3}}T_0, c_1 = c_2 = \beta = 0)$	

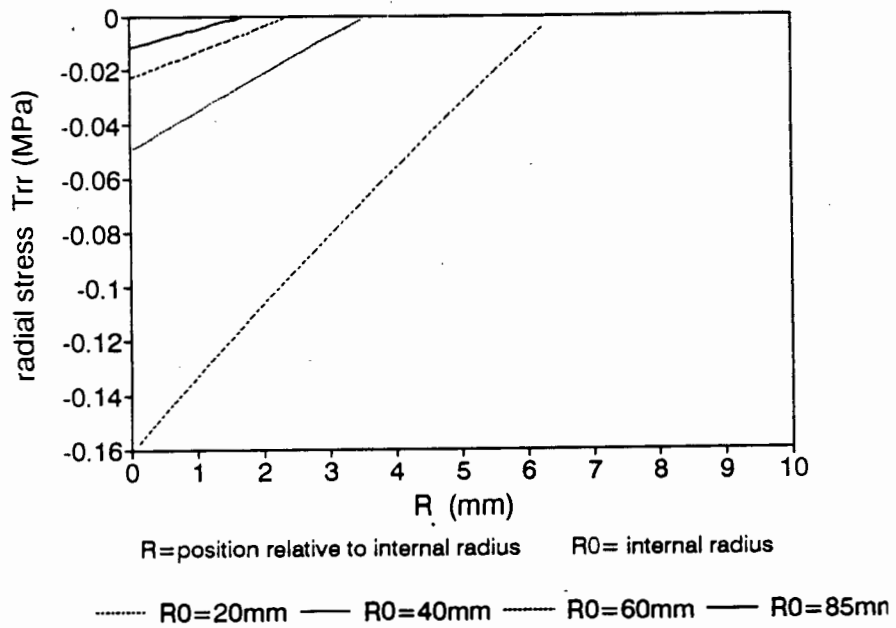
Table 11.1: Material properties for Example 1

With the aim of comparing results with those presented by Simo [107] we concentrate on solutions obtained with the 20-element mesh. We use the same values for the material constants as those used by Simo; these values are given in Table 11.1. These values were chosen so as to replicate rigid-plastic behaviour and therefore allow comparison with an exact solution to the rigid-plastic problem; the exact solution of the rigid-plastic problem is given in [26]. The material is almost incompressible so that $J = J^e \approx 1$. In this case the neo-Hookean hyperelastic model described in Chapter VII is almost identical to that used by Simo.

We first present results obtained with $\alpha = 1$ in the midpoint rule. We show the results obtained using 15 equal increments to achieve the required expansion of the cylinder. In Figure 11.2a we show the relationship between internal radius and internal pressure and in Figure 11.2b we show graphs of radial stress, T_{rr} , versus position relative to internal radius, R , for different values of the internal radius R_0 . All these results ap-



(a) internal pressure vs. inner radius



(b) radial stress vs. position relative to inner radius

Figure 11.2: Results for Example 1 ($\alpha = 1$)

pear to be identical to those obtained by Simo, whose internal pressure results are in agreement with the exact solution. It is not necessary to look at refinement of the time discretisation since the results achieved with 15 large time steps are excellent. Using 10 increments we obtain results which are identical to those presented above; however, we have found that further increase in step size, using for example five increments, leads to problems. The solution breaks down in the first iteration of the first increment with negative value of J occurring, this is because for such large steps, on application of the prescribed displacements, the inner nodes pass the outer ones and a physically unrealistic elastic solution results. This limitation in step size is thus associated with the overall geometry of the problem rather than with the solution algorithm; we make this comment only in the context of this simple problem.

The performance of the algorithm for the 15-increment solution of this example is summarised in Tables 11.2 to 11.5. In Table 11.2a we give the number of Newton iterations required in each increment with the convergence tolerance, tol , at a values of 1×10^{-6} . For tighter tolerances convergence breaks down for the reasons discussed above. We compare these with the performance reported by Simo, see Table 11.2b. In

$tol = 1 \times 10^{-6}$																
(a)	increment	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	iterations	5	5	4	4	4	4	4	4	4	4	4	4	4	4	4
Simo																
(b)	increment	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	iterations	6	5	5	5	5	5	5	5	5	5	5	4	4	4	4

Table 11.2: Iterations per increment for Example 1 ($\alpha = 1$)

Table 11.3 the values of the maximum residual at the end of each iteration are given for several typical steps; these clearly show quadratic convergence although there is some deterioration in convergence rate as the solution progresses. For this problem Simo only reports the values of an energy norm at each iteration; by inspection of other of his results where both residual and energy norms are reported it can be seen that the residual values we have calculated at the end of the last iteration reflect a similar result.

increment 1 max residuals	increment 2 max residuals	increment 8 max residuals	increment 14 max residuals
0.906E+5	0.231E+5	0.561E+5	0.752E+5
0.185E+5	0.186E+4	0.214E+3	0.275E+3
0.354E+3	0.501E+1	0.317E-2	0.377E-2
0.612E-1	0.113E-4	0.654E-7	0.275E-6
0.605E-8	0.246E-8		

Table 11.3: Convergence for typical increments; Example 1

In Tables 11.4 and 11.5 we give some indication of the performance of the corrector algorithm. The number of iterations of the corrector algorithm at a typical Gauss point are given for each iteration in typical increments. In Table 11.5 we illustrate how a typical corrector solution proceeds, by giving values of maximum residual, yield function and J^p at the end of each iteration.

Increment 1					
Global iteration	1	2	3	4	5
Corrector iterations	7	5	4	3	1

Increment 14					
Global iteration	1	2	3	4	
Corrector iterations	5	3	2	1	

For solution with $tol = 1 \times 10^{-6}$ and $tolcor = 1 \times 10^{-10}$

Table 11.4: Corrector iterations for Example 1

We repeated the analysis with $\alpha = 0.5$ in the midpoint rule, the constraint $J^p = 1$ applied at time $t_{n+\alpha}$. The 20-element mesh and 15-increment time discretisation were used. The displacement results obtained were in good agreement with those obtained with $\alpha = 1$. Results for internal pressure are given in Figure 11.3, also shown is the solution obtained with $\alpha = 1$ which is the same as the exact solution. These pressures are calculated at $t_{n+\alpha}$. It can be seen that as the solution progresses there is a deterioration in the results. This deterioration can be attributed to errors which accumulate in the

Increment 1 iteration 1 Element 1			
Iteration	max residual	Yield function	J^p
0	0.950E+04	12525.253925922	1.000000
1	0.921E+04	12184.153000511	1.000000
2	0.267E+05	4687.338478648	0.999910
3	0.188E+05	1062.601991149	0.741256
4	0.427E+03	45.822171615	1.013958
5	0.717E-01	1.034887087	0.997804
6	0.200E-05	1.000000068	0.999997
7	0.217E-11	0.999999999	1.000000

Table 11.5: Corrector convergence for Example 1

calculation of \mathbf{F}^p . The updating procedure assumes a linear path from t_n and t_{n+1} which is invalid in the case of large plastic deformations where large increments are used. The constraint $J^p = 1$ is not maintained after update and hence in the next increment the calculation of $\Delta\mathbf{F}^p$ is affected. We have found that this situation is aggravated rather than improved by using very small steps. We note that convergence of the algorithm is also adversely effected, as reflected in an increase in the number of iterations required per increment, (see Table 11.6).

