11

VARIATIONAL FORMULATIONS AND NUMERICAL ANALYSIS OF SOME PROBLEMS IN SMALL STRAIN ELASTOPLASTICITY

by

T B GRIFFIN B.Sc.(Eng), M.Sc.(Eng)

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

Department of Civil Engineering

University of Cape Town

March 1986

ABSTRACT

In this thesis we study the mathematical structure and numerical approximation of two boundary-value problems in small strain elasto-plasticity. The first problem, which we call the incremental holonomic problem, is based on a consistent incremental holonomic constitutive law, which in turn derives from the notion of extremal paths in stress and strain space as originally proposed by PONTER & MARTIN (1972); the second problem which we study is the classical rate problem. We show that both problems can be formulated as variational inequalities, with internal variables being included explicitly in the formulation. Corresponding minimisation problems follow naturally from standard results in convex analysis.

Perturbed minimisation problems are introduced, in which the original functionals J are replaced by perturbed functionals J_ε which depend on a parameter $\varepsilon > 0$. In the rate problem ε is a penalty parameter; here J_ε differs from J by a term $\varepsilon^{-1}j(\cdot)$ where $j(\cdot)$ is a penalty functional which allows the non-negativity constraint on the plastic multipliers to be removed. In the incremental holonomic problem the non-differentiable plastic work function $\hat{W}^p(\cdot)$ is regularised, and replaced by a differentiable function $\hat{W}^p_\varepsilon(\cdot)$. In both problems the perturbed functionals form the basis for finite element approximations, the error in the approximate solutions now depending on both mesh size and on the magnitude of ε .

Numerical algorithms are proposed, and implemented in two computer programs. On the basis of preliminary numerical experiments we conclude that the penalty-rate formulation is useful in a limited class of elastic-plastic problems, and that the incremental holonomic formulation has exceptional potential, without any apparent limitations.

To the memory of my father

DECLARATION

I, Terence Bernard Griffin, 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.

Signed by candidate

Signature removed

T B GRIFFIN

March 1986

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to the following people :

Professor John Martin for providing me with both the opportunity and the motivation to undertake this work.

Dr Daya Reddy, who directed this work, for his patient guidance, and whose contributions I gratefully acknowledge.

The Department of Civil Engineering and my colleagues in the Applied Mechanics Research Unit for allowing me considerable time off from other duties in which to complete this work.

Shirley Breed for a magnificent job in typing the manuscript under considerable duress.

Harold Cable for printing the thesis.

Helena du Toit for her help in drawing the diagrams.

My family and friends for tolerating my neglect of them over this period.

A special word of thanks to Wendy Millroy, whose patience, support, and encouragement are greatly appreciated.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

"Another way of pursuing unification [of the theory of plasticity] ... is to focus on the formal structure of some basic governing relations and utilise the relevant mathematics".

MAIER and NAPPI (1984)

The solution of a boundary-value problem in classical quasi-static rate-independent plasticity consists in seeking the history of response of a body (comprising the displacements, strains, and so on) to a given history of applied loading. For convenience we regard the history of loading and the history of response to be parametrised with respect to a

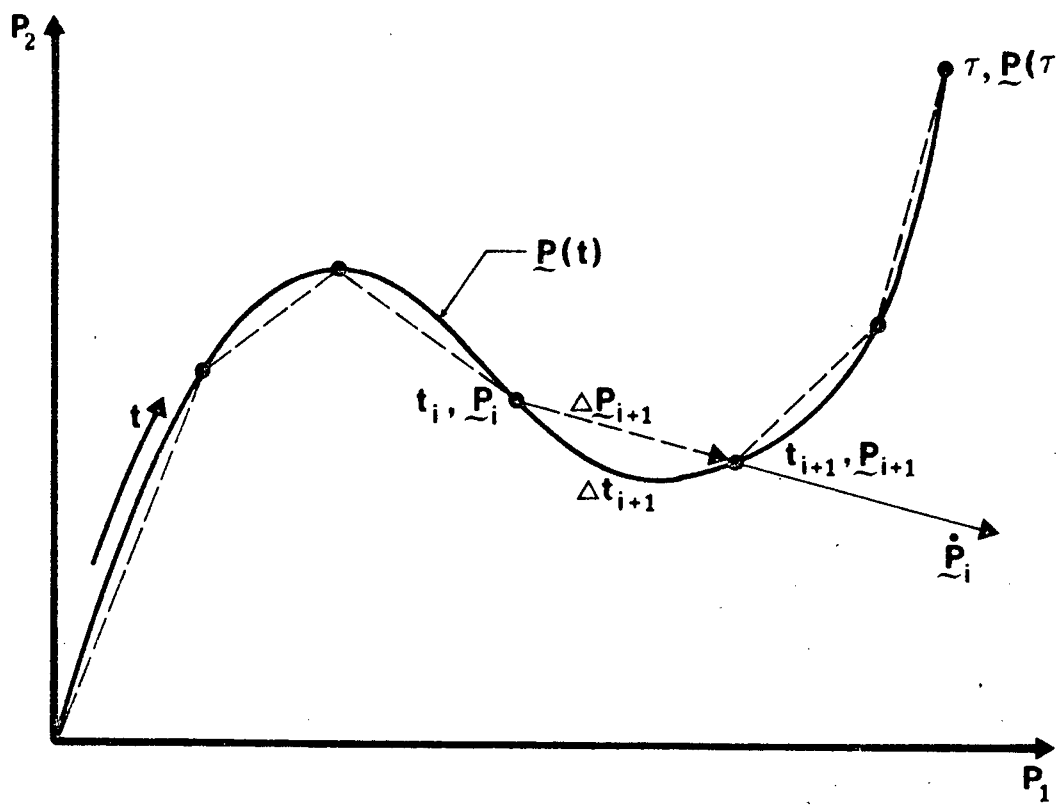


Figure 1.1 Loading history, parametrised with respect to t , $0 \leqslant t \leqslant \tau$, showing the discretisation of the load path.

parameter t ; for example, t could measure distance along the path, as shown in Fig. 1.1 for a typical loading history. Without loss of generality we may assume that the loading and response are zero at the start of their respective histories. Thus, for a given loading function $\underset{\sim}{P}(t)$, $0 \leqslant t \leqslant \tau$, we seek the response of the body, also as a function of t . The response is in general path-dependent, so that different loading paths which terminate at the point $\underset{\sim}{P}(\tau)$ will give rise to different responses at $t = \tau$.

Numerical approximations of the elastic-plastic problem necessarily involve both a spatial discretisation and a discretisation with respect to the parameter t . For the spatial discretisation the finite element method provides a well established procedure which we need not elaborate on at this point. However, the discretisation with respect to t is accomplished on a somewhat ad hoc basis, whereby we sample the history of loading at any n convenient points $\underset{\sim}{P}(t_i) \equiv \underset{\sim}{P}_i$, $1 \leqslant i \leqslant n$, as shown in Fig. 1.1; this procedure is equivalent to a piecewise linearisation of the load path. We may then conceive of the following generic problem : given the response at $t = t_i$ and the loading increment $\Delta \underset{\sim}{P}_{i+1} = \underset{\sim}{P}_{i+1} - \underset{\sim}{P}_i$, find the response at $t = t_{i+1}$. We shall refer to this as the <u>incremental elastic-plastic problem</u>; clearly, a sequence of such problems may be expected to provide some sort of approximation to the true continuous response.

The parametrisation of the loading and response histories allows us to define rates of change of these quantities with respect to t . Thus, it is meaningful to speak of the rate of change of the loading, or loading rate, $\dot{\underset{\sim}{P}}(t_i) \equiv \dot{\underset{\sim}{P}}_i$ at some point $\underset{\sim}{P}(t_i)$ on the loading path, where $\dot{\underset{\sim}{P}}(t) = d\underset{\sim}{P}(t)/dt$.

Similarly, we may speak of displacement rates, stress rates, and so forth. For numerical approximations, where the loading path is piecewise-linearised, we usually effect a consistent approximation of the loading rate by choosing $\dot{\underset{\sim}{P}}_i$ to be in the same direction as $\Delta \underset{\sim}{P}_{i+1}$; thus we have $\Delta \underset{\sim}{P}_{i+1} = \dot{\underset{\sim}{P}}_i \, \Delta t_{i+1}$, where $\Delta t_{i+1} = t_{i+1} - t_i$.

The behaviour of elastic-plastic materials under arbitrary loading histories is path-dependent and this behaviour is conveniently expressed in terms of constitutive laws relating rates of change of stress and strain. We are thus led to the formulation of the following <u>rate problem</u> : given the complete history of response up to $t = t_i$ and the loading rates $\dot{\underset{\sim}{P}}(t_i)$, find the response rates at $t = t_i$. Now, in order to obtain the response at $t = t_{i+1}$ it is necessary to integrate the response rates at $t = t_i$ and hence update the known response at $t = t_i$, a procedure which is generally referred to as "state determination". We will refer to the rate problem coupled with a suitable state determination scheme as the <u>incremental rate problem</u>, and recognise this as one way of solving the incremental elastic-plastic problem.

For certain programs of loading, for example, proportional loading, elastic-plastic materials exhibit behaviour which may be regarded as being path-independent. Such material behaviour is called <u>holonomic</u>* , and it is assumed that the response of such materials may be determined

---

* This term was first used by FINZI (1955), and later popularised by Professor Guilio Maier and the Italian School.

without regard to the loading path. These are the assumptions upon which deformation theories of plasticity are based (see MARTIN (1975a)), and they have played an important role in the development of the more general holonomic problem which we describe next.

An alternative approach to solving the incremental elastic-plastic problem is via the following incremental holonomic problem* : given the response at $t = t_i$ and the loading increment $\Delta \underset{\sim}{P}_{i+1}$, and assuming holonomic material behaviour over the interval $\Delta t_{i+1} = [t_i, t_{i+1}]$, find the response at $t = t_{i+1}$. Because of the holonomic assumption the governing equations for this problem may be written in terms of finite increments, thus obviating the need for a state determination scheme to obtain the response at $t = t_{i+1}$ (compare this with the incremental rate problem). Indeed, no restriction is placed on the size of the increments, so that we may immediately define a special case of the incremental holonomic problem, being that problem for which the chosen interval is $\Delta t = [0, \tau]$, and only a single response at $t = \tau$ is of interest. We shall refer to this as the holonomic problem or deformation theory problem.

To provide some perspective on the various problems described above we offer the following summary. Let the interval $[t_n, t_{n+1}]$ define an arbitrary increment in the loading and response paths (Fig. 1.2), and let $\mathcal{X}(t_i)$ = {displacements, strains, ... at $t = t_i$} characterise the response of the body at any t_i ; similarly, let $\dot{\mathcal{X}}(t_i)$ characterise the

---

* Also referred to as the stepwise holonomic problem.

response rates at any t_i . The definitions of the various problems are summarised in Table 1.1.

TABLE 1.1

PROBLEM DEFINITIONS

Problem	Given	Find	
Rate	$\chi(t_i)$, $0 \leqslant i \leqslant n$ $\dot{\underset{\sim}{P}}(t_n)$	$\dot{\chi}(t_n)$	Incremental Rate Problem
State Determination	$\dot{\chi}(t_n)$ $\Delta t_{n+1} = [t_n, t_{n+1}]$	$\chi(t_{n+1})$	
Incremental Holonomic	$\chi(t_n)$ $\Delta \underset{\sim}{P}_{n+1}$, Δt_{n+1}	$\chi(t_{n+1})$	
Holonomic or Deformation Theory	$\chi(0)$ $\underset{\sim}{P}(\tau)$	$\chi(\tau)$	

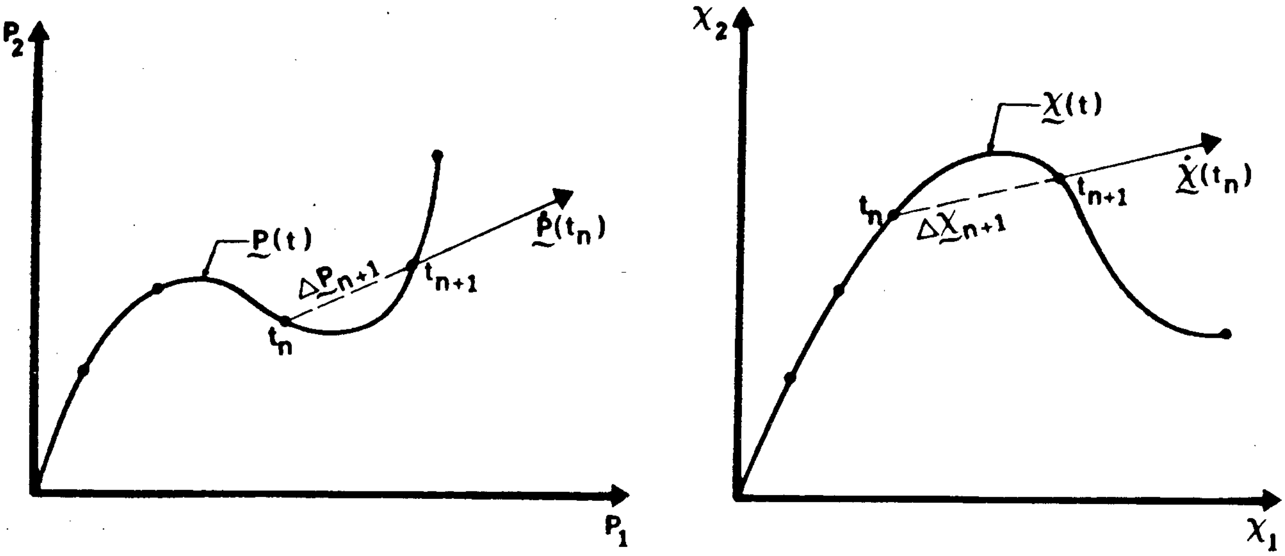


Figure 1.2 Loading and response paths showing the arbitrary increment $[t_n, t_{n+1}]$.

We propose in this thesis to study two elastic-plastic boundary-value problems. We will assume that the loading is quasi-static and the material behaviour is rate-independent, and based on the classical theory of plasticity for small deformations (where the influence of geometry changes on equilibrium equations is negligible). The first problem which we study is the incremental holonomic problem, and the second is the rate problem. In both cases our objectives will be two-fold : to study in detail the mathematical structures of the problems in terms of equivalent variational formulations, and to develop and analyse consistent numerical approximations to the original boundary-value problems using the Galerkin finite element method.

With the above broadly stated objectives in mind we embark now on a review of the relevant literature.

THEORETICAL ASPECTS

Closed-form solutions to the elastic-plastic boundary-value problem are in general unobtainable and recourse must therefore be made to approximate numerical methods, of which those which use the finite element method are most common. It has long been recognised that the elastic-plastic boundary-value problem can be formulated alternatively as a constrained minimisation problem, also commonly known as a minimum principle, and it is these principles which are used as the basis for numerical approximations. Minimum principles thus play a central role in the development of numerical approximations and we devote some time to their discussion here. Recently, interest has been shown in the formal study of the mathematical structure of the elastic-plastic boundary-value problem, particularly within the context of the theory of variational inequalities, and we discuss these developments as well.

Variational formulations for the Rate Problem

Constrained minimisation problems can be formulated either in terms of kinematic variables (for example, displacement rates or strain rates), or static variables (for example, stress rates); thus, we refer to a kinematic minimum principle in the former case, and a static minimum principle in the latter case. PRAGER (1942, 1946) was the first to establish both kinematic and static minimum principles and these were later generalised into their present form for smooth yield surfaces by GREENBERG (1949a,b), and for singular yield surfaces by KOITER (1953).

A brief historical sketch of the subsequent development and discussion of these principles, known as the classical minimum principles, has been given by MARTIN (1975a).

An alternative formulation of the static minimum principle was given by CERADINI (1966), who derived his result directly from the classical minimum principle. MAIER (1968) recognised that, when cast in finite dimensional form, Ceradini's principle could be formulated as a mathematical programming problem; subsequently, he derived an alternative finite dimensional formulation of the kinematic minimum principle (MAIER (1969b)) in the form of a quadratic programming problem, using the decomposition principle of COLONNETTI (1955). HODGE (1968) has given dual minimum principles for the rate problem, and has shown that finite dimensional forms of these minimum principles could be formulated as mathematical programming problems.

By making use of a particular property of the constitutive equations in the form of an inequality concerning an arbitrary division of the strain rate into elastic and plastic parts, MARTIN (1975a,b) showed that the result of MAIER (1969b) could be derived directly from the classical kinematic minimum principle. Martin's principle, known as the extended kinematic minimum principle, allows the rate problem to be expressed as a constrained minimisation problem involving velocities (that is, displacement rates) and plastic multipliers.

Following an initial study of thermodynamically based internal variable theories of plasticity (MARTIN (1975a,c)), CARTER and MARTIN (1977) made use of an internal variable description of the constitutive equations to rederive both the conventional and extended kinematic minimum principles, and later also the conventional static minimum principle (CARTER and MARTIN (1979)). In Carter and Martin's internal variable formulations the internal variables (whose physical meaning varies from one problem to the next) are included amongst the fields of variables which are to be determined by the solution.

An alternative minimum principle, also in terms of velocities and plastic multipliers, but for piecewise linear yield surfaces, has been given by HAVNER and PATEL (1976). The value of this work derives not so much from the principle itself but from the detailed convergence proofs which they give for their finite element approximation.

In a recent paper JIANG (1984) has shown how the rate problem may be formulated as a variational inequality, and has proved existence and uniqueness of a solution when the yield function is piecewise-linear and the hardening matrix is positive-definite. Jiang also considers the problem of regularity and shows that the solution, consisting of the velocity $\underset{\sim}{u}$ and the scalar k-tuple $\underset{\sim}{\lambda}$ of multipliers, is smooth enough to belong to $[H^2(\Omega)]^n \times [H^1(\Omega)]^k$ when the data are in $L_2(\Omega)$. ANZELLOTTI (1983) has given a detailed treatment of a rate boundary-value problem for elastic-perfectly plastic bodies, and gives existence results for stress rates, plastic multipliers, and velocities, the last in the space* $BD(\Omega)$ of functions of bounded deformation.

---

* We expand on the space $BD(\Omega)$ a little later in this chapter.

REDDY, GRIFFIN and MARAIS (1985) have shown the equivalence of the rate problem to a variety of variational formulations, and have discussed penalty-finite element approximations to the problem in some detail. This paper is essentially a summary of part of the work reported in this thesis.

Variational formulations for the Holonomic Problem

Holonomic, or deformation theory, descriptions of mechanical behaviour are capable of yielding exact solutions to the elastic-plastic problem when the stress path follows a radial line in stress space and no unloading or neutral loading occurs. Such behaviour occurs as the result of proportional loading, that is, when the load path follows radial lines in the load space. HENCKY (1924) and later NADAI (1931) both introduced deformation theories, but the widespread belief that such theories were limited strictly to proportional loading (and the fact that proportional loading was of limited practical significance) led to certain amount of hesitation in their being accepted (see, for example, HILL (1950), page 47). BUDIANSKY (1959) succeeded in dispelling this belief to some extent by showing that, using Nadai's theory, quite acceptable results could be obtained for loading paths which deviated considerably from proportional loading.

The first minimum principle for the holonomic problem was that of HAAR and VON KARMAN (1909) who extended by heuristic argument the elastic minimum complementary energy principle (see also MARTIN (1975a) for a discussion of this). Interest in the holonomic problem was revived with the publication of a series of papers (MAIER (1968), (1969a,b)) in which dual minimum principles were established based on quadratic programming arguments. These principles were derived for structures composed of a discrete assemblage of finite elements and a material which obeys Koiter's hardening rule with a number of piecewise linear, independent yield surfaces. The stress point is assumed to remain on the yield surface, once engaged, indicating that no local unloading may occur, and thus satisfying the holonomic assumption. DE DONATO (1968) extended Maier's finite dimensional results to continua for both the rate and holonomic problems. Further dual minimum principles which proved to be more attractive from the computational viewpoint were given by MAIER (1970), again formulated as finite dimensional quadratic programming problems.

From the concept of extremal paths in stress space, introduced by MARTIN (1966a,b), and the various complementary work bounding theorems which followed (PONTER (1968), SOECHTING and LANCE (1969), MARTIN (1970)), PONTER and MARTIN (1972) were able to establish a consistent definition of a holonomic material for continua which exhibit hardening behaviour governed by a smooth yield surface. Dual extremum principles and bounding theorems are given for the holonomic problem defined using this material. The bounding theorems indicate that solutions obtained using this holonomic theory bear a consistent relationship to those obtained using the incremental rate theory. Subsequently, MARTIN and PONTER (1972) showed that the holonomic theorems of Maier, referred to above, could be derived from Ponter and Martin's theorems and that Maier's theorems thus also provided a consistent formulation of the finite dimensional holonomic problem within the context of quadratic programming methods.

Finite dimensional formulations of the incremental holonomic problem using quadratic programming techniques appear to have been derived initially as a direct extension of the corresponding holonomic formulation (see, for example, COHN and MAIER (1979), Chapter 15). Recently, however, MAIER and NAPPI (1984) have given dual minimum principles for the finite dimensional incremental holonomic problem, again using quadratic programming methods. MARTIN, REDDY, GRIFFIN and BIRD (1984) have also given a minimum principle for the finite dimensional problem using an internal variable formulation in terms of displacements and plastic multipliers.

It is clear from the above that the theory of mathematical programming has played an important role in the development of minimum principles for finite dimensional cases of both the rate and holonomic problems, and it is fitting to expand briefly on this role. A mathematical programming problem (to quote from MAIER and MUNRO (1982))

> "consists of the optimisation (say minimisation) of an objective function over a feasible domain singled out in the vector space of the variables by the equality or inequality constraints".

Mathematical programming problems may be divided into three major categories, depending on whether the objective function which is to be minimised is linear, quadratic, or generally nonlinear : thus, we refer to a linear programming problem (provided the constraints are also linear), a quadratic programming problem (again, provided the constraints are linear), and a nonlinear programming problem (the constraints may be nonlinear). Broadly speaking, the use of a rigid-plastic constitutive law will give rise to a linear programming problem, whereas the use of an elastic-perfectly plastic or elastic-strain hardening law will give rise to a quadratic programming problem. The inclusion of second-order geometric effects will generally result in a nonlinear programming problem. It is therefore clear that the elastic-plastic problems in which we are interested here will all give rise to quadratic programming problems.

Powerful theoretical tools and numerical algorithms are provided for dealing with mathematical programming problems, provided that the problem is cast in finite dimensional (or discrete) form, and the yield surface is piecewise-linearised beforehand. Thus, in the case of the discrete elastic-plastic problem, for example, one organises the yield condition, compatibility equations and equilibrium equations into the form of a set of Kuhn-Tucker conditions (these provide the optimality conditions for the problem), and then, subject to certain convexity conditions, one may infer the corresponding mathematical programming problem. Suitable interpretation of the problem in mechanical terms then allows the inference of an extremum principle. More than this, the dual extremum principle usually follows naturally from the duality theory of mathematical programming.

Mathematical programming also provides a unified framework for the study of finite dimensional elastic-plastic problems. A simple example of this is the formal analogy between the quadratic programming problems for the rate and holonomic problems, as demonstrated by MAIER (1969b). Recently, MAIER and NAPPI (1984) have formulated the finite dimensional incremental holonomic problem as a pair of dual quadratic programming problems on some finite interval Δt of the loading path, and shown that under certain specific assumptions a variety of well established principles may be recovered by suitable interpretation of these problems. In particular, they show that as $\Delta t \to 0$ these quadratic programming problems become formulations of the minimum principles for the rate problem established by CERADINI (1966) and MAIER (1969b) which we referred to earlier. Their suggestion is clearly that a single unified pair of dual minimum principles suffices to define both the holonomic and rate approaches to the elastic-plastic problem.

A state-of-the-art review of mathematical programming applications to engineering plasticity has been given by MAIER and MUNRO (1982), but for an in-depth coverage of all aspects of the theory and application of the method the conference proceedings edited by COHN and MAIER (1979) is essential reading.

To complete our review of the holonomic problem we should mention the work of ODEN and WHITEMAN (1982) who considered the analysis of a holonomic problem formulated in terms of stress only. Variational inequalities are established and an exterior penalty formulation of the problem is presented for which existence, uniqueness, and convergence theorems are given. Finite element approximations based on the penalised problem are discussed together with some convergence criteria.

Other Variational Formulations of the Elastic-Plastic Problem

Up to now we have been concerned with the variational formulations of the rate and holonomic problems, where, in the majority of cases, the primary objective has been the development of suitable numerical approximations. There has also, however, been much interest in the study and analysis of the mathematical structure of elastic-plastic problems, (in particular the questions of existence, uniqueness, and regularity of solutions), which cannot be classified as either rate problems or holonomic problems (according to the definitions which we gave earlier), and it is these particular cases which we wish to review here.

Perhaps the greatest interest has been in the application of the theory of variational inequalities to the elastic-plastic problem. An early contribution in this field was made by TING (1966), in the context of the elastic-plastic torsion problem, who showed that this problem may be formulated as a variational inequality if a stress function is used. However, it was the definitive work of DUVAUT and LIONS (1972) which provided the major impetus for the mathematical study of the elastic-plastic problem.

Duvaut and Lions formulate the dynamic and quasi-static problems for both a visco-plastic and a perfectly plastic material. Briefly, the quasi-static problem is stated as follows : find the stress σ_{ij} and the displacement u_i in Ω which satisfy

(i) the equations of equilibrium

$\sigma_{ij,j} + f_i = 0$ in Ω

where f_i is a force per unit volume,

(ii) a suitable constitutive law, written in terms of stress rates, velocities (i.e. displacement rates), and plastic strain rates,

(iii) suitable boundary conditions on the stress and displacement.

Existence, uniqueness and regularity of solutions is then proved by setting up a weak form of the problem containing a variational inequality. The procedure involves first eliminating the velocities from the formulation, showing the existence of the stress field, and finally proving the existence of a corresponding velocity field. For the perfectly plastic case, the determination of the velocity field presents a problem which Duvaut and Lions left unresolved. JOHNSON (1976a), using the same formulation as that used by Duvaut and Lions, was able to resolve this problem by introducing an additional assumption on the behaviour of the stresses. He subsequently extended his existence results to include hardening (JOHNSON (1978)), by introducing an additional solution variable called a hardening parameter.

Based on his earlier work, Johnson has described finite element approximations for a perfectly plastic material (JOHNSON (1976b)) and a hardening material (JOHNSON (1977)). In the former case an error estimate is derived, and in the latter case, in which a mixed method is used, he proves convergence of an iterative scheme, based on Uzawa's method, for obtaining the finite element solution, . Although Johnson himself gives no numerical results, SAMUELSSON and FROIER (1978) have done so for Johnson's hardening material, using his suggested numerical scheme. HLAVACEK (1980) considers a mixed finite element approximation in which only stresses and hardening parameters are approximated (the velocities being eliminated) and gives error estimates (for three different types of boundary conditions) and a proof of convergence of his numerical approximation.

The problem of determining the velocities for a perfectly plastic material (left unresolved by Duvaut and Lions, and resolved only with the aid of additional assumptions by Johnson) was formally resolved by SUQUET (1978a,b), by recognising that the conventional Sobolev spaces provide too restrictive a setting in which to seek a solution. Physically, Sobolev spaces do not admit the possibility of slip lines across which the velocity is discontinuous. This deficiency has been overcome by requiring that the displacements belong to a space of integrable functions for which the corresponding strain is a bounded measure. This space, denoted BD(Ω) is called the space of functions of bounded deformation, and was first introduced by MATTHIES, STRANG and CHRISTIANSEN (1979), SUQUET (1978a), and TEMAM and STRANG (1978). It was within the space BD(Ω) that Suquet succeeded in proving the existence of both stresses and velocities. Subsequently, SUQUET (1981) extended this work to consider the dynamic problem for a large class of dissipative materials.

COMPUTATIONAL ASPECTS

The most successful numerical solutions for the elastic-plastic problem are based on the classical kinematic minimum principle for the rate problem, formulated in terms of velocities. Discretisation in space is accomplished by using the finite element method, and discretisation in the parameter t is usually done on an ad hoc basis. Thus, if at some point $t = t_i$ in the history of response the discrete displacements $\underset{\sim}{u}_i$ corresponding to discrete loading $\underset{\sim}{P}_i$ are known, then we can write

$$\underset{\sim}{K}_T^i \, \dot{\underset{\sim}{u}} = \dot{\underset{\sim}{P}} \qquad (1.1)$$

where $\dot{\tilde{u}}$ and $\dot{P}$ are corresponding rates at $t = t_i$, and K_T^i is the tangent stiffness matrix, which depends on the solution $\tilde{u}_i$. From (1.1) we may easily solve for the rates $\dot{\tilde{u}}$ and then integrate these rates forward over some finite chosen interval $\Delta t = t_{i+1} - t_i$ to find the response at $t = t_{i+1}$. This is called the tangent stiffness (or tangent modulus) method* and represents a well-established and powerful technique for solving a wide variety of elastic-plastic problems (see, for example, ZIENKIEWICZ (1977), OWEN and HINTON (1980), and BATHE (1982)).

Various state determination schemes (for forward integration and updating of the various response quantities) have been suggested, all of which are based on heuristic arguments. The earliest, and simplest, scheme involves an Euler forward integration in which $\dot{\tilde{u}}$ and $\dot{\tilde{P}}$ are replaced by $\Delta\tilde{u}_{i+1}$ and $\Delta\tilde{P}_{i+1}$ respectively (MARCAL and KING (1967)). Subsequently, a number of implicit iterative schemes based on Newton-Raphson methods were introduced and these remain the most effective schemes in use today (see, in particular, OWEN and HINTON (1980) and MARQUES and OWEN (1984)). Recent studies include an analysis of accuracy and stability of various state determination schemes (ORTIZ and POPOV (1984)), and a proposal for the notion of consistency between the tangent stiffness operator and the state determination algorithm (SIMO and TAYLOR (1985)); the latter study was motivated by the desire to maintain the quadratic convergence characteristics of the Newton-Raphson method.

---

* There are several variants of the method of which the best known is the initial stress method for which $\tilde{K}_T^i$ remains constant throughout the sequence of incremental problems.

Although studies of this type lend credibility to certain state determination schemes they do not address the fundamental problem of attempting to measure the accuracy of a sequence of incremental solutions; that is, in what way, if any, the full solution can be expected to improve as the number of increments is increased.

In the search for a mechanical principle which incorporates both spatial discretisation and discretisation of the parameter t , MARTIN (1986) has provided a partial answer to the above question. Spatial discretisation is assumed to be adequately represented by a suitable finite element mesh, but the choice of a consistent algorithm for the integration of the constitutive equations (with respect to t) remained unresolved until recently when MARTIN, REDDY, GRIFFIN and BIRD (1984) showed that the form of the governing equations of plasticity, written in discrete internal variable form, dictates the choice of a backward difference algorithm for the integration of the internal variables. In mechanical terms this choice is equivalent to the assumption of extremal work and complementary work paths in strain and stress space respectively, which in turn is the foundation of the consistent holonomic theory of PONTER and MARTIN (1972). Thus, the assumption of holonomic material behaviour implies a consistent choice of algorithm for the integration of the constitutive equations.

Now solutions obtained via the consistent holonomic theory (or deformation theory) are related to the continuous solutions via the deformation theory complementary work bounding principle (HODGE (1966)). What MARTIN (1986) has attempted to show is that a similar principle can be established for the incremental holonomic problem, which will bound the complementary work computed as a function of the total number of increments used to approximate the continuous solution. To date only preliminary results in this direction have been obtained and we return to a discussion and numerical confirmation of these in Chapter 7. Suffice it to say that this work has provided a strong motivation for our present detailed study of the incremental holonomic problem.

Numerical approximations based on the extended kinematic minimum principle of MARTIN (1975b) have been given by MARTIN and REDDY (1977) for trusses, REDDY and MITCHELL (1983) for plates, and DITTMER, GRIFFIN and MARTIN (1985) for two-dimensional continua. Although the problem here is a quadratic programming one, the numerical solutions are not obtained using conventional quadratic programming algorithms. Instead, a tangent stiffness approach is used in conjunction with a simple ad-hoc algorithm for handling the inequality constraints. The subsequent recognition that the inequality constraints could be handled formally using an exterior penalty algorithm led to the present work, part of which has been reported by REDDY, GRIFFIN and MARAIS (1985).

Mathematical programming techniques provide alternative methods for the numerical solution of elastic-plastic problems. There appear to be a bewildering array of quadratic programming algorithms available, a trend which has been actively encouraged in the interests of exploiting particular features of different problems. This trend has been recognised, however, as being not in the interests of the general engineering user since the selection of an appropriate algorithm requires extensive experience. This has led to the development of a general purpose computer program STRUPL (COHN and FRANCHI (1979)), designed to automatically select the appropriate algorithm for any given problem. Nevertheless, even with such programs available, of which there appears to be only one, quadratic programming methods must compete with tangent stiffness methods which have a very wide and well-established base within the engineering community. Moreover, the chances of quadratic programming methods being generally accepted are not helped by the fact that the method remains fairly restricted with respect to the classes of problems to which it may be easily and effectively applied. An extensive review of mathematical programming applications is given in COHN and MAIER (1979), and again we mention the recent state-of-the-art survey by MAIER and MUNRO (1982) for further remarks.

It is particularly interesting that numerical quadratic programming solutions are based almost exclusively on the incremental holonomic formulation of the elastic-plastic problem*, with piecewise linearised yield surfaces. Certainly those working with quadratic programming methods recognise inherent advantages in this formulation (see COHN and MAIER (1979), Chapter 15) : for example, the elimination of the requirement of numerical forward integration and its associated error, being the major cause of concern in the incremental rate problem.

---

* This contrasts with the tangent stiffness method which is based exclusively on the rate formulation. The present work appears to be the first application of the incremental holonomic problem which is not based on quadratic programming methods.

FRANCHI and GENNA (1984) have shown that the initial stress tangent stiffness algorithm (ZIENKIEWICZ (1977)) can be cast in the form of a nonlinear programming problem which uses an incremental holonomic constitutive law with a backward difference integration method. This work parallels to a large extent that reported by MARTIN, REDDY, GRIFFIN and BIRD (1984) referred to earlier. Again, the useful insights provided by mathematical programming methods into the consistent formulation of elastic-plastic problems are apparent.

OBJECTIVES

We propose to study first an incremental holonomic boundary-value problem based on a constitutive law which is an extension of that given by PONTER and MARTIN (1972), and second, a rate boundary-value problem based on a conventional rate constitutive law. In both cases we start with the partial differential equations and inequalities which describe the problems, and show that both these problems have a common variational setting. In particular, we show that both these problems are naturally formulated as variational inequalities : in the case of the incremental holonomic problem the inequality is due to the presence of a non-differentiable function in the original boundary-value problem and is known as a variational inequality of the second kind; in the case of the rate problem the inequality is due to the presence of inequality constraints in the original boundary-value problem and is known as a variational inequality of the first kind. The minimum principles first given by PONTER and MARTIN (1972) and MARTIN (1975b) then arise automatically from standard results in convex analysis. Current interest in a unified formulation of the elastic-plastic problem has also provided a motivation for studying these two problems in parallel, although, apart from confirming numerically results given by MARTIN (1986), we do not propose to make a definitive contribution in this area.

Our study of these two problems follows a parallel development within a variational framework*. We begin by establishing variational inequalities which are equivalent statements of the original boundary-value problems. In the case of the rate problem we generalise the treatment of JIANG (1984) by distinguishing between elastic and plastic zones (Jiang considers bodies whch are everywhere plastic), and by dealing with an arbitrary convex, continuously differentiable yield function. Minimum principles, involving the constrained minimisation of a functional J , then follow in a natural way, as mentioned above. We then introduce perturbed variational principles in which the original functionals J are replaced by perturbed functionals J_ε which depend on a parameter $\varepsilon > 0$. In the rate problem ε is a penalty parameter : here J_ε differs from J by a term $\varepsilon^{-1}j(\cdot)$ where $j(\cdot)$ is a penalty functional which allows the non-negativity constraint on the internal variables (plastic multipliers in this case) to be removed. We also discuss a saddle-point formulation of the rate problem. In the incremental holonomic problem the non-differentiable plastic work function $\hat{W}^p(\cdot)$ is

---

* Although our studies in this field were motivated by the work of DUVAUT and LIONS (1976), our ideas have been strongly influenced by the numerous studies of variational principles and numerical analysis of problems in mechanics by Professor J T Oden and his co-workers.

regularised and replaced by a differentiable function $\hat{W}^p_\varepsilon(\cdot)$. For both problems we show that the perturbed solutions converge to the exact solutions as the parameter ε approaches zero.

The perturbed functionals form the basis for finite element approximations, leading to a system of algebraic equations for each problem. For the rate problem these equations represent the discrete approximations of the displacement rates and plastic multipliers, and we show that a condensed form of these equations, from which the plastic multipliers have been eliminated, is identical to that used in the conventional tangent stiffness approach. Our work on the rate problem constitutes a formalisation of that reported earlier by DITTMER, GRIFFIN and MARTIN (1985). In the case of the incremental holonomic problem the algebraic equations represent the discrete approximations of the displacement and plastic strain increments; unlike the rate problem, these equations are nonlinear in the plastic strain increments. We solve these equations using Newton's method, thus providing a direct solution for the incremental problem. For both problems we provide a full analysis of the convergence of solutions to the perturbed problems, and give estimates of the errors in the numerical approximations in terms of the penalty (or regularisation) parameter ε and the finite element mesh size h (for regular mesh refinements).

We discuss several worked examples to illustrate the effectiveness of our numerical solutions as compared to solutions obtained using alternative methods, both analytical and numerical. We suggest that both the penalty-rate and incremental holonomic formulations offer viable alternatives for the solution of elastic-plastic problems, and that the penalty-rate formulation in particular has advantages for certain applications.

To summarise, we regard our work as being of an essentially investigative nature where we attempt to clarify the variational structures of two important elastic-plastic problems. In addition we use parts of these respective structures as the basis for numerical approximations which we analyse for convergence, and whose efficacy we investigate via numerical examples. We regard the fundamental study of these problems as relatively complete and hope that this work provides a foundation for their successful exploitation, particularly in the case of the incremental holonomic problem.

PLAN OF THIS THESIS

In Chapter 2 we discuss in detail constitutive laws for elastic-plastic hardening materials. We develop first the classical rate constitutive equations and then, extending the work of PONTER and MARTIN (1972), we develop a constitutive law for an incremental holonomic material.

Chapters 3 and 4 are devoted to the incremental holonomic problem. In Chapter 3 we discuss theoretical aspects : statement of the boundary-value problem, various variational principles, statement of the problem on a finite dimensional subspace(s), and an estimate of the error in the solution of the finite dimensional approximation. In Chapter 4 we discuss computational aspects : finite element approximations and numerical procedures for obtaining the solution.

Chapters 6 and 7 are devoted to the rate problem with the development being identical to that described above for Chapters 3 and 4.

In Chapter 7 we discuss numerical examples for both the incremental holonomic and rate problems and in Chapter 8 we present our conclusions.

Appendices A and B and the list of References will be found following Chapter 8.

CHAPTER 2

CONSTITUTIVE EQUATIONS FOR ELASTIC-PLASTIC CONTINUA

We propose to treat in this thesis two boundary-value problems each of which may be used under appropriate conditions to describe the behaviour of elastic-plastic continua. Both boundary-value problems have the same physical foundations : they employ the same stress and strain measures and obey the same equilibrium equations. They differ, however, in the constitutive equations which govern the material behaviour, although even here both sets of constitutive equations are based on what is commonly regarded as the classical theory of plasticity for small strains. The first set of constitutive equations with which we will deal are the well known rate constitutive equations which relate rates of change of stress and strain along paths in stress and strain space which remain a priori unspecified. The second set of constitutive equations are based on the assumption that the stress and strain paths are known in principal beforehand, and are extremal paths in a sense which we will describe later. These equations define a nonlinear elastic material which is equivalent, under appropriate conditions, to the original elastic-plastic material.

We begin this chapter with a brief description in Section 2.1 of the stress and strain measures, and the equilibrium equations which together will form the foundation of our boundary-value problems. Since the rate constitutive equations are now well-established we present in Section 2.2 an overview of their development which follows the monograph of MARTIN (1975a). In our later numerical work we will restrict our attention to materials which exhibit linear kinematic hardening and which obey the von Mises yield criterion, so that Section 2.3 is devoted to a discussion of these.

The remainder of the chapter is devoted to a detailed development of the constitutive equations based on extremal paths. We discuss first the fundamental constitutive equations in Section 2.4, and follow this with a development of suitable criteria for determining the extremal paths themselves in Section 2.5.

Before proceeding with the body of the chapter we define the notation which we will be using throughout this work.

NOTATION

Throughout this thesis we will use coordinate-free notation as far as possible, but indicial notation will also be used wherever additional clarity is necessary.

We denote by R^N the set of all ordered n-tuples of real numbers. Let $\underset{\sim}{e}_i$ $(i = 1,2, \ldots, N)$ constitute a fixed orthonormal basis for R^N. Then any vector $\underset{\sim}{v}$ has the representation

$$\underset{\sim}{v} = v_i \underset{\sim}{e}_i \qquad (1)$$

and we identify vectors with elements of R^N.

We adopt the summation convention throughout, unless indication to the contrary is given : when a letter subscript is repeated in a term it denotes the sum of all the terms obtained by giving the letter subscript the values 1,2, ..., N .

The orthonormal basis $\underset{\sim}{e}_i$ satisfies

$$\underset{\sim}{e}_i \cdot \underset{\sim}{e}_j = \delta_{ij} \tag{2}$$

$$\text{where} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{3}$$

is the Kronecker delta. The scalar product of any two vectors $\underset{\sim}{u},\underset{\sim}{v}$ is given by

$$\begin{aligned} \underset{\sim}{u} \cdot \underset{\sim}{v} &= u_i v_j \underset{\sim}{e}_i \cdot \underset{\sim}{e}_j \\ &= u_i v_i \quad , \quad \text{using (2)} . \end{aligned} \tag{4}$$

If the scalar product of two non-zero vectors $\underset{\sim}{u},\underset{\sim}{v}$ is zero, that is, $\underset{\sim}{u} \cdot \underset{\sim}{v} = \underset{\sim}{0}$, then $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are orthogonal vectors. Here $\underset{\sim}{0}$ indicates the zero vector.

A second-order tensor is a linear map of the space of vectors into itself. We can define a basis $\underset{\sim}{e}_i \otimes \underset{\sim}{e}_j$, (i,j = 1, ..., N) for this space of tensors where $(\underset{\sim}{e}_i \otimes \underset{\sim}{e}_j)\underset{\sim}{e}_k = \underset{\sim}{e}_i \delta_{jk}$. Then any second-order tensor $\underset{\sim}{T}$ has the representation

$$\underset{\sim}{T} = T_{ij} \, \underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \tag{5}$$

and we have, for any vectors $\underset{\sim}{u}, \underset{\sim}{v}$

$$\underset{\sim\sim}{T}\underset{\sim}{u} = \underset{\sim}{v} \quad \text{or} \quad T_{ij}u_j = v_i \quad . \tag{6}$$

We identify second-order tensors with elements of R^{NxN} . The scalar product of two tensors $\underset{\sim}{T}, \underset{\sim}{S}$ is defined by

$$\underset{\sim}{T} \cdot \underset{\sim}{S} = T_{ij}S_{ij} \quad . \tag{7}$$

If the scalar product of two non-zero second-order tensors $\underset{\sim}{T}, \underset{\sim}{S}$ is zero, then $\underset{\sim}{T}$ and $\underset{\sim}{S}$ are orthogonal. The inner product on the space of second order tensors is defined by

$$(\underset{\sim}{T}, \underset{\sim}{S}) = \underset{\sim}{T} \cdot \underset{\sim}{S}$$

and the norm generated by the inner product is

$$|\underset{\sim}{T}| = \sqrt{\underset{\sim}{T} \cdot \underset{\sim}{T}} \quad .$$

The Schwarz inequality is

$$|\underset{\sim}{T} \cdot \underset{\sim}{S}| \leq |\underset{\sim}{T}| \; |\underset{\sim}{S}| \quad .$$

A fourth-order tensor is a linear map of the space of second-order tensors to itself. We can define a basis $\underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \otimes \underset{\sim}{e}_k \otimes \underset{\sim}{e}_\ell$ (i,j,k,l = 1 , ..., N) for this space of tensors where

$$(\underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \otimes \underset{\sim}{e}_k \otimes \underset{\sim}{e}_\ell)\underset{\sim}{e}_r \otimes \underset{\sim}{e}_s = \underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \delta_{kr} \delta_{\ell s} \quad .$$

Then any fourth-order tensor $\underset{\sim}{C}$ has the representation

$$\underset{\sim}{C} = C_{ijk\ell}\underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \otimes \underset{\sim}{e}_k \otimes \underset{\sim}{e}_\ell$$

and we have, for any second-order tensors $\underset{\sim}{T}, \underset{\sim}{S}$

$$\underset{\sim}{C}\underset{\sim}{T} = \underset{\sim}{S} \quad \text{or} \quad C_{ijk\ell}T_{k\ell} = S_{ij} \quad .$$

We identify fourth-order tensors with elements of $R^{NxNxNxN}$.

It should be pointed out that we make no distinction between upper-case and lower-case letters for the naming of vectors and tensors. Whether a given quantity is a vector or tensor will be made clear when the quantity is first mentioned in the text.

Vectors and Tensor Fields

Let $\underset{\sim}{x}$ be a point in a bounded domain $\Omega \subset R^n$, and let $\underset{\sim}{u}(\underset{\sim}{x})$ denote a vector field on Ω . The gradient of a vector field is a second-order tensor field defined by

$$\text{grad } \underset{\sim}{u} \equiv \nabla \underset{\sim}{u} = \frac{\partial u_i}{\partial x_j} \underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \quad . \tag{10}$$

The transpose of this tensor is given by

$$(\nabla \underset{\sim}{u})^T \equiv \nabla^T \underset{\sim}{u} = \frac{\partial u_j}{\partial x_i} \underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \quad . \tag{11}$$

We may on occasion make use of the more concise comma notation for writing derivatives; thus we may use either of the following forms :

$$\frac{\partial u_i}{\partial x_j} \equiv u_{i,j}$$

$$\text{or} \quad \frac{\partial T_{ij}}{\partial x_j} \equiv T_{ij,j} \quad . \tag{12}$$

Let $\underset{\sim}{T}(\underset{\sim}{x})$ denote a second-order tensor field on Ω . The divergence of a second-order tensor field is a vector field defined by

$$\text{div } \underset{\sim}{T} = \frac{\partial T_{ij}}{\partial x_j} \underset{\sim}{e}_i \quad . \tag{13}$$

Let $\phi(\underset{\sim}{T})$ be a scalar field on Ω which is a function of the second-order tensor field $\underset{\sim}{T}(\underset{\sim}{x})$. The gradient of ϕ with respect to $\underset{\sim}{T}$ is a second-order tensor defined by

$$\frac{\partial \phi}{\partial \underset{\sim}{T}} = \frac{\partial \phi}{\partial T_{ij}} \underset{\sim}{e}_i \otimes \underset{\sim}{e}_j \quad . \tag{14}$$

If we imagine an (NxN) dimensional subspace defined by the components of $\underset{\sim}{T}$, then $\phi = 0$ may be regarded as a surface in this subspace; we will then refer to the gradient $\partial\phi/\partial\underset{\sim}{T}$ as the outward normal to this surface at the point $\underset{\sim}{T}$.

2.1 FRAMEWORK FOR THE PROBLEM

We consider a homogeneous, isotropic material body which occupies an open bounded domain Ω in R^N , $N \leq 3$ (Fig. 2.1). Each material point in the body is identified by its position vector $\underset{\sim}{x}$, or its coordinates

x_i, $1 \leq i \leq N$, with respect to a fixed cartesian set of axes, X_i . The vector field $\underset{\sim}{u}(\underset{\sim}{x})$ represents the displacement, and since displacements are assumed to be small, we make use of the (infinitesimal) strain tensor

$$\underset{\sim}{\varepsilon} = \frac{1}{2} (\nabla \underset{\sim}{u} + \nabla^T \underset{\sim}{u}) \quad . \tag{1.1}$$

The strain tensor is a second-order symmetric tensor by definition.

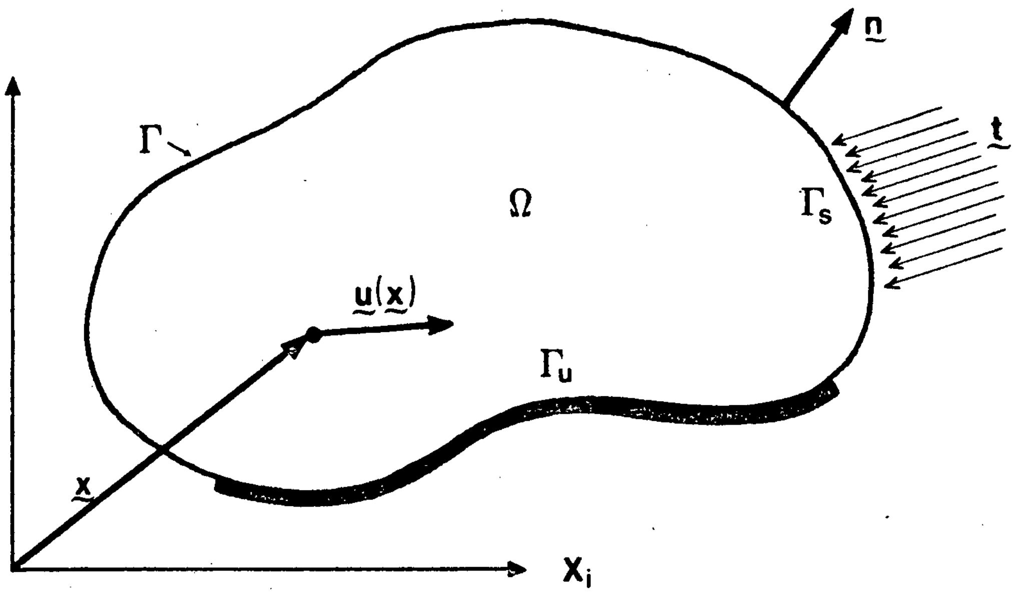


Figure 2.1 The material body Ω .

At each material point in the body we require that the equations of equilibrium are satisfied :

$$\text{div}\, \underset{\sim}{\sigma} + \underset{\sim}{f} = \underset{\sim}{0} \qquad \text{on } \Omega \tag{1.2}$$

where $\underset{\sim}{\sigma}(\underset{\sim}{x})$ is the Cauchy stress tensor field and $\underset{\sim}{f}$ is the body force per unit volume, assumed constant over Ω . The stress $\underset{\sim}{\sigma}$ is a symmetric second-order tensor by definition, and the force $\underset{\sim}{f}$ is a vector.

The boundary Γ of the domain Ω is assumed to be Lipschitz and divided into two non-overlapping parts, Γ_u and Γ_s, such that $\Gamma = \Gamma_u \cup \Gamma_s$. The displacement field $\underset{\sim}{u}(\underset{\sim}{x})$ will be assumed to be given over Γ_u, while a traction vector $\underset{\sim}{t}$ is prescribed over the remainder of the boundary, Γ_s. If the outward normal vector at any point on Γ_s is $\underset{\sim}{\nu}$ then the traction vector $\underset{\sim}{t}$ is related to the stress tensor $\underset{\sim}{\sigma}$ at that point by

$$\underset{\sim\sim}{\sigma\nu} = \underset{\sim}{t} \quad . \tag{1.3}$$

The above relations constitute the basic framework for the boundary-value problem in which we are interested. To complete the framework, however, we require one or more constitutive equations which govern the relationship between stress and strain. The remainder of this chapter is devoted to this important topic.

2.2 THE CLASSICAL RATE CONSTITUTIVE EQUATIONS

We introduce in this section the fundamental relationships governing the behaviour of our idealised elastic-plastic material. At the outset we restrict our discussion to behaviour which is time-independent, path-dependent, and takes place under isothermal conditions. The types of materials which we have in mind as falling within the framework to be discussed here are polycrystalline metals at room temperature. We present the results in this section in the form of a review, and the interested reader may refer to the monograph of MARTIN (1975a) for an in-depth analysis of the subject.

We distinguish between two types of idealised elastic-plastic materials : hardening materials, in which changes in strain at constant stress do not occur, and materials which exhibit flow, which means that changes in strain at constant stress can occur. Unless otherwise specifically stated, we shall assume throughout that we are dealing with a hardening material.

Yield Surfaces

A material point in a given state of stress may exhibit either elastic or elastic-plastic behaviour. Thus we assume that the strain is divisible into two parts,

$$\underset{\sim}{\varepsilon} = \underset{\sim}{e} + \underset{\sim}{p} \; ; \tag{2.1}$$

the elastic strain tensor $\underset{\sim}{e}$ is related to the Cauchy stress $\underset{\sim}{\sigma}$ through

$$\underset{\sim}{e} = \underset{\sim\sim}{D\sigma} \tag{2.2a}$$

where $\underset{\sim}{D}$ is a symmetric, positive-definite fourth-order tensor of elastic constants. The inverse relationships may be written as

$$\underset{\sim}{\sigma} = \underset{\sim\sim}{Ce} \tag{2.2b}$$

where $\underset{\sim}{C}$ is the inverse of $\underset{\sim}{D}$ and is assumed to exist. The plastic strain tensor $\underset{\sim}{p}$ is defined by (2.1).

The presence of plastic strains $\underset{\sim}{p}$ causes the behaviour of the material to be history-dependent. We choose to characterise this history-dependence by a set of internal variables H_α , $\alpha = 1,2, \ldots, \eta$, whose nature and number η we leave unspecified for the present. The state of an element of material is then defined by the stress $\underset{\sim}{\sigma}$ and the internal variables H_α . We may then imagine a multi-dimensional $(\underset{\sim}{\sigma},H_\alpha)$ space which is defined in such a way that the state of the material is uniquely represented by a point in this space.

In order to distinguish between elastic and elastic-plastic behaviour in an element of the material we assume that there exists a convex region in $(\underset{\sim}{\sigma},H_\alpha)$ space, bounded by a hypersurface called the yield surface, such that if the material state point lies within this region the behaviour is elastic and path-independent, and the plastic strain $\underset{\sim}{p}$ does not change. The yield surface is characterised by a yield function

$$\phi = \phi(\underset{\sim}{\sigma},H_\alpha) = 0 \qquad (2.3)$$

such that $\phi \leq 0$ for material states which lie within or on the yield surface; material states for which $\phi > 0$ are not admissable.

We will find it more convenient when dealing with the yield surface to define a subspace of the more general $(\underset{\sim}{\sigma},H_\alpha)$ space called the stress space, in which the H_α are assumed fixed and are the current values of the internal variables in the material element. The projection of the yield surface $\phi = 0$, for fixed H_α , into stress space may then be expressed as

$$\phi(\underset{\sim}{\sigma}) = 0 \qquad (2.4)$$

wherein it is tacitly assumed that the H_α are fixed at their current values. The yield surface in stress space is shown schematically in Fig. 2.2.

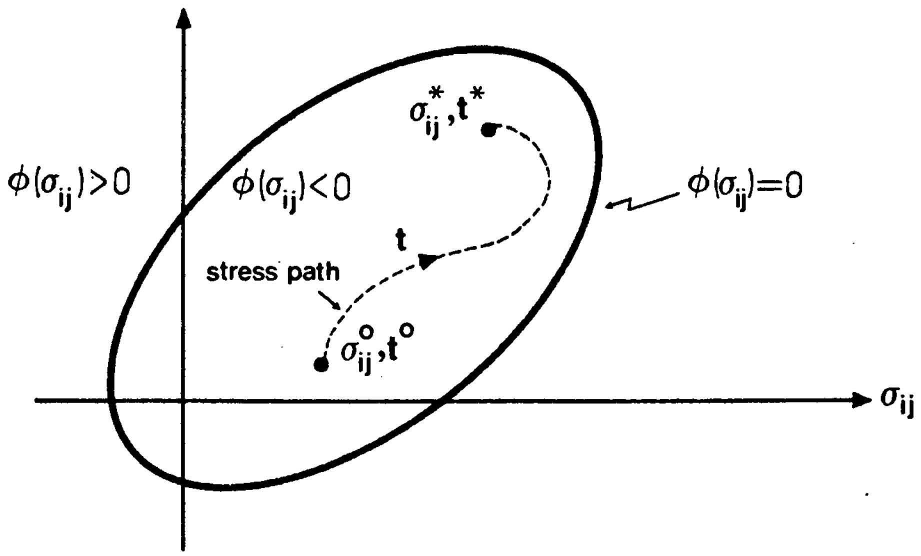


Figure 2.2 The yield surface in stress space.

Any stress state $\underset{\sim}{\sigma}$ may be represented by a point in stress space. As the state of stress in an element of material changes, the stress point traces out a trajectory which we call the stress path (Fig. 2.2). Since we have already noted that the material behaviour is history-dependent, it is apparent that the state of stress at any time during a given history of loading will in general depend on the particular stress path which has been followed to reach this state of stress. Thus, in order to account for the path-dependence of the stress point we introduce, as we did in Chapter 1, a scalar parameter t which parametrises the stress path. We may then define the rate of change of the stress state at any point along the stress path as

$$\dot{\underset{\sim}{\sigma}} = \frac{d}{dt} (\underset{\sim}{\sigma}) \quad . \tag{2.5}$$

Referring again to Fig. 2.2, the line integral along the stress path between some initial state $\underset{\sim}{\sigma}^{\circ}$ at $t = t^{\circ}$ and some terminal state $\underset{\sim}{\sigma}^{*}$ at $t = t^{*}$ is given by

$$\int_{t^{o}}^{t^{*}} \dot{\underset{\sim}{\sigma}} \, dt = \int_{\underset{\sim}{\sigma}^{o}}^{\underset{\sim}{\sigma}^{*}} d\underset{\sim}{\sigma} \tag{2.6}$$

and this will represent the total change in the state of stress between the initial state $\underset{\sim}{\sigma}^{\circ}$ and the terminal state $\underset{\sim}{\sigma}^{*}$. Definitions similar to (2.5) may be written for the strain rates $\dot{\underset{\sim}{\varepsilon}}$, $\dot{\underset{\sim}{e}}$, $\dot{\underset{\sim}{p}}$, displacement rates $\dot{\underset{\sim}{u}}$, internal variable rates $\dot{H}_{\alpha}$, and the rate of change of the yield function $\dot{\phi}$.

Let us now consider a material state $(\underset{\sim}{\sigma},H_{\alpha})$ for which $\phi(\underset{\sim}{\sigma},H_{\alpha}) = 0$, and consider changes in the stress state and internal variables represented by $\dot{\sigma}\,dt$ and $\dot{H}_{\alpha}\,dt$ respectively. Changes in the yield function are then represented by

$$\dot{\phi} = \frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} + \frac{\partial\phi}{\partial H_{\alpha}} \cdot \dot{H}_{\alpha} \quad . \tag{2.7}$$

If the new stress point $(\underset{\sim}{\sigma} + \dot{\underset{\sim}{\sigma}}dt)$ lies within the yield surface then clearly $\dot{\phi} < 0$; the behaviour will be entirely elastic and no change in the internal variables will occur. We refer to such a stress change as representing unloading, characterised by

$$\phi(\underset{\sim}{\sigma}) = 0 \quad \text{and} \quad \frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} < 0 \quad . \tag{2.8}$$

During loading the internal variables and the yield surface in stress space must change, but the stress point $(\underset{\sim}{\sigma} + \dot{\underset{\sim}{\sigma}}dt)$ must continue to lie on the yield surface so that $\dot{\phi} = 0$. Thus, loading is characterised by the condition

$$\phi(\underset{\sim}{\sigma}) = 0 \quad \text{and} \quad \frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} > 0 \quad . \tag{2.9}$$

A third possibility exists in which neither the internal variables nor the yield surface in stress space change, but the stress point remains on the yield surface. We refer to this as neutral loading, characterised by

$$\phi(\underset{\sim}{\sigma}) = 0 \quad \text{and} \quad \frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} = 0 \quad . \tag{2.10}$$

We assume that changes in plastic strain occur only during loading, that is, only when there are changes in the internal variables H_α . The magnitude of the change in internal variables is governed by $\dot{\underset{\sim}{\sigma}}$, and the direction of the change by the stress state $(\underset{\sim}{\sigma},H_\alpha)$ at the start of loading. The plastic strain rate $\dot{\underset{\sim}{p}}$ is assumed also to be homogeneous and of degree one in the stress rate $\dot{\underset{\sim}{\sigma}}$. Together with these assumptions we postulate a plastic potential $g(\underset{\sim}{\sigma},H_\alpha)$, such that the plastic strain rate satisfies the following relations :

$$\dot{\underset{\sim}{p}} = 0 \quad \text{if} \quad \phi(\underset{\sim}{\sigma}) < 0 \tag{2.11a}$$
$$\text{or} \quad \phi(\underset{\sim}{\sigma}) = 0 \quad \text{and} \quad \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} \leq 0 \quad ;$$

$$\dot{\underset{\sim}{p}} = G(\underset{\sim}{\sigma},H_\alpha) \frac{\partial g}{\partial \underset{\sim}{\sigma}} (\frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}}) \tag{2.11b}$$
$$\text{if } \phi(\underset{\sim}{\sigma}) = 0 \quad \text{and} \quad \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} > 0$$

where $G(\underset{\sim}{\sigma},H_\alpha)$ is a scalar hardening parameter, as yet unspecified.

Bearing in mind that we are dealing exclusively with hardening materials for which the projection of the yield surface in stress space may change during some program of loading, it is necessary to distinguish between certain basic projections. We will refer to the yield surface at the start of a program of loading as the virgin yield surface, defined by

$$\bar{\phi}(\underset{\sim}{\sigma},0) = 0 \quad . \tag{2.12}$$

Since a change in the projection of the yield surface in stress space must be accompanied by a change in the values of the internal variables H_α , any subsequent yield surface, (including the current yield surface), will be uniquely defined by the current values of H_α . Thus subsequent yield surfaces will be defined by

$$\phi(\underset{\sim}{\sigma},H_\alpha) = 0$$

$$\text{or} \quad \phi(\underset{\sim}{\sigma}) = 0 \tag{2.13}$$

where it is to be understood in the second form that the H_α are fixed.

Uniqueness and Stability Postulates

We now review the two well-known postulates due to D.C. Drucker, which together constitute the definition of a stable plastic material. For a more detailed discussion of these postulates the reader is referred to DRUCKER (1951), or MARTIN (1975a), Section 2.4.

The First Postulate

If the stress state is changed from $\underset{\sim}{\sigma}^{\circ}$ at $t = t^{\circ}$ to $\underset{\sim}{\sigma}^{*}$ at $t = t^{*}$ in such a way that the stress point moves monotonically along a straight-line path in stress space then we require that

$$(\underset{\sim}{\sigma}^{*} - \underset{\sim}{\sigma}^{\circ}) \cdot (\underset{\sim}{\varepsilon}^{*} - \underset{\sim}{\varepsilon}^{\circ}) \geqslant 0 \tag{2.14}$$

where $\underset{\sim}{\varepsilon}^{\circ}$ and $\underset{\sim}{\varepsilon}^{*}$ are the strains associated with $\underset{\sim}{\sigma}^{\circ}$ and $\underset{\sim}{\sigma}^{*}$ respectively.

Consequently, since rates of change occur by definition along straight line paths, we require that

$$\dot{\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\varepsilon}} \geqslant 0 \quad . \tag{2.15}$$

The first postulate has the following two consequences : the net work and the net complementary work along a straight line path in stress space are non-negative, that is

$$\int_{t^{o}}^{t^{*}} (\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^{o}) \cdot \dot{\underset{\sim}{\varepsilon}} \, dt \geqslant 0 \quad , \tag{2.16a}$$

and

$$\int_{t^o}^{t^*} (\underset{\sim}{\varepsilon} - \underset{\sim}{\varepsilon}^o) \cdot \dot{\underset{\sim}{\sigma}}\, dt \geq 0 \quad . \tag{2.16b}$$

Here, $\underset{\sim}{\sigma}$ and $\underset{\sim}{\varepsilon}$ are the stress and strain states at any point t along the respective paths.

Before stating the second postulate we introduce the concept of a cycle in stress as a stress program in which the initial and final values of the stress are identical. We place no restrictions on the strains during such a cycle and assume that the cycle may be either entirely elastic or may also include elastic-plastic behaviour.

<u>The Second Postulate</u>

For a cycle in stress the complementary work is non-positive, that is,

$$\oint \underset{\sim}{\varepsilon} \cdot d\underset{\sim}{\sigma} \leq 0 \quad . \tag{2.17}$$

Furthermore, we distinguish two forms of the second postulate : if the changes in plastic strain are infinitesimal or zero during the cycle we refer to (2.17) as the weak form of the second postulate; alternatively, if no restriction is placed on the size of the plastic strain changes then we refer to (2.17) as the second postulate in its strong form.

If we combine the first postulate and the strong form of the second postulate we obtain the result that the net work associated with any path in stress space is non-negative; that is,

$$\int_{t^o}^{t^*} (\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^o) \cdot \dot{\underset{\sim}{\varepsilon}} \, dt \geq 0 \quad . \tag{2.18}$$

It can be shown that this result implies both the first postulate and the second postulate in its strong form and can thus be used in their place (see MARTIN (1975a), page 96).

The second postulate in its weak form can be used directly to establish a result of fundamental importance in small strain classical plasticity. This is the principle of maximum plastic work which we now state without proof (see MARTIN (1975a), page 101).

The Principle of Maximum Plastic Work

Let the current yield surface be defined by $\phi = 0$, and let $\underset{\sim}{\sigma}^*$ be any stress state for which $\phi(\underset{\sim}{\sigma}^*) = 0$, and at which the plastic strain rate is $\dot{\underset{\sim}{p}}^*$. If $\underset{\sim}{\sigma}$ is any stress state for which $\phi(\underset{\sim}{\sigma}) \leq 0$, then

$$(\underset{\sim}{\sigma}^* - \underset{\sim}{\sigma}) \cdot \dot{\underset{\sim}{p}}^* \geq 0 \quad . \tag{2.19}$$

As we stated above, this result follows from the second postulate in its weak form. However, it may be physically interpreted as requiring that for a given plastic strain rate $\dot{\underset{\sim}{p}}^*$ the associated stress $\underset{\sim}{\sigma}^*$ is distinguished from all other admissable stress states by the requirement that the plastic work $\underset{\sim}{\sigma}^* \cdot \dot{\underset{\sim}{p}}^*$ during the increment takes its greatest value.

The principle of maximum plastic work has the following two important consequences :

(i) the yield surface is convex;

(ii) the plastic strain rate is normal to the yield surface, which may be expressed as

$$\dot{\underset{\sim}{p}} = \lambda \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \tag{2.20}$$

where λ is a positive scalar parameter called the plastic multiplier.

Equation (2.20) expresses the so called normality rule for the plastic strain rate $\dot{\underset{\sim}{p}}$. If we now compare this result with (2.11b) we see that if both results are to be true we may put

$$\lambda = G(\underset{\sim}{\sigma}, H_\alpha) \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} \quad , \quad \text{and} \quad \frac{\partial g}{\partial \underset{\sim}{\sigma}} = \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \tag{2.21}$$

thus identifying the plastic potential g with the yield function ϕ . Since $(\partial\phi/\partial\underset{\sim}{\sigma}) \cdot \dot{\underset{\sim}{\sigma}} > 0$ for loading, (eqn (2.9)), it follows that $G > 0$. On the basis of the above results, the plastic constitutive equations may be written in the following form :

$$\dot{\underset{\sim}{p}} = \lambda \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \tag{2.22}$$

$$\lambda = 0 \qquad \text{if } \phi(\underset{\sim}{\sigma}) < 0 \quad \text{or } \phi(\underset{\sim}{\sigma}) = 0 \text{ and } \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} \leqslant 0 \tag{2.23a}$$

$$\lambda = G \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} \quad \text{if } \phi(\underset{\sim}{\sigma}) = 0 \text{ and } \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} > 0 \ . \tag{2.23b}$$

Equations (2.22) and (2.23) are the rate form of the plastic constitutive equations. To complete the description of an elastic-plastic material it remains only to add the rate form of the elastic constitutive equations, together with the rate form of the subdivision of the strain into its elastic and plastic parts. These results follow immediately from (2.1) and (2.2) and are written as follows :

$$\dot{\underset{\sim}{\varepsilon}} = \dot{\underset{\sim}{e}} + \dot{\underset{\sim}{p}} \tag{2.24}$$

$$\dot{\underset{\sim}{e}} = \underset{\approx}{D}\dot{\sigma} \tag{2.25a}$$

$$\dot{\underset{\sim}{\sigma}} = \underset{\approx}{C}\dot{e} \quad . \tag{2.25b}$$

Equations (2.22) through (2.25) are the complete set of constitutive equations for an elastic-plastic hardening material, written in rate form.

<u>Elastic-Perfectly Plastic Materials</u>

Our development up to now has been restricted to hardening materials. Nevertheless we would also like to make allowance in the formulation for materials which exhibit flow, that is, changes of strain at constant stress. We may do this by introducing a limit function $\phi(\underset{\sim}{\sigma})$ (see MARTIN (1975a), Section 2.6) which for the purposes of the present work we will take to coincide with the virgin yield function $\bar{\phi}(\underset{\sim}{\sigma},0)$. In this case changes in plastic strain occur as a result of flow only, and the yield surface remains fixed in stress space. It will become apparent when we develop explicit expressions for such quantities as the plastic work and the yield function that the quantities required for the description of flow arise quite naturally when considering hardening materials, and that the case of flow may be considered simply as a special case of hardening when some suitable hardening parameter tends to zero. For the sake of completeness, however, we include here the plastic constitutive equations for the case of flow (or for what we shall also refer to as an elastic-perfectly plastic material) :

$$\dot{\underset{\sim}{p}} = \lambda \frac{\partial \psi}{\partial \underset{\sim}{\sigma}} \tag{2.26}$$

$$\begin{aligned} \lambda = 0 \quad & \text{if} \quad \psi(\underset{\sim}{\sigma}) < 0 \\ & \text{or} \quad \psi(\underset{\sim}{\sigma}) = 0 \quad \text{and} \quad \frac{\partial \psi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} < 0 \end{aligned} \tag{2.27a}$$

$$\lambda > 0 \quad \text{if} \quad \psi(\underset{\sim}{\sigma}) = 0 \quad \text{and} \quad \frac{\partial \psi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} = 0 \quad . \tag{2.27b}$$

Here λ is a non-negative but otherwise unspecified scalar.

The relations (2.26) and (2.27) replace (2.22) and (2.23) when the material is assumed to be elastic-perfectly plastic. In the remainder of this thesis we shall not find it necessary to treat the elastic-perfectly plastic case separately since the mechanism of flow will arise naturally as a special case in our consideration of hardening materials.

Summary of the Classical Rate Constitutive Equations

The constitutive equations for a hardening material given in (2.22) through (2.25) may be written more concisely by combining (2.22), (2.24) and (2.25b) to obtain

$$\dot{\underset{\sim}{\sigma}} = \underset{\sim}{C}\,[\dot{\underset{\sim}{\varepsilon}} - \lambda \frac{\partial\phi}{\partial\underset{\sim}{\sigma}}] \tag{2.28}$$

where the plastic multiplier λ is given by (2.23). Later on when we come to discuss a variational setting for the rate problem we will find it convenient to express eqns (2.23) in a slightly different form which we now describe.

If we assume that the initial stress at $t = t^\circ$ is known then the region Ω may be divided into two non-overlapping open regions Ω^e and Ω^p defined by

$$\begin{aligned} \Omega^e &= \{\underset{\sim}{x} \in \Omega \;:\; \phi(\underset{\sim}{\sigma}(\underset{\sim}{x})) < 0\} \\ \Omega^p &= \{\underset{\sim}{x} \in \Omega \;:\; \phi(\underset{\sim}{\sigma}(\underset{\sim}{x})) = 0\} \\ \Omega^e \cup \Omega^p &= \Omega \;, \quad \bar{\Omega}^e \cap \bar{\Omega}^p = \Gamma^{ep} \end{aligned} \tag{2.29}$$

where $\bar{\Omega}^e, \bar{\Omega}^p$ denote the closures of Ω^e and Ω^p , and Γ^{ep} is the elastic-plastic interface. Thus, purely elastic behaviour will take place in Ω^e while either elastic or plastic behaviour may occur in Ω^p . Continuity considerations dictate that elastic behaviour occurs on Γ^{ep} .

Let us now define the scalar parameter κ by

$$\kappa = \frac{1}{G}\lambda - \frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} \;. \tag{2.30}$$

Then with the definitions (2.29) in mind it is readily verified that eqns (2.23) may be written as

$$\lambda = 0 \qquad \text{on } \Omega^e$$

$$\lambda \geqslant 0 \ , \quad \kappa \geqslant 0 \ , \quad \lambda\kappa = 0 \quad \text{on } \Omega^p \ . \tag{2.31}$$

Eqns (2.23), or alternatively eqns (2.31), refer of course to the loading-unloading conditions for the elastic-plastic material. It is in this context that eqns (2.31) are sometimes referred to as the Kuhn-Tucker form of the loading-unloading conditions, being the form commonly used in the mathematical programming formulation of the elastic-plastic problem.

2.3 THE VON MISES YIELD SURFACE WITH LINEAR KINEMATIC HARDENING

The elastic-plastic constitutive equations which were established in the preceding section depend explicitly on the yield function ϕ and the hardening parameter G . In the later stages of our development of constitutive equations based on extremal paths (Sections 2.4 and 2.5) we will find it necessary to have at hand explicit expressions for ϕ and G . We propose now to develop the necessary expressions, and we choose to do so using a von Mises yield function with linear kinematic hardening.

We assume that for the materials in which we are interested, namely polycrystalline metals, there is no volume change associated with plastic strain, so that

$$p_{kk} = 0 \ . \tag{3.1}$$

It follows that the yield surface must be independent of the mean hydrostatic tension $\sigma_{kk}/3$, and consequently it is a function of the stress deviator $\underset{\sim}{s}$, so that

$$\phi = \phi(\underset{\sim}{s}, H_\alpha) \tag{3.2}$$

where $\underset{\sim}{s}$ is a symmetric second-order tensor with components

$$s_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij} \quad . \tag{3.3}$$

We introduce the plastic strain deviator $\underset{\sim}{r}$ whose components are

$$r_{ij} = p_{ij} - \frac{1}{3}p_{kk}\delta_{ij} \quad . \tag{3.4}$$

In view of (3.1) it follows that

$$\underset{\sim}{r} = \underset{\sim}{p} \quad . \tag{3.5}$$

For linear kinematic hardening the internal variables H_α are the plastic strains themselves so that we may replace (3.2) by

$$\phi = \phi(\underset{\sim}{s}, \underset{\sim}{p}) = \phi(\underset{\sim}{s}, \underset{\sim}{r}) \quad . \tag{3.6}$$

In accordance with (3.6) we adopt the so called J_2 form of the von Mises yield function for linear kinematic hardening :

$$\phi = \frac{1}{2}(\underset{\sim}{s} - h\underset{\sim}{r})(\underset{\sim}{s} - h\underset{\sim}{r}) - k^2 \tag{3.7}$$

where h is a positive hardening constant and $k = \sigma_o/\sqrt{3}$, σ_o being the virgin yield stress obtained from a uniaxial tension test. Equation (3.7) defines a hypersphere in the deviatoric stress space whose centre is given by the quantity $h\underset{\sim}{r}$; thus the virgin yield surface is given by (3.7) with $h\underset{\sim}{r} = \underset{\sim}{0}$.

During loading the stress point must remain on the yield surface so that we require

$$\dot{\phi} = \frac{\partial \phi}{\partial \underset{\sim}{s}} \cdot \dot{\underset{\sim}{s}} + \frac{\partial \phi}{\partial \underset{\sim}{r}} \cdot \dot{\underset{\sim}{r}} = 0 \tag{3.8}$$

where the rate of change of the plastic strain deviator is given by

$$\dot{\underset{\sim}{r}} = G \frac{\partial \phi}{\partial \underset{\sim}{s}} \left(\frac{\partial \phi}{\partial \underset{\sim}{s}} \cdot \dot{\underset{\sim}{s}}\right) . \tag{3.9}$$

Substituting (3.9) into (3.8) and rearranging, we have

$$G = - \left(\frac{\partial \phi}{\partial \underset{\sim}{r}} \cdot \frac{\partial \phi}{\partial \underset{\sim}{s}}\right)^{-1} . \tag{3.10}$$

Noting from (3.7) that

$$\frac{\partial \phi}{\partial \underset{\sim}{r}} = - h \frac{\partial \phi}{\partial \underset{\sim}{s}} \quad \text{and} \quad \frac{\partial \phi}{\partial \underset{\sim}{s}} \cdot \frac{\partial \phi}{\partial \underset{\sim}{s}} = 2k^2 \tag{3.11}$$

and making use of these results in (3.10) we have

$$G = \frac{1}{2hk^2} = \frac{3}{2h\sigma_o^2} . \tag{3.12}$$

For numerical purposes it is useful to relate the parameters G and h to data which is directly available from uniaxial stress-strain curves. Referring to Fig. 2.3 and using (2.24) we may define a plastic modulus E_p by

$$\frac{1}{E_p} = \frac{1}{E_T} - \frac{1}{E} \tag{3.13}$$

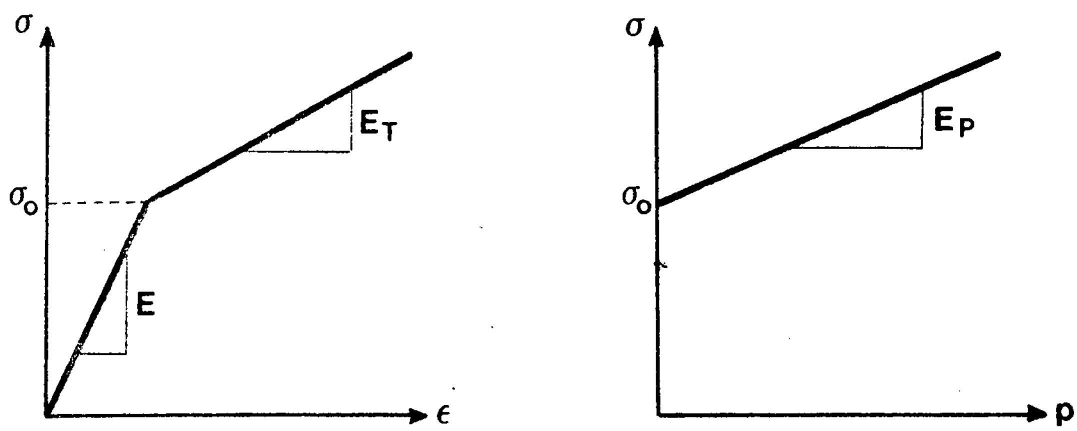


Figure 2.3 Uniaxial stress-strain curves defining E, E_T and E_p .

where E is Young's modulus and E_T is the tangent modulus. It is a simple matter to then show that

$$h = E_p \quad \text{and} \quad G = \frac{3}{2E_p\sigma_o^2} \quad . \tag{3.14}$$

We will also have occasion to use the von Mises yield function written in terms of total stresses and plastic strains. Writing $\tau_{ij} = \sigma_{ij} - hp_{ij}$ we have for a general state of stress

$$\phi = \phi(\tau_{ij})$$

$$= \frac{1}{6}\,[(\tau_{11} - \tau_{22})^2 + (\tau_{22} - \tau_{33})^2 + (\tau_{33} - \tau_{11})^2] + \tau_{12}^2 + \tau_{23}^2 + \tau_{31}^2 - k^2 \quad . \tag{3.15}$$

This form is equivalent to the deviatoric form (3.7) so that the parameters h and k have the same meaning as before. Again, it is easy to show that

$$\frac{\partial\phi(\underset{\sim}{\sigma},\underset{\sim}{p})}{\partial\underset{\sim}{\sigma}} = \frac{\partial\phi(\underset{\sim}{s},\underset{\sim}{r})}{\partial\underset{\sim}{s}} \tag{3.16a}$$

and that $\dfrac{\partial\phi(\underset{\sim}{\sigma},\underset{\sim}{p})}{\partial\underset{\sim}{p}} = \dfrac{\partial\phi(\underset{\sim}{s},\underset{\sim}{r})}{\partial\underset{\sim}{r}}$ (3.16b)

so that $(3.11)_1$ may also be written as

$$\frac{\partial\phi}{\partial\underset{\sim}{p}}\,(\underset{\sim}{\sigma},\underset{\sim}{p}) = -\,h\,\frac{\partial\phi(\underset{\sim}{\sigma},\underset{\sim}{p})}{\partial\underset{\sim}{\sigma}} \quad . \tag{3.17}$$

2.4 CONSTITUTIVE EQUATIONS BASED ON EXTREMAL PATHS

In Chapter 1 we discussed the motivation for a holonomic theory which makes use of extremal paths in stress and strain space, and pointed out that the solution obtained from such a theory provides a consistent bound on the exact rate solution. PONTER and MARTIN (1972) have derived constitutive equations for the consistent holonomic problem where the body is assumed to be in its virgin state at the start of the application of the loading. We propose here to extend these constitutive equations to the case where initial stress and strain fields are present in the body at the start of a finite increment of loading. Thus, we will generalise Ponter and Martin's equations to a form which is suitable for use in a consistent incremental holonomic problem.

As we have already mentioned, the concept of extremal paths was introduced by MARTIN (1966a,b) for elastic-perfectly plastic materials. The idea was subsequently extended to isotropically hardening materials by PONTER (1968) and by MARTIN (1970), and to kinematically hardening materials by SOECHTING and LANCE (1969). In the developments we describe here we draw upon and extend the more recent treatment of extremal paths in both stress and strain space given by MARTIN (1975a).

Let us assume that at $t = t^\circ$ there exists at an arbitrary point in the body an initial stress $\underset{\sim}{\sigma}^\circ$ and an initial strain $\underset{\sim}{\varepsilon}^\circ$, and that at $t = t^* > t^\circ$ the terminal stress and strain are $\underset{\sim}{\sigma}^*$ and $\underset{\sim}{\varepsilon}^*$ respectively. We assume that the terminal states of stress and strain are reached by following an extremal path (which will be defined below) from the initial states, unless we specifically state otherwise.

We use the symbol Δ to denote a change or finite increment in a given quantity along a specified path in either stress or strain space. The change in stress along any path between the two states $\underset{\sim}{\sigma}^\circ$ and $\underset{\sim}{\sigma}^*$ is written as

$$\Delta\underset{\sim}{\sigma}^* = \underset{\sim}{\sigma}^* - \underset{\sim}{\sigma}^\circ \quad . \tag{4.1}$$

The superscript attached to the change (for example, the * in $\Delta\underset{\sim}{\sigma}^*$ above) identifies the terminal state with which the change is associated. Similarly, the change in strain along any path between the states $\underset{\sim}{\varepsilon}^\circ$ and $\underset{\sim}{\varepsilon}^*$ is written as

$$\Delta\underset{\sim}{\varepsilon}^* = \underset{\sim}{\varepsilon}^* - \underset{\sim}{\varepsilon}^\circ \quad . \tag{4.2}$$

It is of course clear that since $\underset{\sim}{\sigma}^\circ$ and $\underset{\sim}{\varepsilon}^\circ$ will always be given, the changes $\Delta\underset{\sim}{\sigma}$ and $\Delta\underset{\sim}{\varepsilon}$ are functions of only one independent variable each, namely, the terminal states. We will therefore treat the changes $\Delta\underset{\sim}{\sigma}$ and $\Delta\underset{\sim}{\varepsilon}$ as independent variables in their own right, it being thereby understood that it is the terminal states which are of real importance.

Throughout the development of this theory we will work alternately in stress and strain space. Suppose that a material point whose initial stress state is $\underset{\sim}{\sigma}^\circ$ is subjected to a history of stress which terminates in a stress state $\underset{\sim}{\sigma}^*$; the complementary work along any path between these two stress states in stress space is defined by

$$\bar{\Omega}(\Delta\underset{\sim}{\sigma}^*) = \int_{\underset{\sim}{\sigma}^o}^{\underset{\sim}{\sigma}^o + \Delta\underset{\sim}{\sigma}^*} \underset{\sim}{\varepsilon}(\underset{\sim}{\sigma}) \cdot d\underset{\sim}{\sigma}$$

or

$$\bar{\Omega}(t^*) = \int_{t^o}^{t^*} \underset{\sim}{\varepsilon}(t) \cdot \dot{\underset{\sim}{\sigma}}(t)dt \quad . \tag{4.3}$$

Similarly, for an element of material whose initial state of strain is $\underset{\sim}{\varepsilon}^\circ$ and which is subjected to a history of strain which terminates in a strain state $\underset{\sim}{\varepsilon}^*$, the work along any such path in strain space is defined by

$$\bar{W}(\Delta\underset{\sim}{\varepsilon}^*) = \int_{\underset{\sim}{\varepsilon}^o}^{\underset{\sim}{\varepsilon}^o + \Delta\underset{\sim}{\varepsilon}^*} \underset{\sim}{\sigma}(\underset{\sim}{\varepsilon}) \cdot d\underset{\sim}{\varepsilon}$$

or

$$\bar{W}(t^*) = \int_{t^o}^{t^*} \underset{\sim}{\sigma}(t) \cdot \dot{\underset{\sim}{\varepsilon}}(t)\, dt \quad . \tag{4.4}$$

<u>Definition 4.1</u> An extremal path in stress space is a particular stress path such that the complementary work $\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*)$ along this path is not less than the complementary work between the initial and terminal stress states $\underset{\sim}{\sigma}^\circ$ and $\underset{\sim}{\sigma}^*$ along any other path; thus

$$\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*) \geq \bar{\Omega}(\Delta\underset{\sim}{\sigma}^*) \quad . \tag{4.5}$$

<u>Definition 4.2</u> An extremal path in strain space is a particular strain path such that the work $\hat{W}(\Delta\underset{\sim}{\varepsilon}^*)$ along this path is not greater than the work between the initial and terminal strain states $\underset{\sim}{\varepsilon}^\circ$ and $\underset{\sim}{\varepsilon}^*$ along any other path; thus

$$\hat{W}(\Delta\underset{\sim}{\varepsilon}^*) \leq \bar{W}(\Delta\underset{\sim}{\varepsilon}^*) \quad . \tag{4.6}$$

The paths described in the above definitions are illustrated in Fig. 2.4; from the definitions it is clear that the extremal paths depend only on the initial and terminal states of stress or strain. These definitions do not imply any specific connection between extremal paths in stress and strain space, although we shall show later that such a connection does indeed exist.

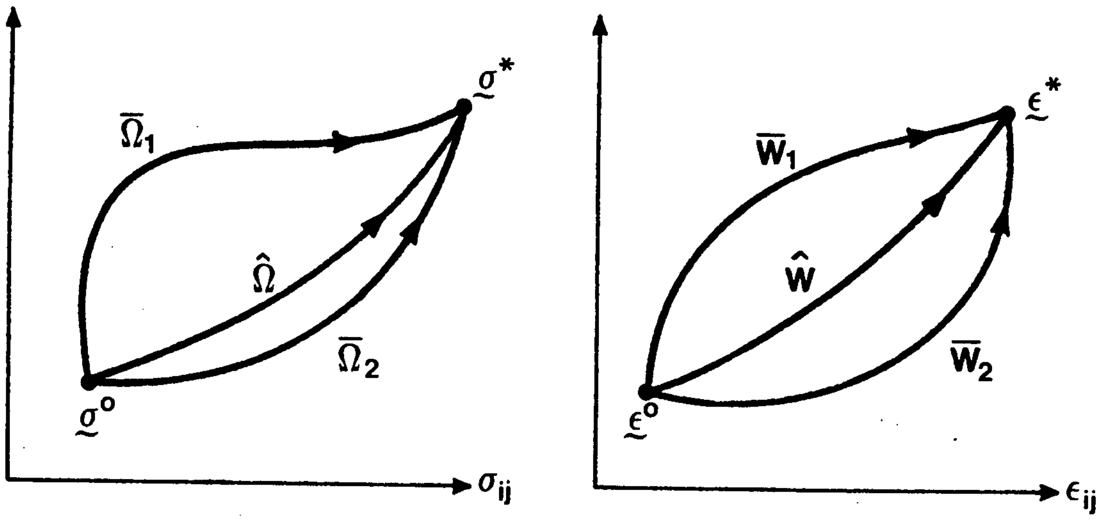


Figure 2.4 Paths in stress and strain space.

We now present a result which plays a decisive role in establishing the duality of $\hat{\Omega}$ and $\hat{W}$. The proof (and indeed the result itself) is similar to that already given by MARTIN (1975a), page 724, to which the reader is referred.

Proposition 4.1 (The Complementary Work Inequality)

Consider a material point which has an initial stress state $\underset{\sim}{\sigma}^o$. Suppose that the stress state is changed along an extremal path to $\underset{\sim}{\sigma}^*$. Alternatively, let the stress state be changed along some unspecified path to a stress state $\underset{\sim}{\sigma}^a \neq \underset{\sim}{\sigma}^*$ (Fig. 2.5). Then

$$\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*) - \underset{\sim}{\sigma}^* \cdot \underset{\sim}{\varepsilon}^a > \bar{\Omega}(\Delta\underset{\sim}{\sigma}^a) - \underset{\sim}{\sigma}^a \cdot \underset{\sim}{\varepsilon}^a \; ; \qquad (4.7)$$

this inequality may be expressed in an alternative form as

$$\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*) + \bar{W}(\Delta\underset{\sim}{\varepsilon}^a) \geq \underset{\sim}{\sigma}^* \cdot \underset{\sim}{\varepsilon}^a \quad . \quad \Box \qquad (4.8)$$

The proof of the above proposition relies on the introduction of a straight line path between $\underset{\sim}{\sigma}^a$ and $\underset{\sim}{\sigma}^*$; as a consequence the first postulate in the form of (2.16b) may be invoked to obtain the desired result.

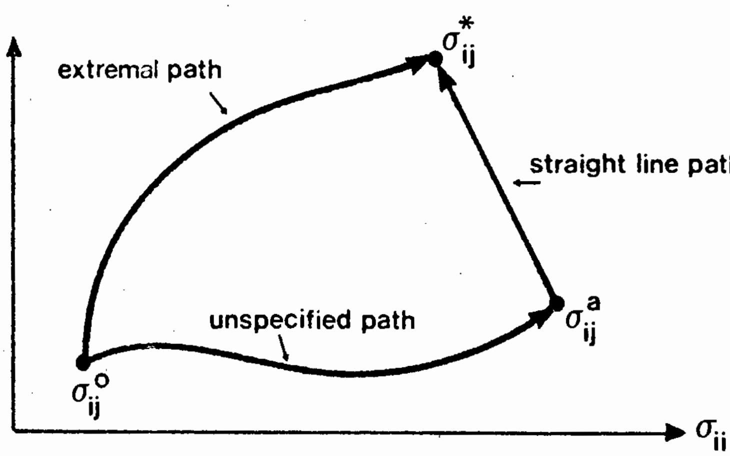


Figure 2.5 Alternative stress paths.

An important class of extremal paths arises when the behaviour of an element of material is elastic, that is, when the stress path lies entirely within the current yield surface, and the change in plastic strain associated with this stress path is zero. It is easily shown that for such paths the complementary energy $\Omega(\Delta\underset{\sim}{\sigma})$ (see eqn (4.21)) and the maximum complementary work $\hat{\Omega}(\Delta\underset{\sim}{\sigma})$ are identical and that any elastic path is consequently an extremal path. We summarise this result in the following proposition.

Proposition 4.2

If an elastic path can be found between the stress states $\underset{\sim}{\sigma}^{\circ}$ and $\underset{\sim}{\sigma}^{*}$ it is an extremal path, and

$$\hat{\Omega}(\Delta\underset{\sim}{\sigma}^{*}) = \Omega(\Delta\underset{\sim}{\sigma}^{*}) \quad . \quad \square \tag{4.9}$$

This result is of particular significance for elastic-perfectly plastic materials when the limit function ϕ coincides with the virgin yield function $\bar{\Phi}$, since $\hat{\Omega}(\Delta\underset{\sim}{\sigma}^{*})$ is then given by the elastic complementary energy $\Omega(\Delta\underset{\sim}{\sigma}^{*})$. It is of course implicitly assumed throughout that any extremal path must be realisable; it must lie either within or on the yield surface at all times.

We turn our attention now to the properties of extremal paths. Let us assume that for every terminal stress state $\underset{\sim}{\sigma}^{*}$ an extremal stress path exists in stress space. We may then associate with each $\underset{\sim}{\sigma}^{*}$ a state of total strain $\underset{\sim}{\varepsilon}^{*}$ which is the terminal strain when the stress path follows an extremal path from the initial stress state $\underset{\sim}{\sigma}^{\circ}$ to $\underset{\sim}{\sigma}^{*}$. Let $\bar{\Omega}(\Delta\underset{\sim}{\sigma}^{*})$, as defined in (4.3), denote the complementary work along some path between $\underset{\sim}{\sigma}^{\circ}$ and $\underset{\sim}{\sigma}^{*}$, and let $\hat{\Omega}(\Delta\underset{\sim}{\sigma}^{*})$ denote the maximum complementary work along an extremal path between $\underset{\sim}{\sigma}^{\circ}$ and $\underset{\sim}{\sigma}^{*}$. Then we have the following proposition.

Proposition 4.3

A sufficient condition that

$$\underset{\sim}{\varepsilon}^* = \left. \frac{\partial \bar{\Omega}(\Delta \underset{\sim}{\sigma}^*)}{\partial \Delta \underset{\sim}{\sigma}} \right|_{\Delta \underset{\sim}{\sigma}^*} \tag{4.10}$$

for a class of stress paths between $\underset{\sim}{\sigma}^\circ$ and $\underset{\sim}{\sigma}^*$, and defined in terms of $\Delta \underset{\sim}{\sigma}^*$, is that the path makes $\bar{\Omega}$ an extremum.

Proof

Consider two adjacent states of stress $\underset{\sim}{\sigma}^*$ and $\underset{\sim}{\sigma}$ which we assume may be reached by following extremal paths AB and AC from the initial stress state $\underset{\sim}{\sigma}^\circ$ (Fig. 2.6). Let $\hat{\Omega}(\Delta \underset{\sim}{\sigma}^*)$ and $\hat{\Omega}(\Delta \underset{\sim}{\sigma})$ be the maximum complementary work associated with these two extremal paths. Then, from the complementary work inequality (Proposition 4.1) we have

$$\hat{\Omega}(\Delta \underset{\sim}{\sigma}^*) - \hat{\Omega}(\Delta \underset{\sim}{\sigma}) - \underset{\sim}{\varepsilon}(\underset{\sim}{\sigma}) \cdot (\underset{\sim}{\sigma}^* - \underset{\sim}{\sigma}) \geq 0 \tag{4.11a}$$

where $\underset{\sim}{\varepsilon}(\underset{\sim}{\sigma})$ is the terminal strain associated with the extremal stress path between $\underset{\sim}{\sigma}^\circ$ and $\underset{\sim}{\sigma}$. Now set $\underset{\sim}{\sigma} = \theta \underset{\sim}{\sigma}^* + (1-\theta)\underset{\sim}{\sigma}$, $0 \leq \theta \leq 1$, in (4.11a) to obtain

$$\hat{\Omega}(\Delta \underset{\sim}{\sigma}^*) - \hat{\Omega}(\theta \Delta \underset{\sim}{\sigma}^* + (1-\theta)\Delta \underset{\sim}{\sigma}) - \underset{\sim}{\varepsilon}(\theta \underset{\sim}{\sigma}^* - (1-\theta)\underset{\sim}{\sigma}) \cdot (\underset{\sim}{\sigma}^* - \theta \underset{\sim}{\sigma}^* - (1-\theta)\underset{\sim}{\sigma}) \geq 0 \quad . \tag{4.11b}$$

Similarly, set $\underset{\sim}{\sigma}^* = \underset{\sim}{\sigma}$ and $\underset{\sim}{\sigma} = \theta \underset{\sim}{\sigma}^* + (1-\theta)\underset{\sim}{\sigma}$ in (4.11a) to obtain

$$\hat{\Omega}(\Delta \underset{\sim}{\sigma}) - \hat{\Omega}(\theta \Delta \underset{\sim}{\sigma}^* + (1-\theta)\Delta \underset{\sim}{\sigma}) - \underset{\sim}{\varepsilon}(\theta \underset{\sim}{\sigma}^* - (1-\theta)\underset{\sim}{\sigma}) \cdot (\underset{\sim}{\sigma} - \theta \underset{\sim}{\sigma}^* - (1-\theta)\underset{\sim}{\sigma}) \geq 0 \quad . \tag{4.11c}$$

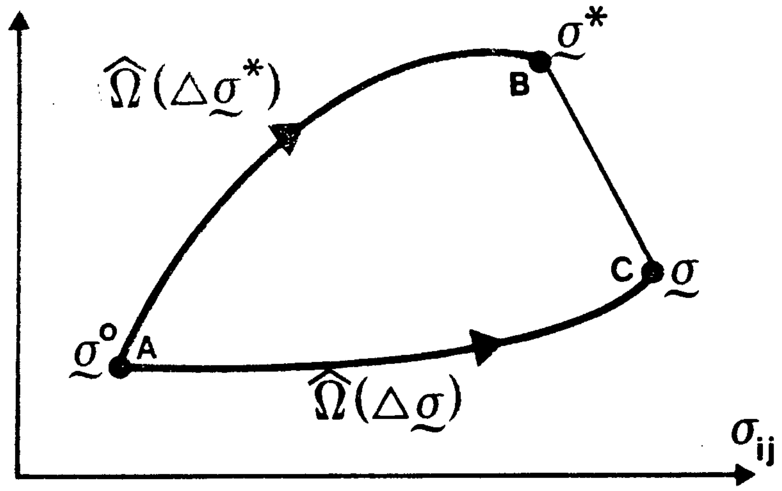


Figure 2.6 Extremal stress paths from $\tilde{\sigma}^o$ to $\tilde{\sigma}^*$ and $\tilde{\sigma}$.

Multiplying (4.11b) by θ and (4.11c) by $(1-\theta)$ and adding the results, we obtain

$$\theta\hat{\Omega}(\Delta\tilde{\sigma}^*) + (1-\theta)\hat{\Omega}(\Delta\tilde{\sigma}) \geq \hat{\Omega}(\theta\Delta\tilde{\sigma}^* + (1-\theta)\Delta\tilde{\sigma}) \quad . \tag{4.12}$$

Eqn (4.12) expresses the convexity of $\hat{\Omega}$. Thus, from eqn (4.11a) and making use of a standard result from convex analysis (see Section 2.5, Definition 5.1), we have

$$\tilde{\varepsilon}(\tilde{\sigma}) \in \partial\hat{\Omega}(\Delta\tilde{\sigma})$$

where $\partial\hat{\Omega}(\Delta\tilde{\sigma})$ is the subdifferential of $\hat{\Omega}$ at $\Delta\tilde{\sigma}$. Further, if $\hat{\Omega}$ is differentiable at $\Delta\tilde{\sigma}$ then

$$\partial\hat{\Omega}(\Delta\tilde{\sigma}) = \left\{ \left. \frac{\partial\hat{\Omega}(\Delta\tilde{\sigma})}{\partial\Delta\tilde{\sigma}} \right|_{\Delta\tilde{\sigma}} \right\}$$

from which it follows that

$$\tilde{\varepsilon}(\tilde{\sigma}) = \left. \frac{\partial\hat{\Omega}(\Delta\tilde{\sigma})}{\partial\Delta\tilde{\sigma}} \right|_{\Delta\tilde{\sigma}} \quad .$$

Similarly,

$$\tilde{\varepsilon}(\tilde{\sigma}^*) \equiv \tilde{\varepsilon}^* = \left. \frac{\partial\hat{\Omega}(\Delta\tilde{\sigma}^*)}{\partial\Delta\tilde{\sigma}} \right|_{\Delta\tilde{\sigma}^*} \quad . \tag{4.13}$$

This result implies that $\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*)$ is a potential function from which the total strain $\underset{\sim}{\varepsilon}^*$ resulting from an extremal stress history may be derived. The implication is that we may associate with a path-dependent elastic-plastic material a new material which is path-independent (or elastic) and whose constitutive equation is given by (4.13).

A dual result may be obtained for paths in strain space between an initial state of strain $\underset{\sim}{\varepsilon}^\circ$ and a terminal state of strain $\underset{\sim}{\varepsilon}^*$. Let $\underset{\sim}{\sigma}^*$ be the stress state associated with $\underset{\sim}{\varepsilon}^*$ when the strain path follows an extremal path in strain space. The work along some path in strain space between $\underset{\sim}{\varepsilon}^\circ$ and $\underset{\sim}{\varepsilon}^*$ is denoted by $\bar{W}(\Delta\underset{\sim}{\varepsilon}^*)$, as defined in (4.4), and the minimum work along an extremal path between $\underset{\sim}{\varepsilon}^\circ$ and $\underset{\sim}{\varepsilon}^*$ is denoted by $\hat{W}(\Delta\underset{\sim}{\varepsilon}^*)$. We then have the following proposition, the proof of which follows identical arguments to those of Proposition 4.3.

<u>Proposition 4.4</u>

A sufficient condition that

$$\underset{\sim}{\sigma}^* = \frac{\partial\bar{W}(\Delta\underset{\sim}{\varepsilon}^*)}{\partial\Delta\underset{\sim}{\varepsilon}}\Bigg|_{\Delta\underset{\sim}{\varepsilon}^*} \qquad (4.14)$$

for the class of paths between $\underset{\sim}{\varepsilon}^\circ$ and $\underset{\sim}{\varepsilon}^*$, defined in terms of $\Delta\underset{\sim}{\varepsilon}^*$, is that the path makes $\bar{W}$ an extremum.

It follows from this result that

$$\underset{\sim}{\sigma}^* = \left.\frac{\partial \hat{W}(\Delta\underset{\sim}{\varepsilon}^*)}{\partial \Delta \underset{\sim}{\varepsilon}}\right|_{\Delta\underset{\sim}{\varepsilon}^*} \tag{4.15}$$

and that $\hat{W}$ is thus a potential function for the stress $\underset{\sim}{\sigma}^*$ at the end of an extremal path in strain space. Furthermore, recalling (4.2), $\hat{W}(\Delta\underset{\sim}{\varepsilon}^*)$ depends only on the initial and terminal states of strain on the extremal path and is thus path-independent. In (4.15) we again have a constitutive equation which defines a path-independent (or elastic) material.

We propose now to present a result which establishes a relationship between maximum complementary work paths in stress space and minimum work paths in strain space. The proof of this result has been given by MARTIN (1975a), Section 22.3, for the holonomic problem. Since its extension to the case under discussion is trivial we will briefly sketch the salient ideas.

Using the complementary work inequality (Proposition 4.1) it is easy to show that $\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*)$ and $\hat{W}(\Delta\underset{\sim}{\varepsilon}^*)$ are convex functions (we have already shown this for $\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*)$ in Proposition 4.3). It can then be shown that the maximum complementary work path between the initial state $\underset{\sim}{\sigma}^{\circ}$ and terminal state $\underset{\sim}{\sigma}^*$ maps into strain space the minimum work path between the initial state $\underset{\sim}{\varepsilon}^{\circ}$ and terminal state $\underset{\sim}{\varepsilon}^*$, where $\underset{\sim}{\varepsilon}^*$ and $\Delta\underset{\sim}{\sigma}^*$ are related by (4.13). The converse can also be argued to be true except that the maximum complementary work paths which are mapped by minimum work paths need not necessarily be unique. We can show, further, that the terminal stress obtained from (4.15) is uniquely determined, but that the terminal strain obtained from (4.13) is not necessarily uniquely determined; a situation where the latter case is obvious is when the material is elastic-perfectly plastic. We summarise these observations in the following proposition.

Proposition 4.5 The duality of $\hat{\Omega}$ and $\hat{W}$.

The maximum complementary work function $\hat{\Omega}(\Delta\underset{\sim}{\sigma}^*)$ and the minimum work function $\hat{W}(\Delta\underset{\sim}{\varepsilon}^*)$, as defined in (4.5) and (4.6) respectively, are dual potential functions in the sense that (4.13) and (4.15) are inverses with respect to each other. Furthermore, both $\hat{\Omega}$ and $\hat{W}$ are convex functions.

It is apparent that Propositions 4.3 and 4.4 provide simple yet powerful tools for determining the terminal strain $\underset{\sim}{\varepsilon}^*$ associated with an extremal stress path in stress space, and the terminal stress $\underset{\sim}{\sigma}^*$ associated with an extremal strain path in strain space. Furthermore, from the duality result above, these two results provide two forms of the constitutive equation for the same path-independent material. All that remains is to determine explicit expressions for $\hat{\Omega}$ and $\hat{W}$ such that these quantities represent maximum and minimum values of $\bar{\Omega}$ and $\bar{W}$ respectively. Before doing so, however, we shall investigate a consequence of such a procedure which allows us to frame an alternative but equivalent set of constitutive equations which involve work quantities only.

Let us first discuss the special case in which the extremal paths are elastic paths. As before we assume that the total strain is divisible into an elastic part $\underset{\sim}{e}$ and a plastic part $\underset{\sim}{p}$ such that

$$\underset{\sim}{\varepsilon} = \underset{\sim}{e} + \underset{\sim}{p} \quad ; \tag{4.16}$$

recall from eqns (2.2) that the elastic part $\underset{\sim}{e}$ is given by

$$\underset{\sim}{e} = \underset{\approx}{D}\underset{\sim}{\sigma} \tag{4.17a}$$

and we assume also that the inverse relationship

$$\underset{\sim}{\sigma} = \underset{\approx}{C}\underset{\sim}{\varepsilon} \tag{4.17b}$$

exists as well. We assume that the initial stress state $\underset{\sim}{\sigma}^{\circ}$ satisfies the inequality

$$\phi(\underset{\sim}{\sigma}^{\circ},\underset{\sim}{p}^{\circ}) \leq 0 \quad ; \tag{4.18}$$

then the states of stress which can be reached by elastic paths are those states of stress which satisfy the inequality

$$\phi(\underset{\sim}{\sigma},\underset{\sim}{p}^{\circ}) \leq 0 \quad . \tag{4.19}$$

From Proposition 4.2 such a path is an extremal path, which, according to Proposition 4.5, maps into strain space a minimum work path which begins at a strain state $\underset{\sim}{\varepsilon}^{\circ}$ (which may include initial plastic strains $\underset{\sim}{p}^{\circ}$) and terminates at the state $\underset{\sim}{\varepsilon}$. Since the path is elastic the changes in stress and strain along these extremal paths must be related by the elastic equations (4.17), and the change in plastic strain along the strain path must be zero. The minimum work along an extremal path in strain space for which the change in plastic strain $\Delta\underset{\sim}{p} = \underset{\sim}{0}$ is then given by

$$\hat{W}(\Delta\underset{\sim}{\varepsilon}) = W(\Delta\underset{\sim}{\varepsilon}) \equiv \frac{1}{2} (\Delta\underset{\sim}{\varepsilon})^T \underset{\sim}{C}\, \Delta\underset{\sim}{\varepsilon} + \underset{\sim}{\sigma}^{\circ} \cdot \Delta\underset{\sim}{\varepsilon} \tag{4.20}$$

where $\Delta\underset{\sim}{\varepsilon} = \underset{\sim}{\varepsilon} - \underset{\sim}{\varepsilon}^{\circ}$ and $W(\Delta\underset{\sim}{\varepsilon})$ denotes the elastic strain energy associated with $\Delta\underset{\sim}{\varepsilon}$. Similarly for an extremal path in stress space between the initial stress state $\underset{\sim}{\sigma}^{\circ}$ and terminal stress state $\underset{\sim}{\sigma}$, which satisfy (4.18) and (4.19) respectively, the maximum complementary work is given by

$$\hat{\Omega}(\Delta\underset{\sim}{\sigma}) = \Omega(\Delta\underset{\sim}{\sigma}) \equiv \frac{1}{2} (\Delta\underset{\sim}{\sigma})^T \underset{\sim}{D}\, \Delta\underset{\sim}{\sigma} + \underset{\sim}{\varepsilon}^{\circ} \cdot \Delta\underset{\sim}{\sigma} \tag{4.21}$$

where $\Delta\underset{\sim}{\sigma} = \underset{\sim}{\sigma} - \underset{\sim}{\sigma}^{\circ}$ and $\Omega(\Delta\underset{\sim}{\sigma})$ is the elastic complementary energy associated with $\Delta\underset{\sim}{\sigma}$. It is clear that both $W(\Delta\underset{\sim}{\varepsilon})$ and $\Omega(\Delta\underset{\sim}{\sigma})$ are path-independent since they depend only on initial and terminal states of strain and stress respectively.

Suppose now that, following the division of the total strain into elastic and plastic parts, we divide the complementary work along any path in stress space between initial state $\underset{\sim}{\sigma}^{\circ}$ and terminal state $\underset{\sim}{\sigma}^{*}$ into two parts and write

$$\bar{\Omega}(\Delta\underset{\sim}{\sigma}^{*}) = \int_{t^{o}}^{t^{*}} \underset{\sim}{e} \cdot \dot{\underset{\sim}{\sigma}}\, dt + \int_{t^{o}}^{t^{*}} \underset{\sim}{p} \cdot \dot{\underset{\sim}{\sigma}}\, dt$$

$$= \Omega(\Delta\underset{\sim}{\sigma}^{*}) + \bar{\Omega}^{p}(\Delta\underset{\sim}{\sigma}^{*}) \quad . \tag{4.22}$$

The elastic part Ω is already determined from (4.21), so that in order to determine the maximum value of $\bar{\Omega}$ we need to find the maximum value of $\bar{\Omega}^p$. In doing so we may ignore the elastic behaviour, so that the determination of the extremal value of $\bar{\Omega}^p$ is equivalent to finding the maximum complementary work path for a rigid-plastic material.

Our development so far has placed no limitation on the type of material which may be considered, so that we may conclude that the duality result (Proposition 4.5) is valid also for rigid-plastic materials. Thus, paths in stress space which maximise $\bar{\Omega}^p$ will map into plastic strain space paths which minimise $\bar{W}^p$, where $\bar{W}^p$ is the plastic part of the total work $\bar{W}$ along any path in strain space between initial state $\underset{\sim}{\varepsilon}^\circ = \underset{\sim}{e}^\circ + \underset{\sim}{p}^\circ$ and terminal state $\underset{\sim}{\varepsilon}^* = \underset{\sim}{e}^* + \underset{\sim}{p}^*$, with $\bar{W}$ defined by

$$\bar{W}(\Delta\underset{\sim}{\varepsilon}^*) = W(\Delta\underset{\sim}{e}^*) + \bar{W}^p(\Delta\underset{\sim}{p}^*)$$

$$= \int_{t^o}^{t^*} \underset{\sim}{\sigma} \cdot \dot{\underset{\sim}{e}}\, dt + \int_{t^o}^{t^*} \underset{\sim}{\sigma} \cdot \dot{\underset{\sim}{p}}\, dt \quad . \qquad (4.23)$$

From (4.20) the minimum value of the elastic part W is given by $W(\Delta\underset{\sim}{e}^*) = \frac{1}{2} (\Delta\underset{\sim}{e}^*)^T\underset{\sim}{C}\, \Delta\underset{\sim}{e}^* + \underset{\sim}{\sigma}^\circ \cdot \Delta\underset{\sim}{e}^*$, so that in order to minimise $\bar{W}$ we need to determine the minimum value of $\bar{W}^p$.

Let us assume for the present that we can determine this minimum value and call it $\hat{W}^p(\Delta\underset{\sim}{p}^*)$. Now it is evident from the dual nature of the extremal paths in stress and plastic strain space that the terminal state $\underset{\sim}{\varepsilon}^*$ cannot be arbitrarily chosen, but must be considered, for the purposes of minimising $\bar{W}^p$, as a given state. The elastic and plastic parts of $\underset{\sim}{\varepsilon}^*$, on the other hand, can be considered to be as yet undetermined. Since the initial state $\underset{\sim}{\varepsilon}^{\circ}$ is also by definition a given state this implies that the change in strain $\Delta\underset{\sim}{\varepsilon}^{*}$ along the path must remain fixed, whilst the changes in the elastic and plastic parts may vary.

Let us now assume that we have expressions for $W(\Delta\underset{\sim}{e}^{*})$ and $\hat{W}^{P}(\Delta\underset{\sim}{p}^{*})$ so that the minimum work along the extremal path in strain space is

$$\hat{W}(\Delta\underset{\sim}{\varepsilon}^{*}) = W(\Delta\underset{\sim}{e}^{*}) + \hat{W}^{P}(\Delta\underset{\sim}{p}^{*}) \quad . \tag{4.24}$$

Let $\Delta\dot{\underset{\sim}{e}}^{*}$ and $\Delta\dot{\underset{\sim}{p}}^{*}$ be the rates of change of $\Delta\underset{\sim}{e}^{*}$ and $\Delta\underset{\sim}{p}^{*}$ along the extremal path. Then, by definition we have

$$\dot{\hat{W}} = \left.\frac{\partial\hat{W}}{\partial\Delta\underset{\sim}{\varepsilon}}\right|_{\Delta\underset{\sim}{\varepsilon}^{*}} \cdot (\Delta\dot{\underset{\sim}{e}}^{*} + \Delta\dot{\underset{\sim}{p}}^{*}) \tag{4.25a}$$

and from (4.24) we also have

$$\dot{\hat{W}} = \left(\left.\frac{\partial W}{\partial\Delta\underset{\sim}{e}}\right|_{\Delta\underset{\sim}{e}^{*}} \cdot \Delta\dot{\underset{\sim}{e}}^{*}\right) + \left(\left.\frac{\partial\hat{W}^{P}}{\partial\Delta\underset{\sim}{p}}\right|_{\Delta\underset{\sim}{p}^{*}} \cdot \Delta\dot{\underset{\sim}{p}}^{*}\right) \quad . \tag{4.25b}$$

Equating the coefficients of $\Delta\dot{\underset{\sim}{e}}^{*}$ and $\Delta\dot{\underset{\sim}{p}}^{*}$ in eqns (4.25) we conclude that

$$\left.\frac{\partial W}{\partial\Delta\underset{\sim}{e}}\right|_{\Delta\underset{\sim}{e}^{*}} = \left.\frac{\partial\hat{W}^{P}}{\partial\Delta\underset{\sim}{p}}\right|_{\Delta\underset{\sim}{p}^{*}} \quad . \tag{4.26}$$

From Propositions 4.2 and 4.5 the elastic strain energy $W(\Delta \underset{\sim}{e}^*)$ is always an extremum and thus we have from Proposition 4.4 that

$$\underset{\sim}{\sigma}^* = \frac{\partial W}{\partial \Delta \underset{\sim}{e}} \bigg|_{\Delta \underset{\sim}{e}^*} \tag{4.27}$$

and hence, from (4.26)

$$\underset{\sim}{\sigma}^* = \frac{\partial \hat{W}^p}{\partial \Delta \underset{\sim}{p}} \bigg|_{\Delta \underset{\sim}{p}^*} \; . \tag{4.28}$$

Equations (4.27) and (4.28) represent respectively the elastic and plastic constitutive equations for a nonlinear path-independent (or elastic) material. Such a material, as we have seen, can be associated with the classical path-dependent elastic-plastic material.

Since we already have an explicit expression for the elastic strain energy W we treat eqn (4.27) immediately. Replacing $\Delta \underset{\sim}{\varepsilon}$ by $\Delta \underset{\sim}{e}$ in (4.20), we have

$$W(\Delta \underset{\sim}{e}) = \tfrac{1}{2} \; (\Delta \underset{\sim}{e})^T \underset{\sim}{C} \; \Delta \underset{\sim}{e} + \underset{\sim}{\sigma}^\circ \cdot \Delta \underset{\sim}{e} \; .$$

Making use of this result in (4.27) we obtain

$$\underset{\sim}{\sigma}^* = \underset{\sim}{C} \Delta \underset{\sim}{e}^* + \underset{\sim}{\sigma}^\circ$$

$$= \underset{\sim}{C} \Delta \underset{\sim}{e}^* + \underset{\sim}{C} \underset{\sim}{e}^\circ \; , \quad \text{from (4.17b)}$$

$$= \underset{\sim}{C}[\underset{\sim}{e}^{\circ} + \Delta\underset{\sim}{e}^{*}]$$

$$= \underset{\sim}{C}[\underset{\sim}{\varepsilon}^{\circ} - \underset{\sim}{p}^{\circ} + \Delta\underset{\sim}{\varepsilon}^{*} - \Delta\underset{\sim}{p}^{*}] \,, \quad \text{using (4.16)} \,. \qquad (4.29)$$

This result establishes an explicit expression for the elastic constitutive equation for our nonlinear elastic material.

We defer further discussion of the plastic work contribution until the following section where we will develop an explicit expression for $\hat{W}^P$ for the particular case of the von Mises yield function with linear kinematic hardening.

2.5 THE DETERMINATION OF EXTREMAL PATHS FOR A VON MISES YIELD FUNCTION WITH LINEAR KINEMATIC HARDENING

Our first objective is to establish appropriate conditions which must be satisfied by an extremal path, based on the simultaneous maximisation of the plastic complementary work and minimisation of the plastic work. Let us suppose that at $t = t^{\circ}$ the state of a material point is defined by initial stresses $\underset{\sim}{\sigma}^{\circ}$ and initial strains $\underset{\sim}{\varepsilon}^{\circ} = \underset{\sim}{e}^{\circ} + \underset{\sim}{p}^{\circ}$. Recall that for kinematic hardening the internal variables are the plastic strains themselves so that the current yield surface at some $t > t^{\circ}$ is given by

$$\phi(\underset{\sim}{\sigma},\underset{\sim}{p}) = 0 \quad . \qquad (5.1)$$

We assume as before that ϕ is a convex, continuously differentiable function. Since we are interested only in plastic behaviour we may also assume without limiting the generality of the intended result that the initial stress state $\underset{\sim}{\sigma}^{\circ}$ lies on the yield surface at $t = t^{\circ}$, so that

$$\phi(\underset{\sim}{\sigma}^{\circ}, \underset{\sim}{p}^{\circ}) = 0 \tag{5.2}$$

and that the stress state remains on the yield surface at all times along the extremal path.

We adopt the classical rate form of the plastic constitutive equations, (eqns (2.22) and (2.23)), which we write here as

$$\begin{aligned} \dot{\underset{\sim}{p}} &= G \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \left(\frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}}\right), \quad \text{if } \phi(\underset{\sim}{\sigma}) = 0 \text{ and } \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}} > 0 \\ \dot{\underset{\sim}{p}} &= \underset{\sim}{0}, \quad \text{otherwise} \end{aligned} \tag{5.3}$$

where $G > 0$ is a hardening parameter, for which an explicit expression has already been given in (3.12).