$tol = 1 \times 10^{-6}$															
increment	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
iterations	4	5	5	5	5	5	5	5	5	5	5	6	5	6	6

Table 11.6: Iterations per increment for Example 1 ($\alpha = 0.5$)

Pressure vs internal radius ($\alpha=0.5$)

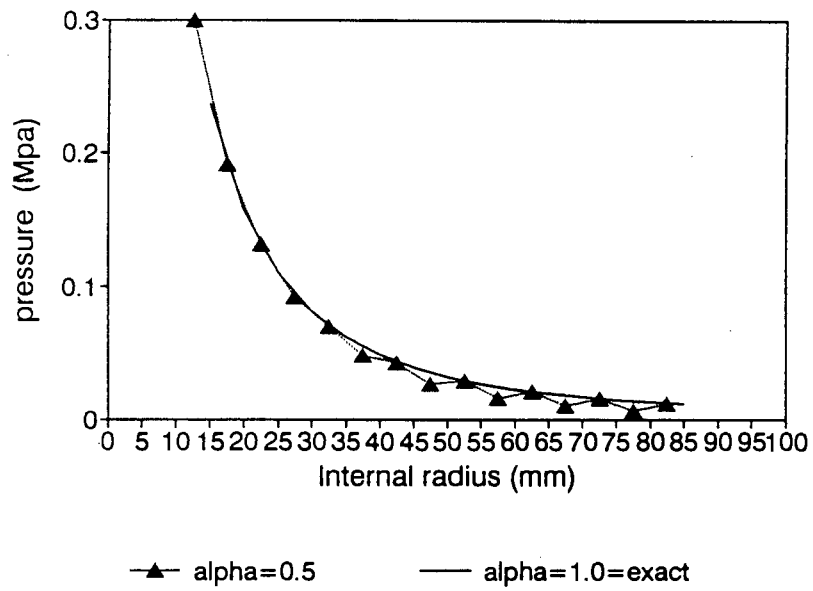


Figure 11.3: Internal pressure vs. inner radius

11.5.2 EXAMPLE 2: Upsetting of an axisymmetric billet

The objective is to achieve 60% upsetting of a cylindrical billet. The problem is shown

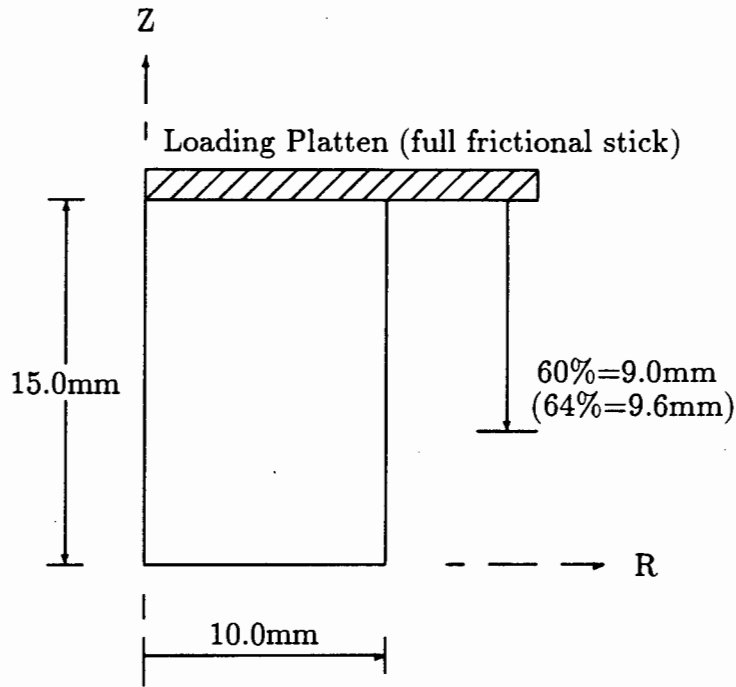


Figure 11.4: *Example 2* Upsetting of a axisymmetric billet

in Figure 11.4. The loading plattens are assumed to be perfectly rough and the steel billet is assumed to harden linearly; the material properties used are given in Table 11.7. This example has been proposed as a severe test for models of finite-strain plasticity. The problem was used in a joint examination of various numerical methods for analysis of metal forming problems; the results of this work are summarised and presented by Kudo and Matsubara [54]. Subsequently it has been used by many other researchers, see for example [107,115,113,118]; it also used as a verification example by the developers of commercial nonlinear finite element software, see for example ABAQUS users manual. (We note however, that the material properties used by some do not match those used here, in particular [107,115].)

The first example we have looked at has given some insight into the performance of the algorithm. It is however a very simple example. The upsetting of an axisymmetric

Elastic properties:	
Youngs modulus	$E = 200.0 \text{ GPa}$
Bulk modulus	$K = 0.3$
Plastic properties:	
Flow stress	$T_0 = 450.0 \text{ MPa}$
linear isotropic hardening coefficient	$h = 300.0 \text{ MPa}$
$(c_0 = \sqrt{\frac{2}{3}}T_0, c_1 = \sqrt{\frac{2}{3}}h \text{ and } c_2 = \beta = 0)$	

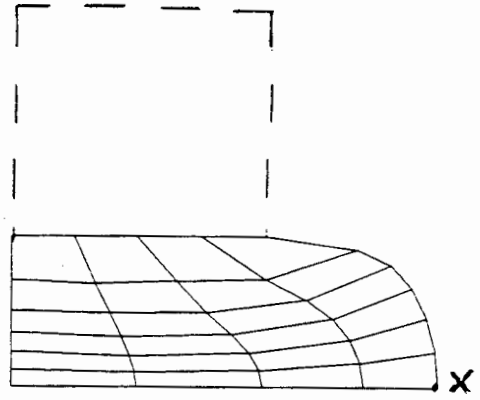
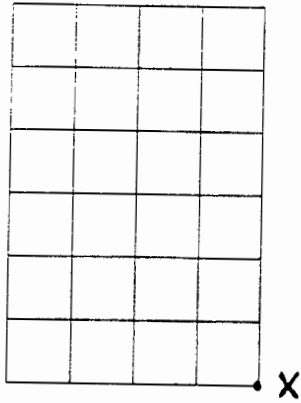
Table 11.7: Material properties for Example 2

billet is considered to be a considerably more taxing problem.