Let the plastic work computed along some path in plastic strain space between the initial state $\underset{\sim}{p}^{\circ}$ and terminal state $\underset{\sim}{p}^{*}$ be given by

$$\bar{W}^{p}(\Delta \underset{\sim}{p}^{*}) = \int_{t^{o}}^{t^{*}} \underset{\sim}{\sigma} \cdot \dot{\underset{\sim}{p}} \, dt \,. \tag{5.4}$$

The first variation of $\bar{W}^{p}$, keeping the initial and terminal values of the plastic strain fixed, is

$$\delta \bar{W}^{p}(\Delta \underset{\sim}{p}^{*}) = \int_{t^{o}}^{t^{*}} (\dot{\underset{\sim}{p}} \cdot \delta \underset{\sim}{\sigma} + \underset{\sim}{\sigma} \cdot \delta \dot{\underset{\sim}{p}}) \, dt$$

which after integrating by parts becomes

$$\delta\bar{W}^p(\Delta\underset{\sim}{p}^*) = \int_{t^o}^{t^*} (\dot{\underset{\sim}{p}} \cdot \delta\underset{\sim}{\sigma} - \dot{\underset{\sim}{\sigma}} \cdot \delta\underset{\sim}{p})\, dt \tag{5.5}$$

in which the variations $\delta\underset{\sim}{\sigma}$ and $\delta\underset{\sim}{p}$ are related via (5.3).

Let the complementary plastic work along some path in stress space between the initial stress $\underset{\sim}{\sigma}^o$ and the terminal stress $\underset{\sim}{\sigma}^*$ be given by

$$\bar{\Omega}^p(\Delta\underset{\sim}{\sigma}^*) = \int_{t^o}^{t^*} \underset{\sim}{p} \cdot \dot{\underset{\sim}{\sigma}}\, dt \quad . \tag{5.6}$$

The first variation of $\bar{\Omega}^p$, keeping the initial and terminal values of the stress fixed, is

$$\delta\bar{\Omega}^p(\Delta\underset{\sim}{\sigma}^*) = \int_{t^o}^{t^*} (\underset{\sim}{p} \cdot \delta\dot{\underset{\sim}{\sigma}} + \dot{\underset{\sim}{\sigma}} \cdot \delta\underset{\sim}{p})\, dt$$

which after integrating by parts becomes

$$\delta\bar{\Omega}^p(\Delta\underset{\sim}{\sigma}^*) = - \int_{t^o}^{t^*} (\dot{\underset{\sim}{p}} \cdot \delta\underset{\sim}{\sigma} - \dot{\underset{\sim}{\sigma}} \cdot \delta\underset{\sim}{p})\, dt \quad . \tag{5.7}$$

In order to extremise $\bar{W}^p$ and $\bar{\Omega}^p$ we set $\delta\bar{W}^p = \delta\bar{\Omega}^p = 0$, whence we see that the resulting two problems become identical except for their respective terminal conditions; (note that since the initial conditions are assumed to be given, $\delta\underset{\sim}{p}^o = \delta\underset{\sim}{\sigma}^o = 0$ by definition). Thus, if we can solve the broader variational problem

$$\int_{t^o}^{t^*} (\dot{\underset{\sim}{p}} \cdot \delta\underset{\sim}{\sigma} - \dot{\underset{\sim}{\sigma}} \cdot \delta\underset{\sim}{p})\, dt = 0 \tag{5.8}$$

such that at $t = t^*$ no restrictions are placed on the values of the plastic strain nor on stress, we would then have coincident solutions for the original problems

$$\delta\bar{W}^p = \delta\bar{\Omega}^p = 0 \quad .$$

We assume that during the interval $[t^\circ, t^*]$ the stress point remains on the current yield surface $\phi(\underset{\sim}{\sigma}, \underset{\sim}{p}) = 0$ so that the first variation of ϕ must be zero :

$$\delta\phi = \frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \delta\underset{\sim}{\sigma} + \frac{\partial\phi}{\partial\underset{\sim}{p}} \cdot \delta\underset{\sim}{p} = 0 \quad . \tag{5.9}$$

Substituting for $\dot{\underset{\sim}{p}}$ in (5.8) from (5.3) we get

$$\int_{t^o}^{t^*} [G\,(\frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}})(\frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \delta\underset{\sim}{\sigma}) - (\dot{\underset{\sim}{\sigma}} \cdot \delta\underset{\sim}{p})]\, dt = 0 \quad ;$$

now substitute for $(\partial\phi/\partial\underset{\sim}{\sigma}) \cdot \delta\underset{\sim}{\sigma}$ from (5.9) to get

$$\int_{t^o}^{t^*} [-\,G\,\frac{\partial\phi}{\partial\underset{\sim}{p}}\,(\frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}}) - \dot{\underset{\sim}{\sigma}}] \cdot \delta\underset{\sim}{p}\, dt = 0 \quad . \tag{5.10}$$

From the fundamental lemma of the calculus of variations the extremal path must satisfy the Euler equation

$$\dot{\underset{\sim}{\sigma}} = -\,G\,\frac{\partial\phi}{\partial\underset{\sim}{p}}\,(\frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}}) \quad . \tag{5.11}$$

Equation (5.11) yields a set of extremal paths in stress space emanating from various stress states on the yield surface $\phi(\underset{\sim}{\sigma}^\circ, \underset{\sim}{p}^\circ) = 0$.

Recall from Section 2.3, eqn (3.17), that for linear kinematic hardening with a von Mises yield function

$$\frac{\partial \phi}{\partial \underset{\sim}{p}} = - h \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} , \qquad (5.12)$$

where h is a positive scalar hardening parameter. Substituting (5.12) into (5.11) yields

$$\dot{\underset{\sim}{\sigma}} = hG \frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \left(\frac{\partial \phi}{\partial \underset{\sim}{\sigma}} \cdot \dot{\underset{\sim}{\sigma}}\right)$$

$$= h\dot{\underset{\sim}{p}} \qquad (5.13)$$

where we have used (5.3) to obtain the final result. It is now clear that the extremal paths have the same direction as the plastic strain rate $\dot{\underset{\sim}{p}}$. Since $\dot{\underset{\sim}{p}}$ is always normal to the yield surface, and since we have chosen to work with yield surfaces whose normals are uniquely defined at every point, $\dot{\underset{\sim}{p}}$ also defines the point on the yield surface from which the extremal stress path emanates : we call this stress state $\hat{\underset{\sim}{\sigma}}$. An extremal path and the various associated stress states are illustrated in Fig. 2.7.

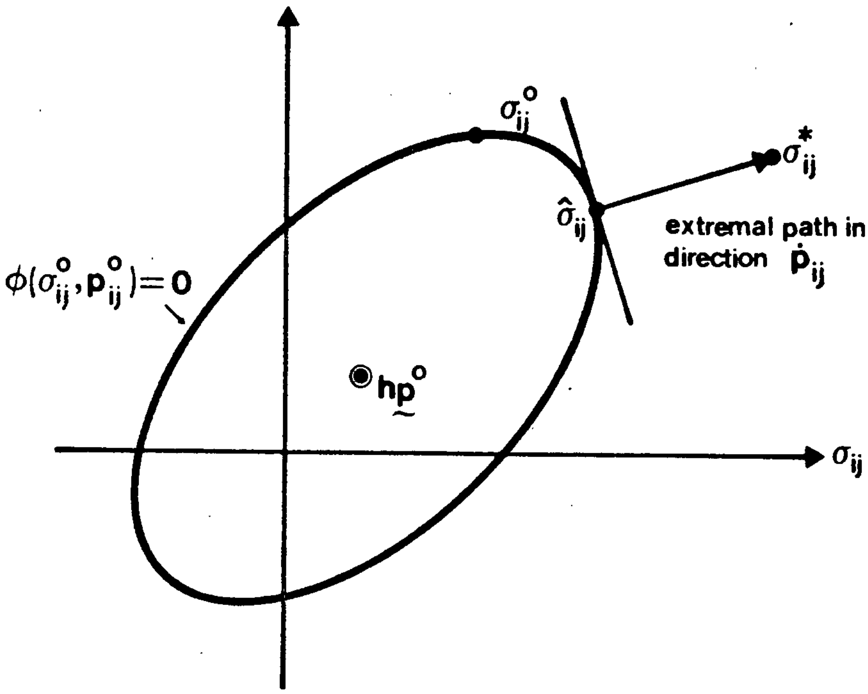


Figure 2.7 Extremal path in stress space.

We will show later that the stress state $\hat{\underset{\sim}{\sigma}}$ depends only on the direction of $\dot{\underset{\sim}{p}}$ and not on its magnitude. To emphasise this we will write $\hat{\underset{\sim}{\sigma}} = \hat{\underset{\sim}{\sigma}}(\Delta \underset{\sim}{p})$, where $\Delta \underset{\sim}{p}$ has the same direction as $\dot{\underset{\sim}{p}}$ but is of unrestricted magnitude. We shall also demand of $\hat{\underset{\sim}{\sigma}}$ that it lies on the yield surface defined by

$$\phi(\hat{\underset{\sim}{\sigma}}(\Delta \underset{\sim}{p}), \underset{\sim}{p}^{\circ}) = 0 \qquad (5.14)$$

assuming, of course, that $\Delta \underset{\sim}{p} \neq 0$.

We must emphasise that we are still working under the assumption that the material behaviour is rigid-plastic. This allows us to assume that $\underset{\sim}{\sigma}^{\circ}$ is on the yield surface, as shown in Fig. 2.7. However, it should be noted that $\underset{\sim}{\sigma}^{\circ}$ and $\hat{\underset{\sim}{\sigma}}$ need not be coincident stress states. Furthermore, when the elastic part of the behaviour is reintroduced later it will become apparent that $\underset{\sim}{\sigma}^{\circ}$ may be anywhere within or on the yield surface at $t = t^{\circ}$.

Returning to eqn (5.13) and integrating over the extremal path to the terminal state $\underset{\sim}{\sigma}^{*}$ we have

$$\int_{t^{o}}^{t^{*}} \dot{\underset{\sim}{p}}\, dt = \frac{1}{h} \int_{t^{o}}^{t^{*}} \dot{\underset{\sim}{\sigma}}\, dt$$

$$\Rightarrow \int_{\underset{\sim}{p}^{o}}^{\underset{\sim}{p}^{*}} d\underset{\sim}{p} = \frac{1}{h} \int_{\hat{\underset{\sim}{\sigma}}}^{\underset{\sim}{\sigma}^{*}} d\underset{\sim}{\sigma}$$

$$\Rightarrow \Delta\underset{\sim}{p}^{*} = \frac{1}{h} (\underset{\sim}{\sigma}^{*} - \hat{\underset{\sim}{\sigma}}) \qquad (5.15)$$

where $\Delta\underset{\sim}{p}^{*} = \underset{\sim}{p}^{*} - \underset{\sim}{p}^{\circ}$. Obviously the extremal paths in stress space are straight line paths, as illustrated in Fig. 2.7. For the case of the von Mises yield function defined in (3.7) the extremal paths in stress space will also be proportional loading paths.

Plastic Work and Complementary Plastic Work

Having explicitly defined the extremal paths we may proceed with the development of explicit expressions for the maximum complementary plastic work $\hat{\Omega}^{P}(\Delta\underset{\sim}{\sigma}^{*})$ and minimum plastic work $\hat{W}^{P}(\Delta\underset{\sim}{p}^{*})$ along these paths. For the plastic work we have

$$\hat{W}^P(\underset{\sim}{\Delta p}^*) = \int_{t^o}^{t^*} \underset{\sim}{\sigma} \cdot \dot{\underset{\sim}{p}}\, dt$$

$$= \frac{1}{h} \int_{t^o}^{t^*} \underset{\sim}{\sigma} \cdot \dot{\underset{\sim}{\sigma}}\, dt \;, \quad \text{using (5.13)}$$

$$= \frac{1}{2h} \int_{\hat{\underset{\sim}{\sigma}}}^{\underset{\sim}{\sigma}^*} d(\underset{\sim}{\sigma} \cdot \underset{\sim}{\sigma})$$

$$= \frac{1}{2h} [\underset{\sim}{\sigma} \cdot \underset{\sim}{\sigma}]_{\hat{\underset{\sim}{\sigma}}}^{\underset{\sim}{\sigma}^*}$$

$$= \frac{1}{2} h(\underset{\sim}{\Delta p}^* \cdot \underset{\sim}{\Delta p}^*) + (\hat{\underset{\sim}{\sigma}} \cdot \underset{\sim}{\Delta p}^*) \qquad (5.16)$$

where we have used (5.15) to eliminate $\underset{\sim}{\sigma}^*$ and so obtain the final result. Treating $\underset{\sim}{\Delta p}^*$ as a variable in (5.16) for the purpose of differentiating the plastic work, we have

$$\frac{\partial \hat{W}^P}{\partial \underset{\sim}{\Delta p}^*} = h\underset{\sim}{\Delta p}^* + \hat{\underset{\sim}{\sigma}} + (\underset{\sim}{\Delta p}^* \cdot \frac{\partial \hat{\underset{\sim}{\sigma}}}{\partial \underset{\sim}{\Delta p}^*}) \quad . \qquad (5.17)$$

The last term in this expression involves the scalar product of two orthogonal second-order tensors, since $\hat{\underset{\sim}{\sigma}}$ must always lie on the yield surface and $\underset{\sim}{\Delta p}^*$ must always be normal to the yield surface at $\hat{\underset{\sim}{\sigma}}$; this term then falls away. Substituting for $\underset{\sim}{\Delta p}^*$ from (5.15) leaves us with

$$\frac{\partial \hat{W}^P}{\partial \underset{\sim}{\Delta p}^*} = \underset{\sim}{\sigma}^* \qquad (5.18)$$

confirming that the terminal stress may be derived from the minimum plastic work function.

The maximum complementary plastic work may be obtained by arguments similar to those above and is given by

$$\hat{\Omega}^p(\Delta\underset{\sim}{\sigma}^*) = \frac{1}{2}\,\Delta\underset{\sim}{p}^* \cdot (\underset{\sim}{\sigma}^* - \hat{\underset{\sim}{\sigma}}) \quad . \tag{5.19}$$

Similarly, we may confirm that the terminal strain $\underset{\sim}{\varepsilon}^*$ may be obtained from this function.

We shall find it essential for the later development to separate the minimum plastic work function $\hat{W}^p$ into two parts. Let us then write

$$\hat{W}^p = \hat{W}^p_1 + \hat{W}^p_2 \tag{5.20}$$

where we define the new functions

$$\hat{W}^p_1(\Delta\underset{\sim}{p}^*) = \frac{1}{2}\,h(\Delta\underset{\sim}{p}^* \cdot \Delta\underset{\sim}{p}^*) \tag{5.21}$$

and $$\hat{W}^p_2(\Delta\underset{\sim}{p}^*) = \hat{\underset{\sim}{\sigma}} \cdot \Delta\underset{\sim}{p}^* \tag{5.22}$$

and we emphasise that $\hat{\underset{\sim}{\sigma}} = \hat{\underset{\sim}{\sigma}}(\Delta\underset{\sim}{p}^*)$.

The Plastic Work Function $\hat{W}^p_2$

We now wish to develop an explicit expression for the plastic work function $\hat{W}^p_2$. Recall from (3.7) that the yield function in deviatoric stress space at $t = t^\circ$ is given by

$$\phi = \frac{1}{2}(\underset{\sim}{s} - h\underset{\sim}{r}^\circ)(\underset{\sim}{s} - h\underset{\sim}{r}^\circ) - k^2 = 0 \quad . \tag{5.23}$$

We have already seen from (5.15) that the change in plastic strain $\Delta\underset{\sim}{p}$ follows a straight line path in the direction of the normal to the yield surface at the stress state $\hat{\underset{\sim}{\sigma}}$ in stress space. Transferring this to deviatoric stress space and noting from (3.5) that $\Delta\underset{\sim}{p} = \Delta\underset{\sim}{r}$ it follows that

$$\Delta\underset{\sim}{p} = \beta \left. \frac{\partial\phi}{\partial\underset{\sim}{s}} \right|_{\hat{\underset{\sim}{s}}}$$

$$= \beta\,(\hat{\underset{\sim}{s}} - h\underset{\sim}{r}^\circ) \quad , \quad \text{from (5.23)} \tag{5.24}$$

where β is a positive but otherwise undetermined constant. Since $\hat{\underset{\sim}{s}}$ lies on the yield surface at $t = t^\circ$ we may replace $\underset{\sim}{s}$ by $\hat{\underset{\sim}{s}}$ in (5.23); then combining (5.23) and (5.24) we obtain

$$\beta = \frac{1}{\sqrt{2}k}\sqrt{\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{p}} \quad . \tag{5.25}$$

Substituting (5.25) into (5.24) and rearranging, we get

$$\hat{\underset{\sim}{s}} = \frac{\sqrt{2}k\Delta\underset{\sim}{p}}{\sqrt{\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{p}}} + h\underset{\sim}{r}^{o} \quad . \tag{5.26}$$

Taking the scalar product of both sides with $\underset{\sim}{\Delta p}$ and noting that $\underset{\sim}{p}^{\circ} = \underset{\sim}{r}^{\circ}$, we obtain

$$\hat{\underset{\sim}{s}} \cdot \underset{\sim}{\Delta p} = \sqrt{2}k \sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}} + h\underset{\sim}{p}^{o} \cdot \underset{\sim}{\Delta p} \quad . \tag{5.27}$$

In order to write the left-hand-side of the above equation in terms of stress $\underset{\sim}{\sigma}$ we note that, using eqn (3.3)

$$\begin{aligned}
\hat{\underset{\sim}{s}} \cdot \underset{\sim}{\Delta p} &= \hat{s}_{ij}\Delta p_{ij} \\
&= \hat{\sigma}_{ij}\Delta p_{ij} - \frac{1}{3}\hat{\sigma}_{kk}\delta_{ij}\Delta p_{ij} \\
&= \hat{\sigma}_{ij}\Delta p_{ij} - \frac{1}{3}\hat{\sigma}_{kk}\Delta p_{kk} \\
&= \hat{\sigma}_{ij}\Delta p_{ij} \\
&= \hat{\underset{\sim}{\sigma}} \cdot \underset{\sim}{\Delta p}
\end{aligned} \tag{5.28}$$

since $\Delta p_{kk} = 0$ from (3.1). Using (5.27) and (5.28) in (5.22) we obtain the following expression for the plastic work function $\hat{W}_2^p$:

$$\hat{W}_2^p(\underset{\sim}{\Delta p}) = \sqrt{2}k \sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}} + h\underset{\sim}{p}^{o} \cdot \underset{\sim}{\Delta p} \quad . \tag{5.29}$$

It is apparent from the definition of $\hat{W}_2^p$ given in (5.22) that

$$\frac{\partial \hat{W}_2^p}{\partial \underset{\sim}{\Delta p}} = \hat{\underset{\sim}{\sigma}} \quad . \tag{5.30}$$

Differentiating (5.29) with respect to $\underset{\sim}{\Delta p} \neq \underset{\sim}{0}$ and using (5.30) we obtain

$$\hat{\underset{\sim}{\sigma}} = \frac{\sqrt{2}k\ \underset{\sim}{\Delta p}}{\sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}}} + h\underset{\sim}{p}^{o} \quad . \tag{5.31}$$

It is worth remarking that the components of $\hat{\underset{\sim}{\sigma}}$ are clearly homogeneous and of degree zero in the components of $\underset{\sim}{\Delta p}$, from which we may conclude that the stress state $\hat{\underset{\sim}{\sigma}}$ depends only on the direction of the plastic strain increment $\underset{\sim}{\Delta p}$ and not on its magnitude.

Before proceeding we comment briefly on the plastic work functions $\hat{W}^p_1$ and $\hat{W}^p_2$. The quantity $\hat{W}^p_1$, eqn (5.21), is clearly a quadratic function of $\underset{\sim}{\Delta p}$, but because of its dependence on the hardening parameter h , will only be present for materials which exhibit hardening behaviour. The quantity $\hat{W}^p_2$, eqn (5.29) , consists of two terms, the second of which also depends on the hardening parameter h and on the initial plastic strain state $\underset{\sim}{p}^{\circ}$, and may as a result not always be present. The first term on the right-hand-side of (5.29) will always be present, but suffers from the disadvantage of being <u>non-differentiable</u> when $\underset{\sim}{\Delta p} = 0$. We show in Fig. 2.8 a schematic representation of the function $\hat{W}^p_2$ and its derivative with respect to $\underset{\sim}{\Delta p}$. The function clearly has a discontinuous derivative at $\underset{\sim}{\Delta p} = 0$.

The Subdifferential $\partial\hat{W}^p_2$

The subdifferential $\partial F(u)$ of a function F at u is a well known concept in convex analysis, and one which we aim to make use of here in connection with the plastic work function $\hat{W}^p_2$. Before doing so, however, we must ensure that the plastic work $\hat{W}^p_2$ is a convex function.

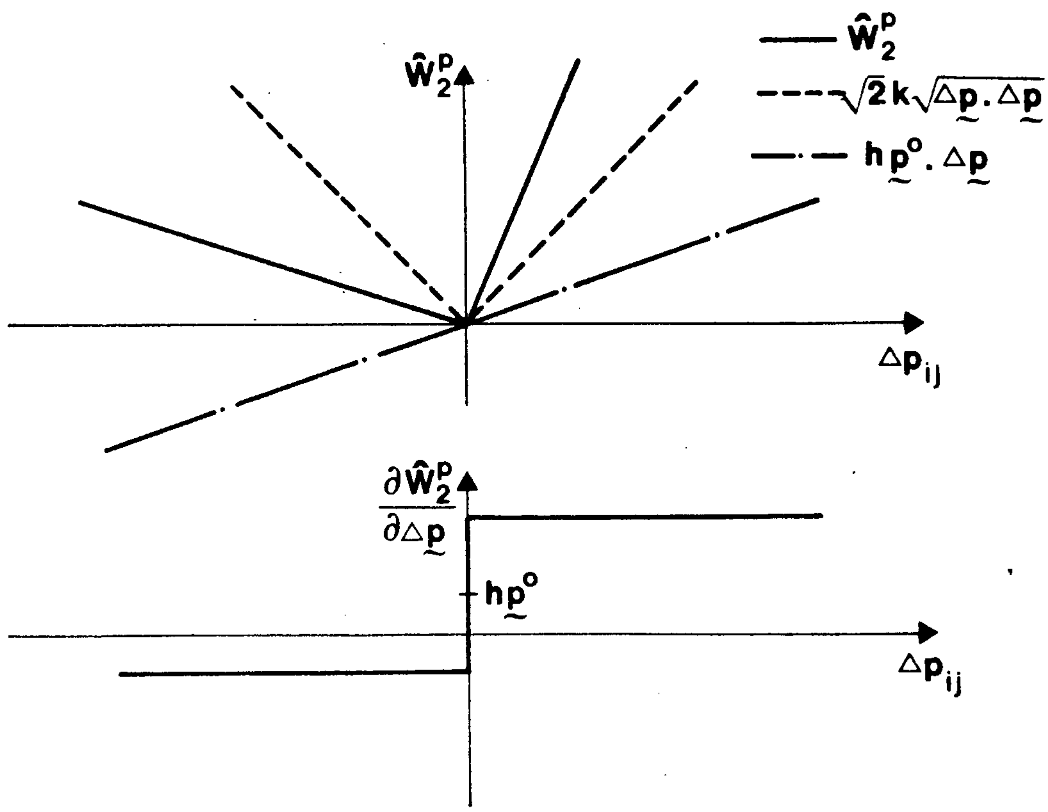


Figure 2.8 Plastic work function $\hat{W}^p_2$ and its derivative.

Lemma 2.1

The plastic work function $\hat{W}^p_2$ as defined in (5.29) is convex.

Proof

For the purposes of this proof let us write

$$\hat{W}^p_2 = W_a + W_b$$

$$= \sqrt{2}k \sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}} + \underset{\sim}{hp}^o \cdot \underset{\sim}{\Delta p} \quad , \quad \text{from (5.29)} .$$

We show first that W_a is convex, and then, using the linearity of W_b, that $\hat{W}^p_2$ must be convex.

(i) We have

$$[W_a(\underset{\sim}{\Delta p})]^2 = c^2 \underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} \text{ , where } c = \sqrt{2}k \ .$$

Then for $0 \leqslant \theta \leqslant 1$, and for all admissible $\underset{\sim}{\Delta p}, \underset{\sim}{\Delta q}$, we have

$$[W_a(\theta\underset{\sim}{\Delta p} + (1-\theta)\underset{\sim}{\Delta q})]^2 = c^2(\theta\underset{\sim}{\Delta p} + (1-\theta)\underset{\sim}{\Delta q})(\theta\underset{\sim}{\Delta p} + (1-\theta)\underset{\sim}{\Delta q})$$

$$= c^2\theta^2(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}) + 2c^2\theta(1-\theta)(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q}) + c^2(1-\theta)^2(\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}) \ . \tag{5.32}$$

For $\underset{\sim}{\Delta p}$, $\underset{\sim}{\Delta q} \in R^{NxN}$ the Schwarz inequality gives

$$c^2|\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q}| \leqslant c^2 \ |\underset{\sim}{\Delta p}||\underset{\sim}{\Delta q}|$$

$$= c\sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}} \ \ c\sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}$$

$$= W_a(\underset{\sim}{\Delta p}) \ \ W_a(\underset{\sim}{\Delta q}) \ . \tag{5.33}$$

Substituting (5.33) into (5.32) we get

$$[W_a(\theta\Delta \underset{\sim}{p} + (1-\theta)\Delta \underset{\sim}{q})]^2 < \theta^2 W_a^2(\Delta \underset{\sim}{p}) + 2\theta(1-\theta)W_a(\Delta \underset{\sim}{p})W_a(\Delta \underset{\sim}{q})$$

$$+ (1-\theta)^2 \, W_a^2(\Delta \underset{\sim}{q})$$

$$= [\theta W_a(\Delta \underset{\sim}{p}) + (1-\theta)W_a(\Delta \underset{\sim}{q})]^2 \quad .$$

Taking the square-root of both sides, we get

$$W_a(\theta\Delta \underset{\sim}{p} + (1-\theta)\Delta \underset{\sim}{q}) \leqslant \theta W_a(\Delta \underset{\sim}{p}) + (1-\theta)W_a(\Delta \underset{\sim}{q}) \qquad (5.34)$$

which establishes the convexity of W_a .

(ii) The linearity of $W_b(\Delta \underset{\sim}{p})$ is expressed as

$$W_b(\theta\Delta \underset{\sim}{p} + (1-\theta)\Delta \underset{\sim}{q}) = \theta W_b(\Delta \underset{\sim}{p}) + (1-\theta)W_b(\Delta \underset{\sim}{q}) \qquad (5.35)$$

Adding (5.34) and (5.35), and recalling that $\hat{W}_2^p = W_a + W_b$, we have

$$\hat{W}_2^p(\theta\Delta \underset{\sim}{p} + (1-\theta)\Delta \underset{\sim}{q}) \leqslant \theta\hat{W}_2^p(\Delta \underset{\sim}{p}) + (1-\theta)\hat{W}_2^p(\Delta \underset{\sim}{q})$$

and the result is established. □

We introduce the subdifferential via the following definition (see EKELAND and TEMAM (1976), Chapter 1, Section 5).

<u>Definition 5.1</u> Let $\hat{W}_2^p : R^{NxN} \to R$ be a proper function. The subdifferential of $\hat{W}_2^p$ at $\Delta \underset{\sim}{p}$, denoted $\partial \hat{W}_2^p(\Delta \underset{\sim}{p})$, is defined by

$$\partial \hat{W}_2^p(\Delta \underset{\sim}{p}) = \{\underset{\sim}{\tau} : \hat{W}_2^p(\Delta \underset{\sim}{q}) - \hat{W}_2^p(\Delta \underset{\sim}{p}) - \underset{\sim}{\tau} \cdot (\Delta \underset{\sim}{q} - \Delta \underset{\sim}{p}) \geq 0 \quad , \quad \Delta \underset{\sim}{q} \in R^{NxN}\} \tag{5.36}$$

where $\underset{\sim}{\tau}$ is called a subgradient at $\Delta \underset{\sim}{p}$.

Further, if $\hat{W}_2^p$ is convex and differentiable at $\Delta \underset{\sim}{p}$, then it is subdifferentiable at $\Delta \underset{\sim}{p}$ and

$$\partial \hat{W}_2^p(\Delta \underset{\sim}{p}) = \left\{ \frac{\partial \hat{W}_2^p(\Delta \underset{\sim}{p})}{\partial \Delta \underset{\sim}{p}} \right\} \quad . \square \tag{5.37}$$

We have already seen that $\hat{W}_2^p$ is differentiable everywhere except at the origin, $\Delta \underset{\sim}{p} = \underset{\sim}{0}$. Thus, for $\Delta \underset{\sim}{p} \neq \underset{\sim}{0}$ (5.36) reduces to (5.37); for $\Delta \underset{\sim}{p} = \underset{\sim}{0}$ (5.36) may be simplified by making use of the fact that $\hat{W}_2^p(\underset{\sim}{0}) = 0$:

$$\partial \hat{W}_2^p(\underset{\sim}{0}) = \{\underset{\sim}{\tau} : \hat{W}_2^p(\Delta \underset{\sim}{q}) - \underset{\sim}{\tau} \cdot \Delta \underset{\sim}{q} \geq 0 \ , \quad \forall \Delta \underset{\sim}{q}\} \quad . \tag{5.38}$$

The subdifferential $\partial \hat{W}_2^p(\underset{\sim}{0})$ is illustrated schematically in Fig. 2.9.

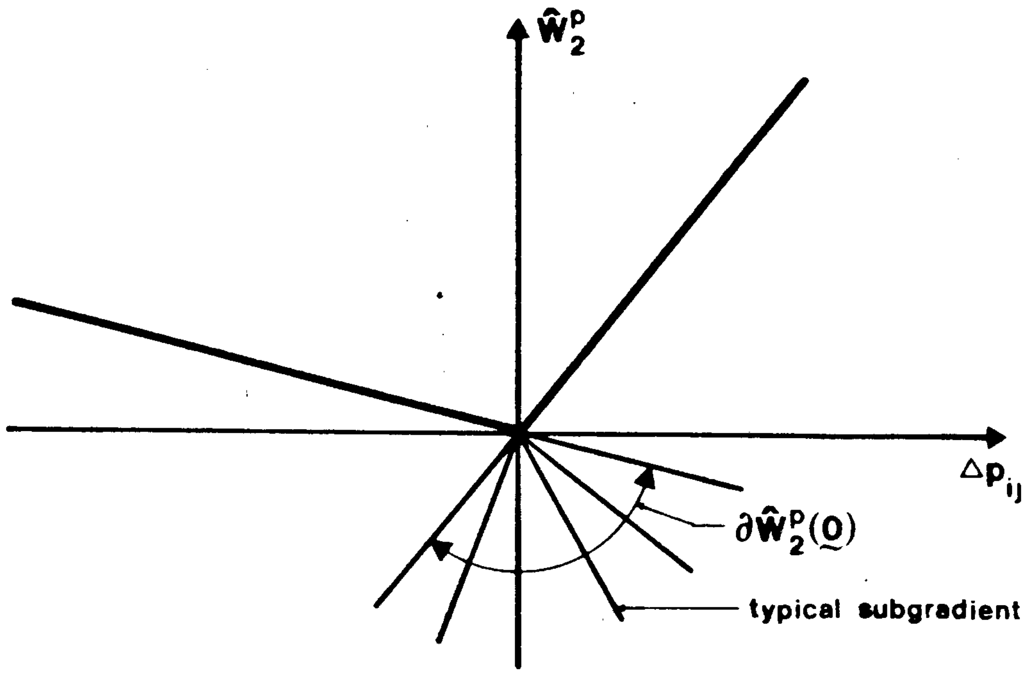


Figure 2.9 Subdifferential of the function $\hat{W}^p_2$ at the origin.

The subgradients $\underset{\sim}{\tau}$ in (5.36) and (5.38) are, of course, stress states; in the case of (5.38) they are all those stress states which are assumed to be attainable when the actual plastic strain $\Delta\underset{\sim}{p} = \underset{\sim}{0}$. It is essential therefore to ensure that the definition of the subdifferential is consistent in the sense that the subgradients are admissible stress states. Thus, for $\Delta\underset{\sim}{p} = \underset{\sim}{0}$ it is necessary that the stress states $\underset{\sim}{\tau}$ lie within or on the yield surface $\phi(\underset{\sim}{\tau},\underset{\sim}{p}^\circ) = \underset{\sim}{0}$, defined at $t = t^\circ$.

From (5.38) the subgradients $\underset{\sim}{\tau}$ are defined by

$$\hat{W}^p_2(\Delta\underset{\sim}{q}) - \underset{\sim}{\tau} \cdot \Delta\underset{\sim}{q} \geq 0 , \quad \forall \Delta\underset{\sim}{q} \quad . \tag{5.39}$$

Substituting for $\hat{W}^p_2$ from (5.22) and rearranging, we get

$$(\hat{\underset{\sim}{\sigma}}(\Delta\underset{\sim}{q}) - \underset{\sim}{\tau}) \cdot \Delta\underset{\sim}{q} \geq 0 \quad , \forall \Delta\underset{\sim}{q} \quad . \tag{5.40}$$

Now since $\underset{\sim}{\Delta q}$ is an arbitrary admissable plastic strain state, and since by definition $\hat{\underset{\sim}{\sigma}}(\underset{\sim}{\Delta q})$ always lies on the yield surface (see eqn (5.14)), it follows from eqn (5.40) and the principle of maximum plastic work (Section 2.2) that

$$\phi(\underset{\sim}{\tau},\underset{\sim}{p}^{\circ}) \leqslant 0 \quad . \tag{5.41}$$

The stress states $\underset{\sim}{\tau}$, or the subgradients of $\hat{W}^p_2$ at $\underset{\sim}{\Delta p} = \underset{\sim}{0}$, are thus seen to be all those stress states which lie within or on the yield surface at $t = t^{\circ}$. We conclude then that the definition of the subdifferential gives a consistent interpretation of the fundamental postulates of Section 2.2.

The Plastic Constitutive Equation

We return now to the plastic constitutive equation which was developed in Section 2.4 based on extremal paths, and rewrite this equation using the subdifferential $\partial\hat{W}^p_2$. Recalling eqn (4.28) and the subdivision of $\hat{W}^p$ into its parts $\hat{W}^p_1$ and $\hat{W}^p_2$, as defined in eqns (5.20) through (5.22), we have

$$\underset{\sim}{\sigma}^{*} = \left.\frac{\partial\hat{W}^p_1}{\partial\underset{\sim}{\Delta p}}\right|_{\underset{\sim}{\Delta p}^{*}} + \left.\frac{\partial\hat{W}^p_2}{\partial\underset{\sim}{\Delta p}}\right|_{\underset{\sim}{\Delta p}^{*}} \quad . \tag{5.42}$$

The second term on the right-hand-side of the above equation is, as we have seen, not everywhere uniquely defined. We therefore introduce the subdifferential $\partial\hat{W}^p_2$, which is everywhere defined, and rewrite (5.42) as

$$\underset{\sim}{\sigma}^* - \left.\frac{\partial \hat{W}_1^p}{\partial \Delta \underset{\sim}{p}}\right|_{\Delta \underset{\sim}{p}^*} \in \partial \hat{W}_2^p(\Delta \underset{\sim}{p}^*) \quad . \tag{5.43}$$

This is the general form of the plastic constitutive equation for a material which is assumed to follow extremal paths in both stress and strain space.

Alternatively, we may rewrite (5.43) in two parts, one part pertaining to the case when $\Delta \underset{\sim}{p} = \underset{\sim}{0}$, and the other pertaining to the case when $\Delta \underset{\sim}{p} \neq \underset{\sim}{0}$. Differentiating the plastic work function $\hat{W}_1^p$ in (5.21) we get

$$\begin{aligned} \frac{\partial \hat{W}_1^p(\Delta \underset{\sim}{p})}{\partial \Delta \underset{\sim}{p}} &= h \Delta \underset{\sim}{p} \\ &= \underset{\sim}{0} \quad \text{when } \Delta \underset{\sim}{p} = \underset{\sim}{0} \quad . \end{aligned} \tag{5.44}$$

Using this result and the fact that the subdifferential $\partial \hat{W}_2^p$ is given by (5.37) when $\Delta \underset{\sim}{p} \neq 0$ we have the following two equations which together are equivalent to (5.43) :

$$\underset{\sim}{\sigma}^* = \left.\frac{\partial \hat{W}_1^p}{\partial \Delta \underset{\sim}{p}}\right|_{\Delta \underset{\sim}{p}^*} + \left.\frac{\partial \hat{W}_2^p}{\partial \Delta \underset{\sim}{p}}\right|_{\Delta \underset{\sim}{p}^*} \quad , \quad \text{if } \Delta \underset{\sim}{p} \neq \underset{\sim}{0} \quad ; \tag{5.45a}$$

$$\underset{\sim}{\sigma}^* \in \partial \hat{W}_2^p(\underset{\sim}{0}) \quad , \quad \text{if } \Delta \underset{\sim}{p} = \underset{\sim}{0} \quad . \tag{5.45b}$$

To summarise, we have developed a description of a new material whose behaviour is governed by two constitutive equations (4.27) and (4.28). These equations express a simple relationship between the state of stress for a material element and two potential functions, W which is the elastic strain energy, and $\hat{W}^p$ which is the plastic work function. The presence of these potential functions means that the material behaviour is path-independent, or elastic, and that the two constitutive equations therefore define a nonlinear elastic material. This material is nevertheless fully capable under certain circumstances of describing conventional elastic-plastic behaviour. For the particular case of a von Mises yield function with linear kinematic hardening we have developed explicit forms of the two constitutive equations; these have been given in (4.29) for the elastic equation and (5.43), or alternatively (5.45), for the plastic equation.

Our development has included the possibility that initial stresses and plastic strains exist in the material element at $t = t^\circ$; moreover, the constitutive equations have been formulated in terms of finite increments in stress and strain. We do not, of course, claim that solutions obtained using these constitutive relations are always exact, although in certain cases they may be; for example, when the loading is proportional. Nevertheless, we will use these equations to formulate a consistent incremental holonomic problem, which subject we take up in the next chapter.

CHAPTER 3

THE INCREMENTAL HOLONOMIC PROBLEM

In this chapter we propose a new statement of the elastic-plastic boundary-value problem which incorporates the constitutive equations described in Sections 2.4 and 2.5. Our objective is then to proceed with the formulation of a variational principle which can be used as a basis for the numerical approximation of this new boundary-value problem. Our formulation makes allowance for the existence of initial stresses and strains, and is written in terms of finite increments in the field variables. For this reason we refer to it as the incremental holonomic problem.

We begin with the statement of the problem and definition of the solution variables and data in Section 3.1. In Section 3.2 we set forth the fundamental variational principle for the problem, which turns out to be a variational inequality. In Section 3.3 we show the existence of an equivalent minimisation problem. However, since these variational problems involve a non-differentiable functional we consider in Section 3.4 a perturbed minimisation problem; we show existence and uniqueness of the solution to the perturbed problem, and the convergence of the perturbed solution to the solution of the original boundary-value problem.

The perturbed minimisation problem is given in the form of a variational principle involving a differentiable functional, and forms a suitable basis for an approximate numerical solution. In Section 3.5 we set up the perturbed minimisation problem on a finite dimensional subspace, and in Section 3.6 we give an estimate of the error in the proposed finite element solution.

3.1 STATEMENT OF THE PROBLEM

Consider a body which occupies an open bounded domain Ω in R^N , $N \leqslant 3$, with Lipschitz boundary Γ . Let $t \in [t^\circ, t^\circ + \Delta t]$ be a real parameter which parametrises a family of stress fields $\underset{\sim}{\sigma}(\underset{\sim}{x},t)$, displacements fields $\underset{\sim}{u}(\underset{\sim}{x},t)$, and plastic strain fields $\underset{\sim}{p}(\underset{\sim}{x},t)$. Assume that when $t = t^\circ$, the stress field $\underset{\sim}{\sigma}^\circ(\underset{\sim}{x})$ is known and is in equilibrium with the prescribed tractions* $\underset{\sim}{t}^\circ$ on Γ_s and body forces $\underset{\sim}{f}^\circ$ on Ω ; the displacements $\underset{\sim}{u}^\circ(\underset{\sim}{x})$, which are assumed to be zero on Γ_u , and the plastic strains $\underset{\sim}{p}^\circ(\underset{\sim}{x})$ are also known and are assumed to consitute a kinematically admissible set.

Suppose that over the interval Δt the tractions are changed by $\Delta \underset{\sim}{t}$ to $\underset{\sim}{t}$ on Γ_s , and the body forces are changed by $\Delta \underset{\sim}{f}$ to $\underset{\sim}{f}$. Suppose further that the stress field $\underset{\sim}{\sigma}(\underset{\sim}{x})$ at the end of this interval is reached by an extremal path for each material point $\underset{\sim}{x}$ in Ω . Then we seek the changes in the displacement field $\Delta\underset{\sim}{u}(\underset{\sim}{x})$ and changes in plastic strain field $\Delta\underset{\sim}{p}(\underset{\sim}{x})$ which satisfy

---

* No confusion should arise between the scalar parameter t and the traction vector $\underset{\sim}{t}$.

(i) the equations of equilibrium

$$\operatorname{div} \underset{\sim}{\sigma} + \underset{\sim}{f} = \underset{\sim}{0} \quad \text{on } \Omega \tag{1.1}$$

(ii) the constitutive equations

$$\left.\begin{aligned} \underset{\sim}{\sigma} &= \underset{\sim}{C}[\nabla \underset{\sim}{u}^\circ + \nabla(\Delta \underset{\sim}{u}) - \underset{\sim}{p}^\circ - \Delta \underset{\sim}{p}] \\ \underset{\sim}{\sigma} - \frac{\partial \hat{W}_1^p(\Delta \underset{\sim}{p})}{\partial \Delta \underset{\sim}{p}} &\in \partial \hat{W}_2^p(\Delta \underset{\sim}{p}) \end{aligned}\right\} \text{ on } \Omega \tag{1.2}$$

(iii) the boundary conditions

$$\begin{aligned} \underset{\sim}{\sigma}\underset{\sim}{\nu} &= \underset{\sim}{t} \quad \text{on } \Gamma_s \\ \underset{\sim}{u} &= \underset{\sim}{0} \quad \text{on } \Gamma_u . \end{aligned} \tag{1.3}$$

We shall refer to the above statement as Problem (S).

The total plastic work $\hat{W}^p(\Delta \underset{\sim}{p})$ is generally the sum of two terms $\hat{W}_1^p(\Delta \underset{\sim}{p})$ and $\hat{W}_2^p(\Delta \underset{\sim}{p})$. The plastic work function $\hat{W}_1^p(\Delta \underset{\sim}{p})$ is a consequence of the hardening behaviour of the material and for linear kinematic hardening is given by, (recall Section 2.5, eqn (5.21)),

$$\hat{W}_1^p(\Delta \underset{\sim}{p}) = \frac{1}{2} E_p(\Delta \underset{\sim}{p} \cdot \Delta \underset{\sim}{p}) \tag{1.4}$$

where E_p ($\equiv h$ in Section 2.5) is a scalar hardening parameter defined in Section 2.3, eqn (3.13). We require that $0 \leq E_p \leq E$, where E is Young's modulus.

The second plastic work function $\hat{W}_2^p(\Delta \underset{\sim}{p})$ is due to both plastic flow behaviour and the translation of the yield surface in stress space, and is always present. From Section 2.5, eqn (5.29), we recall that

$$\hat{W}_2^p(\Delta \underset{\sim}{p}) = \sqrt{2}k \sqrt{\Delta \underset{\sim}{p} \cdot \Delta \underset{\sim}{p}} + E_p \underset{\sim}{p}^{\circ} \cdot \Delta \underset{\sim}{p} \qquad (1.5)$$

where $k = \sigma_o/\sqrt{3}$ is the effective yield stress. The subdifferential $\partial \hat{W}_2^p(\Delta \underset{\sim}{p})$ was defined in Definition 5.1, Section 2.5, and is repeated here for completeness :

$$\partial \hat{W}_2^p(\Delta \underset{\sim}{p}) = \{\underset{\sim}{\tau} : \hat{W}_2^p(\Delta \underset{\sim}{q}) - \hat{W}_2^p(\Delta \underset{\sim}{p}) - \underset{\sim}{\tau} \cdot (\Delta \underset{\sim}{q} - \Delta \underset{\sim}{p}) \geq 0 \, , \ \forall \Delta \underset{\sim}{q}\} \, . \qquad (1.6)$$

The quantity $\nabla \underset{\sim}{u}^{\circ}$ represents the gradient of the initial displacement field $\underset{\sim}{u}^{\circ}(\underset{\sim}{x})$. We assume that the components $u^{\circ}_{i,j}$ of the gradient are in $L_\infty(\Omega)$ and that a positive constant h_1 exists such that

$$\max_{i,j} \| u^{o}_{i,j} \|_\infty \leq h_1 \quad . \qquad (1.7)$$

Similarly, we assume that the components p°_{ij} of the initial plastic strain field $\underset{\sim}{p}^{\circ}(\underset{\sim}{x})$ are contained in $L_\infty(\Omega)$ and that a positive constant h_2 exists such that

$$\max_{i,j} \| p^{o}_{ij} \|_\infty \leq h_2 \quad . \qquad (1.8)$$

Relative to an orthonormal basis the components of $\underset{\sim}{C}$ are $C_{ijk\ell}$, being the elastic constants for the material. These components exhibit the symmetries

$$C_{ijk\ell} = C_{ij\ell k} = C_{jik\ell} = C_{k\ell ij} \tag{1.9}$$

and obey the strong ellipticity condition : there exists a positive constant c_1 such that

$$C_{ijk\ell} A_{k\ell} A_{ij} \geq c_1 A_{ij} A_{ij} \tag{1.10}$$

holds for all symmetric second-order tensors $\underset{\sim}{A}(\underset{\sim}{x})$. We require further that $C_{ijk\ell} \in L_\infty(\Omega)$ and that there exists a positive constant c_2 such that

$$\max_{i,j,k,\ell} \|C_{ijk\ell}\|_\infty \leq c_2 \quad . \tag{1.11}$$

The components f_i of the body force vector $\underset{\sim}{f}$ are assumed to be given as functions in $L_2(\Omega)$, and the components t_i of the surface traction vector $\underset{\sim}{t}$ are assumed to be given in $L_2(\Gamma_s)$.

We will find it convenient at times to make use of the symmetric small strain tensor $\underset{\sim}{\varepsilon}$, defined by

$$\underset{\sim}{\varepsilon} = \tfrac{1}{2} \; (\nabla \underset{\sim}{u} + \nabla^T \underset{\sim}{u}) \quad . \tag{1.12}$$

We define also the anti-symmetric small rotation tensor $\underset{\sim}{\omega}$ by

$$\underset{\sim}{\omega} = \tfrac{1}{2} \; (\nabla \underset{\sim}{u} - \nabla^T \underset{\sim}{u}) \tag{1.13}$$

so that we may write

$$\nabla \underset{\sim}{u} = \underset{\sim}{\varepsilon} + \underset{\sim}{\omega} \quad . \tag{1.14}$$

Operating on both sides with the symmetric fourth-order tensor $\underset{\sim}{C}$, we get

$$\underset{\sim}{C} \, \nabla \underset{\sim}{u} = \underset{\approx}{C\varepsilon} + \underset{\approx}{C\omega}$$

$$= \underset{\approx}{C\varepsilon} \tag{1.15}$$

because of the symmetries in (1.9).

This completes the statement of the problem and the definitions of the functions and data with which we will be working. We proceed now to the formulation of the fundamental variational principle.

3.2 A VARIATIONAL INEQUALITY

Our objective in this section is to construct the fundamental variational principle for the incremental holonomic problem which is in some sense equivalent to the classical statement of the problem discussed in Section 3.1. As we shall see, the principle which we develop here will be in the form of a variational inequality defined on a set of admissible displacements and plastic strains.

Function Spaces

To begin with we will define the spaces of functions within which we will be working. Throughout this chapter we will make use of the notation and conventions for the Sobolev Spaces $H^m(\Omega) \equiv W_2^m(\Omega)$, where $W_2^m(\Omega)$ is the space of functions on Ω whose distributional derivatives of order $\leq m$ are in $L_2(\Omega)$. These Sobolev spaces are Hilbert spaces. We will be concerned exclusively with the case m = 1, for which the inner product is

$$(u,v)_1 = \int_\Omega uv \, dx + \int_\Omega \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i} \, dx \tag{2.1}$$

and for which the norm generated by the inner product is defined by

$$\|u\|_1^2 = (u,u)_1 \quad . \tag{2.2}$$

The following spaces of functions on Ω will be required.

1. The space

$$V = \{\underset{\sim}{v} = (v_1,\ldots,v_N) \; : \; v_i \in H^1(\Omega) \; ; \; v_i = 0 \text{ on } \Gamma_u\} \quad . \tag{2.3}$$

This is a Hilbert space with an inner product

$$(\underset{\sim}{u},\underset{\sim}{v})_V = \int_\Omega \underset{\sim}{u} \cdot \underset{\sim}{v} + \nabla\underset{\sim}{u} \cdot \nabla\underset{\sim}{v} \, dx \tag{2.4}$$

and a norm $\|\underset{\sim}{u}\|_V$ defined by

$$\|\underset{\sim}{u}\|_V^2 = (\underset{\sim}{u},\underset{\sim}{u})_V \quad . \tag{2.5}$$

2. The space

$$L = \{\underset{\sim}{q} : q_{ij} \in L_2(\Omega) \ ; \ i,j = 1,\ldots,N\}$$

$$= (L_2(\Omega))^{N\mathrm{x}N} \quad . \tag{2.6}$$

This is also a Hilbert space with an inner product

$$(\underset{\sim}{p},\underset{\sim}{q})_L = \int_\Omega \underset{\sim}{p} \cdot \underset{\sim}{q} \, dx \tag{2.7}$$

and a norm $\|\underset{\sim}{p}\|_L$ defined by

$$\|\underset{\sim}{p}\|_L^2 = (\underset{\sim}{p},\underset{\sim}{p})_L \quad . \tag{2.8}$$

For convenience we define the pairs

$$\bar{u} = (\Delta\underset{\sim}{u},\Delta\underset{\sim}{p}) \text{ and } \bar{v} = (\Delta\underset{\sim}{v},\Delta\underset{\sim}{q}) \tag{2.9}$$

and the product space

$$\bar{V} = V \times L \tag{2.10}$$

which is a Hilbert space equipped with the product norm $\|\bar{u}\|_{\bar{V}}$ defined by

$$\|\bar{u}\|_{\bar{V}}^2 = \|\Delta\underset{\sim}{u}\|_V^2 + \|\Delta\underset{\sim}{p}\|_L^2 \quad . \tag{2.11}$$

Korn's Inequalities

For proving the V-ellipticity of bilinear forms, Korn's inequalities are essential mathematical tools. Korn's inequalities are actually special cases of Garding's inequality for elliptic systems and are generally stated as follows. Let Ω be an open bounded set with regular boundary. Then we have the following two inequalities.

The First Inequality. Let $\underset{\sim}{v} \in ((H_0^1(\Omega))^N$. Then

$$\int_\Omega \underset{\sim}{\varepsilon} \cdot \underset{\sim}{\varepsilon} \quad dx \geq c_1 \|\underset{\sim}{v}\|^2_{(H^1(\Omega))^N} \tag{2.12}$$

for some constant $c_1 > 0$, independent of $\underset{\sim}{v}$, where $\underset{\sim}{\varepsilon}$ is defined in (1.12).

The Second Inequality. Let $\underset{\sim}{v} \in ((H^1(\Omega))^N$. Then

$$\int_\Omega \underset{\sim}{\varepsilon} \cdot \underset{\sim}{\varepsilon} \, dx + \int_\Omega \underset{\sim}{v} \cdot \underset{\sim}{v} \, dx \geq c_2 \|\underset{\sim}{v}\|^2_{(H^1(\Omega))^N} \tag{2.13}$$

for some constant $c_2 > 0$.

A proof of the first inequality is given by MARSDEN and HUGHES (1983), Chapter 6, Section 1.12. DUVAUT and LIONS (1976) prove the more general second inequality in Section 3.3, Theorem 3.1.

It is clear that the first inequality is suitable for use only in Dirichlet type problems. For mixed problems, having both displacement and traction boundary conditions, the second inequality is not directly applicable because the second term on the left-hand-side does not generally occur in the bilinear form. However DUVAUT and LIONS (1976) have shown (Section 3.3, Theorem 3.3) that for $\underset{\sim}{v} \in V$, as defined in (2.3), and for some $c_o > 0$, that

$$\int_\Omega \underset{\sim}{\varepsilon} \cdot \underset{\sim}{\varepsilon}\, dx \geq c_o \int_\Omega \underset{\sim}{v} \cdot \underset{\sim}{v}\, dx$$

$$\text{or } \int_\Omega \underset{\sim}{\varepsilon} \cdot \underset{\sim}{\varepsilon}\, dx - c_o \int_\Omega \underset{\sim}{v} \cdot \underset{\sim}{v}\, dx \geq 0 \quad . \tag{2.14}$$

If we now multiply eqn (2.13) through by c_o and add the resulting inequality to (2.14) we obtain

$$\int_\Omega \underset{\sim}{\varepsilon} \cdot \underset{\sim}{\varepsilon}\, dx \geq \frac{c_o c_2}{(1 + c_o)} \|\underset{\sim}{v}\|_V^2$$

$$= k\|\underset{\sim}{v}\|_V^2 \quad . \tag{2.15}$$

It is in this form that we will use Korn's inequality later on. The interested reader will find further useful discussion and application of Korn's inequalities in HLAVACEK and NECAS (1970), and CHOU and WANG (1979).

Bilinear Forms, Functionals and their Properties

We turn our attention now to the definition of the bilinear form and functionals which we propose to use in the variational formulation : we define

(i) a symmetric bilinear form $a : \bar{V} \times \bar{V} \to R$,

$$a(\bar{u},\bar{v}) = \int_\Omega [\underset{\sim}{C}(\nabla(\Delta \underset{\sim}{u}) - \Delta \underset{\sim}{p})] \cdot [\nabla(\Delta \underset{\sim}{v}) - \Delta \underset{\sim}{q}]\, dx + E_p \int_\Omega \Delta \underset{\sim}{p} \cdot \Delta \underset{\sim}{q}\, dx$$

$$= \int_\Omega [C_{ijk\ell}(\Delta u_{i,j} - \Delta p_{ij})(\Delta v_{k,\ell} - \Delta q_{k\ell})]\, dx + E_p \int_\Omega \Delta p_{ij}\, \Delta q_{ij}\, dx \tag{2.16}$$

(ii) a linear functional $f : \bar{V} \to R$,

$$f(\bar{v}) = -\int_\Omega (\underset{\sim}{C}[\nabla \underset{\sim}{u}^o - \underset{\sim}{p}^o]) \cdot [\nabla(\Delta \underset{\sim}{v}) - \Delta \underset{\sim}{q}]\, dx$$

$$+ \int_\Omega (\underset{\sim}{f}^o + \Delta \underset{\sim}{f}) \cdot \Delta \underset{\sim}{v}\, dx + \int_{\Gamma_s} (\underset{\sim}{t}^o + \Delta \underset{\sim}{t}) \cdot \Delta \underset{\sim}{v}\, dx$$

$$= -\int_\Omega [C_{ijk\ell}(u^o_{i,j} - p^o_{ij})(\Delta v_{k,\ell} - \Delta q_{k\ell})]\, dx$$

$$+ \int_\Omega (f^o_i + \Delta f_i)\Delta v_i\, dx + \int_{\Gamma_s} (t^o_i + \Delta t_i)\Delta v_i\, dx \tag{2.17}$$

(iii) a non-differentiable functional $j : L \to R$,

$$j(\Delta \underset{\sim}{q}) = \int_\Omega \hat{W}^p_2(\Delta \underset{\sim}{q})\, dx \tag{2.18}$$

where $\hat{W}^p_2(\Delta \underset{\sim}{q})$ is given by (1.5).

The results which are to be presented later in this section depend on certain fundamental properties of the above bilinear form and functionals, which we shall now state and prove in the following four lemmas.

Lemma 3.1

(a) The bilinear form $a : \bar{V} \times \bar{V} \to R$ is $\bar{V}$-elliptic in the sense that there exists a positive constant α such that

$$a(\bar{v},\bar{v}) \geq \alpha \|\bar{v}\|^2_{\bar{V}} \quad . \tag{2.19}$$

(b) The bilinear form $a : \bar{V} \times \bar{V} \to R$ is continuous; that is, there exists a positive constant M such that

$$\left|a(\bar{u},\bar{v})\right| \leq M \|\bar{u}\|_{\bar{V}} \|\bar{v}\|_{\bar{V}} \quad . \tag{2.20}$$

Proof

(a) We follow a similar proof due to JIANG (1984). We have from (2.16), using the identity (1.15),

$$a(\bar{v},\bar{v}) = \int_\Omega (\underset{\sim}{C}[\Delta\underset{\sim}{\varepsilon} - \Delta\underset{\sim}{q}]) \cdot [\Delta\underset{\sim}{\varepsilon} - \Delta\underset{\sim}{q}]\, dx + E_p \int_\Omega \Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}\, dx$$

$$\geq c_1 \int_\Omega [\Delta\underset{\sim}{\varepsilon} - \Delta\underset{\sim}{q}] \cdot [\Delta\underset{\sim}{\varepsilon} - \Delta\underset{\sim}{q}]\, dx + E_p \int_\Omega \Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}\, dx \quad ,$$

using the ellipticity of $\underset{\sim}{C}$, eqn (1.10)

$$= c_1 \int_\Omega [(\Delta\underset{\sim}{\varepsilon} \cdot \Delta\underset{\sim}{\varepsilon}) - 2(\Delta\underset{\sim}{\varepsilon} \cdot \Delta\underset{\sim}{q}) + (\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q})]\, dx + E_p \int_\Omega \Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}\, dx$$

$$= c_1 \int_\Omega [\theta(\Delta\underset{\sim}{\varepsilon} \cdot \Delta\underset{\sim}{\varepsilon}) + (\sqrt{1-\theta}\Delta\underset{\sim}{\varepsilon} - \frac{1}{\sqrt{1-\theta}} \Delta\underset{\sim}{q}) \cdot (\sqrt{1-\theta}\Delta\underset{\sim}{\varepsilon} - \frac{1}{\sqrt{1-\theta}} \Delta\underset{\sim}{q})$$

$$- \frac{\theta}{1-\theta} (\Delta \underset{\sim}{q} \cdot \Delta \underset{\sim}{q})] \, dx + E_p \int_\Omega \Delta \underset{\sim}{q} \cdot \Delta \underset{\sim}{q} \, dx$$

$$\geq \int_\Omega [c_1 \theta (\Delta \underset{\sim}{\varepsilon} \cdot \Delta \underset{\sim}{\varepsilon}) + (E_p - \frac{c_1 \theta}{1-\theta})(\Delta \underset{\sim}{q} \cdot \Delta \underset{\sim}{q})] \, dx$$

where $0 < \theta < 1$. Let us choose $\theta = \dfrac{E_p}{2c_1 + E_p}$; then

$$a(\bar{v},\bar{v}) \geq \int_\Omega (\frac{c_1 E_p}{2c_1 + E_p})(\Delta \underset{\sim}{\varepsilon} \cdot \Delta \underset{\sim}{\varepsilon}) + \frac{1}{2} E_p (\Delta \underset{\sim}{q} \cdot \Delta \underset{\sim}{q}) \, dx$$

$$\geq (\frac{k c_1 E_p}{2c_1 + E_p}) \, \|\Delta \underset{\sim}{v}\|_V^2 + \frac{1}{2} E_p \|\Delta \underset{\sim}{q}\|_L^2 ,$$

using Korn's inequality (2.15), where $k > 0$ is Korn's constant. Hence,

$$a(\bar{v},\bar{v}) \geq \alpha \|\bar{v}\|_{\bar{V}}^2$$

where $\alpha = \min (\dfrac{k c_1 E_p}{2c_1 + E_p} , \dfrac{1}{2} E_p)$. □

Note that $a(\bar{v},\bar{v})$ is $\bar{V}$-elliptic if and only if $E_p > 0$, i.e. if the material exhibits hardening.