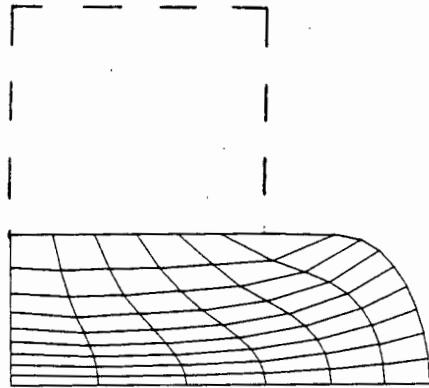
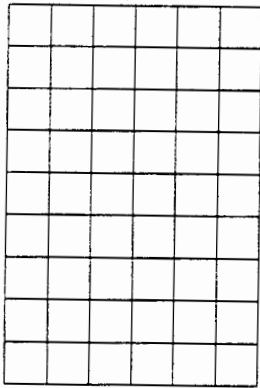
We have restricted attention to the case $\alpha = 1$ and have examined how the solution is affected by the choice of finite element mesh and step size. We have used meshes consisting of 4×6 , 6×9 and 8×12 4-noded bilinear elements; these meshes are shown in Figure 11.5. The displaced shape for each of these meshes, at 60% upsetting, are also shown in Figure 11.5; these results were all obtained using 16 equal increments. We did not take into account the affect of the contact and friction on the material which folded into the loading platten. While for most of the billet the displacement solutions for all meshes are similar to those given in [107] and [118] they differ in the region where contact due to folding takes place. This difference is not great. A comparison of the results we have obtained is made by plotting the displacement of the midside node, labled X in part (a) of Figure 11.5, against percentage upsetting. These plots are presented in Figure 11.6. They show that a reasonable solution is obtained with even the coarsest of these meshes. They also show how the solution converges with mesh refinement as is expected of a finite element discretisation.

In Figure 11.7 we show contour plots of the stresses at 60% upsetting, these stress results were obtained using the 8×12 mesh and 16 increments. Each of the four nonzero components of the Cauchy stress, \bar{T} are plotted. A plot of Mises equivalent stress, that is, $\sigma = \sqrt{\frac{2}{3}\bar{T}_{ij}\bar{T}_{ij}}$ is also given. The stress results obtained by other researchers

(a) 4×6 mesh



(b) 6×9 mesh



(c) 8×12 mesh

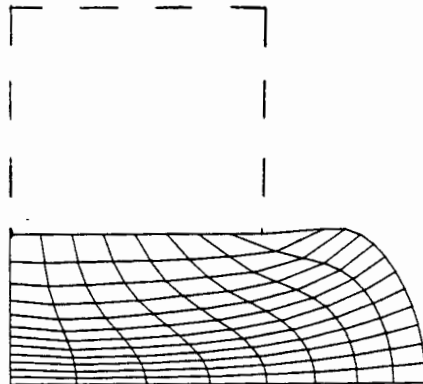
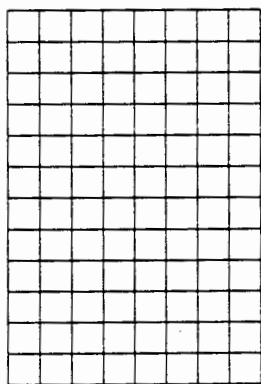


Figure 11.5: Meshes and displacement solutions at 60% upsetting (Example 2)

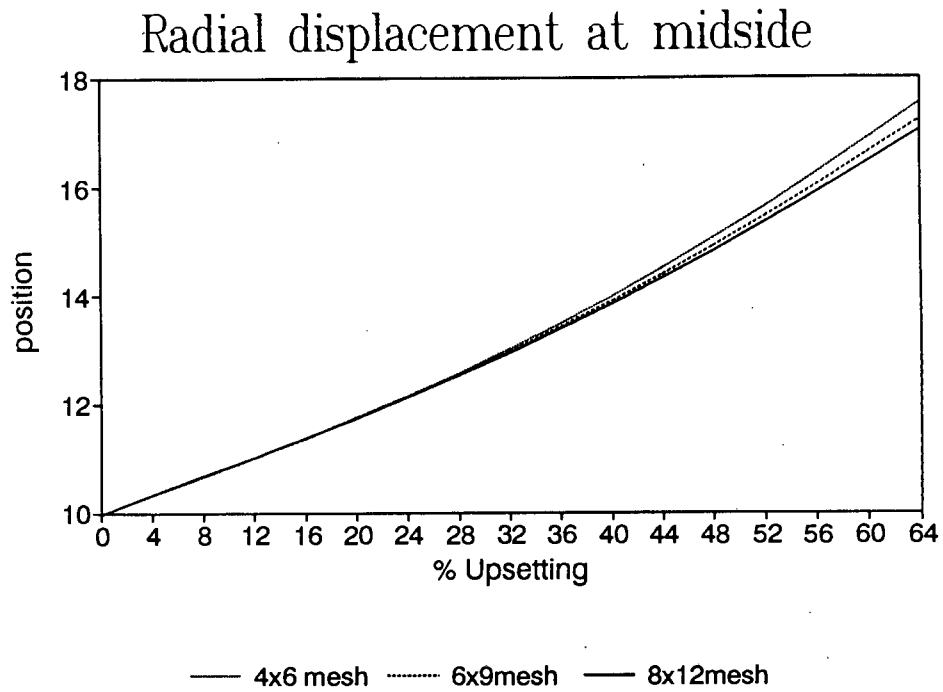
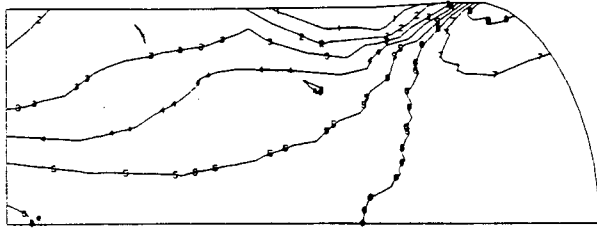


Figure 11.6: Comparison of displacement results for Example 2

have not been recorded in the literature, we therefore compared our stress results with those obtained using the finite element code ABAQUS. We do not present the ABAQUS results but remark that the values of the various stresses are similar, the shape of the contours are also similar except, as expected, near the folding contact. We remark that ABAQUS uses a rate form of the constitutive equations for the solution of this problem.

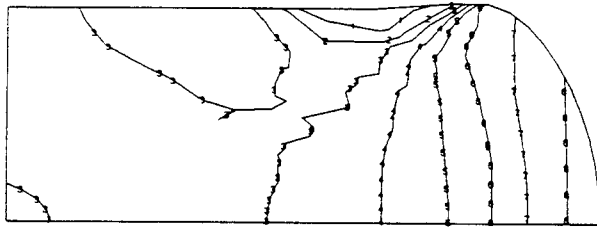
Study of the effect of increment size was performed using the 4×6 mesh. Upsetting of 64% was prescribed in 8, 16, 32, and 64 increments. These results are also compared by plotting the midside node displacement against percentage upsetting, see Figure 11.8. Differences between these solutions diminish with the decrease in step size as expected. It is remarkable that it is possible to obtain a solution in increments as large as 8% upsetting (8 increments); this solution is reasonably close to that obtained with small steps. We note that 150 increments were used to obtain the solution presented by Taylor and Becker [118]; 120 increments (for 64% upsetting) were used by Simo and coworkers [107,115,113].