(b) We will show the proof in two stages. First, we consider only the elastic contribution to $a(\bar{u},\bar{v})$ by defining the bilinear form

$$\bar{a}(\Delta \underset{\sim}{u},\Delta \underset{\sim}{v}) = \int_\Omega C_{ijk\ell} \, \Delta u_{i,j} \, \Delta v_{k,\ell} \, dx \quad .$$

Then,

$$\left|\bar{a}(\Delta\underset{\sim}{u},\Delta\underset{\sim}{v})\right| = \left|\int_{\Omega} C_{ijk\ell}\ \Delta u_{i,j}\ \Delta v_{k,\ell}\ dx\right|$$

$$\leq \left|c_2 \int_{\Omega} (\sum_{i,j} \Delta u_{i,j})(\sum_{k,\ell} \Delta v_{k,\ell})\ dx\ \right| \quad \text{, using (1.11)}$$

$$= c_2 \left|(\sum_{i,j} \Delta u_{i,j}\ ,\ \sum_{k,\ell} \Delta v_{k,\ell})_o\right|$$

$$\leq c_2 \|\sum_{i,j} \Delta u_{i,j}\|_o \quad \|\sum_{k,\ell} \Delta v_{k,\ell}\|_o$$

$$\leq c_2 \sum_{i,j} \|\Delta u_{i,j}\|_o \quad \sum_{k,\ell} \|\Delta v_{k,\ell}\|_o \tag{2.21}$$

using the triangle inequality, $\|\sum\cdot\|_o \leq \sum \|\cdot\|_o$, where $\|\cdot\|_o$ denotes the L_2 norm.

From the definition of the norm $\|\Delta\underset{\sim}{u}\|_V$, eqn (2.5), we easily deduce that

$$\sum_{i,j} \|\Delta u_{i,j}\|_{L_2} \leq N^2 \|\Delta\underset{\sim}{u}\|_V \quad .$$

Substituting this result into (2.21) yields

$$\left|\bar{a}(\Delta\underset{\sim}{u},\Delta\underset{\sim}{v})\right| \leq c_2 N^4 \|\Delta\underset{\sim}{u}\|_V \quad \|\Delta\underset{\sim}{v}\|_V$$

$$= \bar{M} \|\Delta\underset{\sim}{u}\|_V \quad \|\Delta\underset{\sim}{v}\|_V \quad . \tag{2.22}$$

Returning to the bilinear form $a(\bar{u},\bar{v})$ we have

$$|a(\bar{u},\bar{v})| = \left|\int_\Omega C_{ijk\ell}(\Delta u_{i,j} - \Delta p_{ij})(\Delta v_{k,\ell} - \Delta q_{k\ell})\, dx + E_p\int_\Omega \Delta p_{ij}\Delta q_{ij}\, dx\right| .$$

We extend the elastic result (2.22) by replacing $\Delta u_{i,j}$ by $(\Delta u_{i,j} - \Delta p_{ij})$ and $\Delta v_{k,\ell}$ by $(\Delta v_{k,\ell} - \Delta q_{k\ell})$, and so obtain

$$|a(\bar{u},\bar{v})| \leq \bar{M}\|\nabla(\Delta\underset{\sim}{u}) - \Delta\underset{\sim}{p}\|_L \|\nabla(\Delta\underset{\sim}{v}) - \Delta\underset{\sim}{q}\|_L + |E_p(\Delta\underset{\sim}{p},\Delta\underset{\sim}{q})_L| ,$$

$$\leq \bar{M}(\|\Delta\underset{\sim}{u}\|_V + \|\Delta\underset{\sim}{p}\|_L)(\|\Delta\underset{\sim}{v}\|_V + \|\Delta\underset{\sim}{q}\|_L) + E_p\|\Delta\underset{\sim}{p}\|_L \|\Delta\underset{\sim}{q}\|_L ,$$

(using the triangle inequality for the first term, and the Schwarz inequality on L for the second term)

$$\leq (\bar{M} + E_p)[\|\Delta\underset{\sim}{u}\|_V \|\Delta\underset{\sim}{v}\|_V + \|\Delta\underset{\sim}{u}\|_V \|\Delta\underset{\sim}{q}\|_L + \|\Delta\underset{\sim}{v}\|_V \|\Delta\underset{\sim}{p}\|_L + \|\Delta\underset{\sim}{p}\|_L \|\Delta\underset{\sim}{q}\|_L]$$

$$= 2(\bar{M} + E_p)[\frac{1}{\sqrt{2}}(\|\Delta\underset{\sim}{u}\|_V + \|\Delta\underset{\sim}{p}\|_L)][\frac{1}{\sqrt{2}}(\|\underset{\sim}{v}\|_V + \|\Delta\underset{\sim}{q}\|_L)]$$

$$\leq 2(\bar{M} + E_p)[\|\Delta\underset{\sim}{u}\|_V^2 + \|\Delta\underset{\sim}{p}\|_L^2]^{1/2} [\|\Delta\underset{\sim}{v}\|_V^2 + \|\Delta\underset{\sim}{q}\|_L^2]^{1/2}$$

$$= M \|\bar{u}\|_{\bar{V}} \|\bar{v}\|_{\bar{V}}$$

which is the desired result. To obtain the penultimate step in this proof we have made use of the inequality

$$a + b \leq \sqrt{2}(a^2 + b^2)^{1/2} , \quad a,b \in R . \square$$

Lemma 3.2

The functional $f : \bar{V} \to R$ is continuous; that is, there exists a positive constant m such that,

$$\left|f(\bar{v})\right| \leq m\|\bar{v}\|_{\bar{V}} \quad . \tag{2.23}$$

Proof

We give the proof in two stages. Consider the first integral on the right-hand-side of (2.17), which we call I_1 here. Then

$$\left|I_1\right| = \left| \int_\Omega C_{ijk\ell}[u^o_{i,j} - p^o_{ij}][\Delta v_{k,\ell} - \Delta q_{k\ell}]\ dx\right|$$

$$\leq \left|c_2 \int_\Omega \sum_{i,j} \sum_{k,\ell} [u^o_{i,j} - p^o_{ij}][\Delta v_{k,\ell} - \Delta q_{k\ell}]\ dx\right| \quad \text{, using (1.11)}$$

$$= \left|c_2 \int_\Omega (\sum_{i,j} u^o_{i,j})(\sum_{k,\ell} \Delta v_{k,\ell}) - (\sum_{i,j} u^o_{i,j})(\sum_{k,\ell} \Delta q_{k\ell}) - (\sum_{i,j} p^o_{ij})(\sum_{k,\ell} \Delta v_{k,\ell}) + (\sum_{i,j} p^o_{ij})(\sum_{k,\ell} \Delta q_{k\ell})\ dx\right|$$

$$\leq c_2\{\|\sum_{i,j} u^o_{i,j}\|_o \|\sum_{k,\ell} \Delta v_{k,\ell}\|_o + \|\sum_{i,j} u^o_{i,j}\|_o \|\sum_{k,\ell} \Delta q_{k\ell}\|_o + \|\sum_{i,j} p^o_{ij}\|_o \|\sum_{k,\ell} \Delta v_{k,\ell}\|_o + \|\sum_{i,j} p^o_{ij}\|_o \|\sum_{k,\ell} \Delta q_{k\ell}\|_o\}$$

using the Schwarz inequality, where $\|\cdot\|_o$ is the L_2 norm.

From the triangle inequality $\|\sum_{i,j} \cdot\|_o \leq \sum_{i,j} \|\cdot\|_o$, and using (1.7) and (1.8) in the above we get

$$|I_1| \leq c_2 N^2 (h_1 + h_2)\{\sum_{k,\ell} \|\Delta v_{k,\ell}\|_o + \sum_{k,\ell} \|\Delta q_{k\ell}\|_o\} \quad . \tag{2.24}$$

From (2.5) we may easily deduce that

$$\sum_{k,\ell} \|\Delta v_{k,\ell}\|_o \leq N^2 \|\Delta \underset{\sim}{v}\|_V$$

and similarly from (2.8)

$$\sum_{k,\ell} \|\Delta q_{k\ell}\|_o \leq N^2 \|\Delta \underset{\sim}{q}\|_L \quad .$$

Using the above results in (2.24) we get

$$|I_1| \leq c_2 N^4 (h_1 + h_2)(\|\Delta \underset{\sim}{v}\|_V + \|\Delta \underset{\sim}{q}\|_L) \tag{2.25}$$

$$\leq \sqrt{2}\, c_2 N^4 (h \;\; + h_2)(\|\Delta \underset{\sim}{v}\|_V^2 + \|\Delta \underset{\sim}{q}\|_L^2)^{1/2}$$

$$= m_1 \|\bar{v}\|_{\bar{V}} \tag{2.26}$$

where we have used the inequality $(a + b) \leq \sqrt{2}\,(a^2 + b^2)^{1/2}$, $a,b \in R$.

The second and third integrals on the right-hand-side of (2.17), which we will call I_2 , contain the data $\underset{\sim}{f} \in (L_2(\Omega))^N$ and $\underset{\sim}{t} \in (L_2(\Gamma_s))^N$. Using the Trace Theorem on Γ we may write

$$|I_2| = |(\underset{\sim}{f}^o + \Delta f, \Delta \underset{\sim}{v})_{L(\Omega)} + (\underset{\sim}{t}^o + \Delta \underset{\sim}{t}, \Delta \underset{\sim}{v})_{L(\Gamma_s)}|$$

$$\leq \|\underset{\sim}{f}^o + \Delta \underset{\sim}{f}\|_{L(\Omega)} \|\Delta \underset{\sim}{v}\|_V + \|\underset{\sim}{t}^o + \Delta \underset{\sim}{t}\|_{L(\Gamma_s)} \|\Delta \underset{\sim}{v}\|_{L(\Gamma_s)}$$

$$\leq [\|\underset{\sim}{f}^o + \Delta \underset{\sim}{f}\|_{L(\Omega)} + C\|\underset{\sim}{t}^o + \Delta \underset{\sim}{t}\|_{L(\Gamma_s)}] \|\Delta \underset{\sim}{v}\|_V$$

for some $C > 0$, the last term following from the continuity of the trace operator $\gamma : V(\Omega) \to L(\Gamma)$. Hence,

$$|I_2| \leq \frac{m_2}{\sqrt{2}} \|\Delta \underset{\sim}{v}\|_V \quad . \tag{2.27}$$

Adding (2.25) and (2.27) we get

$$|f(\bar{v})| \leq \frac{(m_1 + m_2)}{\sqrt{2}} \|\Delta \underset{\sim}{v}\|_V + \frac{m_1}{\sqrt{2}} \|\Delta \underset{\sim}{q}\|_L$$

$$\leq \frac{m}{\sqrt{2}} [\|\Delta \underset{\sim}{v}\|_V + \|\Delta \underset{\sim}{q}\|_L]$$

where $m = \max\{m_1 + m_2, m_1\} = m_1 + m_2$; hence

$$|f(\bar{v})| \leq m(\|\Delta \underset{\sim}{v}\|_V^2 + \|\Delta \underset{\sim}{q}\|_L^2)^{1/2}$$

$$= m\|\bar{v}\|_{\bar{V}} \quad .$$

which completes the proof. □

Lemma 3.3

(a) The functional j : L → R is convex.

(b) The functional j : L → R is Lipschitz continuous; that is, there exists a positive constant β depending on the domain Ω , such that

$$\left| j(\underset{\sim}{\Delta p}) - j(\underset{\sim}{\Delta q}) \right| \leqslant \beta \| \underset{\sim}{\Delta p} - \underset{\sim}{\Delta q} \|_L \tag{2.28}$$

Proof

(a) Using (1.5) in (2.18) we have

$$j(\underset{\sim}{\Delta q}) = \sqrt{2}k \int_{\Omega} \sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}\, dx + E_p \int_{\Omega} \underset{\sim}{q}^{o} \cdot \underset{\sim}{\Delta q}\, dx \quad .$$

The proof then follows from the convexity of $\sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}$ (Section 2.5, Lemma 2.1), and the linearity of $(\underset{\sim}{q}^{\circ} \cdot \underset{\sim}{\Delta q})$. □

(b) If we fix $\underset{\sim}{x}$ arbitrarily then $\underset{\sim}{\Delta p}(\underset{\sim}{x}), \underset{\sim}{\Delta q}(\underset{\sim}{x}) \in R^{NxN}$ and we have, from (1.5) and the definition of the norm on R^{NxN} (we write simply $\underset{\sim}{\Delta p}$ instead of $\underset{\sim}{\Delta p}(\underset{\sim}{x})$, etc),

$$\hat{W}_2^p(\underset{\sim}{\Delta p}) = \sqrt{2}k\sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}} + E_p(\underset{\sim}{p}^{o} \cdot \underset{\sim}{\Delta p})$$

$$= \sqrt{2}k\left|\underset{\sim}{\Delta p}\right| + E_p(\underset{\sim}{p}^{o} \cdot \underset{\sim}{\Delta p}) \quad . \tag{2.29}$$

Therefore, at $\underset{\sim}{x} \in \Omega$,

$$\left|\hat{W}_2^p(\Delta\underset{\sim}{p}) - \hat{W}_2^p(\Delta\underset{\sim}{q})\right| = \left|\sqrt{2}\left(|\Delta\underset{\sim}{p}| - |\Delta\underset{\sim}{q}|\right) + E_p[p^o \cdot (\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q})]\right|$$

$$\leqslant \sqrt{2}k\left|\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q}\right| + E_p h_2\left|\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q}\right| ,$$

(using (1.8) and the Schwarz inequality for the second term)

$$= (\sqrt{2}k + E_p h_2)\left|\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q}\right| . \qquad (2.30)$$

Here, we have used the inequality $\left||\Delta\underset{\sim}{p}| - |\Delta\underset{\sim}{q}|\right| \leqslant \left|\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q}\right|$, which may be obtained from an application of the Schwarz inequality on R^{NxN} . From (2.18) we have

$$\left|j(\Delta\underset{\sim}{p}) - j(\Delta\underset{\sim}{q})\right| \leqslant \left|\int_\Omega \hat{W}_2^p(\Delta\underset{\sim}{p}) - \hat{W}_2^p(\Delta\underset{\sim}{q})\, dx\right|$$

$$\leqslant \int_\Omega \left|\hat{W}_2^p(\Delta\underset{\sim}{p}) - \hat{W}_2^p(\Delta\underset{\sim}{q})\right| dx$$

$$\leqslant (\sqrt{2}k + E_p h_2) \int_\Omega \left|\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q}\right| dx \quad , \quad \text{using (2.30)}$$

$$\leqslant (\sqrt{2}k + E_p h_2)\, \text{mes}(\Omega)\, \|\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q}\|_L \quad ,$$

$$= \beta \|\Delta\underset{\sim}{p} - \Delta\underset{\sim}{q}\|_L$$

where we used the Schwarz inequality to obtain the penultimate step. □

Lemma 3.4

The functional $j(\underset{\sim}{\Delta q})$ is bounded below in the sense that there exists a positive constant ω such that

$$j(\underset{\sim}{\Delta q}) \geq - \omega \|\bar{v}\|_{\bar{V}} \quad . \tag{2.31}$$

Proof

Using (1.5) in (2.18) we have

$$j(\underset{\sim}{\Delta q}) = \sqrt{2}k\int_{\Omega} \sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}\, dx + E_p\int_{\Omega} \underset{\sim}{q}^o \cdot \underset{\sim}{\Delta q}\, dx \quad . \tag{2.32}$$

Since $k > 0$ always, the first integral is positive. Thus, we focus our attention on the second integral and consider

$$\begin{aligned} \left|E_p\int_{\Omega} \underset{\sim}{q}^o \cdot \underset{\sim}{\Delta q}\, dx\right| &= E_p\left|\int_{\Omega} q^o_{ij}\Delta q_{ij}\, dx\right| \\ &\leq E_p\left|\int_{\Omega} h_2(\sum_{i,j} \Delta q_{ij})\, dx\right| \quad , \text{ using (1.8)} \\ &= E_p\left|(h_2,\sum_{i,j} \Delta q_{ij})_o\right| \\ &\leq E_p \|h_2\|_o \; \|\sum_{i,j} \Delta q_{ij}\|_o \end{aligned}$$

using the Schwarz inequality.

From the triangle inequality $\|\sum_{i,j} \cdot\| \leq \sum_{i,j} \|\cdot\|$; and from the definition of the norm on L we have

$$\sum_{i,j} \|\Delta q_{ij}\|_o \leq N^2 \|\Delta \underset{\sim}{q}\|_L \quad .$$

Hence,

$$\left| E_p \int_\Omega \underset{\sim}{q}^o \cdot \Delta \underset{\sim}{q} \, dx \right| \leq E_p N^2 h_2 \; \text{mes}(\Omega) \; \|\Delta \underset{\sim}{q}\|_L$$

$$= \omega \|\Delta \underset{\sim}{q}\|_L \quad .$$

It follows that

$$- \omega \leq \frac{E_p \int_\Omega \underset{\sim}{q}^o \cdot \Delta \underset{\sim}{q} \, dx}{\|\Delta \underset{\sim}{q}\|_L} \leq \omega$$

from which we see that

$$E_p \int_\Omega \underset{\sim}{q}^o \cdot \Delta \underset{\sim}{q} \, dx \geq - \omega \|\Delta \underset{\sim}{q}\|_L \quad .$$

Dividing both sides by $\|\bar{v}\|_{\bar{V}}$, we get

$$\frac{1}{\|\bar{v}\|_{\bar{V}}} E_p \int_\Omega \underset{\sim}{q}^o \cdot \Delta \underset{\sim}{q} \, dx \geq - \omega \frac{\|\Delta \underset{\sim}{q}\|_L}{\|\bar{v}\|_{\bar{V}}}$$

$$= \frac{- \omega \; \|\Delta \underset{\sim}{q}\|_L}{(\|\Delta \underset{\sim}{v}\|_V^2 + \|\Delta \underset{\sim}{q}\|_L^2)^{1/2}}$$

$$\geq - \omega \quad .$$

Noting that the first integral in (2.32) is positive, the result follows immediately. □

A Variational Inequality

Before proceeding with the statement of the variational principle we first show a simple result, which will prove useful in the subsequent theorem.

Lemma 3.5

The functional $j : L \to R$ satisfies the following inequality :

$$j(\Delta q) - j(\Delta p) - \int_{\Omega} (\sigma - E_p \Delta p) \cdot (\Delta q - \Delta p) \, dx \geq 0 \quad . \tag{2.33}$$

Proof

From (1.2) and (1.6) we have

$$\sigma - \frac{\partial \hat{W}_1^p}{\partial \Delta p} \in \partial \hat{W}_2^p = \{\tau : \hat{W}_2^p(\Delta q) - \hat{W}_2^p(\Delta p) - \tau \cdot (\Delta q - \Delta p) \geq 0\}$$

$$\Rightarrow \quad \hat{W}_2^p(\Delta q) - \hat{W}_2^p(\Delta p) - (\sigma - \frac{\partial \hat{W}_1^p}{\partial \Delta p}) \cdot (\Delta q - \Delta p) \geq 0 \quad .$$

Integrating over Ω and using (2.18) we obtain the result

$$j(\Delta q) - j(\Delta p) - \int_{\Omega} (\sigma(\Delta p) - E_p \Delta p) \cdot (\Delta q - \Delta p) \, dx \geq 0 \quad . \quad \square$$

We are now in a position to state the central result of this section.

Theorem 3.1

Let $\bar{u} = (\Delta\underset{\sim}{u}, \Delta\underset{\sim}{p})$ be the solution of Problem (S) . Then $\bar{u}$ is also the solution of the variational inequality

$$a(\bar{u},\bar{v} - \bar{u}) + j(\Delta\underset{\sim}{q}) - j(\Delta\underset{\sim}{p}) - f(\bar{v} - \bar{u}) \geq 0 , \quad \bar{v} \in \bar{V} . \tag{2.34}$$

Conversely, if $\bar{u}$ is a solution of (2.34) then it also satisfies (S) in a weak sense. We shall refer to this statement as Problem (V).

Proof

(a) (S) => (V) :

We assume the equilibrium equations (1.1) are satisfied. Then if $(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u}) \in V(\Omega)$ is a sufficiently smooth function on Ω , we have*

$$\int_\Omega \{(\mathrm{div}\ \underset{\sim}{\sigma}) + \underset{\sim}{f}\} \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})\ dx = 0 .$$

Using Green's theorem, we obtain

$$\int_{\Gamma_s} \underset{\sim}{t} \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})\ dx - \int_\Omega \underset{\sim}{\sigma} \cdot \nabla(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})\ dx + \int_\Omega \underset{\sim}{f} \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})\ dx = 0 . \tag{2.35}$$

Making the substitution $\Delta\underset{\sim}{\sigma} = \underset{\sim}{C}[\nabla(\Delta\underset{\sim}{u}) - \Delta\underset{\sim}{p}]$ in (2.16), we have

$$a(\bar{u},\bar{v}) = \int_\Omega \Delta\underset{\sim}{\sigma} \cdot [\nabla(\Delta\underset{\sim}{v}) - \Delta\underset{\sim}{q}]dx + E_p\int_\Omega \Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{q}\ dx$$

$$= \int_\Omega [(\Delta\underset{\sim}{\sigma} \cdot \nabla(\Delta\underset{\sim}{v})) - (\Delta\underset{\sim}{\sigma} \cdot \Delta\underset{\sim}{q})]\ dx + E_p\int_\Omega \Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{q}\ dx .$$

---

* For brevity we write $\underset{\sim}{\sigma}$ in place of $\underset{\sim}{\sigma}(\bar{u})$ throughout.

It follows that we may write

$$a(\bar{u},\bar{v} - \bar{u}) = \int_\Omega \Delta\underset{\sim}{\sigma} \cdot \nabla(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx - \int_\Omega \Delta\underset{\sim}{\sigma} \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})dx + E_p\int_\Omega \Delta\underset{\sim}{p} \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})dx$$

$$= \int_\Omega (\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^o) \cdot \nabla(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx - \int_\Omega (\underset{\sim}{\sigma} - \underset{\sim}{\sigma}^o) \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})dx$$

$$+ E_p\int_\Omega \Delta\underset{\sim}{p} \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})dx \quad . \qquad (2.36)$$

Using (2.36) in (2.35) we get

$$- \int_{\Gamma_s} \underset{\sim}{t} \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx + \int_\Omega \underset{\sim}{\sigma}^o \cdot [\nabla(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u}) - (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})]dx$$

$$+ \int_\Omega (\underset{\sim}{\sigma} - E_p\Delta\underset{\sim}{p}) \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})dx + a(\bar{u},\bar{v} - \bar{u}) - \int_\Omega \underset{\sim}{f} \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx = 0 \quad .$$

Using Lemma 3.5 in the above, and noting that $\underset{\sim}{\sigma}^\circ = \underset{\sim}{C}[\nabla\underset{\sim}{u}^\circ - \underset{\sim}{p}^\circ]$, we get

$$a(\bar{u},\bar{v} - \bar{u}) + j(\Delta\underset{\sim}{q}) - j(\Delta\underset{\sim}{p}) - f(\bar{v} - \bar{u}) \geq 0$$

which completes the first part of the proof.

(b) (V) => (S) :

We show this in two parts.

(i) We have from (2.16) and (2.17), replacing $\bar{v}$ by $\bar{v} - \bar{u}$,

$$a(\bar{u},\bar{v} - \bar{u}) = \int_\Omega \Delta\underset{\sim}{\sigma} \cdot [\nabla(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u}) - (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})]dx + E_p\int \Delta\underset{\sim}{p} \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})dx \qquad (2.37)$$

and

$$f(\bar{v} - \bar{u}) = - \int_\Omega \underset{\sim}{\sigma}^o \cdot [\nabla(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u}) - (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})]dx$$

$$+ \int_\Omega (\underset{\sim}{f}^o + \Delta\underset{\sim}{f}) \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx + \int_{\Gamma_s} (\underset{\sim}{t}^o + \Delta\underset{\sim}{t}) \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx \quad . \qquad (2.38)$$

Substituting (2.37) and (2.38) into (2.34) and setting $\underset{\sim}{\Delta q} = \underset{\sim}{\Delta p}$, we get

$$\int_\Omega (\underset{\sim}{\sigma}^o + \Delta\underset{\sim}{\sigma}) \cdot [\nabla(\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})]dx - \int_\Omega (\underset{\sim}{f}^o + \Delta\underset{\sim}{f}) \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx$$

$$- \int_{\Gamma_s} (\underset{\sim}{t}^o + \Delta\underset{\sim}{t}) \cdot (\Delta\underset{\sim}{v} - \Delta\underset{\sim}{u})dx \geq 0 \quad . \tag{2.39}$$

Replacing $\Delta\underset{\sim}{v}$ first by $(\Delta\underset{\sim}{u} + \Delta\underset{\sim}{v})$ and then by $(\Delta\underset{\sim}{u} - \Delta\underset{\sim}{v})$ in (2.39) yields

$$\int_\Omega \underset{\sim}{\sigma} \cdot [\nabla(\Delta\underset{\sim}{v})]dx - \int_\Omega \underset{\sim}{f} \cdot \Delta\underset{\sim}{v}\, dx - \int_{\Gamma_s} \underset{\sim}{t} \cdot \Delta\underset{\sim}{v}\, dx \geq 0$$

and

$$\int_\Omega \underset{\sim}{\sigma} \cdot [\nabla(-\Delta\underset{\sim}{v})]dx - \int_\Omega \underset{\sim}{f} \cdot (-\Delta\underset{\sim}{v})dx - \int_{\Gamma_s} \underset{\sim}{t} \cdot (-\Delta\underset{\sim}{v})dx \geq 0$$

from which it follows that

$$\int_\Omega \underset{\sim}{\sigma} \cdot \nabla(\Delta\underset{\sim}{v})dx - \int_\Omega \underset{\sim}{f} \cdot (\Delta\underset{\sim}{v})dx - \int_{\Gamma_s} \underset{\sim}{t} \cdot (\Delta\underset{\sim}{v})dx = 0 \quad . \tag{2.40}$$

Multiplying (2.40) by -1 and using the reverse Green's theorem yields

$$\int_\Omega (\operatorname{div} \underset{\sim}{\sigma} + \underset{\sim}{f}) \cdot \Delta\underset{\sim}{v}\, dx = 0 \quad , \qquad \Delta\underset{\sim}{v} \in V$$

$$\Rightarrow \quad \operatorname{div} \underset{\sim}{\sigma} + \underset{\sim}{f} = \underset{\sim}{0}$$

in the sense of distributions.

(ii) Again, substituting (2.37) and (2.38) into (2.34) and setting $\Delta\underset{\sim}{v} = \Delta\underset{\sim}{u}$ we get

$$-\int_\Omega \Delta\underset{\sim}{\sigma} \cdot [\Delta\underset{\sim}{q}-\Delta\underset{\sim}{p}]dx + E_p\int_\Omega \Delta\underset{\sim}{p} \cdot (\Delta\underset{\sim}{q}-\Delta\underset{\sim}{p})dx - \int_\Omega \underset{\sim}{\sigma}^o \cdot (\Delta\underset{\sim}{q}-\Delta\underset{\sim}{p})dx + j(\Delta\underset{\sim}{q}) - j(\Delta\underset{\sim}{p}) \geq$$

$$\Rightarrow -\int_\Omega (\underset{\sim}{\sigma} - E_p\Delta\underset{\sim}{p}) \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})dx + \int_\Omega \hat{W}_2^p\,(\Delta\underset{\sim}{q})dx - \int_\Omega \hat{W}_2^p\,(\Delta\underset{\sim}{p})dx \geq 0$$

$$\Rightarrow \int_\Omega \left\{\hat{W}_2^p(\Delta\underset{\sim}{q}) - \hat{W}_2^p(\Delta\underset{\sim}{p}) - (\underset{\sim}{\sigma} - E_p\Delta\underset{\sim}{p}) \cdot (\Delta\underset{\sim}{q} - \Delta\underset{\sim}{p})\right\} dx \geq 0 \quad .$$

Using Lemma 3.5 we deduce the constitutive equation $(1.2)_2$, and this completes the proof of the theorem. □

Theorem 3.1 represents nothing more than a statement of the principle of virtual work for an elastic-plastic body whose plastic material behaviour is governed by the assumption of extremal paths in stress and strain space. Thus $a(\bar{u},\bar{v} - \bar{u})$ represents the work done by the stresses $\underset{\sim}{\sigma}(\bar{u})$ in moving through strains caused by the virtual displacements $\underset{\sim}{v} - \underset{\sim}{u}$, together with the work done in plastic strain hardening. The functional $f(\bar{v} - \bar{u})$ represents the work done by the initial stresses $\underset{\sim}{\sigma}^o$ and the applied body forces and surface tractions. Finally, the terms $j(.)$ represent the plastic work done during flow; this may also be regarded as the work done by the elastic body on internal slip-planes.

The inequality in the variational principle arises from the fact that the functional $j(.)$ is non-differentiable at the origin (recall Section 2.5). This may be interpreted, as can easily be seen from (1.2), that the stresses $\underset{\sim}{\sigma}(\bar{u})$ are not uniquely determined when the plastic strains are zero.

The inequality precludes the direct use of this variational principle as a basis for any conventional approximation procedure, so that we are obliged to investigate alternative formulations for the problem.

3.3 A MINIMISATION PROBLEM

We now make use of a standard result from convex analysis (see, for example, EKELAND and TEMAM (1976)), to establish the connection between the classical problem (S) , the variational inequality (V) , and a minimisation problem which we shall shortly define.

We introduce the functional $J(\bar{v}) : \bar{V} \to R$ defined by

$$J(\bar{v}) = \frac{1}{2} a(\bar{v},\bar{v}) + j(\Delta \underset{\sim}{q}) - f(\bar{v}) \quad . \tag{3.1}$$

For convenience we separate the differentiable and non-differentiable parts of this functional by defining a differentiable functional $I(\bar{v}) : \bar{V} \to R$,

$$I(\bar{v}) = \frac{1}{2} a(\bar{v},\bar{v}) - f(\bar{v}) \tag{3.2}$$

so that $J(\bar{v})$ may then be written as

$$J(\bar{v}) = I(\bar{v}) + j(\Delta \underset{\sim}{q}) \quad . \tag{3.3}$$

The minimisation problem, which we shall refer to as Problem (M), is defined in the following theorem.

Theorem 3.2A

Let $J : \bar{V} \rightarrow R$ be a proper functional of the form $J = I + j$ where I is convex, continuous, and Gateaux-differentiable, and j is convex and continuous. Then if $\bar{u} \in \bar{V}$, the following two conditions are equivalent :

$$J(\bar{u}) \leqslant J(\bar{v}) \quad , \qquad \bar{v} \in \bar{V} \quad \text{(Problem (M))} \tag{3.4}$$

$$\langle DI(\bar{u}),\bar{v} - \bar{u}\rangle + j(\Delta \underset{\sim}{q}) - j(\Delta \underset{\sim}{p}) \geqslant 0 \quad , \qquad \bar{v} \in \bar{V} \tag{3.5}$$

where $DI(\bar{u})$ is the Gateaux differential of I at $\bar{u}$, and $\langle .,.\rangle$ denotes duality pairing on $\bar{V}' \times \bar{V}$, $\bar{V}'$ being the dual space of $\bar{V}$. □

The proof of this theorem is given by EKELAND and TEMAM (1976), Chapter 2, Proposition 2.2, and also by ODEN and KIKUCHI (1980), Theorem 1-5.1, in which they refer to (3.5) as a variational inequality of the second kind. It remains therefore to show that the stated conditions on functionals J,I and j obtain for this particular problem.

(i) $I(\bar{v})$ is Gateaux-differentiable : it is easy to show that $I(\bar{v})$ is differentiable with Gateaux differential $DI(\bar{u})$ characterised by

$$\langle DI(\bar{u}),\bar{v}\rangle = a(\bar{u},\bar{v}) - f(\bar{v}) \tag{3.6}$$

Note that this establishes that (3.5) is an equivalent statement of Problem (V) as defined in Theorem 3.1.

(ii) $I(\bar{v})$ is continuous : this follows immediately from Lemmas 3.1 and 3.2.

(iii) $I(\bar{v})$ is strictly convex : for $0 < \theta < 1$ we have,

$$I(\theta\bar{u} + (1-\theta)\bar{v}) = \frac{1}{2}\, a(\theta\bar{u} + (1-\theta)\bar{v}\ ,\ \theta\bar{u} + (1-\theta)\bar{v}) - f(\theta\bar{u} + (1-\theta)\bar{v})$$

$$= \frac{1}{2}\, \theta^2 a(\bar{u},\bar{u}) + \theta(1-\theta)a(\bar{u},\bar{v}) + \frac{1}{2}\,(1-\theta)^2 a(\bar{v},\bar{v})$$

$$- \theta f(\bar{u}) - (1-\theta)f(\bar{v}) \quad . \qquad (3.7)$$

Now since $a(\bar{u},\bar{u})$ is $\bar{V}$-elliptic (Lemma 3.1) we have

$$a(\bar{u} - \bar{v},\bar{u} - \bar{v}) = a(\bar{u},\bar{u}) - 2a(\bar{u},\bar{v}) + a(\bar{v},\bar{v}) > 0 \quad . \qquad (3.8)$$

Making use of (3.8) in (3.7), we have

$$I(\theta\bar{u} + (1-\theta)\bar{v}) < \frac{1}{2}\,\theta^2 a(\bar{u},\bar{u}) + \frac{1}{2}\,\theta(1-\theta)[a(\bar{u},\bar{u}) + a(\bar{v},\bar{v})]$$

$$+ \frac{1}{2}\,(1-\theta)^2 a(\bar{v},\bar{v}) - \theta f(\bar{u}) - (1-\theta)f(\bar{v})$$

$$< \theta I(\bar{u}) + (1-\theta)I(\bar{v}) \quad . \qquad (3.9)$$

Thus $I(\bar{v})$ is strictly convex.

(iv) $j(\Delta\underset{\sim}{q})$ is convex and continuous : this follows immediately from Lemma 3.3.

(v) $J(\bar{v})$ is strictly convex : this follows immediately from (3.9) above and the convexity of $j(\Delta\underset{\sim}{q})$, Lemma 3.3.

Theorem 3.2A establishes the conditions under which the solution $\bar{u}$ of the variational inequality (V) is also a solution of the minimisation problem (M) . It remains to establish the conditions of existence and uniqueness of the solution $\bar{u}$, which we now do in the following theorem.

Theorem 3.2B

Assume the conditions of Theorem 3.2A are satisfied, and that J is coercive. Then Problem (M) has at least one solution. Furthermore, since J is strictly convex (Theorem 3.2A), the solution $\bar{u}$ is unique. □

The proof of this theorem can be found in EKELAND and TEMAM (1976), Chapter 2, Proposition 1.2. It remains therefore to show that $J(\bar{v})$ is coercive. Using the results of Lemmas 3.1, 3.2 and 3.4, we have

$$J(\bar{v}) \geq \frac{1}{2} \alpha \|\bar{v}\|^2_{\bar{V}} - \omega \|\bar{v}\|_{\bar{V}} - m \|\bar{v}\|_{\bar{V}} \quad .$$

where ω and m are positive constants and α is a non-negative constant.

Dividing both sides by $\|\bar{v}\|_{\bar{V}}$ gives

$$\frac{J(\bar{v})}{\|\bar{v}\|_{\bar{V}}} \geq \frac{1}{2} \alpha \|\bar{v}\|_{\bar{V}} - \omega - m$$

$$\rightarrow \infty \text{ as } \|\bar{v}\|_{\bar{V}} \rightarrow \infty \quad , \quad \text{iff } \alpha > 0 \ . \qquad (3.10)$$

It is important to note that the coercivity result obtains if and only if $\alpha > 0$. Since $\alpha = 0$ when the hardening parameter $E_p = 0$ it is clear that we do not have coercivity for a material which does not harden.

The preceding two theorems, 3.2A and 3.2B, are standard results from convex analysis which establish the conditions under which Problems (V) and (M) are equivalent, and also the conditions under which the solution $\bar{u}$ to Problem (M) exists and is unique. In the light of these theorems and the observations which follow them we may now state the following existence and uniqueness theorem for the incremental holonomic problem.

Theorem 3.3

There exists a unique minimiser $\bar{u}$ of the functional J of (3.3). Moreover, $\bar{u}$ is also the solution of the variational inequality (2.34). □

3.4 A PERTURBED MINIMISATION PROBLEM

The primary objective of developing the variational formulation is to provide a basis for the construction of finite element approximations. However, the way in which the problem is presently formulated does not lend itself to solution using conventional finite element procedures because the functional $j(\Delta \underset{\sim}{q})$ is non-differentiable, a fact which is expressed equivalently in the inequality in (V) .

It seems natural therefore to approximate the non-differentiable functional $j(.)$ by a family of functionals $j_\varepsilon(.)$ which are convex and Gateaux-differentiable, and which are parametrised by a positive real parameter ε . Our objective should be to choose $j_\varepsilon(.)$ in such a way that it approximates $j(.)$ arbitrarily closely as ε approaches zero. We refer to $j_\varepsilon(.)$ as the regularised form of the functional $j(.)$.

Recall from Section 2.5 that the non-differentiability of $j(.)$ arises from the term $\sqrt{2}k\sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}}$ in the plastic work function $\hat{W}_2^p(\Delta\underset{\sim}{q})$, which was given in Section 3.1, eqn (1.5), as

$$\hat{W}_2^p(\Delta\underset{\sim}{q}) = \sqrt{2}k\sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}} + E_p(\underset{\sim}{q}^o \cdot \Delta\underset{\sim}{q}) . \tag{4.1}$$

GLOWINSKI, LIONS and TREMOLIERES (1981) suggest several candidate regularisation procedures of which the most suitable for the present application is to replace the scalar product $\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}$ in (4.1) by the perturbation $\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q} + \varepsilon^2$. Hence, we define the regularised plastic work function

$$\hat{W}_\varepsilon^p(\Delta\underset{\sim}{q}) = \sqrt{2}k\sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q} + \varepsilon^2} + E_p(\underset{\sim}{q}^o \cdot \Delta\underset{\sim}{q}) \tag{4.2}$$

whence the regularised plastic work functional $j_\varepsilon(\Delta\underset{\sim}{q})$ becomes

$$j_\varepsilon(\Delta\underset{\sim}{q}) = \int_\Omega \hat{W}_\varepsilon^p(\Delta\underset{\sim}{q})\, dx \quad . \tag{4.3}$$

The effect of the regularisation is illustrated schematically in Fig. 3.1, where we show both the original curves (recall Fig. 2.8, Section 2.5) and the regularised curves for the plastic work function $\hat{W}_2^p$ and its derivative. In the case of the plastic work it is seen that the regularised function lies everywhere above the original function. It is not zero at the origin, but has a value ε which we can of course make arbitrarily small. In the case of the derivative of the plastic work the curve representing the regularised function is everywhere differentiable.

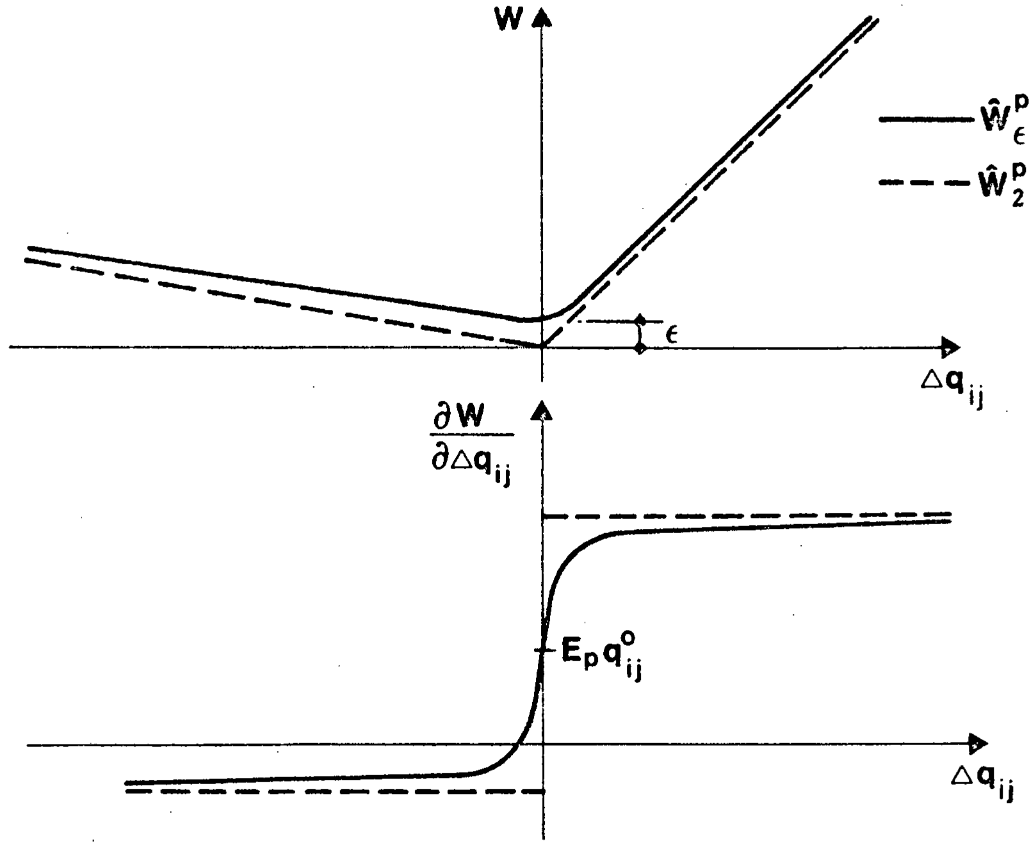


Figure 3.1 Effect of the regularisation of the plastic work function $\hat{W}_2^p$.

Since this derivative represents a stress quantity (see eqn $(1.2)_2$) we have the implication that for every value of the stress a non-zero plastic strain is defined. Needless to say, the curves shown in the figure are highly exaggerated and we would expect that for stress states lying within the current yield surface these plastic strains will be of the order of ε .

We should emphasise that Fig. 3.1 represents only the part $\hat{W}_2^p$ of the total plastic work, which includes the plastic work due to flow of the material and the plastic work which arises out of the presence of initial plastic strains. The plastic work due to hardening behaviour, $\hat{W}_1^p$, is quadratic and well-behaved, and its contribution will be examined later in Chapter 4.

In order for the regularised functional $j_\varepsilon(.)$ to qualify as a suitable candidate for inclusion in the perturbed minimisation problem it is necessary that it possesses certain characteristics. These are established in the following two lemmas.

Lemma 3.6

(a) The functional $j_\varepsilon : L \to R$ is Gateaux-differentiable for all $\varepsilon > 0$, with Gateaux derivative given by

$$\langle Dj_\varepsilon(\underset{\sim}{\Delta p}),\underset{\sim}{\Delta q}\rangle = \sqrt{2}k\int_\Omega (\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^{-1/2}(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q})dx + E_p\int_\Omega \underset{\sim}{p}^o \cdot \underset{\sim}{\Delta q}\, dx \quad . \tag{4.4}$$

(b) The functional $j_\varepsilon : L \to R$ is convex for all $\varepsilon > 0$.

(c) The functional $j_\varepsilon : L \to R$ is Lipschitz continuous; that is, there exists a positive constant β_ε, depending on the domain Ω, such that

$$\left|j_\varepsilon(\underset{\sim}{\Delta p}) - j_\varepsilon(\underset{\sim}{\Delta q})\right| \leqslant \beta_\varepsilon \|\underset{\sim}{\Delta p} - \underset{\sim}{\Delta q}\|_L \quad . \tag{4.5}$$

Proof

(a) From (4.2) and (4.3) we have

$$j_\varepsilon(\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) = \sqrt{2}k \int_\Omega \sqrt{(\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) \cdot (\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) + \varepsilon^2}\; dx$$

$$+ E_p\int_\Omega \underset{\sim}{p}^o \cdot (\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) \; dx \quad .$$

Evaluating $\frac{1}{\theta}[j_\varepsilon(\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) - j_\varepsilon(\underset{\sim}{\Delta p})]$ and taking $\lim_{\theta \to 0}$, with the aid of the Lebesgue dominated convergence theorem we get

$$dj_\varepsilon(\underset{\sim}{\Delta p},\underset{\sim}{\Delta q}) = \sqrt{2}k \int_\Omega \frac{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q}}{\sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2}} dx + E_p \int_\Omega \underset{\sim}{p}^o \cdot \underset{\sim}{\Delta q}\, dx \quad .$$

Clearly $dj(\underset{\sim}{\Delta p},\cdot)$ is linear in $\underset{\sim}{\Delta q}$. Furthermore, we have

$$\left|dj_\varepsilon(\underset{\sim}{\Delta p},\underset{\sim}{\Delta q})\right| = \left|\sqrt{2}k \int_\Omega \frac{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q}}{\sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2}} dx + E_p \int_\Omega \underset{\sim}{p}^o \cdot \underset{\sim}{\Delta q}\, dx\right|$$

$$\leq \sqrt{2}k \left\| \frac{\underset{\sim}{\Delta p}}{\sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2}} \right\|_o \|\underset{\sim}{\Delta q}\|_o + E_p \|\underset{\sim}{p}^o\|_o \|\underset{\sim}{\Delta q}\|_o$$

$$\leq K \|\underset{\sim}{\Delta q}\|_o$$

using the Schwarz inequality on $L_2(\Omega)$, where $\|\cdot\|_o$ denotes the L_2-norm , and K is some positive constant. In arriving at this result we have used the fact that

$$\int_\Omega \frac{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}}{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2} dx \leq \text{mes}(\Omega) < \infty$$

$$\Rightarrow \frac{\underset{\sim}{\Delta p}}{\sqrt{\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2}} \in L(\Omega) \quad .$$

Hence, the operator $Dj_\varepsilon(\underset{\sim}{\Delta p})$, defined by $Dj_\varepsilon(\underset{\sim}{\Delta p})\underset{\sim}{\Delta q}$ $= dj(\underset{\sim}{\Delta p},\underset{\sim}{\Delta q})$ is bounded and is the Gateaux derivative of j_ε at $\underset{\sim}{\Delta p}$ (see Appendix 3A). □

(b) The proof is accomplished in three stages. First we show that the function W_a defined by

$$W_a(\Delta\underset{\sim}{q})(\underset{\sim}{x}) = C\sqrt{\Delta\underset{\sim}{q}(\underset{\sim}{x}) \cdot \Delta\underset{\sim}{q}(\underset{\sim}{x}) + \varepsilon^2}$$

is convex, where we have written $C = \sqrt{2}k$ for brevity.

Squaring both sides and replacing $\Delta\underset{\sim}{q}$ by $(\theta(\Delta\underset{\sim}{p}) + (1-\theta)\Delta\underset{\sim}{q})$, $0 \leq \theta \leq 1$, we get, at $\underset{\sim}{x} \in \Omega$ (it is understood that all functions appearing below are evaluated at $\underset{\sim}{x}$),

$$\begin{aligned}[W_a(\theta\Delta\underset{\sim}{p} + (1-\theta)\Delta\underset{\sim}{q})]^2 &= C^2\{\theta\Delta\underset{\sim}{p} + (1-\theta)\Delta\underset{\sim}{q}\}\{\theta\Delta\underset{\sim}{p} + (1-\theta)\Delta\underset{\sim}{q}\} + C^2\varepsilon^2 \; . \\ &= C^2\theta^2(\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{p}) + 2C^2\theta(1-\theta)(\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{q}) + C^2(1-\theta)^2(\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}) + C^2\varepsilon^2 \\ &= C^2\theta^2(\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{p} + \varepsilon^2) + 2C^2\theta(1-\theta)(\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{q} + \varepsilon^2) \\ &\qquad + C^2(1-\theta)^2(\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q} + \varepsilon^2) \; . \end{aligned} \tag{4.6}$$

Using the Schwarz inequality on $R^{N\mathrm{x}N}$ we have

$$\begin{aligned}|\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{q}| &\leq |\Delta\underset{\sim}{p}||\Delta\underset{\sim}{q}| \\ &= \sqrt{\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{p}}\,\sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}}\end{aligned}$$

and hence,

$$\begin{aligned}C^2(\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{q} + \varepsilon^2) &\leq C^2(\sqrt{\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{p}}\,\sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}} + \varepsilon^2) \\ &\leq C^2\sqrt{\Delta\underset{\sim}{p} \cdot \Delta\underset{\sim}{p} + \varepsilon^2}\,\sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q} + \varepsilon^2} \\ &= C^2 W_a(\Delta\underset{\sim}{p}) W_a(\Delta\underset{\sim}{q})\end{aligned} \tag{4.7}$$

where we have used the inequality $\sqrt{a}\sqrt{b} + \varepsilon^2 \leq \sqrt{a + \varepsilon^2}\sqrt{b + \varepsilon^2}$, $a,b \geq 0$, $\varepsilon > 0$.

Substituting (4.7) into (4.6) and taking the square-root of both sides, we get

$$W_a[\theta \underset{\sim}{\Delta p} + (1-\theta)\underset{\sim}{\Delta q}](\underset{\sim}{x}) \leq \theta\, W_a(\underset{\sim}{\Delta p})(\underset{\sim}{x}) + (1-\theta)W_a(\underset{\sim}{\Delta q})(\underset{\sim}{x})$$

which confirms the convexity of $W_a(\underset{\sim}{\Delta q})$. Using this result it is easy to show that the function $\hat{W}^p_\varepsilon(\underset{\sim}{\Delta q})$ is convex, the proof being identical to that already used to show the convexity of $\hat{W}^p_2(\underset{\sim}{\Delta q})$, (see Lemma 2.1 of Section 2.5).

Finally, we show that $j_\varepsilon(\underset{\sim}{\Delta q})$ is convex : noting that $\hat{W}^p_\varepsilon(\underset{\sim}{\Delta q})$ is everywhere differentiable, from the convexity of $\hat{W}^p_\varepsilon(\underset{\sim}{\Delta p})$ we may write, at $\underset{\sim}{x} \in \Omega$,

$$\hat{W}^p_\varepsilon(\underset{\sim}{\Delta q}) - \hat{W}^p_\varepsilon(\underset{\sim}{\Delta p}) - \left[\frac{\partial \hat{W}^p_\varepsilon(\underset{\sim}{\Delta p})}{\partial \underset{\sim}{\Delta p}} \cdot (\underset{\sim}{\Delta q} - \underset{\sim}{\Delta p})\right] \geq 0 \quad .$$

Integrating this inequality and using (4.3) gives

$$j_\varepsilon(\underset{\sim}{\Delta q}) - j_\varepsilon(\underset{\sim}{\Delta p}) - \langle Dj_\varepsilon(\underset{\sim}{\Delta p}), \underset{\sim}{\Delta q} - \underset{\sim}{\Delta p} \rangle \geq 0$$

which confirms that $j_\varepsilon(\underset{\sim}{\Delta q})$ is convex. □

(c) Define $\underset{\sim}{\Delta q}_\varepsilon(\underset{\sim}{x}) \in R^{(N+1)x(N+1)}$ by

$$\underset{\sim}{\Delta q}_\varepsilon(\underset{\sim}{x}) = \left(\begin{array}{c|c} \underset{\sim}{\Delta q}(\underset{\sim}{x}) & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \varepsilon \end{array}\right)$$

whence we have

$$\underset{\sim}{\Delta q}_{\varepsilon}(\underset{\sim}{x}) \cdot \underset{\sim}{\Delta q}_{\varepsilon}(\underset{\sim}{x}) = |\underset{\sim}{\Delta q}_{\varepsilon}(\underset{\sim}{x})|^2$$

$$= \underset{\sim}{\Delta q}(\underset{\sim}{x}) \cdot \underset{\sim}{\Delta q}(\underset{\sim}{x}) + \varepsilon^2 \quad .$$

Using the above result in (4.2) we obtain, at $\underset{\sim}{x} \in \Omega$ (it is understood that all functions appearing below are evaluated at $\underset{\sim}{x}$) ,

$$\hat{W}^p_\varepsilon(\underset{\sim}{\Delta q}) = \sqrt{2}k\sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q} + \varepsilon^2} + E_p(\underset{\sim}{q}^o \cdot \underset{\sim}{\Delta q})$$

$$= \sqrt{2}k|\underset{\sim}{\Delta q}_\varepsilon| + E_p(\underset{\sim}{q}^o \cdot \underset{\sim}{\Delta q}) \quad .$$

Hence,

$$\left|\hat{W}^p_\varepsilon(\underset{\sim}{\Delta p}) - \hat{W}^p_\varepsilon(\underset{\sim}{\Delta q})\right| = \sqrt{2}k\left||\underset{\sim}{\Delta p}_\varepsilon| - |\underset{\sim}{\Delta q}_\varepsilon|\right| + E_p\left|\underset{\sim}{p}^o \cdot (\underset{\sim}{\Delta p} - \underset{\sim}{\Delta q})\right|$$

$$< \sqrt{2}k\left|\underset{\sim}{\Delta p}_\varepsilon - \underset{\sim}{\Delta q}_\varepsilon\right| + E_p h_2|\underset{\sim}{\Delta p} - \underset{\sim}{\Delta q}| \quad ,$$

(using (1.8) and the Schwarz inequality)

$$= (\sqrt{2}k + E_p h_2)\ |\underset{\sim}{\Delta p} - \underset{\sim}{\Delta q}| \quad .$$

The remainder of the proof follows that of Lemma 3.3(b), whence it is easily shown that the constant β_ε is given by

$$\beta_\varepsilon = (\sqrt{2}k + E_p\ h_2)\ \text{mes}\ (\Omega) > 0 \quad . \ \square$$

Lemma 3.7

The functional $j_\varepsilon : L \to R$ is bounded below in the sense that there exists a positive constant ω such that

$$j_\varepsilon(\Delta\underset{\sim}{q}) \geq -\,\omega\|\bar{v}\|_{\bar{V}} \quad . \tag{4.8}$$

Proof

Subtracting (4.1) from (4.2) we have, at $\underset{\sim}{x} \in \Omega$,

$$\hat{W}^p_\varepsilon(\Delta\underset{\sim}{q}) - \hat{W}^p_2(\Delta\underset{\sim}{q}) = \sqrt{2}k\left(\sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q} + \varepsilon^2} - \sqrt{\Delta\underset{\sim}{q} \cdot \Delta\underset{\sim}{q}}\;\right)$$

$$> 0 \quad , \quad \text{provided } \varepsilon > 0 \quad .$$

Integrating both sides and using (2.18) and (4.3), we get

$$j_\varepsilon(\Delta\underset{\sim}{q}) - j(\Delta\underset{\sim}{q}) > 0 \quad , \quad \text{provided } \varepsilon > 0 \quad .$$

Hence,

$$j_\varepsilon(\Delta\underset{\sim}{q}) > j(\Delta\underset{\sim}{q}) \geq -\,\omega\|\bar{v}\|_{\bar{V}}$$

where we have used Lemma 3.4 to obtain the final result. □

The preceding two lemmas have established the necessary and sufficient conditions for $j_\varepsilon(\cdot)$ to constitute a suitable regularisation of $j(\cdot)$ for inclusion in a perturbed minimisation problem. In the following lemma we show that $j_\varepsilon(\underset{\sim}{q})$ indeed converges to $j(\underset{\sim}{q})$ as $\varepsilon \to 0$.

Lemma 3.8

The functional $j_\varepsilon(.)$ converges to $j(.)$ in the sense that, for $\varepsilon > 0$,

$$\left|j(\underset{\sim}{\Delta q}) - j_\varepsilon(\underset{\sim}{\Delta q})\right| \leqslant 2k \, \text{mes}(\Omega) \, \varepsilon \quad . \qquad (4.9)$$

Proof

Using the definition of $j(.)$ and $j_\varepsilon(.)$ given in (2.18) and (4.3) respectively, we have

$$\left|j_\varepsilon(\underset{\sim}{\Delta q}) - j(\underset{\sim}{\Delta q})\right| = \left|\sqrt{2}k\int_\Omega \left(\sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q} + \varepsilon^2} - \sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}\right) dx\right|$$

$$\leqslant \sqrt{2}k\int_\Omega \left(\sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q} + \varepsilon^2} - \sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}\right) dx$$

$$\leqslant \sqrt{2}k\int_\Omega \left(\sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}} + \varepsilon - \sqrt{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}\right) dx$$

$$= \sqrt{2}k \, \text{mes}(\Omega) \, \varepsilon$$

where we have used the inequality $\sqrt{a^2 + b^2} \leqslant \sqrt{a^2} + \sqrt{b^2}$ to establish the final result. □

It is convenient to introduce at this point the regularised functional $J_\varepsilon(\bar{v})$ which we define as

$$J_\varepsilon(\bar{v}) = \frac{1}{2}\, a(\bar{v},\bar{v}) + j_\varepsilon(\underset{\sim}{\Delta q}) - f(\bar{v}) \quad . \qquad (4.10)$$

An important observation concerning $J_\varepsilon(\bar{v})$ is that it is everywhere differentiable, as follows from the differentiability of $j_\varepsilon(\Delta \underset{\sim}{q})$. Once again we may make use of standard results from convex analysis to establish that a unique solution exists to the perturbed minimisation problem (defined below), and that this solution is characterised by a variational equality. Since these developments parallel closely those of the minimisation problem (M), as given in Theorems 3.2A, 3.2B, and 3.3, we simply summarise the pertinent results in the following theorem.

Theorem 3.4

For each $\varepsilon > 0$ there exists a unique solution $\bar{u}_\varepsilon \in \bar{V}$ of the perturbed minimisation problem

$$J_\varepsilon(\bar{u}_\varepsilon) \leqslant J_\varepsilon(\bar{v}) \ , \qquad \bar{v} \in \bar{V} \tag{4.11a}$$

characterised by the variational equality

$$\langle DJ_\varepsilon(\bar{u}_\varepsilon), \bar{v} \rangle = 0 \ , \qquad \bar{v} \in \bar{V} \ . \tag{4.11b}$$

We shall refer to the minimisation problem defined by (4.11a) as Problem (M_ε) . □

It is easy to show that (4.11b) is equivalent to

$$a(\bar{u}_\varepsilon, \bar{v}) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon), \Delta \underset{\sim}{q} \rangle - f(\bar{v}) = 0 \ , \qquad \bar{v} \in \bar{V} \ . \tag{4.12}$$

We confirm that the following requirements for Theorem 3.4 are also satisfied.

(i) According to Lemmas 3.1, 3.2 and 3.6, $J_\varepsilon(\bar{v})$ is a continuous functional on $\bar{V}$.

(ii) In part (iii) of the demonstration of conditions required for Theorem 3.2A we showed that $I(\bar{v}) = \frac{1}{2}\, a(\bar{v},\bar{v}) - f(\bar{v})$ was strictly convex; thus, since $j_\varepsilon(.)$ is convex (Lemma 3.6), it follows that $J_\varepsilon(\bar{v})$ is strictly convex.

(iii) In part (i) of the demonstration of conditions required for Theorem 3.2A we showed that $I(\bar{v})$ is Gateaux-differentiable; thus, since $j_\varepsilon(.)$ is Gateaux- differentiable (Lemma 3.6) it follows that $J_\varepsilon(\bar{v})$ is differentiable.

(iv) Finally, to show the coercivity of $J_\varepsilon(\bar{v})$ we use the results of Lemmas 3.1, 3.2 and 3.7 in (4.10) to obtain

$$J_\varepsilon(\bar{v}) \geqslant \frac{1}{2}\,\alpha\|\bar{v}\|^2_{\bar{V}} - \omega\|\bar{v}\|_{\bar{V}} - m\|\bar{v}\|_V$$

where ω, m are positive constants and α is a non-negative constant.

Dividing both sides by $\|\bar{v}\|_{\bar{V}}$ we get

$$\frac{J_\varepsilon(\bar{v})}{\|\bar{v}\|_{\bar{V}}} \geqslant \frac{1}{2}\,\alpha\|\bar{v}\|_{\bar{V}} - \omega - m$$

$$\rightarrow +\infty \quad \text{as } \|\bar{v}\|_{\bar{V}} \rightarrow \infty \ , \quad \text{iff } \alpha > 0 \ .$$

The reader is referred at this point to the remark which concludes the demonstration of the coercivity of $J(\bar{v})$ following Theorem 3.2B, concerning the constant α , from which we conclude that the coercivity result obtains if and only if the material exhibits hardening behaviour.

To complete the discussion of the perturbed minimisation problem we enquire whether the solution $\bar{u}_\varepsilon$ converges to the solution $\bar{u}$ of the original minimisation Problem (M). We treat this in the following theorem.

Theorem 3.5

Let $\bar{u}$ be the solution of Problem (M) , and $\bar{u}_\varepsilon$ the solution of Problem (M_ε) , for fixed $\varepsilon > 0$. Then there exists a positive constant C , independent of ε , such that

$$\|\bar{u}_\varepsilon - \bar{u}\|_{\bar{V}} \leqslant C\sqrt{\varepsilon} \quad . \tag{4.13}$$

Proof

From (2.34) we have

$$a(\bar{u},\bar{v} - \bar{u}) + j(\Delta \underset{\sim}{q}) - j(\Delta \underset{\sim}{p}) - f(\bar{v} - \bar{u}) \geqslant 0 \tag{4.14}$$

and from (4.12) we have

$$a(\bar{u}_\varepsilon,\bar{v}) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon),\Delta \underset{\sim}{q}\rangle - f(\bar{v}) = 0 \quad . \tag{4.15}$$

Setting $\bar{v} = \bar{u}_\varepsilon$ in (4.14) and $\bar{v} = \bar{u}_\varepsilon - \bar{u}$ in (4.15) and subtracting (4.14) from (4.15) we obtain

$$a(\bar{u}_\varepsilon - \bar{u}, \bar{u}_\varepsilon - \bar{u}) - j(\Delta \underset{\sim}{p}_\varepsilon) + j(\Delta \underset{\sim}{p}) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon), \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{p}\rangle \leq 0$$

$$\Rightarrow a(\bar{u}_\varepsilon - \bar{u}, \bar{u}_\varepsilon - \bar{u}) \leq j(\Delta \underset{\sim}{p}_\varepsilon) - j(\Delta \underset{\sim}{p}) - \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon), \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{p}\rangle \quad . \qquad (4.16)$$

From the convexity of $j_\varepsilon(\Delta \underset{\sim}{q})$ (Lemma 3.6), we have

$$- \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon), \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{p}) \leq j_\varepsilon(\Delta \underset{\sim}{p}) - j_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) \quad . \qquad (4.17)$$

Substituting (4.17) into (4.16), we get

$$a(\bar{u}_\varepsilon - \bar{u}, \bar{u}_\varepsilon - \bar{u}) \leq j(\Delta \underset{\sim}{p}_\varepsilon) - j(\Delta \underset{\sim}{p}) - j_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) + j_\varepsilon(\Delta \underset{\sim}{p}) \quad . \qquad (4.18)$$

From the $\bar{V}$-ellipticity of $a(\cdot,\cdot)$, (Lemma 3.1), we have

$$\alpha \|\bar{u}_\varepsilon - \bar{u}\|^2_{\bar{V}} \leq a(\bar{u}_\varepsilon - \bar{u}, \bar{u}_\varepsilon - \bar{u})$$

$$\leq j(\Delta \underset{\sim}{p}_\varepsilon) - j(\Delta \underset{\sim}{p}) - j_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) + j_\varepsilon(\Delta \underset{\sim}{p}) \text{ , using (4.18)}$$

$$\leq \left|j(\Delta \underset{\sim}{p}_\varepsilon) - j_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon)\right| + \left|j(\Delta \underset{\sim}{p}) - j_\varepsilon(\Delta \underset{\sim}{p})\right|$$

$$\leq 2k \text{ mes } (\Omega)\ \varepsilon + 2k \text{ mes } (\Omega)\ \varepsilon$$

using Lemma 3.8. Hence, dividing both sides by $\alpha > 0$, we get

$$\|\bar{u}_\varepsilon - \bar{u}\|^2_{\bar{V}} \leqslant 4 \frac{k}{\alpha} \text{ mes } (\Omega) \; \varepsilon$$

from which the result follows with $C = \sqrt{\frac{4k}{\alpha} \text{ mes } (\Omega)}$. □

The variational equality (4.12) which characterises the solution of the perturbed minimisation problem (M_ε) provides a suitable basis for obtaining an approximate solution using conventional finite element procedures. We are assured that the solution $\bar{u}_\varepsilon$ to Problem (M_ε) exists and is unique (provided the material hardens), and moreover, that the solution $\bar{u}_\varepsilon$ converges to the solution $\bar{u}$ of the original minimisation problem (M) as the regularisation parameter $\varepsilon \to 0$.

We now turn our attention to the construction of a suitable finite element approximation for this problem.

3.5 FINITE ELEMENT APPROXIMATIONS AND ERROR ESTIMATES

In the preceding section we derived a variational principle for the original boundary-value problem which provides a suitable basis for a numerical approximation of the problem. We propose now to describe the discrete approximation of the variational formulation using the Galerkin finite element method, following which we will derive an estimate for the error in this approximation.

We introduce two families of finite-dimensional subspaces, $\{V^h\}$ of V and $\{L^h\}$ of L , each member of which is spanned by a finite number of linearly independent basis functions, usually piecewise polynomials. Here $h \in (0,1]$ is a real parameter (such as the mesh size), so that for each h we identify a particular subspace from each of the families. We expect that as $h \to 0$ the corresponding finite-dimensional subspaces V^h and L^h approach the spaces V and L respectively, in some suitable sense.

Recalling Theorem 3.4, we pose the following problem on the above finite-dimensional subspaces : find $\bar{u}^h_\varepsilon \in \bar{V}^h = V^h \times L^h$, such that

$$a(\bar{u}^h_\varepsilon,\bar{v}^h) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}^h_\varepsilon),\Delta \underset{\sim}{q}^h\rangle - f(\bar{v}^h) = 0 \quad , \quad \bar{v} \in \bar{V}^h \quad . \tag{5.1}$$

The subspace $\bar{V}^h$ is a Hilbert space with the same inner product (.,.) as $\bar{V}$ itself. Thus, Theorem 3.4 applies also to the problem defined above and we are assured of the existence and uniqueness of the solution $\bar{u}^h_\varepsilon$ to (5.1).

The remainder of this section is concerned with finding an estimate of the error $\|\bar{u} - \bar{u}^h_\varepsilon\|$ due to both the regularisation and the finite element approximation. Before proceeding with this, however, we will make a brief excursion into interpolation theory in Sobolev spaces. We review the standard interpolation error estimates with a view to making slight modifications to these estimates in order to be able to use them in the present work. Finite element interpolation theory in Sobolev spaces is discussed in detail by CIARLET (1978), and ODEN and CAREY (1983), although the present discussion follows closely that of REDDY (1986).

Interpolation Error Estimates

We begin with a brief introduction to the notation to be used here : much of what we introduce here will be expanded upon in greater detail in Chapters 4 and 6.

Consider a finite element mesh in R^N , where each element Ω_e is generated from a single master element $\hat{\Omega}$ (Fig. 3.2). We assume that Ω_e and $\hat{\Omega}$ are affine-equivalent and that the family of elements $\{\Omega_1,\ldots\ldots,\Omega_E\}$ is thus affine. The relative size and shape of an arbitrary element Ω_e is quantified by defining the constants

$$h_e = \mathrm{diam}(\Omega_e) = \max\{ |\underset{\sim}{x} - \underset{\sim}{y}| \ , \ \underset{\sim}{x},\underset{\sim}{y} \in \Omega_e\} \tag{5.2}$$

and

$$\rho_e = \sup\{\text{diameters of all spheres contained in } \Omega_e\} \ . \tag{5.3}$$

An affine family of elements is said to be regular if

(i) there exists a constant γ such that $h_e/\rho_e \leq \gamma$ for all elements Ω_e , and

(ii) the diameters h_e approach zero.

Let $q \in C(\Omega_e)$ be a continuous scalar function on Ω_e , and let K_e be an operator from $C(\Omega_e)$ to $C(\hat{\Omega})$ which maps q to a function $\hat{q}$ being defined by

$$K_e : C(\Omega_e) \to C(\hat{\Omega}) \quad , \quad K_e q = \hat{q} \quad , \quad \hat{q}(\hat{\underset{\sim}{x}}) = q(\underset{\sim}{x})$$

Affine transformation F_e

$F_e : \hat{\Omega} \to \Omega_e$, $\underset{\sim}{x} = F_e(\underset{\sim}{\hat{x}}) = \underset{\sim}{T}_e \underset{\sim}{\hat{x}} + \underset{\sim}{b}$

for $\underset{\sim}{T}_e$ a constant matrix and $\underset{\sim}{b}$

a constant vector.

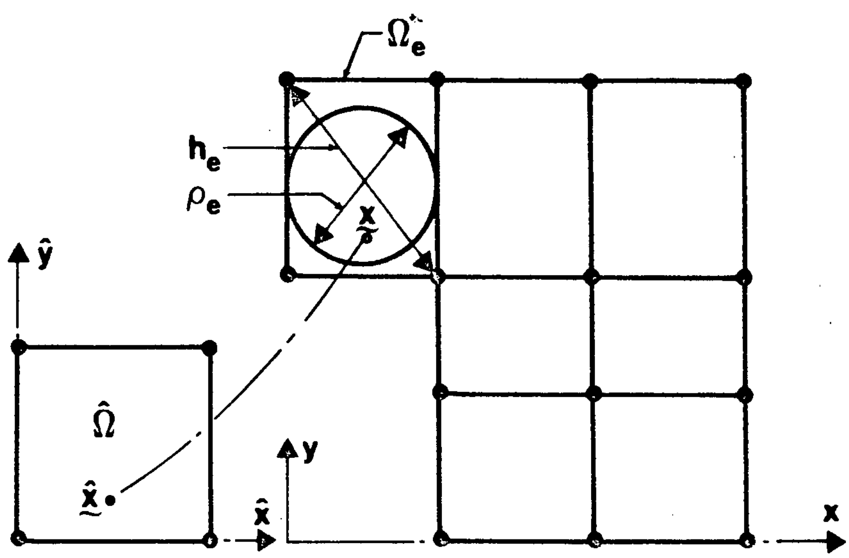


Figure 3.2 Finite element mesh in R^2, generated from $\hat{\Omega}$ using the affine transformation F_e .

where $\underset{\sim}{x} = F_e(\underset{\sim}{\hat{x}})$. Further, let $\hat{X}$ be a finite-dimensional subspace, spanned by local basis functions $\{\hat{\phi}_i\}_{i=1}^{N_e}$, containing polynomials of degree $\leqslant k$, so that $P_k(\hat{\Omega}) \subseteq \hat{X}$; here N_e is the number of nodes on Ω_e , and P_k is the set of polynomials of degree $\leqslant k$. Similarly, let X_e be a finite-dimensional subspace, spanned by local basis functions $\{\psi_i^{(e)}\}_{i=1}^{N_e}$, where the latter are mapped from $\{\hat{\psi}_i\}$ using the inverse operator K_e^{-1} : $C(\hat{\Omega}) \to C(\Omega_e)$, $K_e^{-1}\,\hat{\psi}_i = \psi_i^{(e)}$. We now construct a projection operator $\hat{\Pi}$ which maps any $\hat{q} \in C(\hat{\Omega})$ to its interpolate $\hat{\Pi}\hat{q}$ in $\hat{X}$:

$$\hat{\Pi} : C(\hat{\Omega}) \to \hat{X} \quad , \quad \hat{\Pi}\hat{q} = \sum_{j=1}^{N} \hat{q}(\hat{\underset{\sim}{x}}_j)\,\hat{\psi}_j \quad . \tag{5.4}$$

Similarly, we construct a projection operator Π_e which maps any $q \in C(\Omega_e)$ to its interpolate $\Pi_e q$ in X_e :

$$\Pi_e : C(\Omega_e) \to X_e \quad , \quad \Pi_e q = \sum_{j=1}^{N} q(\underset{\sim}{x}_j)\,\psi_j^{(e)} \quad . \tag{5.5}$$

Our objective now is to derive a local interpolation error estimate in the H^m-norm for a function q which is smooth enough to be in $H^{k+1}(\Omega_e)$. Thus, we require that Π_e maps members of $H^{k+1}(\Omega_e)$ into $H^m(\Omega_e)$, with the range of Π_e being $X_e \subset H^m(\Omega_e)$; that is

$$\Pi_e : H^{k+1}(\Omega_e) \to H^m(\Omega_e) \quad , \quad R(\Pi_e) = X_e \quad ; \tag{5.6}$$

since $P_k(\Omega_e) \subseteq X_e$ we have

$$\Pi_e q = q \quad , \quad \forall\, q \in P_k(\Omega_e) \quad ,$$

and Π_e is a projection operator.

We now state the standard local interpolation error estimate. Note that the references cited with Theorems 3.6 and 3.7 are for the convenience of the reader.

<u>Theorem 3.6</u> [CIARLET (1978), Theorem 3.1.6, or REDDY (1986), Section 45, Theorem 3]

Let k and m be non-negative integers such that

$$H^{k+1}(\hat{\Omega}) \subset C(\hat{\Omega}) \tag{5.7}$$

$$H^{k+1}(\hat{\Omega}) \subset H^{m}(\hat{\Omega}) \tag{5.8}$$

$$P_k(\hat{\Omega}) \subset \hat{X} \subset H^{m}(\hat{\Omega}) \quad . \tag{5.9}$$

Let $\hat{\Pi}$ and Π_e be the projection operators defined in (5.4) and (5.5), and let $\{\Omega_1,\ldots\ldots,\Omega_E\}$ be a regular family of finite elements. Then there exists a constant c such that for any Ω_e in this family and all functions $q \in H^{k+1}(\Omega_e)$,

$$\| q - \Pi_e q \|_{m,\Omega_e} \leq c h_e^{k+1-m} |q|_{k+1,\Omega_e} \tag{5.10}$$

where $|\cdot|_{k+1,\Omega_e}$ denotes the Sobolev semi-norm :

$$|q|^2_{k+1,\Omega_e} = \sum_{|\alpha| = k+1} \int_{\Omega_e} (D^\alpha q(\underset{\sim}{x}))^2 \, dx \quad . \square$$

We turn our attention now to the global interpolation of a function $q \in C(\Omega)$ defined on the entire domain Ω . We define a finite-dimensional subspace X^h , spanned by global polynomial basis functions ϕ_i (which are constructed from the local basis functions $\phi_i^{(e)}$) , and construct a projection operator which maps q to its interpolate $\Pi_h q$, or $\tilde{q}^h$:

$$\Pi_h : C(\Omega) \to X^h \quad , \quad \Pi_h q = \sum_{i=1}^{M} q(\underset{\sim}{x}_i)\, \phi_i(\underset{\sim}{x}) = \tilde{q}^h \tag{5.11}$$

where M is the number of nodes in Ω . Also, since we are now dealing with global quantities we define the mesh parameter h by

$$h = \max_{1 \leq e \leq E} \{h_e\} \quad . \tag{5.12}$$

We now state the global interpolation error estimate in the following theorem.

Theorem 3.7 [CIARLET (1978), Theorem 3.2.1, or REDDY (1986), Section 46, Theorem 1]

Assume that all the conditions of Theorem 3.6 hold. Then there exists a constant C , independent of h , such that for any $q \in H^{k+1}(\Omega)$

$$\| q - \tilde{q}^h \|_{m,\Omega} \leq C h^{k+1-m} |q|_{k+1,\Omega} \quad , \quad m=0 \text{ or } m=1 \quad . \square \tag{5.13}$$

We now wish to examine the situation where $\Omega \in R^N$ and $m = k = 0$. Returning to Theorem 3.6 we see that condition (5.7) cannot be satisfied for $N > 1$, since by the Sobolev Embedding Theorem (CIARLET (1978), Section 3.1) we can only guarantee that for $\hat{\Omega} \in R^N$

$$H^{k+1}(\hat{\Omega}) \subset C(\hat{\Omega}) \quad \text{if} \quad k+1 > \frac{N}{2} \quad . \tag{5.14}$$

We therefore replace eqn (5.7) by the weaker requirement that $q \in H^{k+1}(\Omega_e) \cap C(\Omega_e)$; of course, if $k \geq N/2$ then eqn (5.7) holds anyway. Furthermore, for $k = 0$ we define the operator Π_e by

$$\Pi_e q(\underset{\sim}{x}) = \min_{\underset{\sim}{x} \in \Omega_e} q(\underset{\sim}{x}) \quad . \tag{5.15}$$

It is easy to show that Π_e is a projection with

$$\Pi_e q = q \quad , \qquad q \in P_o(\Omega_e); \tag{5.16}$$

Also, Π_e is bounded from $H^{k+1}(\Omega_e) \to H^m(\Omega_e)$ and the result of Theorem 3.6 remains unchanged. We summarise the above modifications to Theorem 3.6 in the following theorem.

<u>Theorem 3.8</u>

Let the conditions of Theorem 3.6 hold, but with the following modifications : (i) replace eqn (5.7) by the requirement that $q \in H^{k+1}(\Omega_e) \cap C(\Omega_e)$, and (ii) for $k = 0$, define Π_e by eqn (5.15) . Then there exists a positive constant c such that

$$\|q - \Pi_e q\|_{o,\Omega_e} \leqslant c h_e^{k+1} |q|_{k+1,\Omega_e} \quad . \; \square \tag{5.17}$$

We now make the same modifications to Theorem 3.7 as were made to Theorem 3.6. We define the operator Π_h by

$$\Pi_h q\big|_{\Omega_e} = \Pi_e q = \tilde{q}^h(\underset{\sim}{x}) \quad . \tag{5.18}$$

Again, it is easy to show that Π_h is a projection operator with

$$\Pi_h q^* = q^* \quad , \qquad q^* \in P_o(\Omega) \quad . \tag{5.19}$$

The modifications to Theorem 3.7 are contained in the following theorem.

Theorem 3.9

Assume that all the conditions of Theorem 3.8 hold. Then there exists a constant C , independent of h , such that for any $q \in H^{k+1}(\Omega)$,

$$\|q - \tilde{q}^h\|_{o,\Omega} \leq Ch^{k+1}|q|_{k+1,\Omega} \quad . \qquad (5.20)$$

Proof

Making use of (5.18) we have

$$\|q - \tilde{q}^h\|_{o,\Omega} = [\sum_{e=1}^{E} \|q - \Pi_e q\|^2_{o,\Omega_e}]^{1/2}$$

$$\leq [\sum_{e=1}^{E} c^2 h^{2(k+1)} |q|^2_{k+1,\Omega_e}]^{1/2} \quad , \quad \text{using Theorem 3.8}$$

$$\leq ch^{k+1} [\sum_{e=1}^{E} |q|^2_{k+1,\Omega_e}]^{1/2}$$

$$= ch^{k+1}|q|_{k+1,\Omega} \quad . \ \square$$

This completes our discussion of the interpolation error estimates. We now make use of these results in deriving an error estimate for the incremental holonomic problem.

Error Estimates for the Incremental Holonomic Problem

Let V^h be the space spanned by functions whose restrictions to each element Ω_e contain complete polynomials of degree 1 (for example, the 3-

noded triangle or 4-noded quadrilateral in R^2). Suppose further that the solution $\Delta \underset{\sim}{u}$ is smooth enough to belong to $H^2(\Omega)$. Then the following interpolation error estimate holds (Theorem 3.7) :

$$\|\Delta \underset{\sim}{u} - \widetilde{\Delta \underset{\sim}{u}}^h\|_{1,\Omega} \leq C_2 h \|\Delta \underset{\sim}{u}\|_{2,\Omega} \quad , \qquad \widetilde{\Delta \underset{\sim}{u}}^h \in V^h \tag{5.21}$$

where $\widetilde{\Delta \underset{\sim}{u}}^h$ is the interpolate of $\Delta \underset{\sim}{u}$ on Ω .

Let L^h be the space spanned by functions whose restrictions to each element Ω_e are constant. Further, let us define the interpolate $\widetilde{\Delta \underset{\sim}{p}}^h \in L^h$ of $\Delta \underset{\sim}{p}$ by the step-function

$$\widetilde{\Delta \underset{\sim}{p}}^h \Big|_{\Omega_e} = \min_{\underset{\sim}{x} \in \Omega_e} \{\Delta \underset{\sim}{p}(\underset{\sim}{x})\} \quad . \tag{5.22}$$

Then the following interpolation error estimate follows from Theorem 3.9 : assuming $\Delta \underset{\sim}{p} \in H^1(\Omega)$,

$$\|\Delta \underset{\sim}{p} - \widetilde{\Delta \underset{\sim}{p}}^h\|_{o,\Omega} \leq C_1 h \|\Delta \underset{\sim}{p}\|_{1,\Omega} \quad . \tag{5.23}$$

In the absence of an established regularity result for the solution $\bar{u}_\varepsilon = (\Delta \underset{\sim}{u}_\varepsilon \ , \ \Delta \underset{\sim}{p}_\varepsilon)$ we will make the following assumption. Let the data f in eqn (5.1) be given in $(L_2(\Omega))^N$, for $\Omega \in R^N$: then the solution $\bar{u}_\varepsilon$ of eqn (5.1) belongs to $(H^2(\Omega))^N \times (H^1(\Omega))^N$ and satisfies

$$\|\Delta \underset{\sim}{u}_\varepsilon\|_2 + \|\Delta \underset{\sim}{p}_\varepsilon\|_1 \leq \bar{C}\|f\|_o \tag{5.24}$$

where $\|\cdot\|_o$ denotes the L_2-norm.

Before proceeding with the derivation of the estimate $\|\bar{u}_\varepsilon - \bar{u}^h_\varepsilon\|$ we need to establish a preliminary result, as given in the following lemma.

Lemma 3.9

The operator $Dj_\varepsilon : L \to L'(=L)$, as defined in Lemma 3.6(a), satisfies a Lipschitz condition; that is

$$\|Dj_\varepsilon(\underset{\sim}{\Delta p}) - Dj_\varepsilon(\underset{\sim}{\Delta q})\|_L \leq M' \|\underset{\sim}{\Delta p} - \underset{\sim}{\Delta q}\|_L \tag{5.25}$$

where M' is a positive constant.

Proof

Let $Dj_\varepsilon \equiv f : L \to L'$, where $f(\underset{\sim}{\Delta p}) : L \to R$ is defined by

$$f(\underset{\sim}{\Delta p}) = \left\{ \frac{\sqrt{2}k\underset{\sim}{\Delta p}}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^{1/2}} + E_p \underset{\sim}{p}^o \right\} \in L \tag{5.26}$$

(see the proof of Lemma 3.6(a)).

Following Lemma 3A.1 (see Appendix 3A, following this chapter) we need to show that the Gateaux-differential $df(\underset{\sim}{\Delta p},\underset{\sim}{\Delta q})$, defined by

$$df(\underset{\sim}{\Delta p},\underset{\sim}{\Delta q}) = \lim_{\theta \to 0} \frac{1}{\theta} [f(\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) - f(\underset{\sim}{\Delta p})]$$

defines a bounded linear operator $Df(\underset{\sim}{\Delta p})$ on $\underset{\sim}{\Delta q}$.

(i) First we show that $df(\underset{\sim}{\Delta p}, \underset{\sim}{\Delta q})$ is linear in $\underset{\sim}{\Delta q}$. From (5.26) we have

$$df(\underset{\sim}{\Delta p},\underset{\sim}{\Delta q}) = \frac{d}{d\theta}\left[\frac{\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}}{((\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) \cdot (\underset{\sim}{\Delta p} + \theta\underset{\sim}{\Delta q}) + \varepsilon^2)^{1/2}} + E_p\, p^o\right]\Bigg|_{\theta=0}$$

$$= \underset{\sim}{\Delta q}(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^{-1/2} - \frac{\underset{\sim}{\Delta p}(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q})}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^{3/2}}$$

$$= \frac{\underset{\sim}{\Delta q}(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^3) - \underset{\sim}{\Delta p}(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q})}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^{3/2}}$$

$$= \frac{\underset{\sim}{\Delta q}}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^{1/2}} - \frac{\underset{\sim}{\Delta p}(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q})}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^{3/2}} \qquad (5.27)$$

whence it is clear that $df(\underset{\sim}{\Delta p}, \underset{\sim}{\Delta q})$ is linear in $\underset{\sim}{\Delta q}$.

(ii) Next we show that $df(\underset{\sim}{\Delta p},.)$ is a bounded operator. From (5.27) we have

$$\| df(\underset{\sim}{\Delta p}, \underset{\sim}{\Delta q})\|_L^2 = \int_\Omega df(\underset{\sim}{\Delta p}, \underset{\sim}{\Delta q}) \cdot df(\underset{\sim}{\Delta p}, \underset{\sim}{\Delta q})\, dx$$

$$= \int_\Omega \frac{\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q}}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)} + \frac{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p})(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q})^2}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^3} - \frac{2(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q})^2}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^2}\, dx$$

$$= \int_\Omega \frac{(\underset{\sim}{\Delta q} \cdot \underset{\sim}{\Delta q})[(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p})^2 + 2\varepsilon^2(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p}) + \varepsilon^4] + (\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p})(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta q})^2}{(\underset{\sim}{\Delta p} \cdot \underset{\sim}{\Delta p} + \varepsilon^2)^3}\, dx$$

$$- \int_\Omega \frac{2(\Delta\tilde{p} \cdot \Delta\tilde{q})^2[\Delta\tilde{p} \cdot \Delta\tilde{p} + \varepsilon^2]}{(\Delta\tilde{p} \cdot \Delta\tilde{p} + \varepsilon^2)^3}\, dx$$

$$\leqslant \int_\Omega \frac{\varepsilon^2(\Delta\tilde{q} \cdot \Delta\tilde{q})[2(\Delta\tilde{p} \cdot \Delta\tilde{p}) + \varepsilon^2] + (\Delta\tilde{q} \cdot \Delta\tilde{q})(\Delta\tilde{p} \cdot \Delta\tilde{p})^2}{(\Delta\tilde{p} \cdot \Delta\tilde{p} + \varepsilon^2)^3}\, dx$$

$$\leqslant \int_\Omega \frac{\varepsilon^2(\Delta\tilde{q} \cdot \Delta\tilde{q})[2(\Delta\tilde{p} \cdot \Delta\tilde{p}) + \varepsilon^2]}{(\Delta\tilde{p} \cdot \Delta\tilde{p} + \varepsilon^2)^3} + \frac{\Delta\tilde{q} \cdot \Delta\tilde{q}}{\varepsilon^2}\, dx$$

since $(\Delta\tilde{p} \cdot \Delta\tilde{p})^2/(\Delta\tilde{p} \cdot \Delta\tilde{p} + \varepsilon^2)^2 \leqslant 1$, and $1/(\Delta\tilde{p} \cdot \Delta\tilde{p} + \varepsilon^2) \leqslant 1$.
Hence,

$$\|df(\Delta\tilde{p}\ , \Delta\tilde{q})\|_L^2 \leqslant \int_\Omega 2(\Delta\tilde{q} \cdot \Delta\tilde{q})\left[\frac{\varepsilon^2}{(\Delta\tilde{p} \cdot \Delta\tilde{p} + \varepsilon^2)^2} + \frac{1}{2\varepsilon^2}\right] dx$$

$$\leqslant \int_\Omega 2(\Delta\tilde{q} \cdot \Delta\tilde{q})\left[\frac{1}{\varepsilon^2} + \frac{1}{2\varepsilon^2}\right] dx$$

$$= \int_\Omega \frac{3}{\varepsilon^2} (\Delta\tilde{q} \cdot \Delta\tilde{q})\, dx$$

$$= \frac{3}{\varepsilon^2} \|\Delta\tilde{q}\|_L^2 \qquad (5.28)$$

and so $df(\Delta\tilde{p},.) \equiv Df(\Delta\tilde{p})$ is a bounded operator on L , and the operator Dj_ε satisfies a Lipschitz condition. □

<u>Theorem 3.10</u>

Let V^h and L^h be as defined above and assume that $f \in (L_2(\Omega))^N$ and $\bar{u}_\varepsilon \in (H^2(\Omega))^N \times (H^1(\Omega))^N$, so that the interpolation error estimates

(5.21) and (5.23) hold. Then there exists a constant $\hat{C} > 0$, independent of h, such that

$$\|\bar{u}_\varepsilon - \bar{u}^h_\varepsilon\|_{\bar{V}} \leq \hat{C}h\|f\|_0 \quad . \tag{5.29}$$

<u>Proof</u>

From Theorem 3.4 we have

$$a(\bar{u}_\varepsilon,\bar{v}) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) , \Delta \underset{\sim}{q}\rangle - f(\bar{v}) = 0 \quad . \tag{5.30}$$

The counterpart of this result on the finite-dimensional subspace $\bar{V}^h$ is, from eqn (5.1),

$$a(\bar{u}^h_\varepsilon,\bar{v}^h) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}^h_\varepsilon) , \Delta \underset{\sim}{q}^h\rangle - f(\bar{v}^h) = 0 \quad . \tag{5.31}$$

Setting $\bar{v} = \bar{v}^h$ in (5.30) and subtracting (5.31) from (5.30) we get

$$a(\bar{u}_\varepsilon - \bar{u}^h_\varepsilon , \bar{v}^h) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) - Dj_\varepsilon(\Delta \underset{\sim}{p}^h_\varepsilon) , \Delta \underset{\sim}{q}^h\rangle = 0 \quad . \tag{5.32}$$

Now consider the following identities :

$$a(\bar{u}_\varepsilon - \bar{u}^h_\varepsilon , \bar{u}_\varepsilon - \bar{u}^h_\varepsilon) = a(\bar{u}_\varepsilon - \bar{u}^h_\varepsilon , \bar{u}_\varepsilon - \bar{v}^h) + a(\bar{u}_\varepsilon - \bar{u}^h_\varepsilon , \bar{v}^h - \bar{u}^h_\varepsilon) \tag{5.33}$$

and

$$\langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) - Dj_\varepsilon(\Delta \underset{\sim}{p}^h_\varepsilon),\Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{p}^h_\varepsilon \rangle = \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) - Dj_\varepsilon(\Delta \underset{\sim}{p}^h_\varepsilon) , \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{q}^h\rangle$$

$$+ \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) - Dj_\varepsilon(\Delta \underset{\sim}{p}^h_\varepsilon) , \Delta \underset{\sim}{q}^h - \Delta \underset{\sim}{p}^h_\varepsilon \rangle \quad . \tag{5.34}$$

Adding (5.33) and (5.34) and using (5.32), with $\bar{v}^h$ replaced by $\bar{v}^h - \bar{u}_\varepsilon^h$, on the right-hand-side of the result, we get

$$a(\bar{u}_\varepsilon - \bar{u}_\varepsilon^h, \bar{u}_\varepsilon - \bar{u}_\varepsilon^h) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) - Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon^h), \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{p}_\varepsilon^h \rangle$$

$$= a(\bar{u}_\varepsilon - \bar{u}_\varepsilon^h , \bar{u}_\varepsilon - \bar{v}^h) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) - Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon^h) , \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{q}^h \rangle$$

$$\leq M \| \bar{u}_\varepsilon - \bar{u}_\varepsilon^h \|_{\bar{V}} \| \bar{u}_\varepsilon - \bar{v}^h \|_{\bar{V}}$$

$$+ \bar{M} \| Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon) - Dj_\varepsilon(\Delta \underset{\sim}{p}_\varepsilon^h) \|_o \| \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{q}^h \|_o ,$$

(using Lemma 3.1 for the first term and Lemma 3.6(a) for the second term)

$$\leq M \| \bar{u}_\varepsilon - \bar{u}_\varepsilon^h \|_{\bar{V}} \| \bar{u}_\varepsilon - \bar{v}^h \|_{\bar{V}}$$

$$+ \bar{M} M' \| \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{p}_\varepsilon^h \|_o \| \Delta \underset{\sim}{p}_\varepsilon - \Delta \underset{\sim}{q}^h \|_o ,$$

(using Lemma 3.9 for the second term)

$$\leq (M + \bar{M} M') \| \bar{u}_\varepsilon - \bar{u}_\varepsilon^h \|_{\bar{V}} \| \bar{u}_\varepsilon - \bar{v}^h \|_{\bar{V}}$$

$$= \hat{M} \| \bar{u}_\varepsilon - \bar{u}_\varepsilon^h \|_{\bar{V}} \| \bar{u}_\varepsilon - \bar{v}^h \|_{\bar{V}} . \qquad (5.35)$$

We note that from Lemma 3.6 j_ε is convex, and so the differential Dj_ε is monotone, from which it follows that the duality pairing on the left-hand-side of (5.35) is non-negative. Using this result and the $\bar{V}$-

ellipticity of $a(\cdot,\cdot)$ (Lemma 3.1) on the left-hand-side of (5.35) we obtain

$$\alpha \, \|\bar{u}_\varepsilon - \bar{u}^h_\varepsilon\|^2_{\bar{V}} \leq \hat{M} \, \|\bar{u}_\varepsilon - \bar{u}^h_\varepsilon\|_{\bar{V}} \, \|\bar{u}_\varepsilon - \bar{v}^h\|_{\bar{V}} \quad . \tag{5.36}$$

Setting $\bar{v}^h = \tilde{u}^h$ in (5.36), where $\tilde{u}^h$ is the interpolate of $\bar{u}_\varepsilon$, we obtain

$$\|\bar{u}_\varepsilon - \bar{u}^h_\varepsilon\|_{\bar{V}} \leq \frac{\hat{M}}{\alpha} \|\bar{u}_\varepsilon - \tilde{u}^h\|_{\bar{V}}$$

$$= \frac{\hat{M}}{\alpha} \left[\|\Delta \underset{\sim}{u}_\varepsilon - \tilde{\underset{\sim}{u}}^h\|^2_1 + \|\Delta \underset{\sim}{p}_\varepsilon - \tilde{\underset{\sim}{p}}^h\|^2_0 \right]^{1/2} \tag{5.37}$$

where, again, the superposed tildas denote the relevant interpolates. Making use of the interpolation error estimates (5.21) and (5.23) in (5.37) we have

$$\|\bar{u}_\varepsilon - \bar{u}^h_\varepsilon\|_{\bar{V}} \leq \frac{\hat{M}}{\alpha} C_3 h \left[\|\Delta \underset{\sim}{u}_\varepsilon\|^2_2 + \|\Delta \underset{\sim}{p}_\varepsilon\|^2_1 \right]^{1/2}$$

$$\leq \frac{\hat{M}}{\alpha} C_3 h \left[\|\Delta \underset{\sim}{u}_\varepsilon\|_2 + \|\Delta \underset{\sim}{p}_\varepsilon\|_1 \right] \tag{5.38}$$

where $C_3 = \max(C_1, C_2)$. Finally, using the regularity assumption of eqn (5.24) in (5.38) we obtain

$$\|\bar{u}_\varepsilon - \bar{u}^h_\varepsilon\|_{\bar{V}} \leq \frac{\hat{M}}{\alpha} C_3 \bar{C} \, h \, \|f\|_o$$

which is the required result with $\hat{C} = \hat{M} C_3 \bar{C} / \alpha$. □

Since we already have an estimate for $\|\bar{u} - \bar{u}_\varepsilon\|$ from Theorem 3.5, the final error estimate follows as a trivial consequence of the application of the triangle inequality together with the latter result and that of Theorem 3.10. We summarise the final result in the following theorem.

Theorem 3.11

Let the conditions of Theorems 3.5 and 3.10 hold. Then we have

$$\|\bar{u} - \bar{u}_\varepsilon^h\|_{\bar{V}} \leq \hat{C}h\|f\|_o + C\sqrt{\varepsilon} \quad . \quad \square \tag{5.40}$$

APPENDIX 3A

DIFFERENTIATION OF OPERATORS AND FUNCTIONALS

We follow here the work of ODEN and REDDY (1976), and VAINBERG (1964).

Let U and V denote normed linear spaces and let P be an operator from U to V . Let S be a convex subset of U and u an arbitrary element of U , and let $\theta \in [0,1]$. Then we define the Gateaux-differential of the operator $P : S \subset U \to V$ at u in the direction η as the limit

$$dP(u,\eta) = \lim_{\theta \to 0} \frac{1}{\theta} [P(u + \theta\eta) - P(u)] \quad .$$

When $dP(u,\eta)$ is linear and continuous in η it defines for each u a bounded linear operator on U denoted DP(u) which we refer to as the Gateaux derivative of P at u , and $DP(u)\eta = dP(u,\eta)$. When this is the case, P is said to be Gateaux-differentiable (or simply differentiable) at u .

If V = R then DP(u) is a bounded linear functional on U and we write

$$DP(u)\eta \equiv \langle DP(u),\eta \rangle$$

where $\langle \cdot,\cdot \rangle : U' \times U \to R$ denotes duality pairing and U' is the dual space of U . In this case we have $dP(u,\eta) \equiv DP(u)\eta \equiv \langle DP(u),\eta \rangle$, which we refer to as the Gateaux derivative of P at u in the direction η .

The following lemma is required in Lemma 3.9.

Lemma 3A.1

Let U and V denote normed linear spaces and let P be an operator from U to V. If P has a Gateaux derivative DP(u) at each point u in U , then P is Lipschitz-continuous, that is

$$\|P(v) - P(u)\|_V \leqslant K \|v - u\|_U \quad , \qquad v \in U$$

where K is a positive constant.

The above lemma has been given in a more general form by VAINBERG (1964), Lemma 3.3, page 37, to which the reader may refer for the proof.

CHAPTER 4

NUMERICAL SOLUTION OF THE INCREMENTAL HOLONOMIC PROBLEM

In Chapter 3 we developed a variational formulation of the incremental holonomic problem in the form of a perturbed minimisation problem and showed that this problem constituted a suitable basis for the numerical approximation of the original boundary-value problem. In this chapter we wish to continue with the development of the numerical approximation using the Galerkin finite element method.

We continue in Section 4.1 with the description of the discrete approximation of the perturbed minimisation problem begun in Section 3.5. This leads to a system of nonlinear algebraic equations, the solution of which we describe in Section 4.2. In Section 4.3 we provide some useful physical insight into the solution procedure, and end the chapter with some computational details and a summary in Section 4.4.

4.1 DISCRETE APPROXIMATION OF THE INCREMENTAL HOLONOMIC PROBLEM

In Section 3.5 we wrote down the global approximation of the variational problem on the finite-dimensional subspace $\bar{V}^h$, where $\bar{V}^h$ was assumed to be spanned by a finite number of piecewise polynomial global basis functions. It is well known that these global basis functions can be constructed from local basis functions defined on each element. Accordingly, we partition the domain Ω, assumed to be polygonal, into a finite number E of subdomains Ω_e (Fig. 4.1) such that

$$\bar{\Omega} = \bigcup_{e=1}^{E} \bar{\Omega}_e \quad , \quad \Omega_e \cap \Omega_f = \emptyset \quad , e \neq f \quad .$$

We then refer to the connected set $\{\Omega_e\}_{e=1}^{E}$ as the finite element mesh for a particular value of the mesh parameter h , defined by

$$h = \max_{1 \leq e \leq E} \{h_e\} \quad , \quad h_e = \mathrm{dia}(\Omega_e) \quad .$$

The discrete global approximation of the variational problem (Section 3.5, eqn (5.1)) takes the following form : find $\bar{u}^h_\varepsilon = (\Delta \underset{\sim}{u}^h_\varepsilon, \Delta \underset{\sim}{p}^h_\varepsilon)$ such that

$$a(\bar{u}^h_\varepsilon, \bar{v}^h) + \langle Dj_\varepsilon(\Delta \underset{\sim}{p}^h_\varepsilon), \Delta \underset{\sim}{q}^h \rangle - f(\bar{v}^h) = 0 \ , \qquad \bar{v}^h \in \bar{V}^h \quad . \tag{1.1}$$

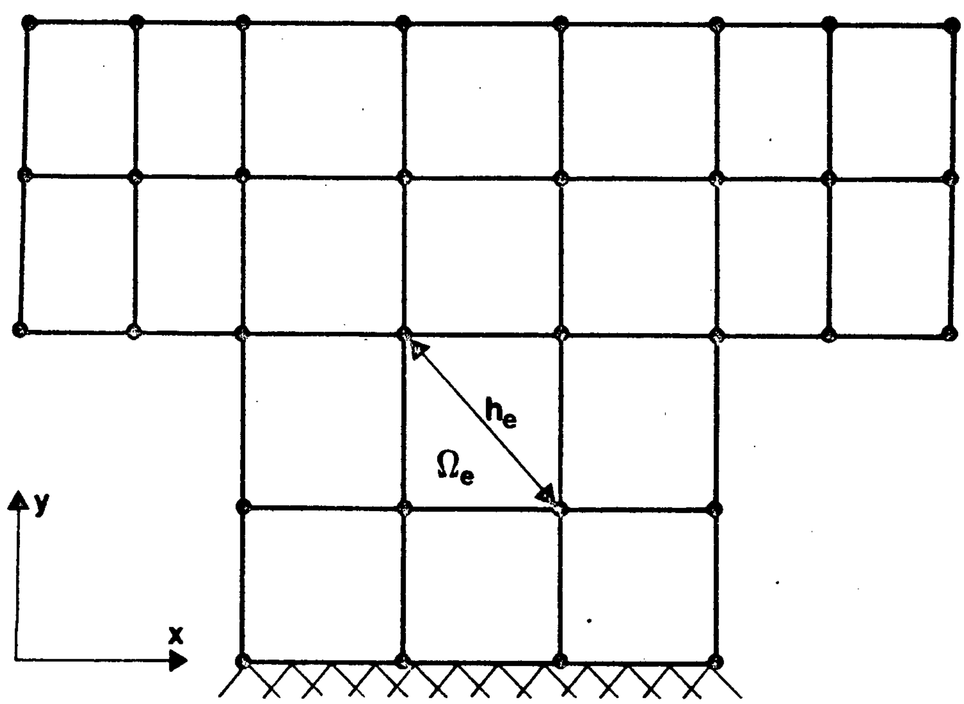


Figure 4.1 A typical finite element mesh in R^2 .

Hereafter we drop the subscript ε for brevity, it being then understood that we are dealing throughout with the approximation of the perturbed minimisation problem.

The global approximations of the bilinear form and functionals which appear in (1.1) are constructed by adding contributions from each element. Thus, noting that

$$\int_{\Omega} \ldots = \sum_{e=1}^{E} \int_{\Omega_e} \ldots$$

we may rewrite (1.1) in the following form :

$$\sum_{e=1}^{E} a^{(e)}(\bar{u}_e^h, \bar{v}_e^h) + \sum_{e=1}^{E} \langle Dj_\varepsilon(\Delta \underset{\sim}{p}_e^h), \Delta \underset{\sim}{q}_e^h \rangle \Big|_{\Omega_e} - \sum_{e=1}^{E} f^{(e)}(\bar{v}_e^h) = 0 \tag{1.2}$$

where (e) and e denote restrictions to Ω_e, with $(\cdot)_e$ being interpreted in the sense $(\cdot)_e = (\cdot)|_{\Omega_e}$; likewise $\langle \cdot, \cdot \rangle|_{\Omega_e}$ denotes the restriction of the duality pairing to Ω_e. With the problem defined in this form we may proceed with the approximation at the level of the individual elements.

Let $\{\phi_i\}_{i=1}^{N_e}$ be a suitable family of local basis functions defined on Ω_e, and having the property

$$\phi_i(\underset{\sim}{x}_j) = \delta_{ij} \quad , \quad 1 \leq i,j \leq N_e \tag{1.3}$$

where $\underset{\sim}{x}_j$ is the position vector of the j-th node on Ω_e and N_e is the number of nodes on Ω_e. Then the restriction of a typical element Δu_e^h

of the finite-dimensional space $H^h \subset H^1(\Omega_e)$ of displacements* to Ω_e may be approximated by

$$\Delta u_e^h(\underset{\sim}{x}) = \sum_{i=1}^{N_e} \Delta a_i \phi_i(\underset{\sim}{x}) \tag{1.4}$$

where $\Delta a_i = \Delta u_e^h(\underset{\sim}{x}_i)$ is the value of Δu_e^h at node i on Ω_e . If $\Delta \underset{\sim}{u}_e^h \in V^h$ is a vector whose components are approximated as in (1.4) then we may write the approximation of this vector in matrix form as

$$\Delta \underset{\sim}{u}_e^h(\underset{\sim}{x}) = \underset{\sim}{\Psi}(\underset{\sim}{x}) \Delta \underset{\sim}{a}_e \quad ; \tag{1.5}$$

similarly, an arbtirary vector $\Delta \underset{\sim}{v}_e^h \in V^h$ may be approximated by

$$\Delta \underset{\sim}{v}_e^h(\underset{\sim}{x}) = \underset{\sim}{\Psi}(\underset{\sim}{x}) \Delta \underset{\sim}{a}_e^* \quad . \tag{1.6}$$

Here, $\underset{\sim}{\Psi}$ is a matrix of shape functions $\phi_i(\underset{\sim}{x})$, and $\Delta \underset{\sim}{a}_e$ and $\Delta \underset{\sim}{a}_e^*$ are ordered lists of the discrete nodal values of the displacement functions.

The strain vector is related to the displacement vector by

$$\Delta \underset{\sim}{\varepsilon}(\Delta \underset{\sim}{u}_e^h) = \underset{\sim}{D}\, \Delta \underset{\sim}{u}_e^h \tag{1.7}$$

---

* For brevity we will omit the word "increment" when referring to the quantities prefixed by Δ .

where $\underset{\sim}{D}$ is an appropriate matrix of differential operators (see Appendix A), eqn (A.4)). This may be approximated using (1.5) by

$$\underset{\sim}{\Delta\varepsilon}(\Delta\underset{\sim}{u}_e^h)(\underset{\sim}{x}) = \underset{\sim}{D}\,\underset{\sim}{\Psi}(\underset{\sim}{x})\Delta\underset{\sim}{a}_e$$

$$= \underset{\sim}{B}(\underset{\sim}{x})\Delta\underset{\sim}{a}_e \quad . \tag{1.8}$$

Here, $\underset{\sim}{B}(\underset{\sim}{x})$ is the element strain-displacement matrix*, consisting of the partial derivatives of the shape functions $\phi_i(\underset{\sim}{x})$ with respect to $\underset{\sim}{x}$. Similarly, we have

$$\underset{\sim}{\Delta\varepsilon}(\Delta\underset{\sim}{v}_e^h)(\underset{\sim}{x}) = \underset{\sim}{B}(\underset{\sim}{x})\Delta\underset{\sim}{a}_e^* \quad . \tag{1.9}$$

For the approximation of the space L^h of plastic strains we adopt a slightly less conventional approach. Let L^h be the space spanned by piecewise polynomial basis functions, with the i-th function having a value of 1 at the i-th quadrature point and a value of 0 at every other quadrature point. Thus for a (2x2) Gaussian quadrature rule for Ω_e in R^2 , for example, the basis functions will be bilinear polynomials over Ω_e , but globally-discontinuous; for a single-point rule the basis functions will be constants. The restrictions to Ω_e of typical basis functions for L^h are illustrated in Fig. 4.2.

---

* Matrices $\underset{\sim}{D}$, $\underset{\sim}{B}$, $\underset{\sim}{B}_p$, etc., are only ever defined at element level and so we omit the subscript e for brevity.

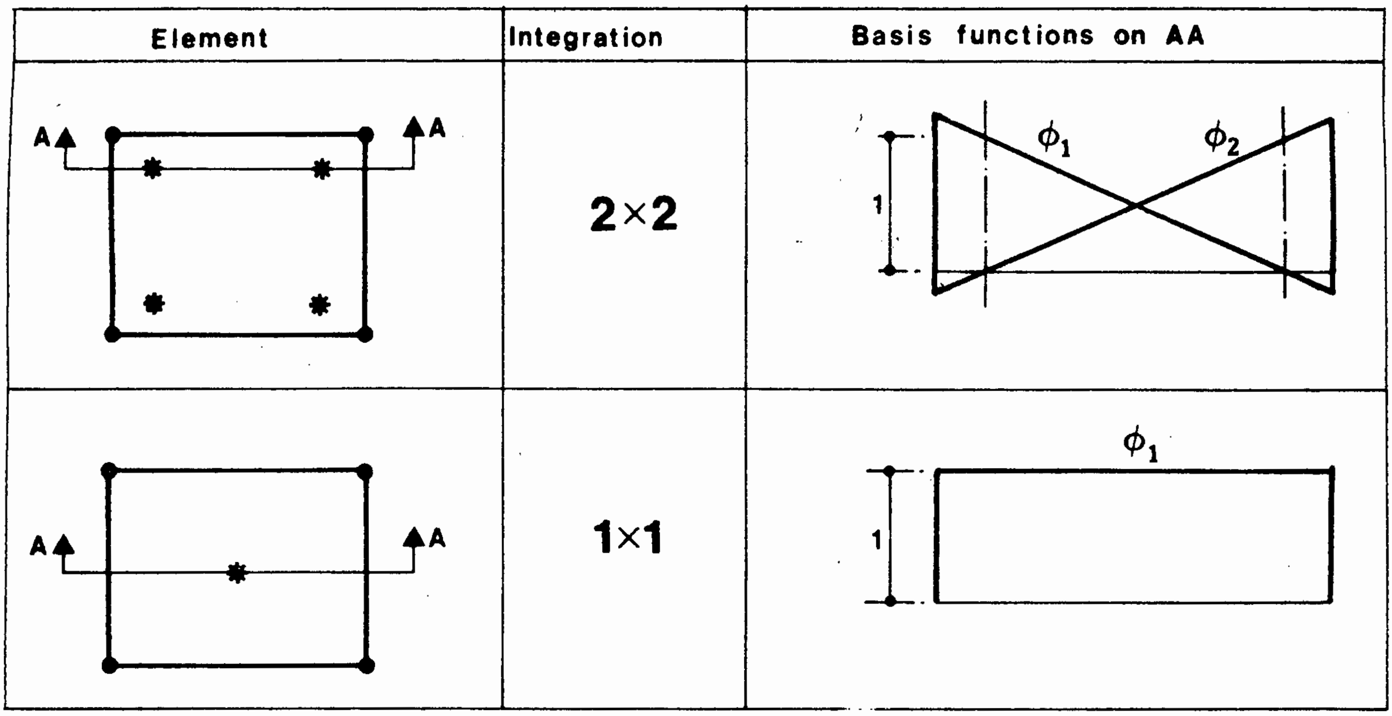


Figure 4.2 Typical basis functions for L^h .

Let $(\underset{\sim}{x}_i)$, $1 \leqslant i \leqslant N_G$, be the position vector of Gauss point i on Ω_e, N_G being the number of Gauss points defining the chosen quadrature rule. Let $\{\phi_j\}_{j=1}^{N_G}$ be a suitable family of local basis functions on Ω_e with the property

$$\phi_j(\underset{\sim}{x}_i) = \delta_{ij} \quad , \quad 1 \leqslant i,j \leqslant N_G \quad . \qquad (1.10)$$

Then the restriction of a typical element Δp_e^h of the finite-dimensional space L_2^h of plastic strains to Ω_e may be approximated by

$$\Delta p_e^h(\underset{\sim}{x}) = \sum_{j=1}^{N_G} \Delta\alpha_j \, \phi_j(\underset{\sim}{x}) \qquad (1.11)$$

where $\Delta\alpha_j = \Delta p_e^h(\underset{\sim}{x}_j)$ is the value of Δp_e^h at Gauss point j on Ω_e . If $\underset{\sim}{\Delta p}_e^h \in L^h$ is the solution vector whose components are approximated as in (1.11) then we write the approximation of this vector in matrix form as

$$\Delta \underset{\sim}{p}_e^h(\underset{\sim}{x}) = \underset{\sim}{B}_p(\underset{\sim}{x}) \Delta \underset{\sim}{\alpha}_e \tag{1.12}$$

and similarly, for an arbitrary vector $\Delta \underset{\sim}{q}_e^h \in L^h$ we have

$$\Delta \underset{\sim}{q}_e^h(\underset{\sim}{x}) = \underset{\sim}{B}_p(\underset{\sim}{x}) \Delta \underset{\sim}{\alpha}_e^* \quad . \tag{1.13}$$

Assuming, for example, that there are three plastic strain components at each Gauss point, $\underset{\sim}{B}_p(\underset{\sim}{x})$ will be a $(3x3N_G)$ matrix of shape functions $\phi_i(\underset{\sim}{x})$ having the following form :

$$\underset{\sim}{B}_p = \left[\begin{array}{ccc|ccc|c|ccc} \phi_1 & & & \phi_2 & & & & \phi_{N_G} & & \\ & \phi_1 & & & \phi_2 & & \cdots\cdots & & \phi_{N_G} & \\ & & \phi_1 & & & \phi_2 & & & & \phi_{N_G} \end{array} \right] . \tag{1.14}$$

The vectors $\Delta \underset{\sim}{\alpha}_e$ and $\Delta \underset{\sim}{\alpha}_e^*$ are ordered lists of discrete plastic strain components at the chosen Gauss points. We note for future reference that according to (1.10) the matrix $\underset{\sim}{B}_p$ has a particularly simple form when it is evaluated at any of the chosen N_G Gauss points. For example, when evaluated at Gauss point 2 , $\phi_2 = 1$ and $\phi_i = 0$, for $1 \leq i \leq N_G$, $i \neq 2$. There will thus be a distinct advantage in integrating all functions which include the matrix $\underset{\sim}{B}_p(\underset{\sim}{x})$ using the same number of Gauss points as were used to define the functions ϕ_i .

The approximation of the initial strains $\underset{\sim}{\varepsilon}^\circ$ is done in the same way as for the strains $\underset{\sim}{\varepsilon}$, so that following (1.8) we write

$$\underset{\sim}{\varepsilon}^{\circ h}_{e}(\underset{\sim}{x}) = \underset{\sim}{B}(\underset{\sim}{x})\underset{\sim}{a}^{\circ}_{e} \tag{1.15}$$

where $\underset{\sim}{a}^{\circ}_{e}$ is the vector of discrete values of the initial displacement components at the nodes on Ω_e . Similarly, the approximation of the initial plastic strains follows according to (1.12) and we write

$$\underset{\sim}{p}^{\circ h}_{e}(\underset{\sim}{x}) = \underset{\sim}{B}_p(\underset{\sim}{x})\underset{\sim}{\alpha}^{\circ}_{e} \tag{1.16}$$

where $\underset{\sim}{\alpha}^{\circ}_{e}$ is the vector of discrete values of the initial plastic strain components at the Gauss points on Ω_e .

We may now substitute the approximations described above into eqn (1.2). In doing so we make use of the original definitions given in Section 3.2, eqns (2.16) and (2.17) for a(•,•) and f(•) , and Section 3.4, eqn (4.4) for <•,•> . Thus, we obtain :

$$\begin{aligned} a^{(e)}(\bar{u}^h_e,\bar{v}^h_e) = &\int_{\Omega_e} (\Delta\underset{\sim}{a}^*_e)^T\underset{\sim}{B}^T\underset{\sim}{C}\,\underset{\sim}{B}\,\Delta\underset{\sim}{a}_e\,d\underset{\sim}{x} - \int_{\Omega_e} (\Delta\underset{\sim}{a}^*_e)^T\underset{\sim}{B}^T\underset{\sim}{C}\,\underset{\sim}{B}_p\Delta\underset{\sim}{\alpha}_e\,d\underset{\sim}{x} \\ &- \int_{\Omega_e} (\Delta\underset{\sim}{\alpha}^*_e)^T\underset{\sim}{B}^T_p\,\underset{\sim}{C}\,\underset{\sim}{B}\,\Delta\underset{\sim}{a}_e\,d\underset{\sim}{x} + \int_{\Omega_e} (\Delta\underset{\sim}{\alpha}^*_e)^T\underset{\sim}{B}^T_p\,\underset{\sim}{C}\,\underset{\sim}{B}_p\Delta\underset{\sim}{\alpha}_e\,d\underset{\sim}{x} \\ &+ E_p\int_{\Omega_e} (\Delta\underset{\sim}{\alpha}^*_e)^T\underset{\sim}{B}^T_p\,\underset{\sim}{B}_p\Delta\underset{\sim}{\alpha}_e\,d\underset{\sim}{x} \end{aligned} \tag{1.17}$$

$$\begin{aligned} \langle Dj_\varepsilon(\Delta\underset{\sim}{p}^h_e),\Delta\underset{\sim}{q}^h_e\rangle\Big|_{\Omega_e} = &\sqrt{2}k\int_{\Omega_e} [(\Delta\underset{\sim}{\alpha}^*_e)^T\underset{\sim}{B}^T_p\,\underset{\sim}{B}_p\Delta\underset{\sim}{\alpha}_e + \varepsilon^2]^{-1/2}(\Delta\underset{\sim}{\alpha}^*_e)^T\underset{\sim}{B}^T_p\,\underset{\sim}{B}_p\Delta\underset{\sim}{\alpha}_e\,d\underset{\sim}{x} \\ &+ E_p\int_{\Omega_e} (\Delta\underset{\sim}{\alpha}^*_e)^T\underset{\sim}{B}^T_p\,\underset{\sim}{B}_p\underset{\sim}{\alpha}^o_e\,d\underset{\sim}{x} \end{aligned} \tag{1.18}$$

$$f(\bar{v}_e^h) = -\int_{\Omega_e} (\Delta \underset{\sim}{a}_e^*)^T \underset{\sim}{B}^T \underset{\sim}{C}\, \underset{\sim}{B}\, \underset{\sim}{a}_e^o \, d\underset{\sim}{x} + \int_{\Omega_e} (\Delta \underset{\sim}{a}_e^*)^T \underset{\sim}{B}^T \underset{\sim}{C}\, \underset{\sim}{B}_p \underset{\sim}{\alpha}_e^o \, d\underset{\sim}{x}$$

$$+ \int_{\Omega_e} (\Delta \underset{\sim}{\alpha}_e^*)^T \underset{\sim}{B}_p^T \underset{\sim}{C}\, \underset{\sim}{B}\, \underset{\sim}{a}_e^o \, d\underset{\sim}{x} - \int_{\Omega_e} (\Delta \underset{\sim}{\alpha}_e^*)^T \underset{\sim}{B}_p^T \underset{\sim}{C}\, \underset{\sim}{B}_p \underset{\sim}{\alpha}_e^o \, d\underset{\sim}{x}$$

$$+ \int_{\Gamma_{se}} (\Delta \underset{\sim}{a}_e^*)^T \underset{\sim}{\Psi}^T \underset{\sim}{t}_e \, d\Gamma_{se} + \int_{\Omega_e} (\Delta \underset{\sim}{a}_e^*)^T \underset{\sim}{\Psi}^T \underset{\sim}{f}_e \, d\underset{\sim}{x} \tag{1.19}$$

Here, Γ_{se} is that part of the element boundary over which tractions $\underset{\sim}{t}_e$ are applied.