(a) T_{11}



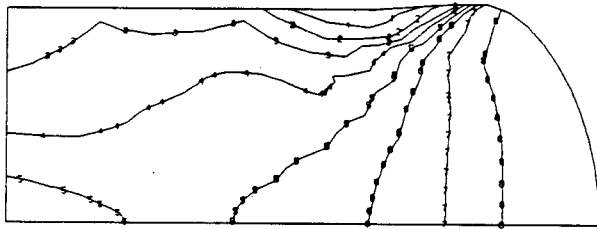
1	-1200	MPa
2	-1000	
3	-800	
4	-600	
5	-400	
6	-200	
7	0	
8	+200	

(b) T_{22}



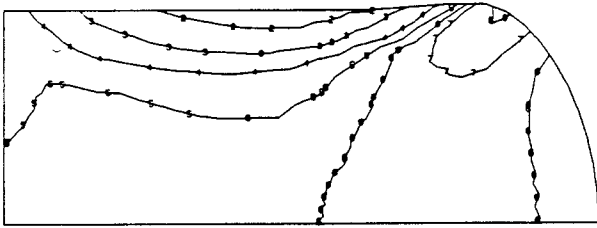
1	-2100	MPa
2	-1800	
3	-1500	
4	-1200	
5	-900	
6	-600	
7	-300	
8	0	

(c) T_{33}



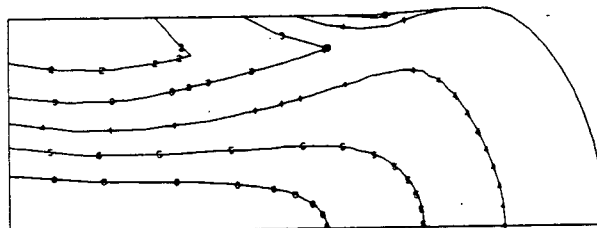
1	-1500	MPa
2	-1200	
3	-900	
4	-600	
5	-300	
6	0	
7	+300	
8	+600	

(d) T_{12}



1	-500	MPa
2	-400	
3	-300	
4	-200	
5	-100	
6	0	
7	+100	
8	+200	

(e) von Mises equivalent stress σ



1	+700	MPa
2	+800	
3	+900	
4	+1000	
5	+1100	
6	+1200	

Figure 11.7: Stress results 8×12 -mesh at 60% upsetting (Example 2)

Radial displacement at midside 4x6 mesh

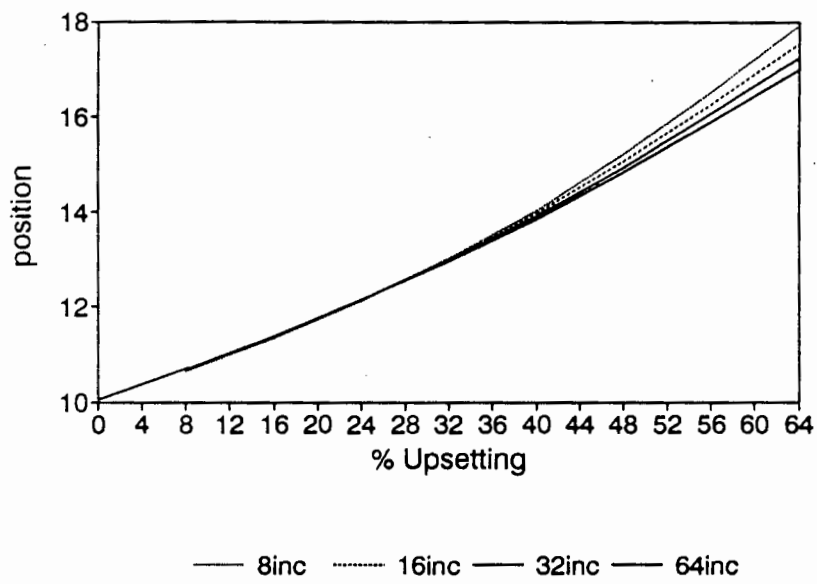


Figure 11.8: Investigation of response to increment size Example 2

11.6 Discussion of results

The objective of the numerical work presented in this Chapter has been to show that the theory we have developed provides an alternative approach to the construction of algorithms for the solution of practical problems. In our corrector algorithm we simply apply the well known Newton method to solve evolution equations given in terms of the dissipation function. This approach requires no additional assumptions such as, for example, the way in which the stresses return to the yield surface as is used in conventional return mapping algorithms. Return of the stress to the yield surface takes place as a consequence of solving the evolution equations. An outstanding feature of the results we have obtained is that good solutions have been found using *large time steps*. Some algorithms based on a conventional return mapping approach have also achieved a degree of success with large steps; in particular that developed by Simo [107]. The algorithm developed by Simo has been shown to perform well with large time steps for the simple thick-walled cylinder example, but does not appear to have performed as well for more complicated problems such as the upsetting problem we have analysed. We have not included the contact problem in our analysis; it is not clear to what extent the contact problem affects the convergence rate of the solution.

We have indicated the performance of the algorithm by giving the number of iterations used. While these numbers in some way reflect the amount of computational effort required in the solution process, it is imprecise, since the computational effort expended within each iteration is not taken into account. We have coded the calculation of tangent stiffnesses and the corrector algorithm in such a way as to allow the most general form of hyperelastic model of the type described in Chapter VII to be used. Considerable saving can be made if the code is adapted specifically for a neo-Hookean material such as the one we have used in the example problems. Aside from the burden of this feature we believe that within each Newton iteration the algorithm we have used is more expensive than conventional return mappings. A detailed comparison would be needed to test this hypothesis; however, any additional cost per iteration is well compensated for by the savings which can be made by using fewer increments to solve a given problem.

We have found the upsetting problem to be sensitive to parameters used in the control

of hourglassing. The implementation of recently published methods for the stabilisation of hourglass modes for nonlinear problems [8] may well alleviate these problems.

The use of the midpoint rule is an aspect of the formulation which we have investigated by numerical experiment. Our results indicate that in the case of finite strain plasticity the midpoint rule does not perform well except for the case $\alpha = 1$, that is, for the case of a simple Euler backward difference. The reason for this is that the nonlinear constraint on the plastic deformation gradient, imposed at times $t_{n+\alpha}$, is not maintained in the linear update procedure used to obtain \mathbf{F}^p_{n+1} . The errors arising from this propagate as the solution proceeds. Further investigation is required to find a consistent method of eliminating these errors.

CHAPTER XII

SUMMARY AND CONCLUSIONS

The prime objective of most studies of plasticity is the solution of practical problems; in general it is procedures applicable to the solution of broad classes of problems that are sought. The development of a suitable constitutive theory for elastoplastic materials is the first step in building a model which serves to achieve this objective. It is necessary to make simplifying assumptions about the material behaviour in order to establish the constitutive theory. Plastic behaviour is history-dependent; this important feature must be taken into account. For practical purposes it is the macroscopic rather than the microscopic behaviour that is of interest; internal variable methods are therefore attractive. A material model for elastoplasticity essentially comprises a characterisation of the elastic response together with a choice of yield criteria and hardening laws. It is the theory of plasticity that determines how these are related and interact to mimic the desired behaviour; in particular the theory sets out evolution laws for internal variables used to describe the plastic deformation and hardening.

In this thesis we have described the development of an internal variable theory of finite-strain plasticity and have shown how this theory may be used as a basis for the formulation and solution of practical problems involving finite deformations. In this final Chapter we highlight points of interest and state what we believe to be the main contributions and shortcomings of the work.