With a view to simplifying the above expressions we define the following matrices on Ω_e :

$$\underset{\sim}{K}^{(e)} = \int_{\Omega_e} \underset{\sim}{B}^T \underset{\sim}{C}\, \underset{\sim}{B}\, d\underset{\sim}{x} \tag{1.20}$$

$$\underset{\sim}{L}^{(e)} = \int_{\Omega_e} \underset{\sim}{B}^T \underset{\sim}{C}\, \underset{\sim}{B}_p \, d\underset{\sim}{x} \quad , \quad (\underset{\sim}{L}^{(e)})^T = \int_{\Omega_e} \underset{\sim}{B}_p^T \underset{\sim}{C}\, \underset{\sim}{B}\, d\underset{\sim}{x} \tag{1.21}$$

$$\underset{\sim}{S}_1^{(e)} = \int_{\Omega_e} \underset{\sim}{B}_p^T \underset{\sim}{C}\, \underset{\sim}{B}_p \, d\underset{\sim}{x} \quad , \quad \underset{\sim}{S}_2^{(e)} = E_p \int_{\Omega_e} \underset{\sim}{B}_p^T \underset{\sim}{B}_p \, d\underset{\sim}{x} \quad ; \tag{1.22}$$

we also define the following vectors :

$$\underset{\sim}{g}^{(e)}(\Delta \underset{\sim}{\alpha}_e) = \sqrt{2}k \int_{\Omega_e} \{\gamma^2(\Delta \underset{\sim}{\alpha}_e) + \varepsilon^2\}^{-1/2} \underset{\sim}{B}_p^T \underset{\sim}{B}_p \Delta \underset{\sim}{\alpha}_e \, d\underset{\sim}{x} \tag{1.23}$$

$$\underset{\sim}{P}^{(e)} = \int_{\Omega_e} \underset{\sim}{\Psi}^T \underset{\sim}{f}_e \, d\underset{\sim}{x} + \int_{\Gamma_{se}} \underset{\sim}{\Psi}^T \underset{\sim}{t}_e \, d\Gamma_{se} \quad , \tag{1.24}$$

where the scalar $\gamma(\Delta \underset{\sim}{\alpha}_e)$ is defined by

$$\gamma(\Delta \underset{\sim}{\alpha}_e) = \sqrt{(\Delta \underset{\sim}{\alpha}_e)^T \underset{\sim}{B}_p^T \underset{\sim}{B}_p \Delta \underset{\sim}{\alpha}_e} \quad . \tag{1.25}$$

In anticipation of a future requirement we also define the following derivative :

$$\frac{\partial \underset{\sim}{g}^{(e)}}{\partial \Delta \underset{\sim}{\alpha}_e} = - \sqrt{2}k\int_{\Omega_e} [\gamma^2(\Delta \underset{\sim}{\alpha}_e) + \varepsilon^2]^{-3/2} (\underset{\sim}{B}_p^T \underset{\sim}{B}_p \Delta \underset{\sim}{\alpha}_e)(\underset{\sim}{B}_p^T \underset{\sim}{B}_p \Delta \underset{\sim}{\alpha}_e)^T \, d\underset{\sim}{x}$$

$$+ \sqrt{2}k\int_{\Omega_e} [\gamma^2(\Delta \underset{\sim}{\alpha}_e) + \varepsilon^2]^{-1/2} \underset{\sim}{B}_p^T \underset{\sim}{B}_p \, d\underset{\sim}{x} \quad . \tag{1.26}$$

Using (1.20) through (1.24) in (1.17) through (1.19) and then substituting the resulting expressions into (1.2), we obtain the discrete global approximation of the perturbed minimisation problem :

$$\{(\Delta \underset{\sim}{a}^*)^T \mid (\Delta \underset{\sim}{\alpha}^*)^T\} \left[\begin{array}{c|c} \underset{\sim}{K} & -\underset{\sim}{L} \\ \hline -\underset{\sim}{L}^T & \underset{\sim}{S} \end{array}\right] \left\{\begin{array}{c} \Delta \underset{\sim}{a} \\ \hline \Delta \underset{\sim}{\alpha} \end{array}\right\}$$

$$- \{(\Delta \underset{\sim}{a}^*)^T \mid (\Delta \underset{\sim}{\alpha}^*)^T\} \left\{\begin{array}{c} \underset{\sim}{P} + \underset{\sim}{L}\underset{\sim}{\alpha}^o - \underset{\sim}{K}\underset{\sim}{a}^o \\ \hline -\underset{\sim}{g}(\Delta \underset{\sim}{\alpha}) - \underset{\sim}{S}\underset{\sim}{\alpha}^o + \underset{\sim}{L}^T \underset{\sim}{a}^o \end{array}\right\}$$

$$= \underset{\sim}{0} \quad . \tag{1.27}$$

where we have written $\underset{\sim}{S} = \underset{\sim}{S}_1 + \underset{\sim}{S}_2$ for brevity. The matrices in (1.27) now refer to the assembled element matrices, whilst the vectors represent ordered lists of all displacements and plastic strains in the finite element model.

Since the starred quantities represent arbitrary displacements and plastic strains we must have

$$\begin{bmatrix} K & -L \\ -L^T & S \end{bmatrix} \begin{Bmatrix} \Delta a \\ \Delta \alpha \end{Bmatrix} = \begin{Bmatrix} P + L\alpha^o - Ka^o \\ -g(\Delta \alpha) - S\alpha^o + L^T a^o \end{Bmatrix} \qquad (1.28)$$

which represents a system of simultaneous nonlinear algebraic equations in the global unknowns Δa and $\Delta \alpha$. For convenience we shall occasionally refer to the global system of equations (1.28) in the form

$$[K^*] \begin{Bmatrix} \Delta a \\ \Delta \alpha \end{Bmatrix} = P^*(\Delta \alpha) \qquad (1.29)$$

Some remarks concerning the system of equations (1.28), or alternatively (1.29), are now in order.

1. The matrix K^* is constant in the sense that none of its submatrices depend on the solution variables Δa or $\Delta \alpha$. K is the conventional elastic stiffness matrix.

2. The global matrices S_1 and S_2 are extremely sparse : in the case of plane stress, for example, their only non-zero elements occur in (3x3) submatrices arranged along their diagonals.

3. The first of the two matrix equations in (1.28) represents the discrete system equilibrium equation and is linear in both Δa and $\Delta \alpha$. Noting that $P = P^o + \Delta P$ we may write this equation as

$$\underset{\sim}{K}\Delta\underset{\sim}{a} - \underset{\sim}{L}\Delta\underset{\sim}{\alpha} - \Delta\underset{\sim}{P} = -\underset{\sim}{K}\underset{\sim}{a}^{o} + \underset{\sim}{L}\underset{\sim}{\alpha}^{o} + \underset{\sim}{P}^{o} = \underset{\sim}{0}$$

since the initial state of the body is assumed to be one of equilibrium. Hence, the terms involving initial quantities do not contribute to the first matrix equation and may be discarded.

4. The second matrix equation in (1.28) represents the discrete form of the system constitutive equations; it is nonlinear since it includes on the right-hand-side a term $\underset{\sim}{g}$ which depends on the current plastic strains $\Delta\underset{\sim}{\alpha}$. We expand on this point in some detail in Section 4.3.

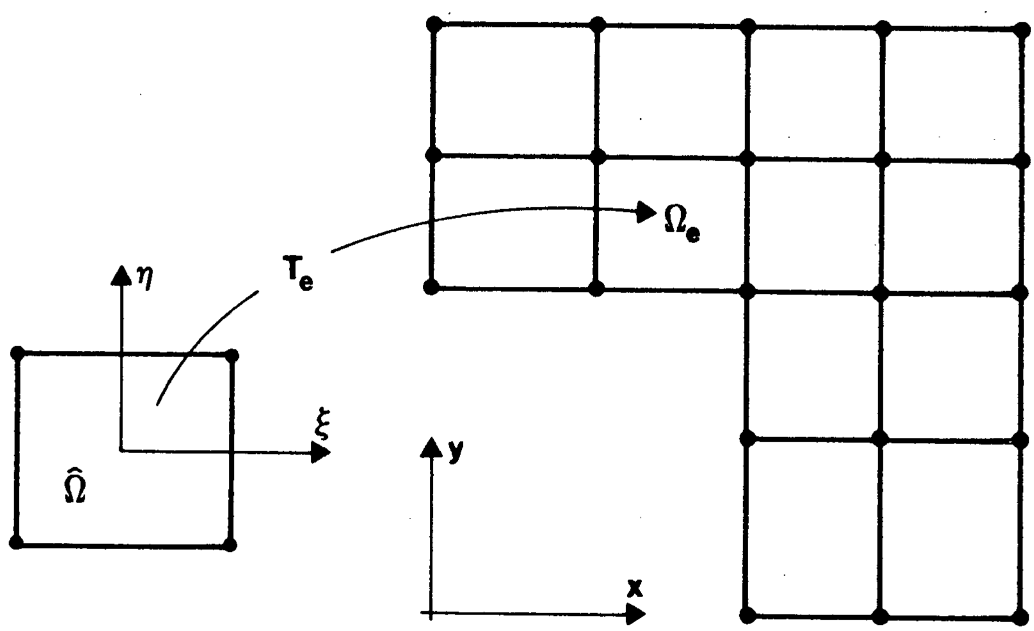


Figure 4.3 The discretised domain and master element defining the coordinate map T_e .

Calculation of Element Matrices and Vectors*

Let us assume that the body occupies a domain Ω in R^2 . The position vector of each point in Ω is $\underset{\sim}{x} = (x,y)$, relative to the cartesian axes X,Y (Fig. 4.3).

We adopt the notion of a master element $\hat{\Omega}$ having a natural coordinate system (ξ,η) , and an invertible coordinate map T_e from $\hat{\Omega}$ to Ω_e . We assume that T_e is an isoparametric map and we let $|\underset{\sim}{J}(\xi,\eta)|$ denote the Jacobian of the transformation T_e . We restrict our attention to a family of quadrilateral, conforming Lagrangian elements having $N_e = 4,8$, or 9, where N_e is the number of nodes on the element.

Let $\{\hat{\psi}_i(\xi,\eta)\}_{i=1}^{N_e}$ be a family of basis functions defined on $\hat{\Omega}$ (see, for example, BECKER, CAREY and ODEN, page 198). Then the family of basis functions $\{\psi_i(x,y)\}_{i=1}^{N_e}$ are obtained from

$$\psi_i(x,y) = \hat{\psi}_i(\xi(x,y),\eta(x,y)) \quad , \quad 1 \leqslant i \leqslant N_e \quad . \tag{1.30}$$

The matrix $\underset{\sim}{B}(x,y)$, which relates strains to displacements (eqn (1.8)), is transformed into $\underset{\sim}{\hat{B}}(\xi,\eta)$ in the usual way using the Jacobian $|\underset{\sim}{J}|$, (see BECKER, CAREY and ODEN, page 189).

---

* The calculations which we describe here follow closely those described in Chapter 5 of BECKER, CAREY and ODEN (1981) to which the reader is referred for further details.

We assume that, given a quadrature rule of order N_G on $\hat{\Omega}$, we can construct a suitable family of basis functions $\{\hat{\phi}_i(\xi,\eta)\}_{i=1}^{N_G}$, which are defined at the Gauss points, as shown in Fig. 4.2. Then the family $\{\phi_i(x,y)\}_{i=1}^{N_G}$ are obtained in the manner indicated in (1.30), and similarly, the matrix $\hat{\underset{\sim}{B}}_p(\xi,\eta)$ may be obtained from $\underset{\sim}{B}_p(x,y)$ in the same way as $\hat{\underset{\sim}{B}}$ is obtained from $\underset{\sim}{B}$. It is worthwhile noting, however, that if we numerically integrate functions containing $\hat{\phi}_i(\xi,\eta)$ using the same quadrature rule as was used to define the $\hat{\phi}_i$, then the shape function $\hat{\phi}_i$ need not be explicitly defined.

Let $\underset{\sim}{f}(x,y)$ be any matrix of functions defined on $\Omega_e \in R^2$ and $\hat{\underset{\sim}{f}}(\xi,\eta)$ the transformation of this matrix of functions to $\hat{\Omega}$ under the inverse map T_e^{-1}. Then, writing $d\underset{\sim}{x} = dxdy$, it is clear that

$$\int_{\Omega_e} \underset{\sim}{f}(x,y)\, d\underset{\sim}{x} = \int_{\hat{\Omega}} \hat{\underset{\sim}{f}}(\xi,\eta) \left|\underset{\sim}{J}(\xi,\eta)\right| \, d\xi d\eta \quad . \tag{1.31}$$

Now let $I(\cdot)$ denote the numerical quadrature formula which is to be used to evaluate the right-hand-side of (1.31), given by

$$I(\hat{\underset{\sim}{f}}) = \sum_{i=1}^{N_G} \hat{\underset{\sim}{f}}(\xi_i,\eta_i) \left|\underset{\sim}{J}(\xi_i,\eta_i)\right| w_i \tag{1.32}$$

where N_G is the chosen number of quadrature points, (ξ_i,η_i) are the coordinates of each quadrature point, and $w_i > 0$ is the quadrature weight associated with point i . We use the above formula to integrate each of the element matrices and vectors, for which the relevant $\hat{\underset{\sim}{f}}$ functions are given in Table 4.1.

TABLE 4.1

Functions $\hat{\underset{\sim}{f}}$ in (1.32) for integration of element matrices and vectors

Matrix or Vector	$\hat{\underset{\sim}{f}}(\xi_i, \eta_i)$
$\underset{\sim}{K}^{(e)}$	$[\hat{\underset{\sim}{B}}(\xi_i,\eta_i)]^T \underset{\sim}{C}\, \hat{\underset{\sim}{B}}(\xi_i,\eta_i)$
$\underset{\sim}{L}^{(e)}$	$[\hat{\underset{\sim}{B}}(\xi_i,\eta_i)]^T \underset{\sim}{C}\, \hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)$
$\underset{\sim}{S}_1^{(e)}$	$[\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)]^T \underset{\sim}{C}\, \hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)$
$\underset{\sim}{S}_2^{(e)}$	$E_p[\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)]^T\, \hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)$
$\underset{\sim}{P}^{(e)}$	$\underset{\sim}{\Psi}^T(\xi_i,\eta_i)\underset{\sim}{f}_e$ and $\underset{\sim}{\Psi}^T(\xi_i,\eta_i)\underset{\sim}{t}_e$

Note: $\underset{\sim}{C}$ is the matrix of elastic constants; see, for example, Appendix A, eqn (A.2).

As far as the orders of integration are concerned we always integrate $\underset{\sim}{K}^{(e)}$ exactly (see BATHE (1982), Table 5.5), but allow the option of a different order of integration for $\underset{\sim}{L}^{(e)}$, $\underset{\sim}{S}_1^{(e)}$ and $\underset{\sim}{S}_2^{(e)}$. However, the computations are considerably simplified if we use the same quadrature rule to integrate the latter matrices as was used to define the basis functions $\phi_j(\xi,\eta)$ in (1.10) . In this way the matrix $\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)$ takes the trivial form described in the text following eqn (1.14), with the result that, assuming there are three plastic strain components at each Gauss point, the matrix $\underset{\sim}{S}_1^{(e)}$ is made up of (3x3) submatrices arranged along its diagonal, and the matrix $\underset{\sim}{S}_2^{(e)}$ is diagonal.

The vector $\underset{\sim}{g}^{(e)}(\Delta\underset{\sim}{\alpha}_e)$ in (1.23), although dependent on the current values of the discrete plastic strains, is integrated using the same procedure as described above. We divide the function $\hat{\underset{\sim}{f}}$ to be used in the formula (1.32) into two parts :

$$\hat{f}_1(\xi_i,\eta_i) = (\Delta\underset{\sim}{\alpha}_e)^T[\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)]^T\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)\Delta\underset{\sim}{\alpha}_e + \varepsilon^2 \tag{1.33}$$

which is a scalar quantity, and

$$\hat{\underset{\sim}{f}}_2(\xi_i,\eta_i) = [\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)]^T\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)\Delta\underset{\sim}{\alpha}_e \tag{1.34}$$

which is a vector. Then, using (1.23), we may write

$$\hat{\underset{\sim}{f}}(\xi_i,\eta_i) = \sqrt{2}k(\hat{f}_1)^{-1/2}\hat{\underset{\sim}{f}}_2 \quad . \tag{1.35}$$

A similar procedure to that described above can be used to evaluate the matrix $\partial\underset{\sim}{g}^{(e)}/\partial\Delta\underset{\sim}{\alpha}_e$ which appears in (1.26). Both the latter matrix and the vector $\underset{\sim}{g}^{(e)}$ are integrated using the same quadrature rule as is used for $\underset{\sim}{L}^{(e)}$ and $\underset{\sim}{S}^{(e)}$.

4.2 SOLUTION PROCEDURES

The system of nonlinear equations in (1.28) represents the following incremental problem : given the initial loads $\underset{\sim}{P}^\circ$ acting on a body, its initial displacements $\underset{\sim}{a}^\circ$ and initial plastic strains $\underset{\sim}{\alpha}^\circ$, find the changes in displacement $\Delta\underset{\sim}{a}$ and plastic strain $\Delta\underset{\sim}{\alpha}$ due to a change in the loads $\Delta\underset{\sim}{P}$. In this section we consider the numerical solution of this problem and the subsequent calculation of the stresses.

An Iterative Solution Procedure

Newton's method is a well-established procedure for solving systems of nonlinear equations of the form

$$\underset{\sim}{F}(\underset{\sim}{u}) = \underset{\sim}{0} \qquad (2.1a)$$

or, using indicial notation,

$$F_i(u_j) = 0 \quad , \quad 1 \leqslant i,j \leqslant N \ . \qquad (2.1b)$$

Taking the first two terms of the Taylor's series of $\underset{\sim}{F}(\underset{\sim}{u})$ we may write

$$\underset{\sim}{F}(\underset{\sim}{u}) \simeq \underset{\sim}{F}(\underset{\sim}{u}_o) + \underset{\sim}{J}(\underset{\sim}{u}_o)(\underset{\sim}{u} - \underset{\sim}{u}_o) = \underset{\sim}{0} \qquad (2.2)$$

where $\underset{\sim}{u}_\circ$ is the initial estimate of the solution and $\underset{\sim}{u}$ is the improved estimate which is to be computed. The matrix $\underset{\sim}{J}$ is called the Jacobian matrix and is given by

$$\underset{\sim}{J}(\underset{\sim}{u}) = [\frac{\partial F_i}{\partial u_j}] \quad , \quad 1 \leqslant i,j \leqslant N \quad . \tag{2.3}$$

Eqn (2.2) is easily generalised to form an iterative procedure which may be written as

$$\underset{\sim}{F}(\underset{\sim}{u}_n) + \underset{\sim}{J}(\underset{\sim}{u}_n)(\underset{\sim}{u}_{n+1} - \underset{\sim}{u}_n) = \underset{\sim}{0} \tag{2.4}$$

where n refers to the iteration number (n=0, 1, 2, ...).

Let us now apply the above procedure to the system of equations given in eqn (1.28). Following (2.1b) we define the variables

$$u_1 \equiv \Delta\underset{\sim}{a} \equiv \underset{\sim}{\bar{a}} \quad , \quad u_2 \equiv \Delta\underset{\sim}{\alpha} \equiv \underset{\sim}{\bar{\alpha}} \tag{2.5}$$

and the functions

$$F_1 = \underset{\approx}{K}\bar{a} - \underset{\approx}{L}\bar{\alpha} - \Delta\underset{\sim}{P} \tag{2.6a}$$

$$F_2 = - \underset{\sim}{L}^T\underset{\sim}{\bar{a}} + \underset{\approx}{S}\bar{\alpha} + \underset{\sim}{g}(\underset{\sim}{\bar{\alpha}}) + \underset{\approx}{S}\alpha^\circ - \underset{\sim}{L}^T\underset{\sim}{a}^\circ \tag{2.6b}$$

where in writing (2.6a) have made use of Remark 3 following eqn (1.29). Using the above notation and definitions in (2.3) and (2.4) we obtain the following iterative procedure :

$$\left[\begin{array}{c|c} \underset{\sim}{K} & -\underset{\sim}{L} \\ \hline -\underset{\sim}{L}^T & \underset{\sim}{S} + \left.\dfrac{\partial \underset{\sim}{g}}{\partial \bar{\underset{\sim}{\alpha}}}\right|_{\bar{\underset{\sim}{\alpha}}_n} \end{array}\right] \left\{\begin{array}{c} \Delta\bar{\underset{\sim}{a}}_n \\ \hline \Delta\bar{\underset{\sim}{\alpha}}_n \end{array}\right\} = \left\{\begin{array}{c} -\underset{\sim}{K}\bar{\underset{\sim}{a}}_n + \underset{\sim}{L}\bar{\underset{\sim}{\alpha}}_n + \Delta\underset{\sim}{P} \\ \hline \underset{\sim}{L}^T\bar{\underset{\sim}{a}}_n - \underset{\sim}{S}\bar{\underset{\sim}{\alpha}}_n - \underset{\sim}{g}(\bar{\underset{\sim}{\alpha}}_n) - \underset{\sim}{S}\underset{\sim}{\alpha}^o + \underset{\sim}{L}^T\underset{\sim}{a}^o \end{array}\right\} \tag{2.7}$$

where $\Delta\bar{\underset{\sim}{a}}_n = \bar{\underset{\sim}{a}}_{n+1} - \bar{\underset{\sim}{a}}_n$, and $\Delta\bar{\underset{\sim}{\alpha}}_n = \bar{\underset{\sim}{\alpha}}_{n+1} - \bar{\underset{\sim}{\alpha}}_n$. We anticipate that a convergent solution to the original nonlinear problem (1.28) will be obtained as $n \to \infty$.

The matrix $\partial \underset{\sim}{g}/\partial \bar{\underset{\sim}{\alpha}}$ in (2.7) has already been defined in (1.26), where we note that because of its dependence on $\underset{\sim}{B}_p$ it has the same degree of sparsity as the matrix $\underset{\sim}{S}_1$.

Static Condensation of eqn (2.7)

There are certain features of the system of equations in (2.7) which suggest that static condensation prior to the solution would improve not only the numerical efficiency but also the ease with which the numerical computations can be handled. The features which we have in mind are the extreme sparseness of the lower right submatrix $\bar{\underset{\sim}{S}}_n$ which we define by

$$\bar{\underset{\sim}{S}}_n = \underset{\sim}{S} + \left.\frac{\partial \underset{\sim}{g}}{\partial \bar{\underset{\sim}{\alpha}}}\right|_{\bar{\underset{\sim}{\alpha}}_n} , \tag{2.8}$$

and the linearity of the first matrix equation.

We propose first to eliminate the plastic strain increments $\Delta\bar{\underset{\sim}{\alpha}}_n$ as follows. From the second matrix equation in (2.7) we have

$$\Delta\bar{\underset{\sim}{\alpha}}_n = (\bar{\underset{\sim}{S}}_n)^{-1}[\underset{\sim}{L}^T\Delta\bar{\underset{\sim}{a}}_n - \underset{\sim}{g}(\bar{\underset{\sim}{\alpha}}_n) + \underset{\sim}{L}^T(\underset{\sim}{a}^o + \bar{\underset{\sim}{a}}_n) - \underset{\sim}{S}(\underset{\sim}{\alpha}^o + \bar{\underset{\sim}{\alpha}}_n)] \quad . \qquad (2.9)$$

Substituting (2.9) into the first matrix equation of (2.7) and rearranging we obtain

$$[\underset{\sim}{K} - \underset{\sim}{L}(\bar{\underset{\sim}{S}}_n)^{-1}\underset{\sim}{L}^T]\Delta\bar{\underset{\sim}{a}}_n = \underset{\sim}{L}(\bar{\underset{\sim}{S}}_n)^{-1}[- \underset{\sim}{g}(\bar{\underset{\sim}{\alpha}}_n) + \underset{\sim}{L}^T(\underset{\sim}{a}^o + \bar{\underset{\sim}{a}}_n)$$
$$- \underset{\sim}{S}(\underset{\sim}{\alpha}^o + \bar{\underset{\sim}{\alpha}}_n)] - \underset{\sim}{K}\bar{\underset{\sim}{a}}_n + \underset{\sim}{L}\bar{\underset{\sim}{\alpha}}_n + \Delta\underset{\sim}{P} \quad . \qquad (2.10)$$

Eqns (2.9) and (2.10) constitute the condensed Newton iterative procedure, which now replaces the original procedure given in (2.7). Thus, we solve (2.10) first for the displacement increments $\Delta\bar{\underset{\sim}{a}}_n$, and then use these in (2.9) to find the plastic strain increments $\Delta\bar{\underset{\sim}{\alpha}}_n$; we continue in this way until an acceptable solution has been obtained.

In applications of Newton's method to problems of the type described here it is standard practice to choose the following initial estimates of the solution

$$\bar{\underset{\sim}{a}}_o = \bar{\underset{\sim}{\alpha}}_o = \underset{\sim}{0} \quad ; \qquad (2.11)$$

(recall that $\underset{\sim}{a}^o$ and $\underset{\sim}{\alpha}^o$ represent the initial conditions at the start of an increment). However, a better choice of initial estimates must clearly be the elastic solution for the increment; in fact, in certain cases, for example, when the body unloads elastically, such estimates will obviously be optimal. Thus, we choose the following initial estimates :

$$\bar{\underset{\sim}{a}}_o = \underset{\sim}{K}^{-1}\Delta\underset{\sim}{P} \quad , \quad \bar{\underset{\sim}{\alpha}}_o = \underset{\sim}{0} \quad . \tag{2.12}$$

Since $\underset{\sim}{K}$ remains constant for a given sequence of incremental solutions the inversion of $\underset{\sim}{K}$ need only be performed once; the initial estimates for increments 2,3,... are obtained by simply scaling the initial estimates for increment 1 according to the current value of $\Delta\underset{\sim}{P}$.

At the end of the n-th iteration we update the displacement and plastic strain changes according to

$$\bar{\underset{\sim}{a}}_{n+1} = \bar{\underset{\sim}{a}}_n + \Delta\bar{\underset{\sim}{a}}_n \quad , \quad \bar{\underset{\sim}{\alpha}}_{n+1} = \bar{\underset{\sim}{\alpha}}_n + \Delta\bar{\underset{\sim}{\alpha}}_n \quad ;$$

at the same time we compute the new residual vector $\underset{\sim}{R}_{n+1}$ (the right-hand-side of (2.10) with n replaced by n+1) using the results $\bar{\underset{\sim}{a}}_{n+1}$ and $\bar{\underset{\sim}{\alpha}}_{n+1}$ and check to see whether $|\underset{\sim}{R}_{n+1}|$ is less than some predetermined tolerance. If it is not, we continue with the (n+1)th iteration; if it is less than the tolerance then the solution is deemed to be acceptable.

It is important to emphasise that the static condensation described above is essential for the viability of the present formulation, since the solution of the original equations (2.7) demands excessive computational effort in all but the most trivial of problems. We are fortunate in this respect that the matrix $\bar{\underset{\sim}{S}}_n$ is so sparse and that the computation of its inverse, as required in (2.10), requires no more than the computation of a sequence of (NxN) matrix inversions, where N is the number of plastic strain components at each Gauss point. Furthermore, the number of equations to be solved in (2.10) is substantially less than the number that would need to be solved in (2.7), even if low order interpolation of the plastic strains were used.

Calculation of Stresses

Since both the displacement and plastic strain increments are immediately available at the end of the iterative procedure described above, the calculation of the stresses is a particularly simple task. Let $\Delta\underset{\sim}{a}_e$ be the restriction to Ω_e of the displacement vector $\Delta\underset{\sim}{a}$ as obtained from a convergent solution. Then the corresponding strains at Gauss point j on Ω_e are, using (1.8),

$$\Delta\underset{\sim}{\varepsilon}_j = \underset{\sim}{B}(\underset{\sim}{x}_j)\Delta\underset{\sim}{a}_e \quad . \tag{2.13}$$

Similarly, let $\Delta\underset{\sim}{p}_j$ be the restriction to the j-th Gauss point on Ω_e of the plastic strain vector $\Delta\underset{\sim}{\alpha}$, as obtained from the preceding convergent solution. Then the corresponding stress state at Gauss point j is computed from

$$\Delta\underset{\sim}{\sigma}_j = \underset{\sim}{C}[\Delta\underset{\sim}{\varepsilon}_j - \Delta\underset{\sim}{p}_j] \tag{2.14}$$

where $\underset{\sim}{C}$ is a matrix of elastic constants (see Appendix A, eqn (A.2)). Thus, if the stress state at the start of the increment is denoted by $\underset{\sim}{\sigma}_j^{\circ}$, then the stress state $\underset{\sim}{\sigma}_j$ at the end of the increment is given by

$$\underset{\sim}{\sigma}_j = \underset{\sim}{\sigma}_j^{\circ} + \Delta\underset{\sim}{\sigma}_j \quad . \tag{2.15}$$

It is particularly interesting to note that there is no explicit requirement for the stresses to be calculated during the iterative solution procedure defined by eqns (2.9) and (2.10). Thus, the stresses need be calculated only once for each loading increment. Furthermore, since there is no implied integration in (2.14) this is the best available estimate of the stresses.

This completes the discussion of the solution procedures. In the next section we will discuss some of the more interesting computational details of these procedures.

4.3 SOME REMARKS CONCERNING THE ROLE OF THE CONSTITUTIVE EQUATIONS IN THE SOLUTION PROCEDURE

We recall from Section 3.1 that the incremental holonomic boundary-value problem includes the following two constitutive equations* :

$$\underset{\sim}{\sigma} = \underset{\sim}{C}[\underset{\sim}{\varepsilon}^{\circ} + \Delta\underset{\sim}{\varepsilon} - \underset{\sim}{p}^{\circ} - \Delta\underset{\sim}{p}] \tag{3.1}$$

$$\underset{\sim}{\sigma} - \frac{\partial \hat{W}_1^p}{\partial \Delta \underset{\sim}{p}} \in \partial \hat{W}_2^p \quad . \tag{3.2}$$

The second constitutive equation includes the non-differentiable functional $\hat{W}_2^p$, which we subsequently regularised in order to be able to formulate the perturbed minimisation problem. The discrete

---

* We make use of eqn (1.15), Section 3.1, in writing down the first of these equations; for convenience we refer to them as the first and second constitutive equations respectively.

approximation of the perturbed minimisation problem gave rise to a system of nonlinear algebraic equations, the nonlinearity arising as a result of the original non-differentiability of $\hat{W}_2^p$. As we shall see here, both of the above equations appear in discrete form in the second (nonlinear) matrix equation in (1.28). It is, therefore, instructive to examine the role which these equations, and in particular the regularisation procedure, play in the solution of the discrete incremental holonomic problem.

We begin by recalling the second matrix equation in (1.28), which we repeat here for convenience :

$$-\underset{\sim}{L}^T \Delta \underset{\sim}{a} + (\underset{\sim}{S}_1 + \underset{\sim}{S}_2)\Delta \underset{\sim}{\alpha} = -\underset{\sim}{g}(\Delta \underset{\sim}{\alpha}) - (\underset{\sim}{S}_1 + \underset{\sim}{S}_2)\underset{\sim}{\alpha}^o + \underset{\sim}{L}^T \underset{\sim}{a}^o \quad . \tag{3.3}$$

Now we may imagine without loss of generality a body consisting of a single element Ω_e : then, using the definitions of $\underset{\sim}{L}$, $\underset{\sim}{S}$ and $\underset{\sim}{g}$ in (1.21) through (1.23), and the various strain approximations proposed earlier, we may rewrite (3.3) in the form

$$\int_{\Omega_e} \underset{\sim}{B}_p^T \underset{\sim}{C}[\underset{\sim}{\varepsilon}^o + \Delta \underset{\sim}{\varepsilon} - \underset{\sim}{p}^o - \Delta \underset{\sim}{p}]\, d\underset{\sim}{x} = \sqrt{2}k \int_{\Omega_e} (\Delta \underset{\sim}{p} \cdot \Delta \underset{\sim}{p} + \varepsilon^2)^{-1/2}\, \underset{\sim}{B}_p^T \Delta \underset{\sim}{p}\, d\underset{\sim}{x}$$
$$+ \int_{\Omega_e} E_p\, \underset{\sim}{B}_p^T\, \underset{\sim}{p}^o\, d\underset{\sim}{x} + \int_{\Omega_e} E_p\, \underset{\sim}{B}_p^T \Delta \underset{\sim}{p}\, d\underset{\sim}{x} \quad . \tag{3.4}$$

It is readily apparent that the expression $\underset{\sim}{C}[\underset{\sim}{\varepsilon}^\circ + \Delta\underset{\sim}{\varepsilon} - \underset{\sim}{p}^\circ - \Delta\underset{\sim}{p}]$ in the integrand on the left-hand-side of this equation is the stress in the element as given by the first constitutive equation (3.1). The complete integrand represents work per unit volume which when integrated over the element volume yields the work done by the stresses, as computed via the first constitutive equation, when the stresses move through unit plastic strains. Similarly, the right-hand-side of eqn (3.4) represents the work done by the stresses, as computed via the second constitutive equation in its regularised form. Eqn (3.4) clearly provides the link between the two constitutive equations of the holonomic theory.

Let us now rewrite eqn (3.3) on Ω_e with the terms in the same order as they occur in (3.4), so that the left-hand-side contains the first constitutive equation and the right-hand-side contains the second :

$$\underset{\sim}{L}^T(\underset{\sim}{a}^o + \Delta\underset{\sim}{a}) - S_1(\underset{\sim}{\alpha}^o + \Delta\underset{\sim}{\alpha}) = \underset{\sim}{g}(\Delta\underset{\sim}{\alpha}) + \underset{\sim}{S}_2(\underset{\sim}{\alpha}^o + \Delta\underset{\sim}{\alpha}) \quad . \tag{3.5}$$

Each of the terms appearing in (3.5) is a vector with the same number of components as the number of discrete plastic strains on Ω_e . As we intimated above, this equation may be regarded as posing the following problem : given $\underset{\sim}{L}$, $\underset{\sim}{S}_1$, $\underset{\sim}{S}_2$, $\underset{\sim}{a}^\circ$ and $\underset{\sim}{\alpha}^\circ$, find $\Delta\underset{\sim}{a}$ and $\Delta\underset{\sim}{\alpha}$ such that the work done by the stresses calculated from the constitutive equation implicit in the left-hand-side is equal to the work done by the stresses as calculated from the constitutive equation implicit on the right-hand-side. (Of course, the stresses must also satisfy the equilibrium equations, but this is not directly relevant to the present discussion.) The difficulty in solving this problem lies in the fact that the right-hand-side of (3.5) depends on the current solution $\Delta\underset{\sim}{\alpha}$.

For simplicity let us consider a single component of each of the vectors in (3.5) and write the expression for this component, as computed from the right-hand-side of (3.5), as

$$G(\Delta\alpha) = g(\Delta\alpha) + S_2(\alpha^\circ + \Delta\alpha) \tag{3.6}$$

Each of the work terms in eqn (3.6) has a stress associated with it : $g(\Delta\alpha)$ is the work done in plastic flow, $S_2\alpha^\circ$ is the work due to the initial plastic strains, and $S_2\Delta\alpha$ is the work due to plastic hardening. The curves representing the stresses associated with each of these work quantities are shown in Fig. 4.4(a), (note that for simplicity we use the same notation to represent work and the stress associated with that work) : here, $g(\Delta\alpha)$ represents the regularised stress, whereas $g'(\Delta\alpha)$, which is a step-function, represents the original unregularised stress. The sum of the stresses shown in Fig. 4.4(a) is represented by the curve $G(\Delta\alpha)$, in Fig. 4.4(b), for the regularised case, and by the curve $G'(\Delta\alpha)$ for the original unregularised case. Thus, the effect of the regularisation, or the parameter ε , is to smooth the nondifferentiable curve $G'(\Delta\alpha)$; alternatively $G(\Delta\alpha) \to G'(\Delta\alpha)$ as $\varepsilon \to 0$. Returning to the problem stated in the previous paragraph, we may restate it as follows : find $\Delta\alpha$ such that the stress given by the curve $G(\Delta\alpha)$ is equal to the stress which satisfies the first constitutive equation (and the equilibrium equations). The essential cause of the nonlinearity of the discrete incremental holonomic problem is thus clearly evident.

It is apparent that there can be no change in work done (or change in stress) without a corresponding change in plastic strain. However, while the solution point remains on the essentially "vertical" part of the curve G the plastic strains will be of the order of ε .

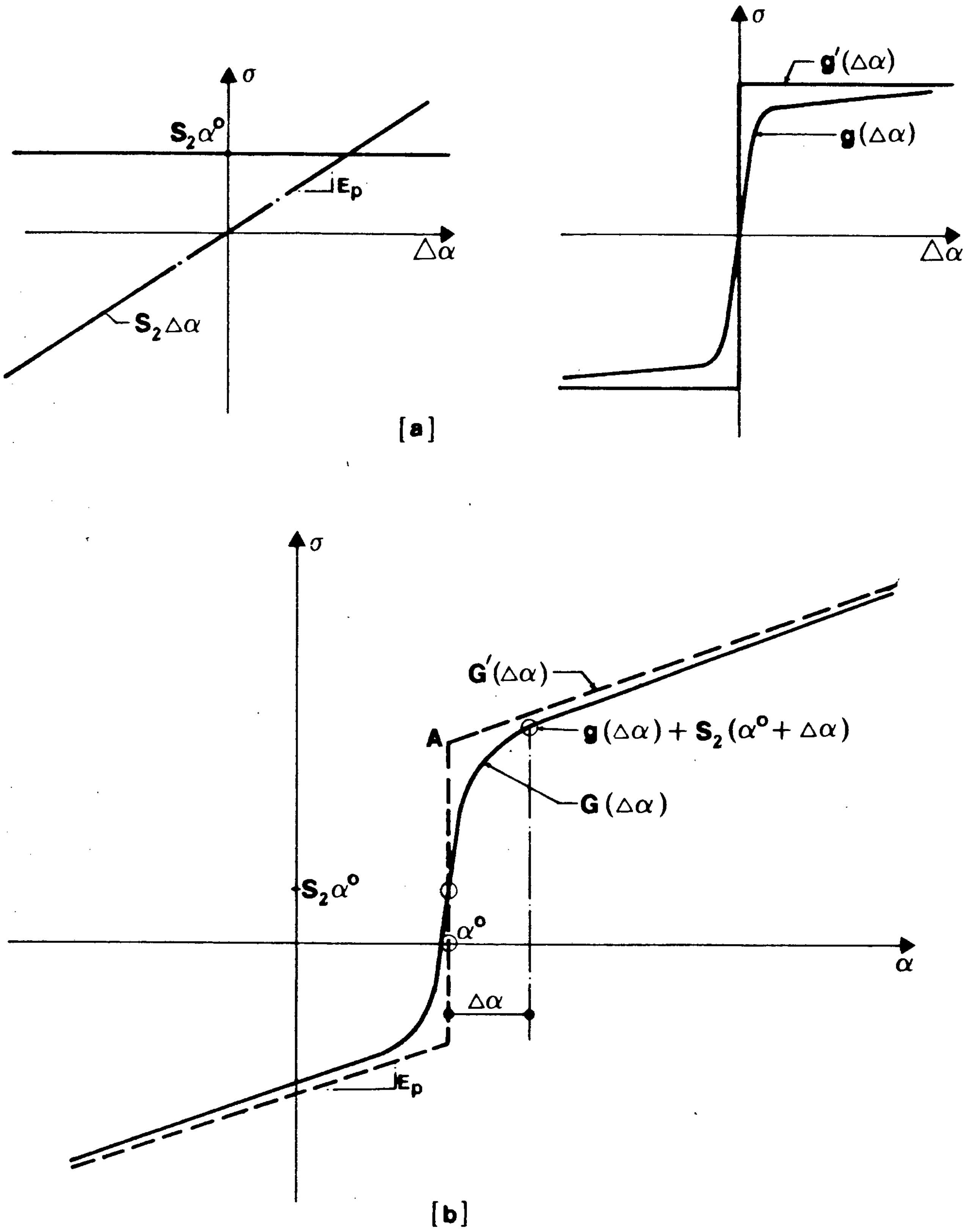


Figure 4.4 Stress-plastic strain curves for the second constitutive equation.

By choosing ε small enough the "vertical" part of the curve G can be made to extend virtually to the point A; since this point corresponds to the current yield stress it is evident that while the stress remains within the yield limits the plastic strains will be of the order of ε. Conversely, it is equally apparent that if ε is chosen too large the curve G will be so "rounded" at the knee that stresses which are well inside the yield limits will not be attainable without significant change in plastic strain. In such a situation we would expect the solution to diverge as in fact we will show in Chapter 7.

The fact that the curve G in Fig. 4.4(b) approaches a step-function as $\varepsilon \to 0$ gives rise to some unusual numerical behaviour in the Newton iterative procedure. Recall from eqn (2.7) that the procedure requires the computation of the term $\partial \underset{\sim}{g}/\partial \Delta \underset{\sim}{\alpha}$ at each iteration. Returning to Fig. 4.4(b) it is clear that this term will have a very large and essentially constant value for values of $\Delta\alpha$ of the order of ε, but will undergo very rapid changes of value in the vicinity of the point A (that is, assuming ε is sufficiently small). Thereafter, the term will again remain essentially constant with a value approaching E_p. There appears to be some cause for concern here in that $\partial \underset{\sim}{g}/\partial \Delta \underset{\sim}{\alpha}$ exhibits pronounced sensitivity to changes in plastic strain over a very limited interval of plastic strain domain. This phenomenon can be easily demonstrated by numerical examples so that we defer further discussion until Chapter 7.

Up to now we have been concerned exclusively with hardening materials. For elastic-perfectly plastic materials $E_p = 0$ and thus $S_2 = 0$, and the curve G becomes antisymmetric about the α axis with horizontal asymptotes. The essential characteristics of the regularised curve remain unchanged. The regularised curve G does however have one additional advantage over the unregularised curve G' in this case : the relationship between stress and plastic strain is uniquely defined for all values of plastic strain, which is clearly not true for the curve G' when $E_p = 0$.

The foregoing discussion suggests an interesting interpretation of the behaviour of the second constitutive equation in a sequence of incremental holonomic problems. For a given sequence of incremental problems the value of S_2 remains constant; hence, since the shape of the $G(\Delta\alpha)$ curve depends only on $\Delta\alpha$ the shape must remain fixed, but may translate (without rotation) in the stress-plastic strain space according to the current value of α° . Since α° is updated at the beginning of each increment we may interpret this as effecting a corresponding update of the second constitutive equation. Thus, we may think of a sequence of incremental holonomic problems as one in which the constitutive equations are continually updated at the beginning of each increment to reflect the current state of plastic strain.

4.4 AN ALGORITHM FOR THE INCREMENTAL HOLONOMIC PROBLEM

We summarise in Table 4.1 an algorithm for computing the solution for the discrete approximation of the incremental holonomic problem. The algorithm is based on the Newton iterative procedure given in eqns (2.9) and (2.10), and assumes the elastic initial estimates of eqn (2.12). It is also assumed that since the matrices $\underset{\sim}{K}$, $\underset{\sim}{L}$ and $\underset{\sim}{S}$ remain constant during the entire iterative procedure, they may be computed and stored at the start of the algorithm for subsequent repetitive use. The solution of the equation in Step 3 is performed using the Cholesky decomposition with forward/backward substitution due to BATHE (1982).

It is worth mentioning briefly the computational advantages of the incremental holonomic formulation. Perhaps its most outstanding feature is the ability to compute the incremental solution within a single iterative procedure, namely, the algorithm described in Table 4.1. This follows from the fact that the integration which is normally required to march the solution forward across the increment is implicit within the formulation. This differs from the rate approach (which we discuss in Chapter 6) where the forward integration must be performed explicitly, subsequent to the solution of the rate problem itself. Furthermore, all the primary matrices required in Table 4.1 are constant (and thus need only be computed once), and the matrix $\bar{\underset{\sim}{S}}_n$ is sufficiently sparse so as to render the computation of its inverse a relatively trivial exercise. It is therefore clear that we have in the incremental holonomic algorithm one which is remarkably compact and simple, and which would appear to offer a computationally viable alternative to the conventional algorithms for solving the incremental elastic-plastic problem.

We return to the incremental holonomic problem in Chapter 7 where we shall perform numerical experiments to test its viability and accuracy.

TABLE 4.2

Solution algorithm for the extended holonomic problem

The following notation is used here :

$$\tilde{K}_n^* = \tilde{K} - \tilde{L}(\bar{\tilde{S}}_n)^{-1}\tilde{L}^T$$

$$\tilde{R}_n = \tilde{L}(\bar{\tilde{S}}_n)^{-1}[-\tilde{g}(\bar{\tilde{\alpha}}_n) + \tilde{L}^T(\tilde{a}^o + \bar{\tilde{a}}_n) - \tilde{S}(\tilde{\alpha}^o + \bar{\tilde{\alpha}}_n)]$$

$$-\tilde{K}\bar{a}_n + \tilde{L}\bar{\alpha}_n + \Delta\tilde{P} \quad .$$

TOLER = a predetermined convergence tolerance.

1. We assume $\tilde{K}$, $\tilde{L}$ and $\tilde{S}$ are already available. Compute the chosen load increment $\Delta\tilde{P}$; compute $\bar{\tilde{S}}_o \equiv \tilde{S}$ and invert. Compute $\bar{\tilde{a}}_o$ using eqn (2.12) (or by scaling $\bar{\tilde{a}}_o$ from a previous increment), and hence compute $\tilde{R}_o$.

Complete the following steps for n = 0,1,2, ...

2. Using $\bar{\tilde{\alpha}}_n$, compute $\bar{\tilde{S}}_n$ and hence $\tilde{L}(\bar{\tilde{S}}_n)^{-1}\tilde{L}^T$. Assemble $\tilde{K}_n^*$.

3. Solve the system of equations $\tilde{K}_n^* \, \Delta\bar{\tilde{a}}_n = \tilde{R}_n$.

4. Using $\Delta\bar{\tilde{a}}_n$ compute $\Delta\bar{\tilde{\alpha}}_n$ from eqn (2.9).

5. Compute $\bar{\tilde{a}}_{n+1} = \bar{\tilde{a}}_n + \Delta\bar{\tilde{a}}_n$

$$\bar{\tilde{\alpha}}_{n+1} = \bar{\tilde{\alpha}}_n + \Delta\bar{\tilde{\alpha}}_n \quad .$$

6. Using the results from Step 5, compute $\tilde{R}_{n+1}$.

7. Compute $|\tilde{R}_{n+1}|$ and perform the following check :
   if $|\tilde{R}_{n+1}| >$ TOLER , go to Step 2 ;
   otherwise proceed to Step 8.

8. Using the results from Step 5, compute the stress increments using eqns (2.13) and (2.14). This completes the solution.

CHAPTER 5

A PENALTY APPROACH TO THE RATE PROBLEM

In this chapter we focus our attention on the quasi-static rate problem : given the complete history of response of a body at $t = t_1$, we are required to find the response rates corresponding to rates of change of data, also at $t = t_1$. We will derive a variety of variational formulations which are equivalent to the rate boundary-value problem, but we will pay particular attention to the penalty formulation since it is our intention to develop a penalty-finite element approximation for the rate boundary-value problem.

First, we formulate the classical rate problem and then show the equivalence of this formulation to a variety of variational formulations; the minimisation problem of MARTIN (1975b), and the variational inequality of JIANG (1984) are two such formulations, and we also discuss saddle-point and penalty formulations. In demonstrating this equivalence we generalise slightly the treatment of Jiang by distinguishing between elastic and plastic zones (Jiang considers a body which is everywhere plastic) and by dealing with an arbitrary convex, continuously differentiable yield function. In the case of the penalty formulation we show that the penalised solution converges to the exact solution as the penalty parameter approaches zero.

Finally, we derive estimates of the error due to the penalisation and finite element approximation.

5.1 STATEMENT OF THE PROBLEM

Consider an elastic-plastic body which occupies an open bounded domain Ω in R^N, $N \leq 3$, with Lipschitz boundary $\Gamma = \Gamma_u \cup \Gamma_s$. Let $t \in [0,t^\circ]$ be a real parameter which defines a family of stress fields $\underset{\sim}{\sigma}(\underset{\sim}{x},t)$ corresponding to a history of given data. Assume that the stress field $\underset{\sim}{\sigma}(\underset{\sim}{x})$ is known at $t = t^\circ$, and that the domain Ω may be divided into two non-overlapping regions Ω^e and Ω^p as defined in Fig. 5.1.

Suppose now that at $t = t^\circ$ the body force rate $\dot{\underset{\sim}{f}}$ is prescribed on Ω, traction* rates $\dot{\underset{\sim}{t}} \equiv \dot{\underset{\sim\sim}{\sigma}}\underset{\sim}{\nu}$ are prescribed on Γ_s, and velocities $\underset{\sim}{u}$ are prescribed to be zero on Γ_u. Then we seek the velocity field $\underset{\sim}{u}(\underset{\sim}{x})$ and a field of plastic multipliers λ which satisfy

(i) the equations of equilibrium

$$\mathrm{div}\, \dot{\underset{\sim}{\sigma}} + \dot{\underset{\sim}{f}} = \underset{\sim}{0} \qquad \text{on } \Omega \tag{1.1}$$

(ii) the constitutive equations

$$\dot{\underset{\sim}{\sigma}} = \underset{\sim}{C}[\underset{\sim}{E} - \lambda \underset{\sim}{M}] \qquad \text{on } \Omega \tag{1.2}$$

$$\left.\begin{array}{ll} \lambda = 0 & \text{on } \Omega^e \\ \lambda \geq 0\,,\ \kappa \geq 0\,,\ \lambda\kappa = 0 & \text{on } \Omega^p \end{array}\right\} . \tag{1.3}$$

---

* No confusion should arise between the scalar parameter t and the traction rate vector $\dot{\underset{\sim}{t}}$.

(iii) the boundary conditions

$$\dot{\underset{\sim}{\sigma}}\underset{\sim}{\nu} = \dot{\underset{\sim}{t}} \qquad \text{on } \Gamma_s$$

$$\underset{\sim}{u} = \underset{\sim}{0} \qquad \text{on } \Gamma_u \quad . \qquad (1.4)$$

We shall refer to the problem defined in (i) through (iii) above as Problem (S).

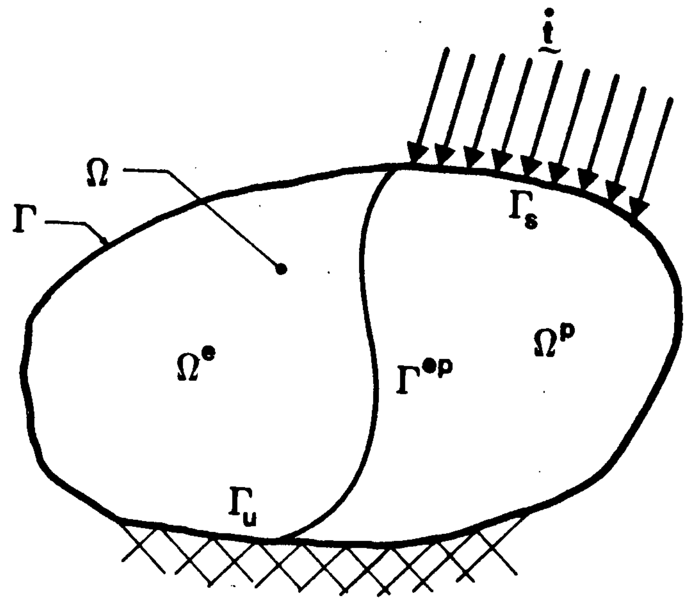
In (1.2) $\underset{\sim}{E} = \frac{1}{2}(\nabla\underset{\sim}{u} + \nabla^T\underset{\sim}{u})$ is the symmetric strain rate tensor, $\underset{\sim}{C}$ is a fourth-order tensor of elastic moduli, and $\underset{\sim}{M}$ is a symmetric second-order tensor defined by

$$\underset{\sim}{M} = \frac{\partial\phi}{\partial\underset{\sim}{\sigma}} \qquad (1.5)$$

where ϕ is the yield function, which is assumed to be convex and continuously differentiable. The reader may wish to review the summary at the end of Section 2.2 for the origins of eqns (1.3) and the definition of the scalar parameter κ, which we repeat here for convenience, using (1.5) :

$$\kappa = \frac{\lambda}{G} - (\underset{\sim}{M} \cdot \dot{\underset{\sim}{\sigma}}) \qquad (1.6)$$

where $G(\underset{\sim}{x})$ is a scalar hardening field which depends on the stress $\underset{\sim}{\sigma}$ and one or more internal variables.



$\Omega^e = \{\underset{\sim}{x} \in \Omega : \phi(\underset{\sim}{\sigma}(\underset{\sim}{x})) < 0\}$

$\Omega^p = \{\underset{\sim}{x} \in \Omega : \phi(\underset{\sim}{\sigma}(\underset{\sim}{x})) = 0\}$

$\Omega = \Omega^e \cup \Omega^p \quad , \quad \Gamma^{ep} = \bar{\Omega}^e \cap \bar{\Omega}^p \quad .$

Figure 5.1 The elastic-plastic body.

Relative to an orthonormal basis the components of $\underset{\sim}{C}$ are $C_{ijk\ell}$; these components exhibit the symmetries

$$C_{ijk\ell} = C_{ij\ell k} = C_{jik\ell} = C_{k\ell ij} \qquad (1.7)$$

and obey the strong ellipticity condition : there exists a positive constant c_1 such that

$$C_{ijk\ell}A_{k\ell}A_{ij} > c_1 A_{ij}A_{ij} \qquad (1.8)$$

for all symmetric second order tensors $\underset{\sim}{A}$. We require further that $C_{ijk\ell} \in L_\infty(\Omega)$ and that there exists a positive constant c_2 such that

$$\max_{i,j,k,\ell} \|C_{ijk\ell}\|_\infty \leq c_2 \quad . \tag{1.9}$$

We assume that the components M_{ij} of $\underset{\sim}{M}$ belong to $L_\infty(\Omega)$, and that positive constants m_1, m_2 exist such that

$$M_{ij}M_{ij} \geq m_1 \quad , \quad \max_{i,j} \|M_{ij}\|_\infty \leq m_2 \quad . \tag{1.10}$$

Returning to the scalar field $G(\underset{\sim}{x})$ in (1.6) we shall require that $G \in L_\infty(\Omega)$ and that $G(\underset{\sim}{x}) \geq 0$ a.e. on Ω . We define a hardening material as one for which positive constants h_1 and h_2 can be found such that

$$\left.\begin{aligned} \|G\|_\infty &\leq 1/h_1 \\ G(\underset{\sim}{x}) &\geq 1/h_2 \end{aligned}\right. \tag{1.11}$$

for all $\underset{\sim}{x} \in \Omega^p$. An elastic-perfectly plastic material is defined as one for which $\|G\|_\infty \to \infty$.

5.2 A VARIATIONAL INEQUALITY

Our objective in this section is to construct a variational principle which is equivalent in some sense to the problem defined in eqns (1.1) through (1.4) of the preceeding section. As we shall see, the variational principle which we develop here will be in the form of a variational inequality, the inequality arising as a consequence of constraints on the plastic multiplier λ defined in eqns (1.3).

Function Spaces

We begin by defining the function spaces within which we propose to work. As in Chapter 3, we adopt the notation and conventions of the Sobolev spaces $H^m(\Omega) \equiv W_2^m(\Omega)$.

The following additional spaces of functions on Ω will be required.

1. The space

$$V = \{\underset{\sim}{v} = (v_1,\dots,v_N) \ : \ v_i \in H^1(\Omega) \ , \quad v_i = 0 \quad \text{on } \Gamma_u\} \ . \tag{2.1}$$

This is a Hilbert space with an inner product

$$(\underset{\sim}{u},\underset{\sim}{v})_V = \int_\Omega \underset{\sim}{u} \cdot \underset{\sim}{v} + \nabla\underset{\sim}{u} \cdot \nabla\underset{\sim}{v} \; dx \tag{2.2}$$

and a norm $\|\underset{\sim}{v}\|_V$ defined by

$$\|\underset{\sim}{v}\|_V^2 = (\underset{\sim}{v},\underset{\sim}{v})_V \ . \tag{2.3}$$

2. The space

$$\Lambda = \{\mu \in L_2(\Omega) \ : \ \mu = 0 \text{ a.e. on } \Omega^e\} \ . \tag{2.4}$$

The space Λ is a Hilbert space with an inner product

$$(\mu,\lambda)_\Lambda = \int_\Omega \mu\lambda \; dx = \int_{\Omega^p} \mu\lambda \; dx \tag{2.5}$$

and a norm $\|\mu\|_\Lambda$ defined by

$$\|\mu\|_\Lambda^2 = (\mu,\mu)_\Lambda \quad . \tag{2.6}$$

3. The convex set

$$K = \{\mu \in \Lambda : \mu \geq 0 \text{ a.e. on } \Omega^p\} \quad . \tag{2.7}$$

For convenience we define the pairs

$$(\underset{\sim}{u},\lambda) = \bar{u} \quad \text{and} \quad (\underset{\sim}{v},\mu) = \bar{v} \quad . \tag{2.8}$$

We shall also make frequent use of the product spaces $\bar{V}$ and $\bar{K}$ defined by

$$\bar{V} = V \times \Lambda \tag{2.9}$$

$$\bar{K} = V \times K \quad . \tag{2.10}$$

The product space $\bar{V}$ is a Hilbert space with norm

$$\|\bar{v}\|_{\bar{V}}^2 = \|\underset{\sim}{v}\|_V^2 + \|\mu\|_\Lambda^2 \tag{2.11}$$

and $\bar{K}$ is a convex subset of $\bar{V}$.

In the definitions of the space Λ and K , and indeed henceforth, we assume meas$(\Omega^e) > 0$, that is, there is always an elastic region present.