A general theory of continuum thermodynamics with internal variables is used as the foundation for the theory of plasticity. This flexible foundation provides scope for generalisation of the constitutive theory to include various hardening laws and possibly other features of more complex materials. The theory of plasticity is developed within a framework of non-smooth convex analysis. A unification of ideas concerning the postulates of plasticity is achieved by using the powerful tools provided by results in this branch of mathematics. Although the new theory is applicable to both small and finite-strain plasticity we have restricted attention to its application to the finite-

strain case. No new yield conditions or hardening laws have been introduced: the theory simply provides a clear mathematical framework within which classical models for plastic materials can be reviewed, and which allows various concepts to be studied in a unified manner.

As we indicated in the introduction to this thesis it is in general not possible to solve exactly problems of this type. The best available alternative is to compute an approximation to the required solution. We have used proven techniques to approximate the problem, namely a generalised midpoint rule for integration of history and the finite element method for spatial discretisation of the problem. We have described how these methods are used to obtain a discrete form of an incremental boundary-value problem which corresponds to a practical problem involving finite plastic deformations. The solution to a discrete problem of this type is assumed to be reasonably close to the solution of the real problem. To obtain an assurance that this assumption is valid a qualitative analysis of the problem and its approximation are required; such information as the existence and uniqueness of solutions, and convergence characteristics of discretisation and solution procedures are of interest. A qualitative analysis of problems involving finite plastic deformations has yet to be accomplished. A clear statement of the problem, amenable to application of known techniques of analysis, is a good starting point for the development of such an analysis.

We believe that the theory of plasticity which we have developed offers a firm foundation for the much needed qualitative analysis of problems of finite-strain plasticity. This claim is made in view of the following observations. We have constructed the theory in such a way that a close relationship to a general theory of hyperelasticity is maintained; in fact, in the absence of plastic deformations our theory reduces precisely to such a model. A complete analysis of problems of hyperelasticity is as yet unavailable. Some aspects of such an analysis have, however, been investigated and a clear mathematical framework for the model has been established. It seems reasonable to expect that an analysis of hyperelastic models will provide results and techniques which will readily translate to the analysis of problems of finite-strain plasticity formulated using the new theory. Clear and mathematically precise relationships between the various variables and functions appearing in this theory are provided; these have been established by using

the powerful tools provided by results in convex analysis. This branch of mathematics offers a store of related results which could prove useful in an analysis of the problem. The generalised midpoint rule has been subjected to a fairly comprehensive analysis, its application to problems involving small-strains is well understood. Some results are also available for a particular formulation of problems in finite-strain plasticity, see [109]. The finite element method is widely used in the solution of problems of mechanics and a variety of techniques for analysing this type of approximation have been developed.

We note that some progress has already been made concerning the analysis of problems of plasticity using the theory developed as part of this thesis; such work may be found in the contribution of Reddy [96], where a small-strain form of the theory has been used. Some aspects of the analysis of the class of small-strain problems considered by Reddy had previously been performed; however, the analysis in [96] offers an alternative approach which has the advantage that it is closely related to procedures used to solve the problem in practice.

While we have provided what we believe to be a significant step towards the development of a complete analysis of problems in finite-strain plasticity there remain major obstacles to the achievement of this goal. We have applied our theory to the solution of some practical problems rather than attempt a qualitative analysis. We have proposed an algorithm for the solution of the discrete incremental boundary-value problem we have described. We have followed a conventional approach to constructing this algorithm in that we have used a Newton-Raphson iterative procedure with calculations in each iteration forming a predictor-corrector scheme. The algorithm does, however, differ in a significant way from those commonly used in computational plasticity. The approach we have followed makes use of evolution equations expressed in terms of the dissipation function; the conventional approach makes use of the dual form in which they are expressed in terms of a yield function. We have used a consistent predictor which is derived using these evolution equations. In practice the novelty of our approach is most apparent in the corrector phase of each iteration of the algorithm. We do not use the typical return mapping approach: in its place we solve, at each point at which yielding is predicted, a system of nonlinear equations which determine the evolution of internal variables. We remark that all the features of a classical theory, including the normality

law which plays a central role in the formulation of conventional return algorithms, are accommodated in the new theory of plasticity. A local Newton procedure is used to obtain the solution to the incremental evolution equations. No assumption about a path for returning predicted stresses to the yield surface is required.

We have shown that the manner in which an assumption of isochoric plastic deformation is included in the incremental problem requires careful consideration. We stress that this assumption is one which should be considered as part of the constitutive theory and not, as it is sometimes presented, as part of kinematic description of finite elastoplastic deformations. We have applied this constraint on the evolution of plastic deformation by way of a Lagrange multiplier. We remark that there are other ways of addressing this part of the problem. An alternative that deserves further attention is that of using the decomposition of deformations into dilational and volume-preserving parts and using J^p as an independent variable. This approach has been successfully exploited by Simo and co-workers [113].

By solving a number of example problems we have shown that the algorithm we have proposed offers a viable alternative to conventional methods. The algorithm performs extremely well in the sense that good results are obtained using relatively large increments. This desirable characteristic is the motivation for using incremental solution procedures of the type we originally set out to investigate. We remark that, in each iteration, the computational cost of using the corrector algorithm and calculating the consistent tangent that we have derived is probably greater than that of equivalent steps in existing procedures. For the problems we have considered this increase in cost is adequately compensated for by the reduction in the total number of iterations used in obtaining a solution at advanced stages in a program of loading.

We remark that the algorithm we have developed conforms to the standard iterative incremental structure which forms the backbone of many nonlinear finite element codes, and it could therefore easily be added into such codes. There are, however, many considerations which must be taken into account when assessing the desirability of any given algorithm. We note that many practical problems which involve finite plastic deformations also involve, for example, contact conditions. The size of increments which

can be used to solve such problems is often governed by the contact algorithm rather than the material model. In such applications the fact that large increments may be used is of little value and hence, in view of the additional computational effort required, the algorithm we have developed may not be particularly attractive as a general purpose method.

We have restricted study of the application of the new theory to a simple example, that of Von Mises plasticity. The theory is, however, developed in such a way as to be applicable to any model of plasticity that involves a convex region of admissible stresses. This generality extends to regions with corners since no assumptions of smoothness of the yield surface are incorporated in the theory. The study of further examples, in particular of an example of this type, would be a worthwhile extension of this work. The use of conventional return mapping methods for problems involving nondifferentiable yield functions is complicated; we have avoided using this approach. The approach we have used lends itself to the application of existing methods for optimisation of nondifferentiable functions. An interesting method is one which belongs to the class of Davidon methods, a special case of which is the Newton method we have used (see [123]).

In developing the theory of plasticity provision is made for hardening by the use of internal variables. A nonlinear isotropic hardening law is included in the example we have studied. We have not fully investigated the application of the new internal variable theory to the modelling of kinematic hardening for the case of finite-strains. Nonlinear kinematic hardening is not easily described; the exact nature of this behaviour appears to be an open question. It seems likely that nonlinear hardening may best be accounted for by some form of coupled isotropic and kinematic hardening law. A valuable extension to this study would be a comprehensive investigation of the application of the theory to the modelling of nonlinear kinematic hardening.

Finally we remark that the full potential of the new internal variable theory of plasticity that we have developed has still to be realised. In particular, in its present form the theory is restricted to the case of rate-independent plasticity. An important direction for further research is the extension of the theory to take into account viscoplastic

behaviour. An indication of the nature of dissipation functions which should be used in a viscoplastic theory has been put forward by Carter and Martin [11]; this work provides a starting point for an investigation of this extension.

BIBLIOGRAPHY

- [1] Alducin G, Subdifferential and variational formulations of boundary value problems. *Comp Meth Appl Mech Engng* **72** (1989) 173–186.
- [2] Atkin R J and Fox N, *An Introduction to the Theory of Elasticity*. Longman, London (1980).
- [3] Aubin J P, *Applied Functional Analysis*. Wiley, New York (1979).
- [4] Backman M E, Form for relation between stress and finite elastic and plastic strains under impulsive loading. *Jour Appl Phys* **35** (1964) 2524–2533.
- [5] Ball J M, Convexity conditions and existence theorems in nonlinear elasticity. *Arch Rat Mech Anal* **63** (1977) 337–406.
- [6] Bell J F, A physical basis for continuum theories of finite strain plasticity: Part I. *Arch Rat Mech Anal* **70** (1979) 319–338.
- [7] Bell J F, A physical basis for continuum theories of finite strain plasticity: Part II. *Arch Rat Mech Anal* **75** (1981) 103–126.
- [8] Belytschko T and Bindeman L P, Assumed strain stabilisation of 4-noded quadrilaterals with 1-point quadrature for nonlinear problems. *Comp Meth Appl Mech Engng* **88** (1991) 311–340.
- [9] Bird W W and Martin J B, Consistent predictors and the solution of the piecewise holonomic incremental problem in elasto-plasticity. *Engng Struct* **12** (1990) 9–14.
- [10] Carroll M M, A rate-independent constitutive theory for finite inelastic deformation. *Jour Appl Mech* **54** (1987) 15–21.
- [11] Carter P and Martin J B, Work bounding functions for plastic materials. *Jour Appl Mech* **43** (1976) 434–438.
- [12] Casey J and Naghdi P M, A correct definition of elastic and plastic deformation

- and its computational significance. *Jour Appl Mech* **48** (1981) 983–985.
- [13] Casey J and Naghdi P M, A remark on the use of the decomposition $\mathbf{F} = \mathbf{F}^e \mathbf{F}^p$ in plasticity. *Jour Appl Mech* **47** (1980) 672–675.
- [14] Casey J and Tseng M, A constitutive restriction related to convexity of yield surfaces in plasticity. *Zeit Angew Math Phys* **35** (1984) 478–496.
- [15] Chadwick P, *Continuum Mechanics: Concise Theory and Problems*. George Allen & Unwin Ltd, London (1976).
- [16] Ciarlet G P, *Mathematical Elasticity: Volume 1. Three Dimensional Elasticity*. North Holland, Amsterdam (1988).
- [17] Ciarlet P G and Geymont G, Sur les lois de comportement en élasticité non-linéaire compressible. *C R Acad Sci Paris Sér II* **295** (1982) 423–426.
- [18] Coleman B D and Gurtin M E, Thermodynamics with internal state variables. *Jouř Chem Phys* **47** (1967) 597–613.
- [19] Dafalias F Y, Co-rotational rates of kinematic hardening at large plastic deformations. *Jour Appl Mech* **50** (1983) 561–565.
- [20] Dafalias F Y, Elasto-plasticity coupling within a thermodynamic strain space formulation of plasticity. *Int Jour Non-linear Mech* **12** (1977) 327–337.
- [21] Dafalias F Y, The plastic spin concept and the simple illustration of its role in finite plastic transformation. *Mech Mat* **3** (1984) 223–233.
- [22] Dashner P A, Large strain inelastic state variable theory. *Int Jour Solids Structs* **22** (1986) 571–592.
- [23] Dashner P A, Plastic potential theory in large strain elasto-plasticity. *Int Jour Solids Structs* **22** (1986) 593–623.
- [24] Dhatt G and Touzot G, *The Finite Element Method Displayed*. Wiley, Chichester (1984).

- [25] Drucker D C, A more fundamental approach to plastic stress–strain relations. In *Proceedings of the 1st United States National Congress of Applied Mechanics* 487–491. (1951) .
- [26] Durban D, Large strain solution for pressurised elasto/plastic tubes. *Jour Appl Mech* **46** (1979) 228–230.
- [27] Eterovic A L and Bathe K–J, A hyperelastic–based large strain elasto–plastic constitutive formulation with combined isotropic–kinematic harding using logarithmic stress and strain measures. *Int Jour Num Meth Engng* **30** (1990) 1099–1114.
- [28] Eve R A, Gültop T and Reddy B D, An internal variable finite–strain theory of plasticity. *Quart Appl Math* **48** (1990) 625–693.
- [29] Eve R A, Reddy B D and Rockafellar R T, An internal variable theory of elasto–plasticity based on the maximum plastic work inequality. *Quart Appl Math* **48** (1990) 59–83.
- [30] Feijóo R A and Zouain N, Formulations in rates and increments for elastic–plastic analysis. *Int Jour Num Meth Engng* **26** (1988) 2031–2048.
- [31] Flanagan D P and Belytschko T, A uniform strain hexahedron and quadrilateral with orthogonal hourglass control. *Int Jour Num Meth Engng* **17** (1981) 679–706.
- [32] Gelin J C and Picart P, Use of Quasi-Newton methods for large strain elastic–plastic finite element computations. *Comm Appl Num Meth* **4** (1988) 457–469.
- [33] Green A E and Naghdi P M, A general theory of an elastic–plastic continuum. *Arch Rat Mech Anal* **18** (1965) 251–281.
- [34] Green A E and Naghdi P M, Some remarks on elastic–plastic deformation at finite strain. *Int Jour Engng Sci* **9** (1971) 1219–1229.
- [35] Griffin T B, *Variational formulations and numerical analysis of some problems in small strain elastoplasticity*. PhD thesis, University of Cape Town (1986) .

- [36] Griffin T B, Reddy B D and Martin J B, A numerical study of holonomic approximations to problems in plasticity. *Int Jour Num Meth Engng* **26** (1988) 1449–1466.
- [37] Gurtin M E, Modern continuum thermodynamics. In *Mechanics Today* (ed. S Nemat-Nasser) 168–210. Pergamon, Oxford (1974).
- [38] Halphen B and Nguyen Q S, Sur les matériaux standards généralisés. *Journal de Mécanique* **14** (1975) 39–63.
- [39] Hibbitt H D, Marcal P V and Rice J R, A finite element formulation for problems of large strain and large displacement. *Int Jour Solids Structs* **6** (1970) 1069–1086.
- [40] Hill R, Constitutive dual potentials in classical plasticity. *Jour Mech Phys Solids* **35** (1987) 23–33.
- [41] Hill R, Constitutive inequalities for isotropic elastic solids under finite element strain. *Proc R Soc* **A314** (1970) 457–472.
- [42] Hill R, The essential structure of constitutive laws for metal composites and polycrystals. *Jour Mech Phys Solids* **15** (1967) 79–95.
- [43] Hill R, On constitutive inequalities for simple materials. Part I. *Jour Mech Phys Solids* **16** (1968) 315–322.
- [44] Hill R, On constitutive inequalities for simple materials. Part II. *Jour Mech Phys Solids* **16** (1968) 229–242.
- [45] Hinton E and Owen D R J, *An Introduction to Finite Element Computations*. Pineridge Press, Swansea (1979).
- [46] Hughes T J R, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Prentice-Hall, New Jersey (1987).
- [47] Hughes T J R and Winget J, Finite rotation effects in numerical integration of rate constitutive equations arising in large-deformation analysis. *Int Jour Num*

Meth Engng **15** (1980) 1413–1418.