Bilinear forms, Functionals and their Properties

We will find it convenient to define the following bilinear forms :

$$a : V \times V \to R \ , \ a(\underset{\sim}{u},\underset{\sim}{v}) = \int_\Omega (\underset{\sim}{C}(\nabla \underset{\sim}{u})) \cdot \nabla \underset{\sim}{v} \, dx$$

$$= \int_\Omega C_{ijk\ell} \, u_{i,j} \, v_{k,\ell} \, dx \quad ; \tag{2.12}$$

$$b : \Lambda \times \Lambda \to R \ , \ b(\lambda,\mu) = \int_\Omega [\{(\underset{\sim}{C}\,\underset{\sim}{M}) \cdot \underset{\sim}{M}\} + \frac{1}{G}]\lambda\mu \, dx$$

$$= \int_\Omega [C_{ijk\ell} \, M_{ij} \, M_{k\ell} + \frac{1}{G}]\lambda\mu \, dx \quad ; \tag{2.13}$$

$$c : V \times \Lambda \to R \ , \ c(\underset{\sim}{v},\mu) = \int_\Omega [(\underset{\sim}{C}\,\underset{\sim}{M}) \cdot \nabla \underset{\sim}{v}]\mu \, dx$$

$$= \int_\Omega C_{ijk\ell} \, M_{ij} \, v_{k,\ell} \, \mu \, dx \quad . \tag{2.14}$$

In view of the symmetry properties of $\underset{\sim}{C}$, a(.,.) and b(.,.) are symmetric bilinear forms. The above bilinear forms may now be combined to form the symmetric bilinear form $A : \bar{V} \times \bar{V} \to R$ defined by

$$A(\bar{u},\bar{v}) = a(\underset{\sim}{u},\underset{\sim}{v}) + b(\lambda,\mu) - c(\underset{\sim}{u},\mu) - c(\underset{\sim}{v},\lambda)$$

$$= \int_\Omega [\{\underset{\sim}{C}(\nabla \underset{\sim}{u} - \lambda \underset{\sim}{M})\} \cdot \{\nabla \underset{\sim}{v} - \mu \underset{\sim}{M}\} + \frac{1}{G}\lambda\mu] \, dx$$

$$= \int_\Omega [C_{ijk\ell}(u_{i,j} - \lambda M_{ij})(v_{k,\ell} - \mu M_{k\ell}) + \frac{1}{G}\lambda\mu] \, dx \quad . \tag{2.15}$$

We will also require the linear functional $f : V \to R$ defined by

$$f(\underset{\sim}{v}) = \int_\Omega \dot{\underset{\sim}{f}} \cdot \underset{\sim}{v} \, dx + \int_{\Gamma_s} \dot{\underset{\sim}{t}} \cdot \underset{\sim}{v} \, dx \ . \tag{2.16}$$

The results which are to be presented later in this section depend on certain basic properties of the above bilinear form, which we now state and prove in the following lemma.

Lemma 5.1

(a) The bilinear form $A(\bar{u},\bar{v})$ in (2.15) is continuous in the sense that there exists a positive constant K such that

$$|A(\bar{u},\bar{v})| \leq K\|\bar{u}\|_{\bar{V}} \|\bar{v}\|_{\bar{V}} \quad . \tag{2.17}$$

(b) The bilinear form $A(\bar{u},\bar{v})$ is $\bar{V}$-elliptic in that there exists a positive constant α such that

$$A(\bar{v},\bar{v}) \geq \alpha\|\bar{v}\|^2_{\bar{V}} \quad . \tag{2.18}$$

Proof

(a) The bilinear form $a(\underset{\sim}{u},\underset{\sim}{v})$ in (2.12) has already been shown to be continuous (see Section 3.2, Lemma 3.1), so that for $K_1 > 0$ we may write

$$|a(\underset{\sim}{u},\underset{\sim}{v})| \leq K_1\|\underset{\sim}{u}\|_V \|\underset{\sim}{v}\|_V \quad .$$

We consider next the bilinear form $b(\lambda,\mu)$: from (2.13) we have

$$|b(\lambda,\mu)| = \left|\int_\Omega C_{ijk\ell}M_{ij}M_{k\ell}\lambda\mu \, dx + \int_\Omega \frac{1}{G} \lambda\mu \, dx\right|$$

$$\leqslant \left| c_2 \int_\Omega \lambda\mu (\sum_{i,j} M_{ij})(\sum_{k,\ell} M_{k\ell})\ dx + h_2 \int_\Omega \lambda\mu\ dx \right| \quad ,$$

using (1.9) and (1.11)

$$= (c_2 N^4 m_2^2 + h_2) \left| \int_\Omega \lambda\mu\ dx \right|$$

$$\leqslant K_2 \|\lambda\|_\Lambda \ \|\mu\|_\Lambda$$

using the Schwarz inequality on Λ , and where $K_2 = c_2 N^4 m_2^2 + h_2$. Now consider the bilinear form $c(\underset{\sim}{v}, \mu)$: from (2.14) we have

$$|c(\underset{\sim}{v}, \mu)| = \left| \int_\Omega C_{ijk\ell} M_{ij} v_{k,\ell} \mu\ dx \right|$$

$$\leqslant \left| c_2 \int_\Omega (\sum_{i,j} M_{ij})(\sum_{k,\ell} v_{k,\ell}) \mu\ dx \right| \quad , \quad \text{using (1.9)}$$

$$\leqslant \left| c_2 N^2 m_2 \int_\Omega \mu (\sum_{k,\ell} v_{k,\ell})\ dx \right| \quad , \quad \text{using (1.10)}$$

$$= c_2 N^2 m_2 \sum_{k,\ell} \left| \int_\Omega \mu\ v_{k,\ell}\ dx \right|$$

$$\leqslant c_2 N^2 m_2 \sum_{k,\ell} \|\mu\|_o \ \|v_{k,\ell}\|_o$$

using the Schwarz inequality on $L_2(\Omega)$ where $\|\cdot\|_o$ is the L_2-norm. It follows from the definition of the norms on V and Λ that

$$|c(\underset{\sim}{v}, \mu)| \leqslant K_3 \|\mu\|_\Lambda \ \|\underset{\sim}{v}\|_V$$

where $K_3 = c_2 N^4 m_2$.

Let $K = 2\max(K_1,K_2,K_3)$. Then, from (2.15) and making use of the above results we have

$$|A(\bar{u},\bar{v})| \le \frac{1}{2} K(\|\underset{\sim}{u}\|_V \|\underset{\sim}{v}\|_V + \|\lambda\|_\Lambda \|\mu\|_\Lambda + \|\underset{\sim}{u}\|_V \|\mu\|_\Lambda + \|\underset{\sim}{v}\|_V \|\lambda\|_\Lambda)$$

$$= \frac{1}{2} K(\|\underset{\sim}{u}\|_V + \|\lambda\|_\Lambda)(\|\underset{\sim}{v}\|_V + \|\mu\|_\Lambda)$$

$$\le K\|\bar{u}\|_{\bar{V}} \|\bar{v}\|_{\bar{V}}$$

where we have used the inequality $\alpha + \beta \le \sqrt{2}(\alpha^2 + \beta^2)^{1/2}$, $\alpha,\beta \in R$, to establish the final result. □

(b) We follow a similar proof due to JIANG (1984). Using the identity $\underset{\sim}{C}(\nabla \underset{\sim}{u}) = \underset{\approx}{C}\underset{\approx}{E}$ (Section 3.1, eqn (1.15)) in (2.15), we have

$$A(\bar{v},\bar{v}) = \int_\Omega \underset{\sim}{C}(\underset{\sim}{E} - \lambda \underset{\sim}{M}) \cdot (\underset{\sim}{E} - \lambda \underset{\sim}{M})\, dx + \int_\Omega \frac{1}{G} \lambda^2\, dx$$

$$\ge c_1 \int_\Omega (\underset{\sim}{E} - \lambda \underset{\sim}{M}) \cdot (\underset{\sim}{E} - \lambda \underset{\sim}{M})\, dx + \int_\Omega \frac{1}{G} \lambda^2\, dx \quad , \quad \text{using (1.8)}$$

$$= c_1 \int_\Omega \Big\{\theta(\underset{\sim}{E} \cdot \underset{\sim}{E}) + (\sqrt{1-\theta}\, \underset{\sim}{E} - \frac{1}{\sqrt{1-\theta}} \lambda \underset{\sim}{M}) \cdot (\sqrt{1-\theta}\, \underset{\sim}{E} - \frac{1}{\sqrt{1-\theta}} \lambda \underset{\sim}{M})$$

$$- \frac{\theta}{1-\theta} \lambda^2 (\underset{\sim}{M} \cdot \underset{\sim}{M})\Big\}\, dx + \int_\Omega \frac{1}{G} \lambda^2\, dx \quad , \quad 0 < \theta < 1$$

$$\ge c_1 \theta \int_\Omega \underset{\sim}{E} \cdot \underset{\sim}{E}\, dx - \frac{c_1 \theta}{1-\theta} \int_\Omega \lambda^2 (\underset{\sim}{M} \cdot \underset{\sim}{M})\, dx + \int_\Omega \frac{1}{G} \lambda^2\, dx$$

$$\ge c_1 \theta \int_\Omega \underset{\sim}{E} \cdot \underset{\sim}{E}\, dx - \frac{c_1 N^2 m_2^2 \theta}{1-\theta} \int_\Omega \lambda^2\, dx + \int_\Omega \frac{1}{G} \lambda^2\, dx \quad , \quad \text{using (1.10)}$$

$$\ge c_1 \theta \int_\Omega \underset{\sim}{E} \cdot \underset{\sim}{E}\, dx + (h_1 - \frac{c_1 N^2 m_2^2 \theta}{1-\theta}) \int_\Omega \lambda^2\, dx \quad , \quad \text{using (1.11)} \quad .$$

Choose $\theta = \dfrac{h_1}{2c_1N^2m_2^2 + h_1}$; then we have

$$A(\bar{v},\bar{v}) \geq \frac{c_1h_1}{2c_1N^2m_2^2 + h_1} \int_\Omega \underset{\sim}{E} \cdot \underset{\sim}{E}\, dx + \frac{1}{2} h_1 \int_\Omega \lambda^2\, dx$$

$$\geq \frac{kc_1h_1}{2c_1N^2m_2^2 + h_1} \|\underset{\sim}{v}\|_V^2 + \frac{1}{2} h_1 \|\lambda\|_\Lambda^2$$

using Korn's inequality (Section 3.2, eqn (2.15)) on the first term on the right-hand-side, where $k > 0$ is Korn's constant. Hence, using (2.11) we obtain

$$A(\bar{v},\bar{v}) \geq \alpha\|\bar{v}\|_{\bar{V}}^2$$

where $\alpha = \min[\dfrac{kc_1h_1}{2c_1N^2m_2^2 + h_1}, \dfrac{1}{2} h_1]$, and the result is established. It should be noted that the result holds if and only if $h_1 > 0$, i.e. if the material exhibits hardening behaviour (see the remark following eqn (1.11)). □

A Variational Inequality

We propose now to establish a relationship between the classical formulation given in (1.1) through (1.4) and a variational inequality. This we do in the following theorem.

Theorem 5.1

Let $\bar{u} = (\underset{\sim}{u}, \lambda)$ be a solution of Problem (S) . Then $\bar{u}$ is also a solution of the variational inequality

$$A(\bar{u}, \bar{v} - \bar{u}) - f(\underset{\sim}{v} - \underset{\sim}{u}) \geq 0 \quad , \qquad \bar{v} \in \bar{K} \quad . \tag{2.19}$$

Conversely, if $\bar{u}$ is a solution of (2.19) then it also satisfies (S) in a weak sense. We refer to eqn (2.19) as Problem (V) .

Proof

(i) (S) $\Rightarrow$ (V) :

We assume that the equilibrium equations (1.1) are satisfied on Ω . Then for arbitrary $\bar{v} \in \bar{K}$ we have

$$\int_\Omega (\mathrm{div}\, \dot{\sigma} + \dot{f}) \cdot (\underset{\sim}{v} - \underset{\sim}{u})\, dx = 0 \quad .$$

Using Green's theorem, we get

$$\int_{\Gamma_s} (\underset{\approx}{\dot{\sigma}} \underset{\sim}{\nu}) \cdot (\underset{\sim}{v} - \underset{\sim}{u})\, dx - \int_\Omega \nabla(\underset{\sim}{v} - \underset{\sim}{u}) \cdot \underset{\approx}{\dot{\sigma}}\, dx + \int_\Omega \underset{\sim}{\dot{f}} \cdot (\underset{\sim}{v} - \underset{\sim}{u})\, dx = 0 \quad . \tag{2.20}$$

Using (1.2) in (2.15), we get

$$A(\bar{u}, \bar{v}) = \int_\Omega \underset{\approx}{\dot{\sigma}} \cdot (\nabla \underset{\sim}{v} - \mu \underset{\approx}{M})\, dx + \int_\Omega \frac{1}{G} \mu \lambda\, dx$$

$$= \int_\Omega \underset{\approx}{\dot{\sigma}} \cdot \nabla \underset{\sim}{v}\, dx + \int_\Omega \mu \kappa\, dx \quad , \quad \text{using (1.6)} \quad .$$

Replacing $\bar{v}$ by $\bar{v} - \bar{u}$ in the above we get

$$A(\bar{u},\bar{v} - \bar{u}) = \int_\Omega \dot{\underset{\sim}{\sigma}} \cdot \nabla(\underset{\sim}{v} - \underset{\sim}{u})\, dx + \int_\Omega (\mu - \lambda)\kappa\, dx \quad .$$

Making use of this result in (2.20) and noting that $\dot{\underset{\sim\sim}{\sigma}}\nu = \dot{\underset{\sim}{t}}$ on Γ_s we get

$$\int_{\Gamma_s} \dot{\underset{\sim}{t}} \cdot (\underset{\sim}{v} - \underset{\sim}{u})\, dx - A(\bar{u},\bar{v} - \bar{u}) + \int_\Omega (\mu - \lambda)\kappa\, dx + \int \dot{\underset{\sim}{f}} \cdot (\underset{\sim}{v} - \underset{\sim}{u})\, dx = 0 \quad .$$

From (1.3) we have $\lambda\kappa = 0$; but since $\mu \geq 0$ is arbitrary, $\mu\kappa \geq 0$ and it follows that

$$A(\bar{u},\bar{v} - \bar{u}) - f(\underset{\sim}{v} - \underset{\sim}{u}) \geq 0 \quad .$$

(ii) (V) => (S) :

We define the space

$$\mathscr{A}(\Omega^P)^+ = \{\eta \in C_0^\infty(\Omega^P) \;:\; \eta(\underset{\sim}{x}) \geq 0 \quad , \quad \underset{\sim}{x} \in \Omega^P\} \quad .$$

In part (i) above we showed that (2.19) may be written as

$$\int_\Omega (\dot{\underset{\sim}{\sigma}} \cdot \nabla(\underset{\sim}{v} - \underset{\sim}{u})\, dx + \int_\Omega (\mu - \lambda)\kappa\, dx - f(\dot{\underset{\sim}{v}} - \underset{\sim}{u}) \geq 0 \quad . \qquad (2.21)$$

Setting $\underset{\sim}{v} = \underset{\sim}{u}$ and $\mu = \lambda + \xi$ in (2.21), we get

$$\int_{\Omega^P} \xi\kappa\, dx \geq 0 \quad , \qquad \xi \in \mathscr{A}(\Omega^P)^+$$

and so $\kappa \geq 0$ in Ω^P .

Next, set $\underset{\sim}{u} = \underset{\sim}{v}$ and $\mu = 0$ in (2.21) : then

$$- \int_{\Omega^P} \lambda\kappa \, dx \geq 0$$

$$\Rightarrow \quad \int_{\Omega^P} \lambda\kappa \leq 0 \quad .$$

But $\lambda \geq 0$ in Ω^P , and $\kappa \geq 0$ in Ω^P . Hence

$$\lambda\kappa = 0 \qquad \text{in } \Omega^P \quad .$$

Finally, we set $\mu = \lambda$ and $\underset{\sim}{v} = \underset{\sim}{u} \pm \underset{\sim}{w}$, $\underset{\sim}{w} \in [C_0^\infty(\Omega)]^N$, in (2.21) to get

$$\int_\Omega \dot{\underset{\sim}{\sigma}} \cdot \nabla\underset{\sim}{w} \, dx - \int_{\Gamma_s} \dot{\underset{\sim}{t}} \cdot \underset{\sim}{w} \, dx - \int_\Omega \dot{\underset{\sim}{f}} \cdot \underset{\sim}{w} \, dx \geq 0$$

and

$$- \int_\Omega \dot{\underset{\sim}{\sigma}} \cdot \nabla\underset{\sim}{w} \, dx + \int_{\Gamma_s} \dot{\underset{\sim}{t}} \cdot \underset{\sim}{w} \, dx + \int_\Omega \dot{\underset{\sim}{f}} \cdot \underset{\sim}{w} \geq 0 \quad .$$

Combining the above two inequalities and using Green's theorem in reverse we get

$$\int_\Omega (\text{div } \dot{\underset{\sim}{\sigma}} - \dot{\underset{\sim}{f}}) \cdot \underset{\sim}{w} \, dx = 0$$

$$\Rightarrow \quad \text{div } \dot{\underset{\sim}{\sigma}} - \dot{\underset{\sim}{f}} = \underset{\sim}{0} \quad .$$

in the sense of distributions. □

Theorem 5.1 represents the principle of virtual work for an elastic-plastic body whose material behaviour is governed by the rate constitutive equations given in (1.2) and (1.3). Thus $A(\bar{u},\bar{v} - \bar{u})$ represents the rate at which work is done by the stresses $\dot{\underset{\sim}{\sigma}}$ in moving through the so-called virtual displacements $\underset{\sim}{v} - \underset{\sim}{u}$ (strictly speaking, displacement rates), part of which is stored as elastic strain energy and part of which is dissipated in plastic straining. Obviously $f(\underset{\sim}{v} - \underset{\sim}{u})$ represents the rate at which work is done by the applied loads. The inequality in (2.19) is directly attributable to the inequality constraints on λ given in (1.3), as can be seen from the proof (S) => (V) above.

We note here for later reference that the variational inequality (2.19) may be written in the following alternative form :

$$a(\underset{\sim}{u},\underset{\sim}{v}) - c(\underset{\sim}{v},\lambda) = f(\underset{\sim}{v}) \quad , \qquad \underset{\sim}{v} \in V \qquad (2.22a)$$

$$- c(\underset{\sim}{u},\mu - \lambda) + b(\lambda,\mu - \lambda) \geq 0 \quad , \qquad \mu \in K \quad . \qquad (2.22b)$$

Eqn (2.22a) is obtained by setting $\mu = \lambda$ and $\underset{\sim}{v} = \underset{\sim}{u} \pm \underset{\sim}{w}$ in eqn (2.19), $\underset{\sim}{w} \in [C_o^\infty(\Omega)]^N$; eqn (2.22b) is obtained by setting $\underset{\sim}{v} = \underset{\sim}{u}$ in eqn (2.19).

Problem (V), involving as it does an inequality, is not suitable as a basis for any of the conventional methods of approximation. We therefore investigate an alternative variational principle in the next section.

5.3 A MINIMISATION PROBLEM

We propose now to make use of a standard result from convex analysis (see, for example, EKELAND and TEMAM (1976)), to establish the connection between the classical problem (S), the variational inequality (V), and the minimisation problem of MARTIN (1975b). In doing so we will define an alternative minimisation problem which we will refer to as Problem (M).

We introduce the functional $J(\bar{v}) : \bar{V} \to R$ defined by

$$J(\bar{v}) = \frac{1}{2} A(\bar{v},\bar{v}) - f(\underset{\sim}{v}) \tag{3.1}$$

where $A(.,.)$ is defined in (2.15) and $f(.)$ in (2.16). Before proceeding with the statement of the minimisation problem we shall establish some fundamental properties of the functional $J(\bar{v})$ in the following lemma.

Lemma 5.2

(a) The functional $J(\bar{v})$ defined in (3.1) is strictly convex.

(b) The functional $J(\bar{v})$ is coercive in the sense that $J(\bar{v}) \to +\infty$ as $\|\bar{v}\|_{\bar{V}} \to \infty$.

Proof

(a) We may write, for $0 < \theta < 1$,

$$J(\theta\bar{u} + (1-\theta)\bar{v}) = \frac{1}{2}\,\theta^2 A(\bar{u},\bar{u}) + \theta(1-\theta)A(\bar{u},\bar{v})$$

$$+ \frac{1}{2}\,(1-\theta)^2 A(\bar{v},\bar{v}) - \theta f(\underset{\sim}{u}) - (1-\theta)f(\underset{\sim}{v}) \quad . \tag{3.2}$$

From the $\bar{V}$-ellipticity of A(.,.), Lemma 5.1, we have

$$A(\bar{u} - \bar{v},\bar{u} - \bar{v}) > 0$$

$$\Rightarrow \quad A(\bar{u},\bar{v}) < \frac{1}{2}[A(\bar{u},\bar{u}) + A(\bar{v},\bar{v})] \quad . \tag{3.3}$$

Using (3.3) in (3.2) we get

$$J(\theta\bar{u} + (1-\theta)\bar{v}) < \frac{1}{2}\,\theta^2 A(\bar{u},\bar{u}) + \frac{1}{2}\,\theta(1-\theta)[A(\bar{u},\bar{u}) + A(\bar{v},\bar{v})]$$

$$+ \frac{1}{2}\,(1-\theta)^2 A(\bar{v},\bar{v}) - \theta f(\underset{\sim}{u}) - (1-\theta)f(\underset{\sim}{v})$$

$$< \theta J(\bar{u}) + (1-\theta)J(\bar{v})$$

which establishes the result. □

(b) Using identical arguments to those used in the second part of Lemma 3.2 (Section 3.2) concerning the continuity of the data $\underset{\sim}{f}$ and $\underset{\sim}{t}$ we may obtain the following result (which parallels that given in eqn (2.27), Section 3.2) :

$$|f(\underset{\sim}{v})| \leqslant m\,\|\underset{\sim}{v}\|_V \quad , \quad m > 0 \quad .$$

It follows that

$$- m \leq \frac{f(\underset{\sim}{v})}{\|\underset{\sim}{v}\|_V} \leq m$$

from which we see that

$$f(\underset{\sim}{v}) \geq - m \|\underset{\sim}{v}\|_V \quad .$$

Dividing both sides by $\|\bar{v}\|_{\bar{V}}$, we get

$$\frac{f(\underset{\sim}{v})}{\|\bar{v}\|_{\bar{V}}} \geq - m \frac{\|\underset{\sim}{v}\|_V}{\|\bar{v}\|_{\bar{V}}}$$

$$= \frac{- m \|\underset{\sim}{v}\|_V}{[\|\underset{\sim}{v}\|_V^2 + \|\mu\|_\Lambda^2]^{1/2}}$$

$$\geq - m \quad .$$

Making use of the above result together with the $\bar{V}$-ellipticity of $A(\cdot,\cdot)$, Lemma 5.1(b), in eqn (3.1), we get

$$J(\bar{v}) \geq \frac{1}{2} \alpha \|\bar{v}\|_{\bar{V}}^2 - m \|\bar{v}\|_{\bar{V}} \quad .$$

Dividing both sides by $\|\bar{v}\|_{\bar{V}}$, we get

$$\frac{J(\bar{v})}{\|\bar{v}\|_{\bar{V}}} \geq \frac{1}{2} \alpha \|\bar{v}\|_{\bar{V}} - m$$

$$\to + \infty \text{ as } \|\bar{v}\|_{\bar{V}} \to \infty \text{ , iff } \alpha > 0 \quad .$$

We emphasise again that the coercivity result is only valid for materials which exhibit hardening (see Lemma 5.1(b)). We now establish the connection between the variational inequality (V) and a minimisation problem (M) in the following theorem.

Theorem 5.2A

Let $J : \bar{V} \to R$ be a proper, convex, Gateaux differentiable functional. Then if $\bar{u} \in \bar{K}$ the following two conditions are equivalent :

$$J(\bar{u}) \leq J(\bar{v}) \quad , \qquad \bar{v} \in \bar{K} \qquad \text{(Problem (M))} \tag{3.5}$$

$$\langle DJ(\bar{u}) \, , \, \bar{v} - \bar{u} \rangle \geq 0 \quad , \qquad \bar{v} \in \bar{K} \tag{3.6}$$

where $DJ(\bar{u})$ is the Gateaux derivative of J at $\bar{u}$ and $\langle .,. \rangle$ denotes duality pairing on $\bar{V}' \times \bar{V}$, $\bar{V}'$ being the dual space of $\bar{V}$. □

The proof of this theorem may be found in EKELAND and TEMAM (1976), Chapter 2, Proposition 2.1. It remains to show that the conditions of the theorem are satisfied. Certainly J is differentiable with Gateaux derivative $DJ(\bar{u}) : \bar{V} \to \bar{V}'$ given by

$$\langle DJ(\bar{u}),\bar{v} \rangle = A(\bar{u},\bar{v}) - f(\underset{\sim}{v}) \quad . \tag{3.7}$$

Furthermore, J is strictly convex (Lemma 5.2(a)), so that the result is established.

Theorem 5.2 establishes the conditions under which the solution $\bar{u} \in \bar{K}$ of the variational inequality (V) is also a solution of the minimisation problem (M) . The conditions for existence and uniqueness of the solution are established in the following theorem.

Theorem 5.2B

Let $J(\bar{v})$ satisfy the conditions of Theorem 5.2A and in addition, let $J(\bar{v})$ be coercive. Then the minimisation problem (M) has at least one solution. Moreover, if $J(\bar{v})$ is strictly convex then this solution is unique. □

The proof of this theorem can be found in EKELAND and TEMAM (1976), Chapter 2, Proposition 1.2. Since $J(\bar{v})$ has already been shown to be coercive provided the material exhibits hardening, (Lemma 5.2(b)), and strictly convex (Lemma 5.2(a)), we are assured of the existence of a unique solution to Problem (M). We summarise these results for the rate problem in the following theorem.

Theorem 5.3

There exists a unique minimiser $\bar{u}$ of the functional J of (3.1). Moreover, $\bar{u}$ is also the solution of the variational inequality (2.19). □

In the following section we digress slightly from the main theme of our development to discuss a saddle-point formulation of the problem. We will find that the ideas expressed in this formulation play an important part in developing error estimates for our numerical approximation.

5.4 A SADDLE-POINT PROBLEM

Yet another way of formulating the weak problem is to use a saddle-point formulation, in which the constraint on λ is removed by introducing a field of Lagrange multipliers. Accordingly we consider the following problem : find $\underset{\sim}{u} \in V$, $\lambda \in \Lambda$ and the Lagrange multiplier $k \in K$ which satisfy

$$L(\bar{u},\ell) \leqslant L(\bar{u},k) \leqslant L(\bar{v},k)$$

for all $\bar{v} \in \bar{V}$ and $\ell \in K$, where the Lagrangian L is defined by

$$L(\bar{v},\ell) = J(\bar{v}) - (\ell,\mu)_\Lambda \tag{4.1}$$

and $(.,.)_\Lambda$ denotes the inner product on Λ (that is, the L_2-inner product on Ω^P) . Note that ℓ here is an arbitrary Lagrange multiplier.

Minimisation of (4.1) in $\bar{V}$ and maximisation over all $\ell \in K$ is equivalent to finding $\bar{u} \in \bar{V}$ and $k \in K$ which satisfy

$$A(\bar{u},\bar{v}) - (k,\mu)_\Lambda = f(\underset{\sim}{v}) \quad , \qquad \bar{v} \in \bar{V} \tag{4.2}$$

or equivalently,

$$a(\underset{\sim}{u},\underset{\sim}{v}) - c(\lambda,\underset{\sim}{v}) = f(\underset{\sim}{v}) \tag{4.3a}$$

and

$$- c(\mu,\underset{\sim}{u}) + b(\lambda,\mu) - (k,\mu)_\Lambda = 0 \quad , \tag{4.3b}$$

together with the inequality

$$(\ell - k,\lambda)_\Lambda \geqslant 0 \quad , \tag{4.4}$$

for all $\underset{\sim}{v} \in V$, $\mu \in \Lambda$ and $\ell \in K$.

The relationship between this problem and the constrained minimisation problem (3.5) follows again from standard results (see, for example, EKELAND and TEMAM (1976)), and is summarised in the following theorem.

Theorem 5.4

The functions $\underset{\sim}{u}$ and λ are the solution of the constrained minimisation problem (M) if and only if $\underset{\sim}{u}$, λ and k are the solution of the mixed problem defined in (4.3) and (4.4). Furthermore λ and k satisfy the conditions

$$\lambda \geqslant 0 \quad , \quad k \geqslant 0 \quad , \quad k\lambda = 0 \quad . \quad \square \tag{4.5}$$

We note that conditions (4.5) are identical to conditions $(1.3)_2$ in the classical statement of the problem when the Lagrange multiplier k are interpreted as the variable κ , defined in (1.6). Lagrange multipliers generally have a physical interpretation and are indeed very often an integral part of the solution; for example, in problems involving incompressible materials the Lagrange multiplier appearing in a mixed formulation is a hydrostatic pressure while in problems of unilateral contact it is the contact pressure. In the present problem, though, the function κ does not appear to be readily interpretable.

We do not propose to use the saddle-point formulation as the basis for a numerical approximation. We return instead to the minimisation problem (M) and effect a perturbation of the formulation which can then be used as a basis for a numerical approximation.

5.5 A PERTURBED MINIMISATION PROBLEM

As discussed earlier, the minimisation problem (M) amounts to seeking a minimum of J in a convex subset $\bar{K}$ of $\bar{V}$. Though there do exist algorithms which carry out this process numerically (for example, the many quadratic programming algorithms), it is clearly of interest to be able to formulate the problem in such a way that a minimum is sought in $\bar{V}$. This may be achieved by penalisation of J , which we now propose to discuss.

The Penalty Functional

We introduce a convex, differentiable functional $j : \Lambda \to R$ with the property that

$$j(\mu) \geq 0 \tag{5.1a}$$

$$j(\mu) = 0 \quad , \quad \text{if and only if } \mu \in K \ . \tag{5.1b}$$

For problems involving constraint sets of the form of (1.3) GLOWINSKI, LIONS and TREMOLIERES (1981) have suggested the penalty functional

$$j(\mu) = \frac{1}{2} \int_{\Omega^P} g(\underset{\sim}{x})[\mu_-(\underset{\sim}{x})]^2 \, dx \tag{5.2}$$

where $\mu_- \in L_2(\Omega^P)$ is the function defined by

$$\mu_-(\underset{\sim}{x}) = \frac{1}{2} [\mu(\underset{\sim}{x}) - |\mu(\underset{\sim}{x})|] \quad , \quad \underset{\sim}{x} \in \Omega^P \tag{5.3}$$

and $g(\underset{\sim}{x}) \geq 0$ is a continuous scalar function. We note that $\mu_-(\underset{\sim}{x}) = \min(0,\mu(\underset{\sim}{x}))$ for $\underset{\sim}{x} \in \Omega^P$ and that this choice of penalty functional clearly satisfies eqns (5.1). In the following lemma we show that $j(\mu)$ is convex and differentiable.

<u>Lemma 5.3</u>

(a) The penalty functional $j(\mu)$ defined in (5.2) is differentiable, with Gateaux derivative given by

$$\langle Dj(\lambda),\mu\rangle = \int_{\Omega^P} g(\underset{\sim}{x})\lambda_-(\underset{\sim}{x})\mu(\underset{\sim}{x}) \, dx \quad . \tag{5.4}$$

(b) The penalty functional $j(\mu)$ is convex.

<u>Proof</u>

(a) Let $\text{sgn} : L_2(\Omega^P) \to L_2(\Omega^P)$ be defined by

$$(\text{sgn}\,\lambda)(\underset{\sim}{x}) = \begin{cases} +1 & \text{if } \lambda(\underset{\sim}{x}) > 0 \\ -1 & \text{if } \lambda(\underset{\sim}{x}) < 0 \end{cases} \qquad \text{for } \underset{\sim}{x} \in \Omega^P \quad .$$

Let us also define for convenience the operator $F : L_2(\Omega^p) \to L_2(\Omega^p)$ by

$$F(\lambda) = 2\lambda_- \tag{5.5}$$

where $\lambda_-(\underset{\sim}{x})$ is defined in (5.3). Then for $\theta \in R$ and some $\underset{\sim}{x} \in \Omega^p$ we have

$$G(\underset{\sim}{x},\theta) \equiv [F(\lambda + \theta\mu)]^2(\underset{\sim}{x}) - [F(\lambda)]^2(\underset{\sim}{x}) = 4\theta\lambda(\underset{\sim}{x})\mu(\underset{\sim}{x}) + 2\theta^2\mu^2(\underset{\sim}{x})$$

$$+ 2\lambda^2(\underset{\sim}{x})\mathrm{sgn}(\lambda(\underset{\sim}{x})) - 2(\lambda(\underset{\sim}{x}) + \theta\mu(\underset{\sim}{x}))^2\mathrm{sgn}(\lambda(\underset{\sim}{x}) + \theta\mu(\underset{\sim}{x})) \quad . \tag{5.6}$$

Now suppose that $\lambda(\underset{\sim}{x}) > 0$: then for sufficiently small θ, $\lambda(\underset{\sim}{x}) + \theta\mu(\underset{\sim}{x}) > 0$ also, and $G(\underset{\sim}{x},\theta) = 0$. Next, suppose that $\lambda(\underset{\sim}{x}) < 0$: then for sufficiently small θ, $\lambda(\underset{\sim}{x}) + \theta\mu(\underset{\sim}{x}) < 0$ also, and we find that $G(\underset{\sim}{x},\theta) = 8\theta\lambda(\underset{\sim}{x})\mu(\underset{\sim}{x}) + 4\theta^2\mu^2(\underset{\sim}{x})$. Finally, if $\lambda(\underset{\sim}{x}) = 0$ then $G(\underset{\sim}{x},\theta) = \theta^2(\mu(\underset{\sim}{x}) - |\mu(\underset{\sim}{x})|)^2$.

Now let $\Omega^p = \Omega^p_1 \cup \Omega^p_2 \cup \Omega^p_3$, where

$$\Omega^p_1 = \{\underset{\sim}{x} \in \Omega^p : \lambda(\underset{\sim}{x}) > 0\}$$

$$\Omega^p_2 = \{\underset{\sim}{x} \in \Omega^p : \lambda(\underset{\sim}{x}) < 0\}$$

$$\Omega^p_3 = \{\underset{\sim}{x} \in \Omega^p : \lambda(\underset{\sim}{x}) = 0\} \quad .$$

Multiplying eqn (5.6) by $g(\underset{\sim}{x})/8$ and integrating over Ω^p , we obtain, for sufficiently small θ ,

$$\frac{1}{8}\int_{\Omega^p} g(\underset{\sim}{x})\, G(\underset{\sim}{x},\theta)\, dx = \int_{\Omega^p_1} 0\, dx + \frac{1}{8}\int_{\Omega^p_2} g(\underset{\sim}{x})[8\theta\lambda(\underset{\sim}{x})\mu(\underset{\sim}{x}) + 4\theta^2\mu^2(\underset{\sim}{x})]\, dx$$

$$+ \frac{1}{8}\int_{\Omega^p_3} g(\underset{\sim}{x})\theta^2(\mu(\underset{\sim}{x}) - |\mu(\underset{\sim}{x})|)^2\, dx \quad .$$

Making use of eqns (5.2) and (5.5) on the left-hand-side, and then dividing both sides by θ and taking $\lim_{\theta \to 0}$ we get

$$\lim_{\theta \to 0} \frac{1}{\theta} [j(\lambda(\underset{\sim}{x}) + \theta\mu(\underset{\sim}{x})) - j(\lambda(\underset{\sim}{x}))] = \int_{\Omega_2^P} g(\underset{\sim}{x})\lambda(\underset{\sim}{x})\mu(\underset{\sim}{x})\, dx$$

$$\Rightarrow \quad dj(\lambda(\underset{\sim}{x}),\mu(\underset{\sim}{x})) = \int_{\Omega^P} g(\underset{\sim}{x})\lambda_-(\underset{\sim}{x})\mu(\underset{\sim}{x})\, dx \tag{5.7}$$

where $dj(\lambda,\mu)$ denotes the Gateaux differential of j, which is clearly linear in $\mu(\underset{\sim}{x})$. Also, we have

$$\left|dj(\lambda(\underset{\sim}{x}),\mu(\underset{\sim}{x}))\right| = \left|\int_{\Omega^P} g(\underset{\sim}{x})\lambda_-(\underset{\sim}{x})\mu(\underset{\sim}{x})\, dx\right|$$

$$\leqslant \|g(\underset{\sim}{x})\lambda_-(\underset{\sim}{x})\|_o \; \|\mu(\underset{\sim}{x})\|_o$$

$$\leqslant C\|\mu(\underset{\sim}{x})\|_o$$

for some constant $C > 0$. Hence, since $j : \Lambda \to R$, $dj(\lambda,\cdot)$ defines a bounded linear functional and we may write $dj(\lambda,\cdot) \equiv \langle Dj(\lambda),\mu\rangle$, where $Dj(\lambda)$ is the Gateaux derivative of j at λ. Thus, from (5.7) we have

$$\langle Dj(\lambda),\mu\rangle = \int_{\Omega^P} g(\underset{\sim}{x})\lambda_-(\underset{\sim}{x})\mu(\underset{\sim}{x})\, dx \quad . \quad \square$$

(b) Using the definition (5.3) it is a simple matter to show that for $\alpha,\beta \in R$

$$(\alpha_- - \beta_-)(\alpha - \beta) - (\alpha_- - \beta_-)^2 = -(\alpha\beta)_- \geqslant 0 \quad . \tag{5.8}$$

From the definition of the Gateaux derivative of j (Lemma 5.3(a)) we have, at $\underset{\sim}{x} \in \Omega^p$ (it is assumed that all functions appearing below are evaluated at $\underset{\sim}{x}$) ,

$$\langle Dj(\lambda) - Dj(\mu), \lambda - \mu\rangle = \langle Dj(\lambda), \lambda - \mu\rangle - \langle Dj(\mu), \lambda - \mu\rangle$$

$$= \int_{\Omega^p} g(\underset{\sim}{x})\lambda_-(\lambda - \mu)\, dx - \int_{\Omega^p} g(\underset{\sim}{x})\mu_-(\lambda - \mu)\, dx$$

$$= \int_{\Omega^p} g(\underset{\sim}{x})(\lambda_- - \mu_-)(\lambda - \mu)\, dx \quad .$$

$$\geqslant \int_{\Omega^p} g(\underset{\sim}{x})(\lambda_- - \mu_-)^2\, dx \quad , \quad \text{using (5.8)}$$

$$\geqslant 0$$

since $g(\underset{\sim}{x}) \geqslant 0$. Thus, the Gateaux derivative $Dj(\mu)$ is a monotone mapping from Λ into Λ' from which it follows that $j(\mu)$ is convex (see EKELAND and TEMAM (1976), Chapter 1, Proposition 5.5). □

<u>The Perturbed Minimisation Problem</u>

We now make use of the penalty functional $j(\mu)$ to construct the perturbed functional $J_\varepsilon : \bar{V} \to R$ defined by

$$J_\varepsilon(\bar{v}) = J(\bar{v}) + \varepsilon^{-1} j(\mu) \qquad (5.9)$$

where $J(\bar{v})$ is defined in (3.1), and $\varepsilon > 0$ is called the penalty parameter. We note that since $J(\bar{v})$ has already been shown to be strictly convex, differentiable and coercive, it follows from Lemma 5.3

that $J_\varepsilon(\bar{v})$ must also be strictly convex, differentiable and coercive. We may thus again make use of the familiar results of convex analysis to establish that a unique solution exists to the perturbed minimisation problem (defined below), and that this solution is characterised by a variational equality. Since these developments parallel closely those of the minimisation problem (M), we simply summarise the pertinent results in the following theorem.

Theorem 5.5

For each $\varepsilon > 0$ there exists a unique solution $\bar{u}_\varepsilon \in \bar{V}$ of the perturbed minimisation problem

$$J_\varepsilon(\bar{u}_\varepsilon) \leq J_\varepsilon(\bar{v}) \quad , \qquad \bar{v} \in \bar{V} \tag{5.10}$$

characterised by the variational equality

$$\langle DJ_\varepsilon(\bar{u}_\varepsilon), \bar{v} \rangle = 0 \quad , \qquad \bar{v} \in \bar{V} \tag{5.11}$$

where $DJ_\varepsilon(\bar{u}_\varepsilon)$ is the Gateaux derivative of J_ε at $\bar{u}_\varepsilon$ and $\langle .,. \rangle$ denotes duality pairing on $\bar{V}' \times \bar{V}$, $\bar{V}'$ being the dual space of $\bar{V}$.

It is easy to show that (5.11) is equivalent to

$$A(\bar{u}_\varepsilon, \bar{v}) + \varepsilon^{-1} \langle Dj(\lambda_\varepsilon), \mu \rangle = f(\underset{\sim}{v}) \quad , \qquad \bar{v} \in \bar{V} \ . \tag{5.12}$$

Substituting for the Gateaux derivative of j at λ_ε from (5.4) we get

$$A(\bar{u}_\varepsilon, \bar{v}) + \varepsilon^{-1} \int_{\Omega^p} g(\underset{\sim}{x}) \lambda_{\varepsilon-}(\underset{\sim}{x}) \ \mu(\underset{\sim}{x}) \ dx = \underset{\sim}{f}(v) \ . \tag{5.13}$$

In the same way as eqns (2.22) are equivalent to eqn (2.19), so also are the following two equations equivalent to (5.13) :

$$a(\underset{\sim}{u}_\varepsilon,\underset{\sim}{v}) - c(\underset{\sim}{v},\lambda_\varepsilon) = f(\underset{\sim}{v}) \tag{5.14a}$$

$$- c(\underset{\sim}{u}_\varepsilon,\mu) + b(\lambda_\varepsilon,\mu) - (k_\varepsilon,\mu)_\Lambda = 0 \tag{5.14b}$$

where $k_\varepsilon(\underset{\sim}{x}) = - \varepsilon^{-1}g(\underset{\sim}{x})\lambda_{\varepsilon-}(\underset{\sim}{x})$, and $(.,.)_\Lambda$ denotes the inner product on Λ , given by

$$(k_\varepsilon,\mu)_\Lambda = \varepsilon^{-1} \int_{\Omega^p} g(\underset{\approx}{x})\lambda_{\varepsilon-}(\underset{\approx}{x}) \; \mu(\underset{\approx}{x}) \; dx \quad . \tag{5.14c}$$

The penalisation procedure has evidently removed the explicit constraints from the perturbed minimisation problem, resulting in the variational equality (5.12). It is clear that this latter form is suitable as a basis for a numerical approximation, a fact which we shall shortly pursue. However, before doing so, it is important to investigate the convergence characteristics of the perturbed minimisation problem.

Convergence of the Perturbed Minimisation Problem

We would like to show that the solution $(\underset{\sim}{u}_\varepsilon,\lambda_\varepsilon)$ of the perturbed minimisation problem converges to the solution $(\underset{\sim}{u},\lambda)$ of the original problem as $\varepsilon \to 0$. In the following theorem we shall attempt this in three stages. First we will show that there exists a sequence $\bar{u}_\varepsilon$ which converges weakly to an element $\bar{u}_o$ in $\bar{V}$. Then we show that $\bar{u}_o \in \bar{K}$, and that $\bar{u}_\varepsilon$ coincides with $\bar{u}$. Finally we show strong convergence of $\bar{u}_\varepsilon$ to $\bar{u}$. The theorem requires the following preliminary result.

Lemma 5.4

Let $\ell_n : U \to R$ be a sequence of bounded linear functionals such that $\ell_n \to \ell_o$ in U' , where U' is the dual space of U . Let u_n be a bounded sequence in U which converges weakly to u_o in U . Then

$$\langle \ell_n, u_n \rangle \to \langle \ell_o, u_o \rangle \quad .$$

Proof

We have

$$\begin{aligned} |\langle \ell_n, u_n \rangle - \langle \ell_o, u_o \rangle| &= |\langle \ell_n, u_n \rangle + \langle \ell_o, u_n \rangle - \langle \ell_o, u_n \rangle - \langle \ell_o, u_o \rangle| \\ &\leq |\langle \ell_n - \ell_o, u_n \rangle| + |\langle \ell_o, u_n - u_o \rangle| \\ &\leq \|\ell_n - \ell_o\|_{U'} \, \|u_n\|_U + |\langle \ell_o, u_n - u_o \rangle| \\ &\to 0 \end{aligned}$$

using the convergence properties of ℓ_n and u_n , and the boundedness of u_n . □

Theorem 5.6

Let $\bar{u}_\varepsilon$ be the minimiser of J_ε over the space $\bar{V}$, and let $\bar{u}$ be the minimiser of J in K . Then $\bar{u}_\varepsilon$ converges to $\bar{u}$ as $\varepsilon \to 0$.

Proof

(i) We show first that the sequence $\bar{u}_\varepsilon$ is bounded independent of ε .

From theorems 5.2A and 5.5 we have respectively

$$J(\bar{u}) \leqslant J(\bar{v}) \quad , \quad \bar{v} \in \bar{K} \tag{5.15}$$

and $$J_\varepsilon(\bar{u}_\varepsilon) \leqslant J_\varepsilon(\bar{v}) \quad , \quad \bar{v} \in \bar{V} \quad . \tag{5.16}$$

Choosing $\bar{v}$ as $\bar{u}$ in (5.16) we get

$$J_\varepsilon(\bar{u}_\varepsilon) \leqslant J_\varepsilon(\bar{u}) \quad . \tag{5.17}$$

Furthermore, using the definitions of J and J_ε , and the fact that $\bar{u}$ and $\bar{u}_\varepsilon$ are their respective minimisers we can show, together with (5.15) and (5.16), that

$$J(\bar{u}_\varepsilon) \leqslant J_\varepsilon(\bar{u}_\varepsilon) \leqslant J_\varepsilon(\bar{u}) = J(\bar{u}) \leqslant J(\bar{v}) \quad . \tag{5.18}$$

Setting $\bar{v} = 0$ in (5.18) and using the coercivity of J , Lemma 5.2, we obtain

$$J(0) \geqslant J(\bar{u}_\varepsilon) \geqslant \frac{\alpha}{2} \|\bar{u}_\varepsilon\|^2_{\bar{V}} \quad . \tag{5.19}$$

Hence, $\|\bar{u}_\varepsilon\|$ is bounded above and so there exists a subsequence, denoted by $\bar{u}_{\varepsilon'}$, which converges weakly to an element $\bar{u}_o$ in $\bar{V}$;

that is, the subsequence $\underset{\sim}{u}_{\varepsilon'}$ converges weakly to $\underset{\sim}{u}_o$ in V , and the subsequence $\lambda_{\varepsilon'}$ converges weakly to λ_o in Λ .

(ii) We show next that $\bar{u}_o$ is in $\bar{K}$. From Lemma 5.3 we know that the Gateaux derivative $Dj(\mu)$ is a monotone mapping from Λ to Λ' . Thus, we have

$$\langle Dj(\lambda_{\varepsilon'}) - Dj(\mu), \lambda_{\varepsilon'} - \mu \rangle \geq 0 \quad . \tag{5.20}$$

In taking the limit as $\varepsilon' \to 0$ of this functional we treat each term separately. We have for the first term

$$\begin{aligned}
\langle Dj(\lambda_{\varepsilon'}), \mu \rangle &= \varepsilon' [-c(\underset{\sim}{u}_{\varepsilon'}, \mu) + b(\lambda_{\varepsilon'}, \mu)] \\
&\leq \varepsilon' [C_1 \|\underset{\sim}{u}_{\varepsilon'}\|_V + C_2 \|\lambda_{\varepsilon'}\|_\Lambda] \|\mu\|_\Lambda \\
&\leq C_3 \varepsilon' \|\bar{u}_{\varepsilon'}\|_{\bar{V}} \|\mu\|_\Lambda \\
&\to 0 \text{ as } \varepsilon' \to 0
\end{aligned}$$

where we have used the boundedness of $\|\bar{u}_{\varepsilon'}\|_{\bar{V}}$.

Before considering the next term in (5.20) we need to show that $Dj(\lambda_{\varepsilon'}) \to 0$ in Λ' . Using the conventional operator norm on Λ' we have

$$\begin{aligned}
\|Dj(\lambda_{\varepsilon'})\|_{\Lambda'} &= \sup_\mu \frac{\|\langle Dj(\lambda_{\varepsilon'}), \mu \rangle\|}{\|\mu\|} \\
&< \sup_\mu \frac{C_4 \varepsilon' \|\mu\|}{\|\mu\|}
\end{aligned}$$

$$= C_4 \varepsilon'$$

$$\to 0 \quad \text{as } \varepsilon' \to 0 \quad .$$

Hence, from Lemma 5.4 we have

$$\lim_{\varepsilon' \to 0} \langle Dj(\lambda_{\varepsilon'}),\lambda_{\varepsilon'}\rangle = 0 \quad .$$

For the third term in (5.20) we have

$$\lim_{\varepsilon' \to 0} \langle Dj(\mu),\lambda_{\varepsilon'}\rangle = \langle Dj(\mu),\lambda_o\rangle$$

since $\lambda_{\varepsilon'} \rightharpoonup \lambda_o$ as $\varepsilon' \to 0$ from (i) above.

Finally, the fourth term in (5.20) remains unchanged in the limit as $\varepsilon' \to 0$, so that collecting the above results together we have

$$\lim_{\varepsilon' \to 0} \langle Dj(\lambda_{\varepsilon'}) - Dj(\mu),\lambda_{\varepsilon'} - \mu\rangle = \langle Dj(\mu),\mu - \lambda_o\rangle$$

$$\geq 0 \quad , \qquad \mu \in \Lambda \quad . \qquad (5.21)$$

Following FUCIK and KUFNER (1980), page 350, we choose $\mu = \lambda_o + t\eta$, where $t > 0$ and $\eta \in \Lambda$; then making this substitution in (5.21) we get

$$\langle Dj(\lambda_o + t\eta) \ , \ t\eta \rangle = \int_\Omega g(\underset{\sim}{x})(\lambda_o + t\eta)_-(\underset{\sim}{x})t\eta(\underset{\sim}{x}) \ dx \geq 0$$

or $\int_\Omega g(\underset{\sim}{x})(\lambda_o + t\eta)_-(\underset{\sim}{x})\eta(\underset{\sim}{x}) \ dx \geq 0$, since $t > 0$.

As $t \to 0$ it follows that

$$\langle Dj(\lambda_o),\eta\rangle \geq 0 \quad , \quad \eta \in \Lambda \quad .$$

But this inequality holds for the element $-\eta$ as well, so that

$$\langle Dj(\lambda_o,\eta\rangle \leq 0 \quad , \qquad \eta \in \Lambda$$

from which it follows that

$$\langle Dj(\lambda_o),\eta\rangle \ = \ 0 \ \ .$$

Thus, from (5.4) we have

$$\langle Dj(\lambda_o),\eta\rangle = \int_{\Omega^P} g(\underset{\sim}{x})\lambda_{o-}(\underset{\sim}{x}) \ \eta(\underset{\sim}{x}) \ dx = 0$$

$$\Rightarrow \ \lambda_{o-} \ = \ 0$$

and so $\lambda_o \in K$, and $\bar{u}_o \in \bar{K}$.

(iii) Next we show that $\bar{u}_o$ solves (3.5), that is, $\bar{u}_o = \bar{u}$.
From (5.18) we have

$$J(\bar{u}_{\varepsilon'}) \leq J(\bar{u}) \quad .$$

Taking $\liminf_{\varepsilon' \to 0}$ of both sides gives

$$\begin{aligned} J(\bar{u}) &\geq \liminf_{\varepsilon' \to 0} \ J(\bar{u}_{\varepsilon'}) \\ &\geq \ J(\bar{u}_o) \end{aligned} \tag{5.22}$$

since J is convex and differentiable, and therefore weakly lower semi-continuous. But since $\bar{u}$ is the unique minimiser of $J(\bar{u})$, we must have $\bar{u}_o = \bar{u}$. Because all the above holds for every convergent subsequence, it follows that the sequence $\bar{u}_\varepsilon$ itself converges to $\bar{u}$.

(iv) We demonstrate strong convergence next. Since (5.22) applies to any member of the sequence $\bar{u}_\varepsilon$, with $\bar{u}_o = \bar{u}$, we have

$$\lim_{\varepsilon \to 0} J(\bar{u}_\varepsilon) = J(\bar{u}) \quad . \tag{5.23}$$

Now, from the definition of J (eqn (3.1)) we have

$$\left|J(\bar{u}_\varepsilon) - J(\bar{u})\right| = \left|A(\bar{u}_\varepsilon - \bar{u},\bar{u}_\varepsilon - \bar{u}) - f(\underset{\sim}{u}_\varepsilon - \underset{\sim}{u}) - 2A(\bar{u},\bar{u} - \bar{u}_\varepsilon)\right|$$

$$\geq \left|A(\bar{u}_\varepsilon - \bar{u},\bar{u}_\varepsilon - \bar{u})\right| - \left|f(\underset{\sim}{u}_\varepsilon - \underset{\sim}{u})\right| - 2\left|A(\bar{u},\bar{u} - \bar{u}_\varepsilon)\right| \quad .$$

Taking the limit as $\varepsilon \to 0$ and using both the weak convergence of $\bar{u}_\varepsilon$ and (5.23) we get

$$\lim_{\varepsilon \to 0} \left|A(\bar{u}_\varepsilon - \bar{u},\bar{u}_\varepsilon - \bar{u})\right| = 0 \quad .$$

Finally, using the continuity and $\bar{V}$-ellipticity of A, (Lemma 5.1) we obtain

$$\lim_{\varepsilon \to 0} \|\bar{u}_\varepsilon - \bar{u}\|_{\bar{V}} = 0 \tag{5.24}$$

which establishes the strong convergence of $\bar{u}_\varepsilon$. □

An Illustration

It is enlightening to review the penalty formulation in one dimension where the physical concepts are more readily appreciated. The classical boundary-value problem corresponding to (5.11), or alternatively (5.12), is

$$\operatorname{div} \dot{\underset{\sim}{\sigma}}(\bar{u}_\varepsilon) + \dot{\underset{\sim}{f}} = 0 \quad , \qquad \text{in } \Omega \tag{5.25}$$

$$\dot{\sigma}(\bar{u}_\varepsilon) = \underset{\sim}{C}[\underset{\sim}{E}(\underset{\sim}{u}_\varepsilon) - \lambda_\varepsilon \underset{\sim}{M}] \quad , \qquad \text{in } \Omega \tag{5.26}$$

$$\lambda_\varepsilon = 0 \; , \quad \text{in } \Omega^e$$

$$\lambda_\varepsilon[\frac{1}{G} + \frac{g}{2\varepsilon}(1 - \operatorname{sgn}\lambda_\varepsilon)] - \underset{\sim}{M} \cdot \dot{\underset{\sim}{\sigma}}(\bar{u}_\varepsilon) = 0 \; , \quad \text{in } \Omega^p \; . \tag{5.27}$$

The term $(\lambda_\varepsilon g/2\varepsilon)(1 - \operatorname{sgn}\lambda_\varepsilon)$ is the penalty function, where g is defined in (5.2), ε is the penalty parameter, the operator sgn is defined in Lemma 5.3, and $\underset{\sim}{M}$ is defined in eqn (1.5).

Consider the case where $\sigma_{11} \neq 0$ and all other stress components are zero. We assume a piecewise linear uniaxial stress-strain relationship (Fig. 5.2) for which the yield function reduces to

$$\phi = |\sigma_{11} - \Delta Y| - Y_o \leq 0 \tag{5.28}$$

where Y_o and Y are the initial and current values of the yield stress respectively, and $\Delta Y = Y - Y_o \operatorname{sgn}(Y)$. The only non-zero component of $\underset{\sim}{M}$ is M_{11} , with

$$M_{11} = \frac{\partial \phi}{\partial \sigma_{11}} = \operatorname{sgn}(\sigma_{11} - \Delta Y) \quad . \tag{5.29}$$

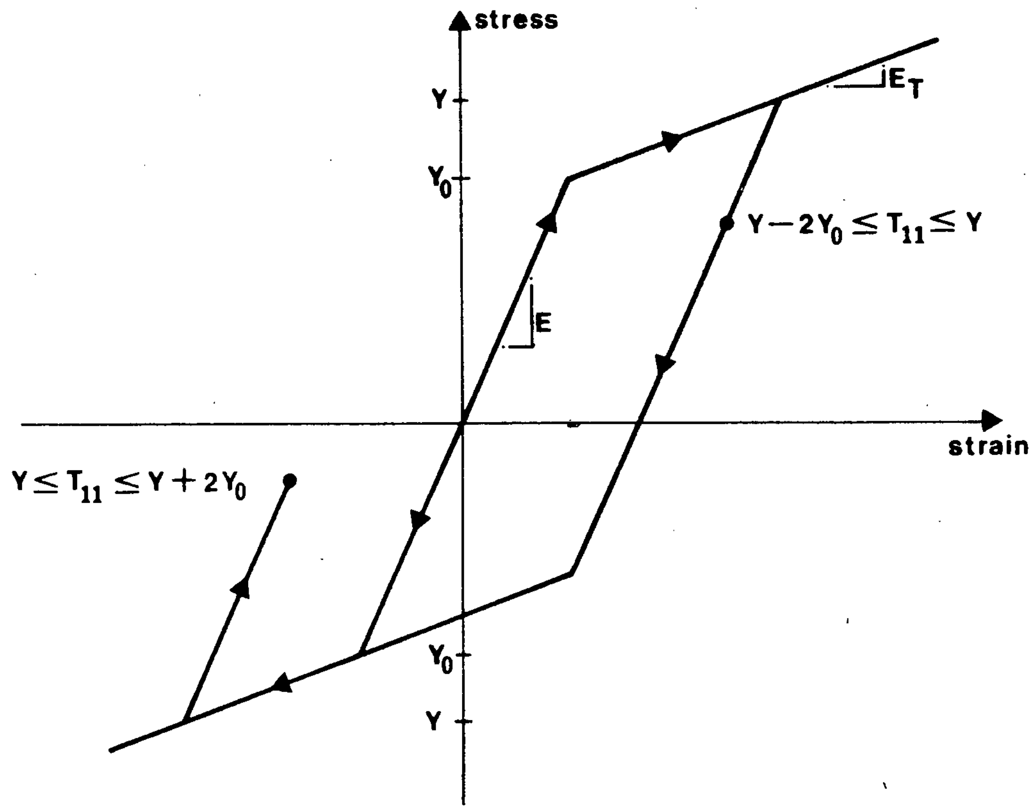


Figure 5.2 Uniaxial stress-strain curve.

We choose g to be equal to 1/G (which is a constant in the uniaxial case) and consider tensile behaviour for which $M_{11} = 1$. Then from (5.27) we have

$$\lambda_\varepsilon = \frac{\dot{\sigma}_{11}}{\frac{1}{G} + \frac{g(1 - \mathrm{sgn}\lambda_\varepsilon)}{2\varepsilon}}$$

$$= \frac{G\dot{\sigma}_{11}}{1 + \frac{(1 - \mathrm{sgn}\lambda_\varepsilon)}{2\varepsilon}} \quad . \tag{5.30}$$

This result indicates that λ_ε and $\dot{\sigma}_{11}$ always have the same sign. If λ_ε is positive then we have plastic loading with $\lambda_\varepsilon = G\dot{\sigma}_{11}$, where G is the inverse slope of the loading curve given by G = cotan θ (Fig. 5.2).

The penalty function gives rise to the possibility of a negative value of λ_ε , in which case

$$\lambda_\varepsilon = \frac{G\dot{\sigma}_{11}}{1 + \varepsilon^{-1}} = \frac{\dot{\sigma}_{11}}{\tan \psi} \quad . \tag{5.31}$$

This behaviour is shown in Fig. 5.3 : whereas elastic unloading normally occurs along AB , the perturbed problem exhibits unloading at a slope $(1/G)(1 + \varepsilon^{-1})$. Thus, as $\varepsilon \to 0$ the unloading slope approaches the vertical corresponding to the exact problem.