- [48] Il'iushin A A, On the postulate of plasticity. *Prikl Mat Mekh* **25** (1961) 503.
- [49] Kestin J and Rice J R, Paradoxes in the application of thermodynamics to strained solids. In *A Critical Review of Thermodynamics* (ed. E B Stuart B Gat-Or and A J Brainard), 275–298. Mono, Baltimore (1970).
- [50] Kim S J and Oden J T, Finite element analysis of a class of problems in finite elastoplasticity based on the thermodynamical theory of materials of type N. *Comp Meth Appl Mech Engng* **53** (1985) 277–302.
- [51] Kim S J and Oden J T, Generalised flow potentials in finite elastoplasticity: II examples. *Int Jour Engng Sci* **23** (1985) 510–530.
- [52] Kim S J and Oden J T, Generalised potentials in finite elastoplasticity. *Int Jour Engng Sci* **22** (1984) 1235–1257.
- [53] Kleiber M, Kinematics of deformation processes in materials subjected to finite elastic-plastic strains. *Int Jour Engng Sci* **13** (1975) 513–525.
- [54] Kudo H and Matsubara S, A joint examination project of validity of various numerical methods for the analysis of metal forming processes. In *Metal Forming Plasticity* (ed. H Lippmann), Springer, Berlin (1979).
- [55] Lai W M, Rubin D and Krempel E, *Introduction to Continuum Mechanics*. Pergamon, Oxford (1978).
- [56] Lee E H, Elastic-plastic deformation at finite strains. *Jour Appl Mech* **36** (1969) 1–6.
- [57] Lee E H, Some comments on elastic-plastic analysis. *Int Jour Solids Structs* **17** (1981) 859–872.
- [58] Lee E H and Lin D T, Finite-strain elastic-plastic theory with applications to plane-strain analysis. *Jour Appl Phys* **38** (1967) 19–27.
- [59] Lee E H and McMeeking R M, Concerning elastic and plastic components of

- deformation. *Int Jour Solids Structs* **16** (1980) 715–721.
- [60] Lee E H and Wierzbicki T, Analysis of the propagation of plane elastic-plastic waves at finite-strain. *Jour Appl Mech* **34** (1967) 931–936.
- [61] Lehmann T, Some remarks on the decomposition of deformations and mechanical work. *Int Jour Engng Sci* **20** (1982) 281–288.
- [62] Lubarda V A and Lee E H, A correct definition of elastic and plastic deformation and its computational significance. *Jour Appl Mech* **48** (1981) 35–40.
- [63] Lubliner J, A maximum-dissipation principle in generalised plasticity. *Acta Mech* **52** (1984) 225–237.
- [64] Lubliner J, Normality rules in large deformation plasticity. *Mech Mat* **5** (1986) 29–34.
- [65] Lubliner J, *Plasticity Theory*. MacMillan, New York (1990).
- [66] Mandel J, Sur la définition de la vitesse de déformation élastique et sa relation avec la vitesse de contrainte. *Int Jour Solids Structs* **17** (1981) 873–878.
- [67] Marsden J E and Hughes T J R, *Mathematical Foundation of Elasticity*. Prentice-Hall, New Jersey (1983).
- [68] Martin J B, An internal variable approach to the formulation of finite element problems in plasticity. In *Physical Non-linearities in Structural Mechanics* (ed. J Halt and J Lemaitre), . Springer, (1981).
- [69] Martin J B, A note on the implications of thermodynamic stability in the internal variable theory of inelastic solids. *Int Jour Solids Structs* **11** (1975) 247–253.
- [70] Martin J B, *Plasticity: Fundamentals and General Results*. MIT Press, Massachusetts (1975).
- [71] Martin J B and Nappi A, An internal variable formulation of perfectly plastic and linear kinematic and isotropic hardening relations with a von Mises yield condition. *Euro Jour Mech* **A9** (1990) 107–131.

- [72] Martin J B and Reddy B D, Variational principles and solution algorithms for internal variable formulations of problems in plasticity. In *Omaggio a Giulio Ceradini Note Scientifiche in Occasione del 70° Compleanno* (ed. U Andreaus et al) , 465–477. Univerità di Roma, ‘La Sapienza’, (Roma) (1988) .
- [73] Matthies H and Strang G, The solution of nonlinear finite element equations. *Int Jour Num Meth Engng* 14 (1979) 1613–1626.
- [74] Moran B, Ortiz M and Shih C F, Formulation of implicit finite element methods for multiplicative finite deformation plasticity. *Int Jour Num Meth Engng* 29 (1990) 483–514.
- [75] Moreau J J, Application of convex analysis to the treatment of elastoplastic systems. In *Applications of Methods of Functional Analysis to Problems in Mechanics, Proceedings of the Joint Symposium IUTAM/IUM, Marseille 1975: Lecture notes in Mathematics* 56–89. (1976) .
- [76] Moreau J J, Sur les lois de frottement de viscosite et plasticite. *C R Acad Sci Paris Sér II* 271 (1970) 608–611.
- [77] Mróz Z, A description of workhardening of metals with application to variable loading. In *Foundations of Plasticity, Proceedings of a symposium 1972*. Noordhoff International Publishing, (1973) .
- [78] Naghdi P M and Trapp J A, On the nature of normality of plastic strain rate and convexity of yield surfaces in plasticity. *Jour Appl Mech* (1975) 61–66.
- [79] Naghdi P M and Trapp J A, The significance of formulating plasticity theory with reference to loading surfaces in strain space. *Int Jour Solids Structs* 13 (1975) 785–797.
- [80] Nagtegaal J C and De Jong J E, Some computational aspects of elastic–plastic large strain analysis. *Int Jour Num Meth Engng* 11 (1981) 817–832.
- [81] Nagtegaal J C, Parks D M and Rice J R, On numerically accurate finite element solutions in the fully plastic range. *Comp Meth Appl Mech Engng* 4 (1974)

1531-177.

- [82] Needleman A, On finite element formulations for large elastic-plastic deformations. *Comp Structs* **20** (1985) 247-257.
- [83] Nemat-Nasser S, Decomposition of strain measures and their rates in finite deformation elasto-plasticity. *Int Jour Solids Structs* **15** (1979) 155-166.
- [84] Nemat-Nasser S, On finite deformation elasto-plasticity. *Int Jour Solids Structs* **18** (1982) 857-872.
- [85] Oden J T, *Qualitative Methods in Non-linear Mechanics*. Prentice-Hall, New Jersey (1986).
- [86] Ogden R W, *Non-linear Elastic Deformations*. Ellis Horwood, Chichester (1984).
- [87] Ortiz M, Some computational aspects of finite deformation plasticity. In *Proceedings of the 1st International Conference on Computational Plasticity: Models Software and Applications* (ed. D R Owen E Hinton and E Oñate) 1717-1756. Pineridge Press, Swansea (1987).
- [88] Ortiz M and Martin J B, Symmetry preserving return mapping algorithms and incrementally extremal paths. A unification of concepts. *Int Jour Num Meth Engng* **28** (1989) 1839-1855.
- [89] Ortiz M and Popov E P, Accuracy and stability of integration algorithms for elasto-plastic constitutive relations. *Int Jour Num Meth Engng* **21** (1985) 1561-1576.
- [90] Owen D R J and Hinton E, *Finite Elements in Plasticity: Theory and Practice*. Pineridge Press, Swansea (1980).
- [91] Pegon P and Guélin P, Finite-strain plasticity in convected frames. *Int Jour Num Meth Engng* **22** (1986) 521-545.
- [92] Perić D and Owen D R J, A model for finite strain elasto-plasticity based on

logarithmic strains: computational issues. *Comp Meth Appl Mech Engng.* to appear.

- [93] Ponter A R S and Martin J B, Some extremal properties and energy theorems for inelastic materials and their relation to the deformation theory of plasticity. *Jour Mech Phys Solids* **20** (1972) 281–300.
- [94] Porta F–A and Ptak V, *Nondiscrete Induction and Iterative Processes*. Pitman, London (1984).
- [95] Rebelo N, Nakazawa S, Wertheimer T B and Nagtegaal J C, . Computational aspects of finite element analysis in manufacturing. Internal report MARC Analysis Research Corporation.
- [96] Reddy B D, Existence of solutions to a quasistatic problem in elastoplasticity. To appear in *Elliptic and Parabolic Problems*. (ed. J Bemelmans and M Grüter).
- [97] Reddy B D, *Functional Analysis and Boundary–value Problems: An Introductory Treatment*. Longman, London (1986).
- [98] Reddy B D and Griffin T B, Variational principles and convergence of finite element approximations of a holonomic elastic-plastic problem. *Numerical Mathematics* **52** (1988) 101–117.
- [99] Reddy B D and Martin J B, Algorithms for the solution of internal variable problems in plasticity. *Comp Meth Appl Mech Engng* **93** (1991) 253–273.
- [100] Rencontré L J, *Time integration schemes for piecewise linear plasticity*. PhD thesis, University of Cape Town (1991).
- [101] Rencontré L J, Caddemi S and Martin J B, The relationship between the generalised mid–point and trapezoidal rules in incremental elastoplasticity. *Comp Meth Appl Mech Engng.* to appear.
- [102] Rice J R, Inelastic constitutive relations for solids: An internal- variable theory and its application to metal plasticity. *Jour Mech Phys Solids* **19** (1971) 433–455.

- [103] Rockafellar R T, *Convex Analysis*. Princeton University Press, New Jersey (1970) .
- [104] Rockafellar R T, On the maximal monotonicity of subdifferential mappings. *Pacific Journal of Mathematics* **33** (1970) 174–222.
- [105] Simo J C, Consistent tangent operators for rate-independent elasto-plasticity. *Comp Meth Appl Mech Engng* **48** (1985) 101–118.
- [106] Simo J C, A framework for finite strain elasto-plasticity based on maximum plastic dissipation and their multiplicative decompositions: Part 1 Continuum formulations. *Comp Meth Appl Mech Engng* **66** (1988) 199–219.
- [107] Simo J C, A framework for finite strain elasto-plasticity based on maximum plastic dissipation and their multiplicative decompositions: Part III Computation Aspects. *Comp Meth Appl Mech Engng* **68** (1988) 1–31.
- [108] Simo J C, Maximum plastic dissipation and the multiplicative decomposition in finite strain plasticity. In *Proceedings of the 1st international on Computational Plasticity: Models Software and Applications* (ed. D R Owen and E Hinton and E Oñate) 1821–1836. Pineridge Press, Swansea (1987) .
- [109] Simo J C, Nonlinear stability of the time–discretised variational problem of evolution in nonlinear heat conduction, plasticity and viscoplasticity. *Comp Meth Appl Mech Engng* **88** (1991) 111–131.
- [110] Simo J C and Govindjee S, Nonlinear B–stability and symmetry preserving Return mapping algorithms for plasticity and viscoplasticity. *Int Jour Num Meth Engng* **31** (1991) 151–176.
- [111] Simo J C, Kennedy J G and Govindjee S, Non-smooth multi-surface plasticity and visco-plasticity loading and unloading conditions and numerical algorithms. *Int Jour Num Meth Engng* **26** (1988) 2161–2185.
- [112] Simo J C, Kennedy J G and Taylor R L, Complementary mixed finite element formulations for elasto-plasticity. *Comp Meth Appl Mech Engng* **74** (1989) 177–

- [113] Simo J C, Kennedy J G and Taylor R L, Variational and projection methods for the volume constraint in finite deformation plasticity. *Comp Meth Appl Mech Engng* **51** (1985) 177-208.
- [114] Simo J C and Ortiz M, A unified approach to finite deformation elasto-plastic analysis based on the use of hyperelastic constitutive equations. *Comp Meth Appl Mech Engng* **49** (1985) 222-235.
- [115] Simo J C and Taylor R C, Consistent tangent operators for rate-independent elasto-plasticity. *Comp Meth Appl Mech Engng* **48** (1983) 101-118.
- [116] Simo J C and Taylor R C, A return mapping algorithm for plane stress elasto-plasticity. *Int Jour Num Meth Engng* **22** (1986) 649-670.
- [117] Strang G, *Introduction to Applied Mathematics*. Wellesley-Cambridge, Wellesley (1986).
- [118] Taylor L M and Becker E B, Some computational aspects of large deformation, rate-dependent plasticity problems. *Comp Meth Appl Mech Engng* **41** (1981) 251-277.
- [119] Teman R and Strang G, Functions of bounded deformation. *Arch Rat Mech Anal* **75** (1980) 7-21.
- [120] Truesdell C and Noll W, Non-linear field theories of mechanics. In *Encyclopedia of Physics Volume III part 3* (ed. S Flügge), Springer, Berlin (1965).
- [121] Voyiadjis G Z and Kiouisis P D, Stress rate and Lagrangian formulation of the finite strain plasticity for a von Mises kinematic hardening model. *Int Jour Solids Structs* **23** (1987) 95-109.
- [122] Wolfe M A, *Numerical Methods for Unconstrained Optimization; An introduction*. van Nostrand Reinhold, London (1978).
- [123] Wolfe P, A method of conjugate subgradients for minimising nondifferentiable

- functions. In *Mathematical Programming Study: Volume 9* (ed. L Balinski and P Wolfe) , 145–173. North Holland, London (1975) .
- [124] Zeidler E, *Nonlinear Functional Analysis and its Applications: Variational Methods and Optimization*. Springer, New York (1985) .
- [125] Ziegler H and Wehrli C, The derivation of constitutive relations from the free energy and the dissipation function. *Adv Appl Mech* **25** (1987) 183–239.