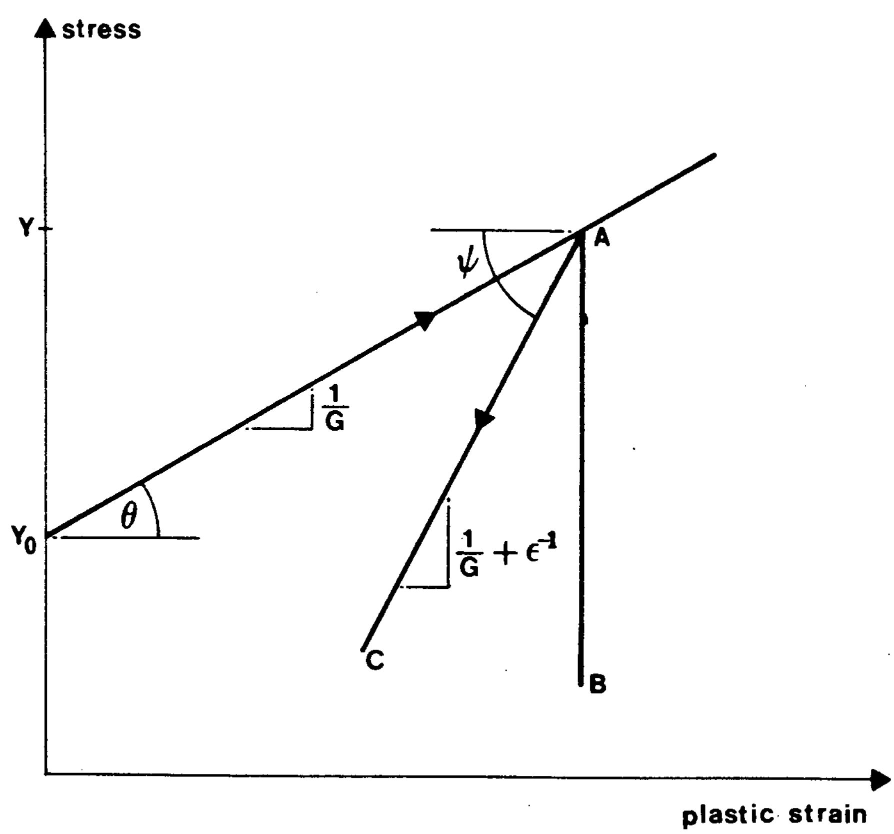


Figure 5.3 Unloading behaviour in the penalised problem.

5.6 FINITE ELEMENT APPROXIMATIONS AND ERROR ESTIMATES

In the preceding sections we have developed three variational principles, each of which may be associated with the original classical formulation of the problem defined in Section 5.1. We propose now to describe numerical approximations which are based on each of these three variational principles. We again choose the Galerkin finite element method to provide the framework for these approximations.

We introduce the families of finite-dimensional subspaces $\{V^h\}$ of V and $\{\Lambda^h\}$ of Λ , using piecewise polynomial basis functions on a sequence of finite element meshes; here $h \in (0,1]$ is a real parameter which identifies a particular subspace $V^h \subset V$, or $\Lambda^h \subset \Lambda$. We expect that $V^h \to V$ and $\Lambda^h \to \Lambda$ in some sense as $h \to 0$. In physical terms we may imagine that the members of the families $\{V^h\}$ and $\{\Lambda^h\}$ are obtained through regular uniform refinements of a finite element mesh defined on Ω , where h then represents the mesh size. We also define a finite-dimensional approximation K^h of the constraint space K . It is not necessary at this stage to be specific about the definition of K^h except to mention that K^h need not necessarily be a subset of K . For example, K^h may be defined to be a set of functions spanned by piecewise polynomials of given degree which satisfy the constraint $\lambda \geq 0$ only at selected points in each element, for example, the integration points. Then clearly $K^h \not\subset K$.

We now define the finite-dimensional counterparts of Theorems 5.1, 5.4, and 5.5 as follows.

A Variational Inequality (see Theorem 5.1, and eqns (2.22))

Find $\bar{u}^h = (\underset{\sim}{u}^h, \lambda^h) \in \bar{K}^h = V^h \times K^h$ such that

$$a(\underset{\sim}{u}^h, \underset{\sim}{v}^h) - c(\underset{\sim}{v}^h, \lambda^h) = f(\underset{\sim}{v}^h)$$

$$- c(\underset{\sim}{u}^h, \mu^h - \lambda^h) + b(\lambda^h, \mu^h - \lambda^h) \geq 0 \qquad (6.1)$$

for all $\bar{v}^h \in \bar{K}^h$.

A Saddle-point Problem (see Theorem 5.4 , and eqns (4.3) and (4.4))

Find $\bar{u}^h \in \bar{V}^h$ and $k^h \in K^h$ such that

$$a(\underset{\sim}{u}^h, \underset{\sim}{v}^h) - c(\lambda^h, \underset{\sim}{v}^h) = f(\underset{\sim}{v}^h)$$

$$- c(\mu^h, \underset{\sim}{u}^h) + b(\lambda^h, \mu^h) - (k^h, \mu^h)_\Lambda = 0 \qquad (6.2)$$

$$(\ell^h - k^h, \lambda^h)_\Lambda \geq 0$$

for all $\bar{v}^h \in \bar{V}^h$ and $\ell^h \in K^h$, and where $(.,.)_\Lambda$ denotes the inner product on Λ .

A Penalised Minimisation Problem (see Theorem 5.5, and eqns (5.14))

Find $\bar{u}^h_\varepsilon \in \bar{V}^h$ such that

$$J_\varepsilon(\bar{u}^h_\varepsilon) \leq J_\varepsilon(\bar{v}^h) \text{ for all } \bar{v}^h \in \bar{V}^h$$

or, equivalently,

$$a(\underset{\sim}{u}^h_\varepsilon, \underset{\sim}{v}^h) - c(\underset{\sim}{v}^h, \lambda^h_\varepsilon) = f(\underset{\sim}{v}^h)$$

$$- c(\underset{\sim}{u}^h_\varepsilon, \mu^h) + b(\lambda^h_\varepsilon, \mu^h) - (k^h_\varepsilon, \mu^h)_\Lambda = 0 \qquad (6.3)$$

for all $\bar{v}^h \in \bar{V}^h$, and where k^h_ε is defined by

$$k^h_\varepsilon(\underset{\sim}{x}) = -\varepsilon^{-1} g(\underset{\sim}{x}) \lambda^h_{\varepsilon-}(\underset{\sim}{x}) \quad . \qquad (6.4)$$

The proofs of the finite-dimensional counterparts to Theorems 5.1, 5.4 and 5.5, as stated above, follow in much the same manner so that we are assured of the existence of unique solutions to the above three problems. Also, the solution $\bar{u}^h_\varepsilon$ converges to $\bar{u}^h$ as $\varepsilon \to 0$.

The remainder of this section is concerned with finding an estimate for the error $\|\bar{u} - \bar{u}^h_\varepsilon\|$ due to both the penalisation and the finite element approximations. This is accomplished in two stages, by establishing estimates for $\|\bar{u} - \bar{u}^h\|$ and for $\|\bar{u}^h - \bar{u}^h_\varepsilon\|$. The general approach follows to some extent that adopted by KIKUCHI (1981) in his study of the obstacle problem.

In developing our error estimates we consider the following specific restrictions of the spaces V^h and Λ^h . Let V^h be the space spanned by functions whose restrictions to each element Ω_e contain complete polynomials of degree 1 (for example, the 3-noded triangle or

4-noded quadrilateral in R^2). Suppose further that the solution $\underset{\sim}{u}$ is smooth enough to belong to $H^2(\Omega)$. Then the following interpolation error estimate holds (Section 3.5, Theorem 3.7) :

$$\|\underset{\sim}{u} - \underset{\sim}{\tilde{u}}^h\|_{1,\Omega} \leq C_2 h \|\underset{\sim}{u}\|_{2,\Omega} \quad , \qquad \underset{\sim}{\tilde{u}}^h \in V^h \tag{6.5}$$

where $\underset{\sim}{\tilde{u}}^h$ is the interpolate of $\underset{\sim}{u}$ on Ω .

Let Λ^h be the space spanned by functions whose restrictions to each element are constant in the plastic region Ω^p , and which are zero in the elastic region Ω^e ; that is, Λ^h consists of piecewise constant functions in Ω^p . Furthermore, let us define $K^h \subset K$ by

$$K^h = \{\mu^h : \mu^h \in \Lambda^h(\Omega) \quad , \quad \mu^h \geq 0\} \tag{6.6}$$

and the interpolate $\tilde{\lambda}^h \in K^h$ of λ by the step-function

$$\tilde{\lambda}^h|_{\Omega_e} = \min_{\underset{\sim}{x} \in \Omega_e} \{\lambda(\underset{\sim}{x})\} \quad \text{for } \lambda \in K \quad . \tag{6.7}$$

Under the assumption that $\lambda \in H^1(\Omega)$ the following interpolation error estimate (Section 3.5, Theorem 3.9) holds :

$$\|\lambda - \tilde{\lambda}^h\|_{0,\Omega} \leq C_1 h \|\lambda\|_{1,\Omega} \quad . \tag{6.8}$$

The choice of piecewise constant basis functions for Λ^h deserves further comment. It can be shown that for higher order polynomial approximations, the estimate $\|\bar{u} - \bar{u}^h_\varepsilon\|$ depends on a positive constant α_h defined by

$$\alpha_h \|\ell^h\|_o \leq \sup_{\mu^h \in \Lambda^h} \frac{I(\ell^h,\mu^h)}{\|\mu^h\|_o} \quad , \qquad \ell^h \in \Lambda^h \tag{6.9}$$

where I(•,•) denotes the operation of numerical integration (see, for example, KIKUCHI (1981)). Condition (6.9) is known as the discrete LBB condition after LADYZHENSKAYA, BABUSKA (1973), and BREZZI (1974). For piecewise constant basis functions the condition reduces to the trivial result where α_h is a constant independent of the mesh parameter h . However, for higher order basis functions the determination of α_h depends on both the choice of element and numerical integration scheme, and can present a formidable computational task. KIKUCHI (1981) has given results for α_h for the three-noded triangle and four-noded quadrilateral elements within the context of the obstacle problem. In each case a simple numerical integration rule was used : for example, the trapezoidal rule in the case of the four-noded quadrilateral.

We record in the following theorem a regularity result for the elastic-plastic problem due to JIANG (1984).

Theorem 5.7

Let the data f in eqn (5.12) be given in $(L_2(\Omega))^N$, for $\Omega \subset R^N$. Then the solution $\bar{u} = (\underset{\sim}{u},\lambda)$ of the variational inequality, Theorem 5.1, belongs to $(H^2(\Omega))^N \times (H^1(\Omega))^N$ and satisfies

$$\|\underset{\sim}{u}\|_2 + \|\lambda\|_1 \leq \bar{C}\|f\|_o \quad . \; \square \tag{6.10}$$

It is important to note that Jiang proves this theorem under the assumption that the entire body is plastic, that is, $\Omega^e = \emptyset$. Although the extension to the case $\Omega^e \neq \emptyset$ is not obvious because of the conditions on the interface between the elastic and plastic parts, we will nevertheless assume that for the case $\Omega^e \neq \emptyset$ the result of Theorem 5.7 remains valid.

We proceed now with the error estimate due to the finite element approximations, as given in the following theorem.

Theorem 5.8

Let V^h and Λ^h be as defined above, and let the interpolation error estimates (6.5) and (6.8) hold. Then there exists a constant $\hat{C} > 0$ such that

$$\|\bar{u} - \bar{u}^h\|_{\bar{V}} \leq \hat{C} h \|f\|_o \quad . \qquad (6.11)$$

Proof

From Theorem 5.1 we have

$$A(\bar{u},\bar{v} - \bar{u}) - f(\underset{\sim}{v} - \underset{\sim}{u}) \geq 0 \qquad (6.12)$$

for which the counterpart on the finite-dimensional subspace $\bar{K}^h$ may be written as

$$A(\bar{u}^h,\bar{v}^h - \bar{u}^h) - f(\underset{\sim}{v}^h - \underset{\sim}{u}^h) \geq 0 \quad . \qquad (6.13)$$

Adding (6.12) and (6.13) we get

$$A(\bar{u} - \bar{u}^h,\bar{u} - \bar{u}^h) \leq A(\bar{u},\bar{v} - \bar{u}^h) + A(\bar{u}^h,\bar{v}^h - \bar{u})$$

$$+ f(\underset{\sim}{u} - \underset{\sim}{v}) + f(\underset{\sim}{u}^h - \underset{\sim}{v}^h) \qquad (6.14)$$

$$= A(\bar{u} - \bar{u}^h,\bar{u} - \bar{v}^h) + A(\bar{u},\bar{v} - \bar{u} + \bar{v}^h - \bar{u}^h)$$

$$- f(\underset{\sim}{v} - \underset{\sim}{u} + \underset{\sim}{v}^h - \underset{\sim}{u}^h) \quad .$$

Since $K^h \subset K \subset \Lambda$ and Λ^h is spanned by piecewise constant basis functions, we may choose $\bar{v} = \bar{u} + \bar{u}^h - \tilde{u}^h$, and $\bar{v}^h = \tilde{u}^h$ (here $\tilde{u}^h$ denotes the interpolate of $\bar{u}^h$) in (6.14) to give

$$A(\bar{u} - \bar{u}^h,\bar{u} - \bar{u}^h) \leq A(\bar{u} - \bar{u}^h,\bar{u} - \tilde{u}^h) \quad . \qquad (6.15)$$

Using the intermediate results of Lemma 5.1(a) we may express the continuity of A as follows :

$$A(\bar{u},\bar{v}) \leq K_1 \|\underset{\sim}{u}\|_V \ \|\underset{\sim}{v}\|_V + K_2 \|\lambda\|_\Lambda \ \|\mu\|_\Lambda$$

$$+ K_3 \|\mu\|_\Lambda \ \|\underset{\sim}{v}\|_V + K_3 \|\lambda\|_\Lambda \ \|\underset{\sim}{u}\|_V$$

$$\leq K(\|\underset{\sim}{u}\|_V + \|\lambda\|_\Lambda)(\|\underset{\sim}{v}\|_V + \|\mu\|_\Lambda) \quad . \qquad (6.16)$$

We now make use of the $\bar{V}$-ellipticity of A (Lemma 5.1(b)) on the left-hand-side of (6.15) and the continuity of A , as expressed by (6.16), on the right-hand-side of (6.15) to get

$$\alpha \|\bar{u} - \bar{u}^h\|^2_{\bar{V}} \leq K(\|\underset{\sim}{u} - \underset{\sim}{u}^h\|_V + \|\lambda - \lambda^h\|_\Lambda)(\|\underset{\sim}{u} - \tilde{\underset{\sim}{u}}^h\|_V + \|\lambda - \tilde{\lambda}^h\|_\Lambda)$$

$$\leq \sqrt{2}\, K[\|\underset{\sim}{u} - \underset{\sim}{u}^h\|^2_V + \|\lambda - \lambda^h\|^2_\Lambda]^{1/2} (\|\underset{\sim}{u} - \tilde{\underset{\sim}{u}}^h\|_V + \|\lambda - \tilde{\lambda}^h\|_\Lambda)$$

$$= \sqrt{2}\, K\|\bar{u} - \bar{u}^h\|_{\bar{V}} (\|\underset{\sim}{u} - \tilde{\underset{\sim}{u}}^h\|_1 + \|\lambda - \tilde{\lambda}^h\|_o) \quad . \qquad (6.17)$$

Using the interpolation error estimates (6.5) and (6.8) in (6.17) we obtain

$$\|\bar{u} - \bar{u}^h\|_{\bar{V}} \leq \frac{\sqrt{2}}{\alpha} KC_3 h(\|\underset{\sim}{u}\|_2 + \|\lambda\|_1) \qquad (6.18)$$

where $C_3 = \max(C_1, C_2)$. Finally, using the regularity result of Theorem 5.7 in (6.18) we get

$$\|\bar{u} - \bar{u}^h\|_{\bar{V}} \leq \frac{\sqrt{2}}{\alpha} KC_3 \bar{C} h \ \|f\|_o$$

which is the required result with $\hat{C} = \sqrt{2}\ KC_3\bar{C}/\alpha$. □

We now proceed with the estimate of the error due to penalisation, in which extensive use of the mixed (saddle-point) formulation is made. The result is given in the following theorem.

<u>Theorem 5.9</u>

Let the conditions of Theorem 5.8 hold. Then there exists a positive constant $\hat{K}$, independent of ε , such that

$$\|\bar{u}^h - \bar{u}^h_\varepsilon\|_{\bar{V}} \leq \hat{K}\varepsilon \ \|k^h\|_\Lambda \qquad (6.19)$$

Proof

Following eqn (5.3) we define $\mu_+(\underset{\sim}{x})$ by

$$\mu_+(\underset{\sim}{x}) = \frac{1}{2}\,[\mu(\underset{\sim}{x}) + |\mu(\underset{\sim}{x})|], \ \mu \in L_2(\Omega^p)\ , \ \underset{\sim}{x} \in \Omega^p\ , \tag{6.20}$$

so that $\mu_+ + \mu_- = \mu$.

Now, subtracting $(6.2)_2$ from $(6.3)_2$ we get

$$- c(\underset{\sim}{u}^h_\varepsilon - \underset{\sim}{u}^h, \mu^h) + b(\lambda^h_\varepsilon - \lambda^h, \mu^h) - (k^h_\varepsilon - k^h, \mu^h)_\Lambda = 0 \ .$$

With the choice of $\mu^h = \lambda^h - \lambda^h_\varepsilon$ in the above, we get

$$- c(\underset{\sim}{u}^h_\varepsilon - \underset{\sim}{u}^h, \lambda^h - \lambda^h_\varepsilon) + b(\lambda^h_\varepsilon - \lambda^h, \lambda^h - \lambda^h_\varepsilon) = - (k^h_\varepsilon - k^h, \lambda^h_\varepsilon - \lambda^h)_\Lambda \ . \tag{6.21}$$

Consider the right-hand-side of the above equation :

$$- (k^h_\varepsilon - k^h, \lambda^h_\varepsilon - \lambda^h)_\Lambda = (k^h - k^h_\varepsilon, \lambda^h_\varepsilon)_\Lambda - (k^h - k^h_\varepsilon, \lambda^h)_\Lambda$$

$$\leq (k^h - k^h_\varepsilon, \lambda^h_\varepsilon)_\Lambda \quad , \text{ using (4.5) and (6.4)}$$

$$= (k^h - k^h_\varepsilon, \lambda^h_{\varepsilon-})_\Lambda + (k^h - k^h_\varepsilon, \lambda^h_{\varepsilon+})_\Lambda \quad , \text{ using (6.20)}$$

$$\leq \varepsilon(k^h - k^h_\varepsilon, k^h_\varepsilon)_\Lambda \quad , \text{ using (4.5) and (6.4)}$$

$$= \varepsilon(k^h - k^h_\varepsilon, k^h_\varepsilon - k^h + k^h)_\Lambda$$

$$= \varepsilon(k^h - k^h_\varepsilon, k^h)_\Lambda - \varepsilon(k^h - k^h_\varepsilon, k^h - k^h_\varepsilon)_\Lambda$$

$$\leq \varepsilon(k^h - k^h_\varepsilon, k^h)_\Lambda$$

$$= \varepsilon \| k^h - k^h_\varepsilon \|_\Lambda \; \| k^h \|_\Lambda \tag{6.22}$$

where the final result follows from the fact that k^h and k^h_ε were chosen to be piecewise constant functions.

Subtracting $(6.1)_1$ from $(6.3)_1$ and choosing $\underset{\sim}{v}^h = \underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon$ we get

$$a(\underset{\sim}{u}^h_\varepsilon - \underset{\sim}{u}^h, \underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon) - c(\underset{\sim}{u}^h_\varepsilon - \underset{\sim}{u}^h, \lambda^h - \lambda^h_\varepsilon) = 0 \quad . \tag{6.23}$$

Substituting (6.22) into (6.21), adding the result to (6.23), and then using (2.15), we get

$$A(\bar{u}^h_\varepsilon - \bar{u}^h, \bar{u}^h_\varepsilon - \bar{u}^h) \leq \varepsilon \| k^h - k^h_\varepsilon \|_\Lambda \; \| k^h \|_\Lambda \quad . \tag{6.24}$$

Also, from the $\bar{V}$-ellipticity of $A(\cdot,\cdot)$, Lemma 5.1, we have

$$A(\bar{u}^h_\varepsilon - \bar{u}^h, \bar{u}^h_\varepsilon - \bar{u}^h) \geq \alpha \| \bar{u}^h_\varepsilon - \bar{u}^h \|^2_{\bar{V}} \quad . \tag{6.25}$$

Now, from $(6.3)_2$ we have

$$|(k^h - k^h_\varepsilon, \mu^h)_\Lambda| = |- c(\underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon, \mu^h) + b(\lambda^h - \lambda^h_\varepsilon, \mu^h)|$$

$$\leq |c(\underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon, \mu^h)| + |b(\lambda^h - \lambda^h_\varepsilon, \mu^h)|$$

$$\leq K_3 \| \underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon \|_V \; \| \mu^h \|_\Lambda + K_2 \| \lambda^h - \lambda^h_\varepsilon \|_\Lambda \; \| \mu^h \|_\Lambda \tag{6.26}$$

which follows from the continuity of $b(\cdot,\cdot)$ and $c(\cdot,\cdot)$, Lemma 5.1(a). But, since k^h and k^h_ε are piecewise constant functions, we may write

$$\left|(k^h - k^h_\varepsilon,\mu^h)_\Lambda\right| = \|k^h - k^h_\varepsilon\|_\Lambda \|\mu^h\|_\Lambda \quad . \tag{6.27}$$

Hence, from (6.26) and (6.27) we have

$$\begin{aligned}
\|k^h - k^h_\varepsilon\|_\Lambda &\leqslant K_3\|\underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon\|_V + K_2\|\lambda^h - \lambda^h_\varepsilon\|_\Lambda \\
&\leqslant C_4[\|\underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon\|_V + \|\lambda^h - \lambda^h_\varepsilon\|_\Lambda] \quad , \; C_4 = \max(K_2, K_3) \\
&\leqslant \sqrt{2}\, C_4[\|\underset{\sim}{u}^h - \underset{\sim}{u}^h_\varepsilon\|^2_V + \|\lambda^h - \lambda^h_\varepsilon\|^2_\Lambda]^{1/2} \\
&= \sqrt{2}\, C_4\|\bar{u}^h - \bar{u}^h_\varepsilon\|_{\bar{V}}
\end{aligned} \tag{6.28}$$

where we have used the definition (2.11), and the inequality $\alpha + \beta \leqslant \sqrt{2}\,(\alpha^2 + \beta^2)^{1/2}$, $\alpha,\beta \in R$. Using (6.28) in (6.24) and then combining the result with (6.25) we get

$$\|\bar{u}^h - \bar{u}^h_\varepsilon\|_{\bar{V}} \leqslant \frac{\sqrt{2}\, C_4\varepsilon}{\alpha} \|k^h\|_\Lambda$$

which is the required result with $\hat{K} = \sqrt{2}\, C_4/\alpha$. □

The final error estimate is a trivial consequence of the application of the triangle inequality together with (6.11) and (6.19), and is summarised in the following theorem.

Theorem 5.10

Let the conditions of Theorems 5.8 and 5.9 hold. Then we have

$$\|\bar{u} - \bar{u}_\varepsilon^h\|_{\bar{V}} \leq \hat{C}h \ \|f\|_o + \hat{K}\varepsilon \ \|k^h\|_\Lambda \quad . \ \square \tag{6.29}$$

We should point out that the result in Theorem 5.10 depends on the indeterminate quantity $\|k^h\|_\Lambda$, where k^h is the discrete approximation of the Lagrange multiplier defined in Section 5.4. We may, of course, assume that $k^h \to 0$ as $h \to 0$, but if this proves to be an invalid assumption (as shown by numerical experiment, for example), then we may be forced to make the penalty parameter ε depend on h .

CHAPTER 6

NUMERICAL SOLUTION OF THE ELASTIC-PLASTIC PROBLEM USING A PENALTY-RATE APPROACH

In Chapter 5 we developed a variational formulation of the rate problem in the form of a penalised minimisation problem, and showed that this problem constituted a suitable basis for the numerical approximation of the original boundary-value problem. In this chapter we wish to continue with the development of the numerical approximation using the Galerkin finite element method.

To set the stage for our discussions we will review the role of the penalty-rate problem within the context of the incremental elastic-plastic problem. Recall that in Chapter 1 we stated that the elastic-plastic problem could be solved numerically by subdividing the loading history into a number of intervals Δt to which there correspond loading increments $\Delta \underset{\sim}{P}$, thus defining a sequence of incremental problems for each of which the corresponding displacement, stress and strain increments were to be sought. The first stage in the solution of a typical incremental problem is the solution of the rate problem, and this is usually followed by a suitable state determination* procedure to complete the incremental solution. We propose to describe in this chapter the complete solution of the incremental problem using the penalty-rate formulation.

---

* By state determination we mean any algorithm by means of which the rates are integrated and the solution is updated.

Assume for the purposes of illustration that a body is subjected to two loads P_1 and P_2 whose magnitudes vary in some predetermined way. Let us define a load space (P_1,P_2) and plot the history of the loads in the form of a load path in this space, as shown in Fig. 6.1. So as to be able to uniquely identify each point on the load path we introduce a scalar parameter t which parameterises the path length. This allows us to define the rate of change of the loads along the load path by

$$\dot{\underset{\sim}{P}}(t) = \frac{d}{dt}\, \underset{\sim}{P}(t) \quad . \tag{a}$$

Due to the nature of numerical approximations it would be impractical to attempt to follow exactly the arbitrary load path shown in Fig. 6.1. Thus, we construct a suitable piecewise-linear approximation of the original load path and, similarly, we effect a consistent approximation of the load rates such that they remain constant over each interval; thus, in Fig. 6.1, $\dot{\underset{\sim}{P}}_i \equiv \dot{\underset{\sim}{P}}(t_i)$ is the approximation of the load rate over the interval $[t_i, t_{i+1}]$.

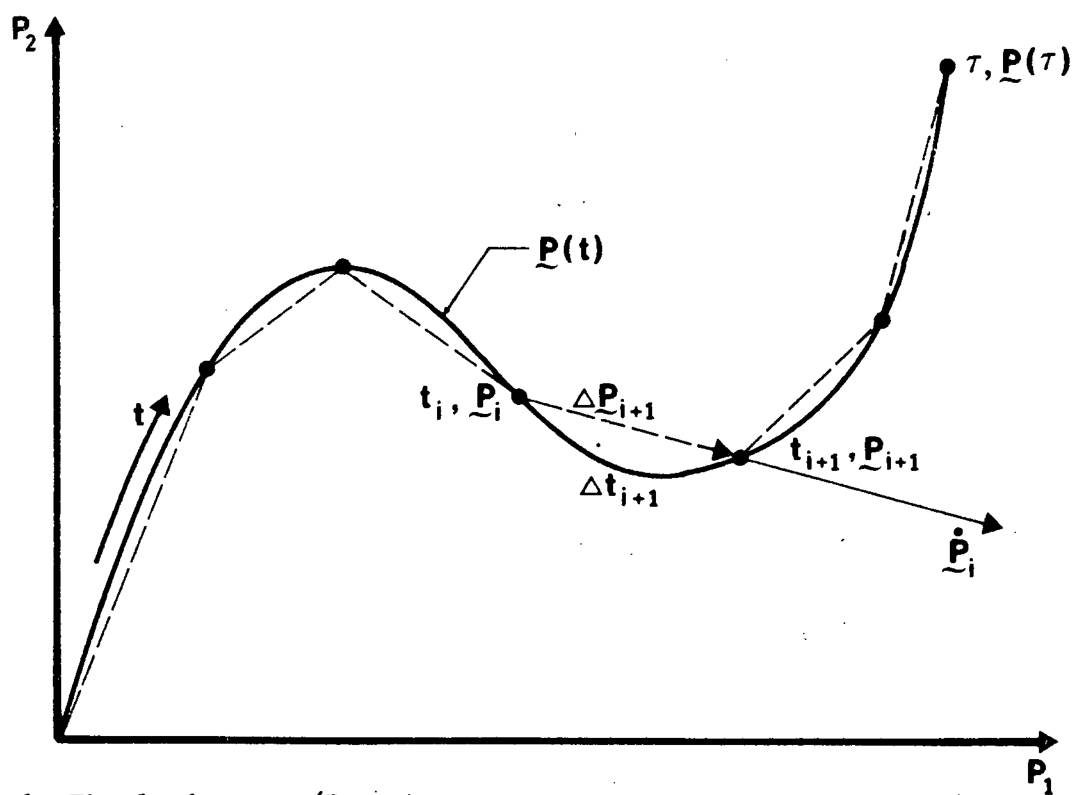


Figure 6.1 The load space (P_1,P_2) showing the load path $\underset{\sim}{P}(t)$ and its piecewise-linear approximation.

The generic incremental rate problem is illustrated schematically in Fig. 6.2 in terms of a single loading variable P and a single solution variable which we characterise as the displacement u . For simplicity we assume a monotonically increasing solution path which is parameterised using the scalar t , and assume also that the path corresponding to the interval [0,t] is known. Our objective is to determine the displacement u at the end of the given interval $[t, t+\Delta t]$.

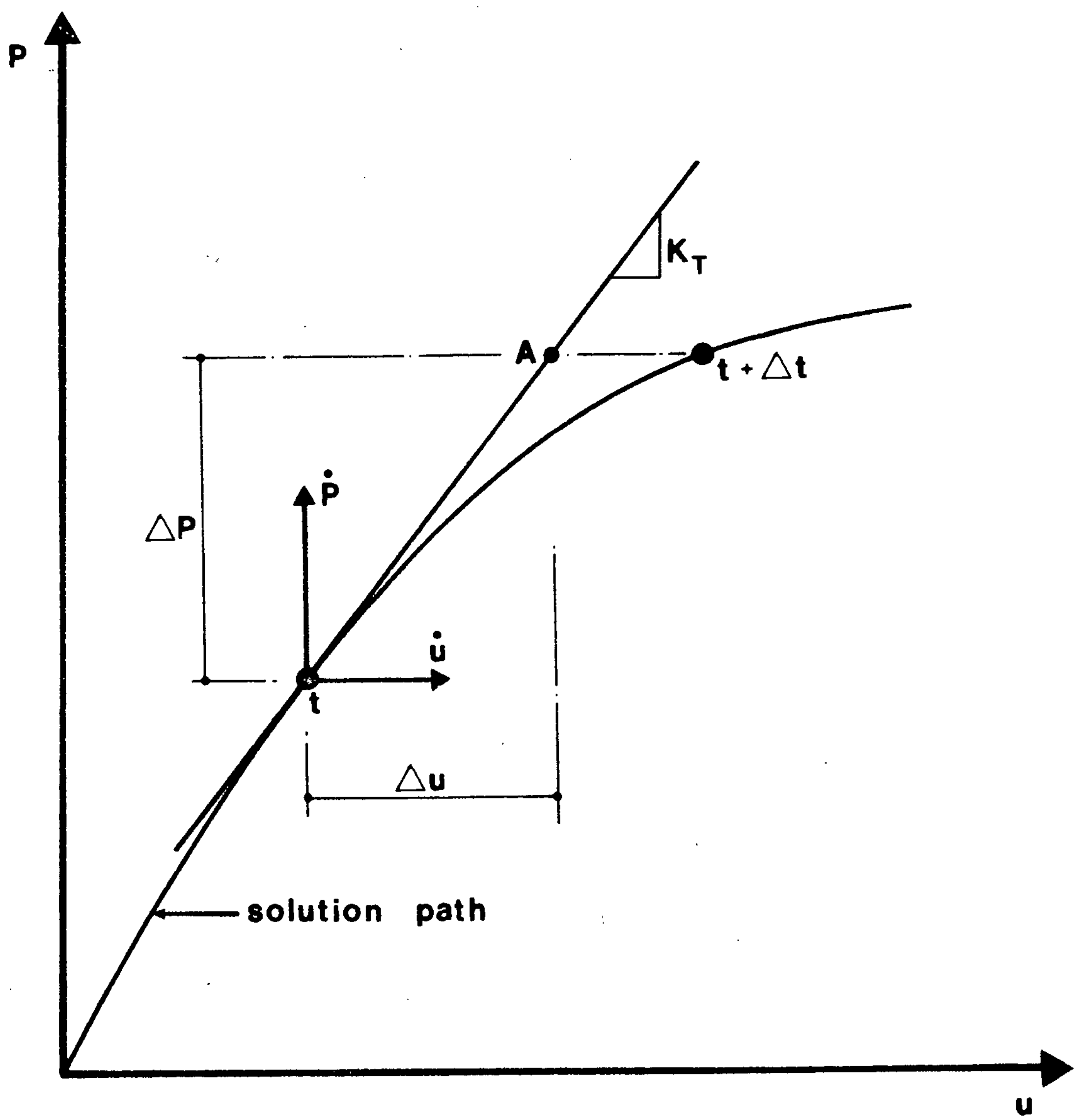


Figure 6.2 Illustrating the incremental rate problem.

As we shall see in this chapter, the rate problem is characterised by a system of algebraic equations of the form

$$\underset{\sim}{K}_T\dot{\underset{\sim}{u}} = \dot{\underset{\sim}{P}} \qquad \text{(b)}$$

or, for the situation depicted in Fig. 6.2,

$$K_T\dot{u} = \dot{P} \quad . \qquad \text{(c)}$$

Here, $\underset{\sim}{K}_T$ is referred to as the tangent stiffness matrix since it represents, by analogy with the situation in Fig. 6.2, the slope of the solution curve at t ; $\dot{\underset{\sim}{P}}$ and $\dot{\underset{\sim}{u}}$ are the load rate and displacement rate vectors at t .

This system of equations, as we shall also see, is homogeneous in the rates : thus, assuming the load rates remain constant over the interval Δt , we may always multiply the load rates and displacement rates by Δt , leaving the original rate equations unchanged. Thus we may rewrite eqn (b) as

$$\underset{\sim}{K}_T\dot{\underset{\sim}{u}}\Delta t = \dot{\underset{\sim}{P}}\Delta t \quad . \qquad \text{(d)}$$

Clearly, in setting up the rate problem the magnitude of the load rate vector may be arbitrarily chosen since we may always effect a linear scaling of both the load rate vector and the corresponding displacement rate vector subsequent to the solution of the rate problem.

Let us now assume that since the rate equations (b) are homogeneous in the rates we may integrate forward along the solution path using a one-step Euler forward method : then from eqn (b) we obtain

$$\underset{\sim}{K}_T \Delta \underset{\sim}{u} = \Delta \underset{\sim}{P} \quad . \tag{e}$$

Eqn (e) constitutes a first-order approximation of the <u>incremental problem</u>, since if we are given the load increment $\Delta \underset{\sim}{P}$ we may immediately solve for the displacement increment $\Delta \underset{\sim}{u}$. However, a comparison of eqns (d) and (e) shows that if we already know the load and displacement rates (from the solution of eqn (b)) we may compute the corresponding increments from

$$\Delta \underset{\sim}{P} = \dot{\underset{\sim}{P}} \Delta t \quad , \qquad \Delta \underset{\sim}{u} = \dot{\underset{\sim}{u}} \Delta t \quad . \tag{f}$$

The first-order approximation of the incremental problem, therefore, may be solved using two slightly different approaches :

(i) given the load rates $\dot{\underset{\sim}{P}}$, solve the rate equations (b) for the displacement rates $\dot{\underset{\sim}{u}}$; hence, for a given interval Δt solve the incremental problem using eqns (f) ;

(ii) given the load rates $\dot{\underset{\sim}{P}}$ and the interval Δt , calculate the load increments $\Delta \underset{\sim}{P}$ and solve the incremental problem directly* using eqn (e) .

---

* This is the approach normally adopted in the literature : see, for example, OWEN and HINTON (1980), BATHE (1982), ZIENKIEWICZ (1977).

We adopt the former approach in our work here.

It is clear that the first-order approximation described above causes the computed solution path to follow the tangent line K_T . Thus, the point A in Fig. 6.2 represents the solution which would be obtained from eqn (e), and this point will not lie on the true solution path, which in general will be nonlinear as a result of our assumptions regarding the material behaviour. Some of these assumptions can be weakened by, for example, assuming a piecewise-linear stress-strain relationship. Others, however, cannot be weakened without seriously compromising the credibility of the resulting solution : for example, we cannot assume that the direction of the normal to the yield surface remains constant over any finite interval Δt . Unfortunately, this is precisely what we are doing with the first-order approximation so that this approximation will in general exhibit some error. However, the error can be arbitrarily reduced by reducing the size of the interval Δt .

Whilst the procedure described above provides a simple and effective method of solving the incremental problem it is by no means the most efficient due to the strict control that must be kept on the size of the interval Δt in order to produce reasonable solutions. To avoid such strict control we could adopt the following alternative approach. Let eqn (e) be an initial solution predictor and use a suitable Newton-Raphson iterative procedure to provide the corrections to this prediction. At the end of each iteration calculate the stresses using a separate predictor/corrector algorithm and hence compute a new tangent modulus $\underset{\sim}{K}_T$. In this way an essentially smooth approximation to the solution path over the chosen interval Δt is built up, without any control over Δt (apart from common sense) being necessary. Although these methods generally provide good accuracy and rapid convergence, they do so at the price of a significant increase in the level of complexity of the numerical computations. For an excellent review of such methods the reader may refer to OWEN and HINTON (1980)). For a discussion of the use of these methods in a formulation of the rate problem very similar to the present one the reader may refer to REDDY and MITCHELL (1983) and DITTMER, GRIFFIN and MARTIN (1985).

We begin in Section 6.1 with the description of the discrete approximation on a typical element* Ω_e of the bilinear forms and functionals which appear in the penalised minimisation problem. The assembly of the element contributions leads to a system of nonlinear algebraic equations which constitute the discrete global approximation of the penalty-rate problem. We then describe how each of the element contributions are constructed using an isoparametric mapping from a master element $\hat{\Omega}$ to Ω_e . In Section 6.2 we discuss the solution procedure, including an algorithm for solving the global equations referred to above, and a scheme for computing the increment Δt such that the yield condition is always satisfied. Finally, in Section 6.3 we summarise the advantages of the penalty-rate formulation and compare this formulation with a conventional approach to the rate problem.

---

* For the purposes of illustrating certain concepts we assume Ω_e is in R^2 ; however, this is not intended to imply that the discussion is in general restricted to R^2 .

6.1 DISCRETE APPROXIMATION OF THE PENALTY-RATE PROBLEM

In Section 5.6 we wrote down the global approximation of the penalised minimisation problem on a finite-dimensional subspace $\bar{V}^h$, where $\bar{V}^h$ is assumed to be spanned by a finite number of piecewise polynomial global basis functions. It is well known that these global basis functions can be constructed from local basis functions defined on each element. Accordingly, we partition the domain Ω, which we will assume is polygonal, into a finite number E of triangular or quadrilateral subdomains Ω_e such that

$$\bar{\Omega} = \bigcup_{e=1}^{E} \bar{\Omega}_e \quad , \quad \Omega_e \cap \Omega_f = \emptyset \quad , \quad e \neq f \quad . \tag{1.1}$$

The mesh parameter h is defined by

$$h = \max_{1 \leqslant e \leqslant E} \{h_e\} \quad , \quad h_e = \mathrm{dia}(\Omega_e) \quad . \tag{1.2}$$

Then we refer to the connected set $\{\Omega_e\}_{e=1}^{E}$ as the finite element mesh for a given value of the mesh parameter h (see Fig. 4.1).

The global approximation of the penalised minimisation problem on $\bar{\Omega}$ takes the following form : find $\bar{u}_\varepsilon^h \equiv (\underset{\sim}{u}_\varepsilon^h, \lambda_\varepsilon^h)$ such that

$$a(\underset{\sim}{u}_\varepsilon^h, \underset{\sim}{v}^h) - c(\underset{\sim}{v}^h, \lambda_\varepsilon^h) = f(\underset{\sim}{v}^h) \tag{1.3a}$$

$$-c(\underset{\sim}{u}_\varepsilon^h, \mu^h) + b(\lambda_\varepsilon^h, \mu^h) - (\underset{\sim}{k}_\varepsilon^h, \mu^h)_\Lambda = 0 \quad , \tag{1.3b}$$

for all $\bar{v}^h = (\underset{\sim}{v}^h, \mu^h) \in \bar{V}^h$.

Note that for brevity we will drop the subscript ε hereafter, it being understood that we are dealing with the penalised minimisation problem.

The global approximations of the terms in (1.3) are constructed by adding contributions from each element. Thus, noting that

$$\int_{\Omega} \cdots = \sum_{e=1}^{E} \int_{\Omega_e} \cdots$$

we may rewrite (1.3) in the following form :

$$\sum_{e=1}^{E} a^{(e)}(\underset{\sim}{u}_e^h, \underset{\sim}{v}_e^h) - \sum_{e=1}^{E} c^{(e)}(\underset{\sim}{v}_e^h, \lambda_e^h) = \sum_{e=1}^{E} f^{(e)}(\underset{\sim}{v}_e^h) \qquad (1.4a)$$

$$- \sum_{e=1}^{E} c^{(e)}(\underset{\sim}{u}_e^h, \mu_e^h) + \sum_{e=1}^{E} b^{(e)}(\lambda_e^h, \mu_e^h) - \sum_{e=1}^{E} (k_e^h, \mu_e^h)_{\Lambda^{(e)}} = 0 \qquad (1.4b)$$

where the superscript (e) and subscript e denote restrictions to Ω_e , with $(\cdot)_e$ being interpreted in the sense $(\cdot)_e \equiv (\cdot)\big|_{\Omega_e}$; likewise $(\cdot,\cdot)_{\Lambda^{(e)}}$ is the restriction of the inner product on Λ to Ω_e . With the problem defined in the above form we may proceed with the approximation at the level of a typical element Ω_e .

Let $\{\phi_i\}_{i=1}^{N_e}$ be a suitable family of local basis functions defined on Ω_e and having the property

$$\phi_i(\underset{\sim}{x}_j) = \delta_{ij} \quad , \quad 1 \leqslant i,j \leqslant N_e$$

where $\underset{\sim}{x}_j$ is the position vector of the j-th node on Ω_e and N_e is the number of nodes on Ω_e . Then the restriction of a typical element of the finite-dimensional space $H^h \subset H^1(\Omega_e)$ to Ω_e may be approximated by

$$u_e^h(\underset{\sim}{x}) = \sum_{i=1}^{N_e} a_i \phi_i(\underset{\sim}{x}) \qquad (1.5)$$

where $a_i = u_e^h(\underset{\sim}{x}_i)$ is the value of u_e^h at node i on Ω_e . If the solution vector $\underset{\sim}{u}_e^h \in V^h$ has components which are approximated as in (1.5) then we may write the approximation of this vector as

$$\underset{\sim}{u}_e^h(\underset{\sim}{x}) = \underset{\sim}{\Psi}(\underset{\sim}{x})\underset{\sim}{a}_e \quad ; \qquad (1.6)$$

similarly, an arbitrary vector $\underset{\sim}{v}_e^h \in V^h$ may be approximated by

$$\underset{\sim}{v}_e^h(\underset{\sim}{x}) = \underset{\sim}{\Psi}(\underset{\sim}{x})\underset{\sim}{a}_e^* \quad . \qquad (1.7)$$

Here, $\underset{\sim}{\Psi}$ is a matrix of shape functions $\phi_i(\underset{\sim}{x})$, and $\underset{\sim}{a}_e$ and $\underset{\sim}{a}_e^*$ are ordered lists of the discrete nodal values of $\underset{\sim}{u}_e^h$ and $\underset{\sim}{v}_e^h$ respectively.

The strain rate vector $\dot{\underset{\sim}{\varepsilon}}$ is related to the velocity vector by

$$\dot{\underset{\sim}{\varepsilon}} = \underset{\sim}{D}\, \underset{\sim}{u}_e^h \qquad (1.8)$$

where $\underset{\sim}{D}$ is an appropriate matrix of differential operators (see Appendix A, eqn (A.4)). This may be approximated using (1.6) by

$$\dot{\underset{\sim}{\varepsilon}}(\underset{\sim}{x}) = \underset{\sim}{D}\, \underset{\sim}{\Psi}(\underset{\sim}{x})\underset{\sim}{a}_e$$

$$= \underset{\sim}{B}(\underset{\sim}{x})\underset{\sim}{a}_e \quad , \quad \underset{\sim}{x} \in \Omega_e \quad . \qquad (1.9)$$

Here, $\mathbf{B}(\mathbf{x})$ is the element strain rate-velocity matrix* and consists of partial derivatives of the shape functions $\phi_i(\mathbf{x})$ with respect to the components of the position vector $\mathbf{x}$. Similary, we have

$$\dot{\boldsymbol{\varepsilon}}(\mathbf{x}) = \mathbf{B}(\mathbf{x})\mathbf{a}_e^* \quad . \tag{1.10}$$

For the approximation of the space Λ^h of plastic multipliers we adopt a similar approach to that used for the approximation of the plastic strains in the incremental holonomic problem. Let Λ^h be the space spanned by piecewise polynomial basis functions, with the i-th function having a value of 1 at the i-th quadrature point and a value of 0 at every other quadrature point. Thus, for a (2x2) Gaussian quadrature rule for Ω_e in R^2 , for example, the basis functions will be bilinear polynomials over Ω_e , but globally-discontinuous; for a single-point quadrature rule the basis functions will be constants. The restrictions to Ω_e of typical basis functions for Λ^h are illustrated in Fig. 6.3.

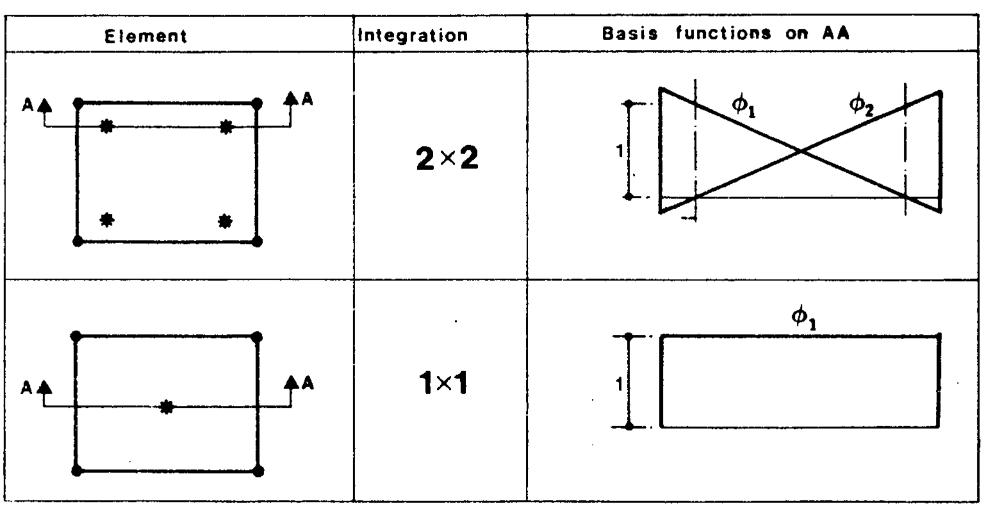


Figure 6.3 Typical basis functions for Λ^h on Ω_e .

---

* Matrices $\mathbf{D}$, $\mathbf{B}$, $\mathbf{B}_p$, etc., are only ever defined at element level and so we omit the subscript e for brevity.

Let $(\underset{\sim}{x}_i)$, $1 \leq i \leq N_G$, be the position vector of Gauss point i on Ω_e , N_G being the chosen number of Gauss points. Let $\{\phi_j\}_{j=1}^{N_G}$ be a suitable family of local basis functions on Ω_e with the property

$$\phi_j(\underset{\sim}{x}_i) = \delta_{ij} \quad , \quad 1 \leq i,j \leq N_G \quad . \tag{1.11}$$

Then the plastic multiplier solution on Ω_e is approximated by

$$\lambda_e^h(\underset{\sim}{x}) = \sum_{j=1}^{N_G} \alpha_j \phi_j(\underset{\sim}{x}) \tag{1.12}$$

where $\alpha_j = \lambda_e^h(\underset{\sim}{x}_j)$ is the value of the function λ_e^h at Gauss point j . We write (1.12) in matrix form as

$$\lambda_e^h(\underset{\sim}{x}) = \underset{\sim}{B}_p(\underset{\sim}{x})\underset{\sim}{\alpha}_e \tag{1.13}$$

and similary, for an arbitrary member of Λ^h we have

$$\mu_e^h(\underset{\sim}{x}) = \underset{\sim}{B}_p(\underset{\sim}{x})\underset{\sim}{\alpha}_e^* \quad . \tag{1.14}$$

Here, $\underset{\sim}{B}_p$ is a row matrix of the form

$$\underset{\sim}{B}_p = [\phi_1, \phi_2, \ldots., \phi_{N_G}] \tag{1.15}$$

and $\underset{\sim}{\alpha}_e$ and $\underset{\sim}{\alpha}_e^*$ are ordered lists of the discrete plastic multipliers at the chosen Gauss points. We note for future reference that according to (1.11) the matrix $\underset{\sim}{B}_p$ has a particularly simple form when it is evaluated at any of the chosen N_G Gauss points. For example, when evaluated at Gauss point 2 on Ω_e it takes the form

$$\underset{\sim}{B}_p(\underset{\sim}{x}_2) = [0, 1, 0, \ldots\ldots, 0] \quad . \tag{1.16}$$

It will thus be a distinct advantage to integrate functions which include λ_e^h or μ_e^h using the same quadrature rule as was used to define the functions ϕ_i .

For completeness we repeat here the definition of the row matrix $\underset{\sim}{M}$ (Section 5.1, eqn (1.5)) :

$$\underset{\sim}{M} = [\frac{\partial\phi}{\partial\underset{\sim}{\sigma}}] \quad . \tag{1.17}$$

Clearly, $\underset{\sim}{M}$ may be evaluated pointwise on Ω_e and depends on the state of stress and plastic strain at each point. We note also that the bilinear form b(•,•) (Section 5.2, eqn (2.13)) contains the term 1/G and that for a von Mises yield function with linear kinematic hardening this term is given by (Section 2.3)

$$1/G = E_p\underset{\sim}{M}^T\underset{\sim}{M} \quad . \tag{1.18}$$

In view of this we will find it convenient to introduce the following normalisation : let $|\underset{\sim}{M}| = (\underset{\sim}{M} \cdot \underset{\sim}{M})^{1/2}$ and define

$$\bar{\underset{\sim}{M}} = \frac{\underset{\sim}{M}}{|\underset{\sim}{M}|} \text{ and } \bar{\underset{\sim}{\alpha}}_e = \underset{\sim}{\alpha}_e|\underset{\sim}{M}| \tag{1.19}$$

so that

$$\bar{\underset{\sim}{\alpha}}_e\bar{\underset{\sim}{M}} = \underset{\sim}{\alpha}_e\underset{\sim}{M} \quad . \tag{1.20}$$

We may now substitute the approximations of $\underset{\sim}{u}_e^h$, $\underset{\sim}{v}_e^h$, λ_e^h and μ_e^h as described above into the local forms of the functionals and bilinear forms given in (1.4), whilst also making use of the normalisations given in (1.19). We treat each bilinear form and functional in turn, using the original definitions given in Section 5.2, eqns (2.12) through (2.16). Thus we obtain

$$a^{(e)}(\underset{\sim}{u}^h_e, \underset{\sim}{v}^h_e) = (\underset{\sim}{a}^*_e)^T \int_{\Omega_e} \underset{\sim}{B}^T \underset{\sim}{C}\, \underset{\sim}{B}\, d\underset{\sim}{x}\, \underset{\sim}{a}_e$$

$$= (\underset{\sim}{a}^*_e)^T\, \underset{\sim}{K}^{(e)}\, \underset{\sim}{a}_e \quad ; \tag{1.21}$$

$$b^{(e)}(\lambda^h_e, \mu^h_e) = (\bar{\underset{\sim}{\alpha}}^*_e)^T \int_{\Omega_e} \underset{\sim}{B}^T_p\, \bar{\underset{\sim}{M}}^T \underset{\sim}{C}\, \bar{\underset{\sim}{M}}\, \underset{\sim}{B}_p\, d\underset{\sim}{x}\, \bar{\underset{\sim}{\alpha}}_e$$

$$+ (\bar{\underset{\sim}{\alpha}}^*_e)^T E_p \int_{\Omega_e} \underset{\sim}{B}^T_p\, \underset{\sim}{B}_p\, d\underset{\sim}{x}\, \bar{\underset{\sim}{\alpha}}_e$$

$$= (\bar{\underset{\sim}{\alpha}}^*_e)^T \underset{\sim}{S}^{(e)}_1 \bar{\underset{\sim}{\alpha}}_e + (\bar{\underset{\sim}{\alpha}}^*_e)^T \underset{\sim}{S}^{(e)}_2 \bar{\underset{\sim}{\alpha}}_e \quad ; \tag{1.22}$$

$$c^{(e)}(\underset{\sim}{v}^h_e, \lambda^h_e) = (\underset{\sim}{a}^*_e)^T \int_{\Omega_e} \underset{\sim}{B}^T \underset{\sim}{C}\, \bar{\underset{\sim}{M}}\, \underset{\sim}{B}_p\, d\underset{\sim}{x}\, \bar{\underset{\sim}{\alpha}}_e$$

$$= (\underset{\sim}{a}^*_e)^T \underset{\sim}{L}^{(e)}\, \bar{\underset{\sim}{\alpha}}_e \quad ; \tag{1.23}$$

$$c^{(e)}(\underset{\sim}{u}^h_e, \mu^h_e) = (\bar{\underset{\sim}{\alpha}}^*_e)^T \int_{\Omega_e} \underset{\sim}{B}^T_p\, \bar{\underset{\sim}{M}}^T \underset{\sim}{C}\, \underset{\sim}{B}\, d\underset{\sim}{x}\, \underset{\sim}{a}_e$$

$$= (\bar{\underset{\sim}{\alpha}}^*_e)^T (\underset{\sim}{L}^{(e)})^T\, \underset{\sim}{a}_e \quad . \tag{1.24}$$

The linear functional $f^{(e)}(\underset{\sim}{v}_e^h)$ is approximated by

$$f^{(e)}(\underset{\sim}{v}_e^h) = (\underset{\sim}{a}_e^*)^T \int_{\Omega_e} \underset{\sim}{\Psi}^T \dot{\underset{\sim}{f}} \, d\underset{\sim}{x} + (\underset{\sim}{a}_e^*)^T \int_{\Gamma_{se}} \underset{\sim}{\Psi}^T \dot{\underset{\sim}{t}}_e \, ds$$

$$= (\underset{\sim}{a}_e^*)^T \dot{\underset{\sim}{P}}^{(e)} \tag{1.25}$$

where Γ_{se} is that part of the boundary Γ_e of Ω_e over which a traction rate $\dot{\underset{\sim}{t}}_e$ is applied.

The approximation of the inner product $(k_e^h, \mu_e^h)_{\Lambda(e)}$ requires a little more care : recalling Section 5.5, eqn (5.14c), we have

$$(k_e^h, \mu_e^h)_{\Lambda(e)} = \varepsilon^{-1} \int_{\Omega_e} g_e (\lambda_e^h)_- \mu_e^h \, d\underset{\sim}{x} \tag{1.26}$$

where $(\cdot)_-$ is defined in Section 5.5, eqn (5.3), and $g_e \geqslant 0$ is some as yet unspecific function, being the restriction of g to Ω_e . Although the inner product in (1.26) is "almost" linear in λ_e^h , it cannot be evaluated exactly because of the presence of the function $(\lambda_e^h)_-$. To overcome this difficulty we resort to numerical quadrature : let $I(\cdot)$ denote the operation of Gaussian numerical quadrature on Ω_e ,

$$I(f) = \sum_{i=1}^{N_{G'}} w_i f(\underset{\sim}{x}_i) \tag{1.27}$$

where $\underset{\sim}{x}_i$ is the position vector of the i-th quadrature point on Ω_e , and $w_i > 0$ is the quadrature weight for the i-th point. Then writing λ_e^h and

μ_e^h in (1.26) in terms of their approximations (1.13) and (1.14), and using (1.27) to perform the integration, we have

$$(k_e^h, \mu_e^h)_{\Lambda(e)} = \varepsilon^{-1} I[g_e(\underset{\sim}{\alpha}_e^*)^T \underset{\sim}{B}_p^T (\underset{\sim}{B}_p \underset{\sim}{\alpha}_e)_-] \quad . \tag{1.28}$$

Replacing $\underset{\sim}{\alpha}_e^*$ and $\underset{\sim}{\alpha}_e$ by their normalised forms, eqn (1.19), and choosing $g_e = |\underset{\sim}{M}|^2$ we get

$$\begin{aligned}(k_e^h, \mu_e^h)_{\Lambda(e)} &= \varepsilon^{-1} \; I[(\underset{\sim}{\bar{\alpha}}_e^*)^T \underset{\sim}{B}_p^T \; (\underset{\sim}{B}_p \underset{\sim}{\bar{\alpha}}_e)_-] \\ &= \varepsilon^{-1} (\bar{\alpha}_e^*)^T [w_1 \bar{\alpha}_{1-} \; , \; w_2 \bar{\alpha}_{2-} \; , \; \ldots , \; w_{N_G} \bar{\alpha}_{N_G-}]^T \\ &= \varepsilon^{-1} (\underset{\sim}{\bar{\alpha}}_e^*)^T \; \underset{\sim}{F}^{(e)} (\underset{\sim}{\bar{\alpha}}_e) \end{aligned} \tag{1.29}$$

where we have made use of the quadrature formula (1.27) with $N_{G'} = N_G$. The constraints $\lambda_e^h \geq 0$ are now controlled at the quadrature points on Ω_e via the nonlinear vector $\underset{\sim}{F}^{(e)}(\underset{\sim}{\bar{\alpha}}_e)$.

We turn now to the assembly of the global approximations from the element contributions. Let $\underset{\sim}{a}^*$ and $\underset{\sim}{a}$ be ordered lists of discrete velocities at the nodes, and $\underset{\sim}{\bar{\alpha}}^*$ and $\underset{\sim}{\bar{\alpha}}$ be ordered lists of discrete normalised plastic multipliers at the Gauss points. Then using (1.21) through (1.25) and (1.29) in (1.4) and assembling in the usual way we obtain the following discrete global approximation of the penalised minimisation problem :

$$\{(\underset{\sim}{a}^*)^T \ (\underset{\sim}{\bar{\alpha}}^*)^T\} \begin{bmatrix} \underset{\sim}{K} & \underset{\sim}{L} \\ \hline -\underset{\sim}{L}^T & \underset{\sim}{S} \end{bmatrix} \begin{Bmatrix} \underset{\sim}{a} \\ \hline \underset{\sim}{\bar{\alpha}} \end{Bmatrix} + \varepsilon^{-1}(\underset{\sim}{\bar{\alpha}}^*)^T \underset{\sim}{F}(\underset{\sim}{\bar{\alpha}})$$

$$- \{(\underset{\sim}{a}^*)^T \ (\underset{\sim}{\bar{\alpha}}^*)^T\} \begin{Bmatrix} \dot{\underset{\sim}{P}} \\ \hline \underset{\sim}{0} \end{Bmatrix} = 0 \ . \qquad (1.30)$$

Since the starred quantities are arbitrary this equation reduces to the following system of algebraic equations :

$$\begin{bmatrix} \underset{\sim}{K} & L \\ \hline -\underset{\sim}{L}^T & \underset{\sim}{S} \end{bmatrix} \begin{Bmatrix} \underset{\sim}{a} \\ \hline \underset{\sim}{\bar{\alpha}} \end{Bmatrix} + \begin{Bmatrix} \underset{\sim}{0} \\ \hline \varepsilon^{-1}\underset{\sim}{F}(\underset{\sim}{\bar{\alpha}}) \end{Bmatrix} = \begin{Bmatrix} \dot{\underset{\sim}{P}} \\ \hline \underset{\sim}{0} \end{Bmatrix} \qquad (1.31)$$

where we have written $\underset{\sim}{S} = \underset{\sim}{S}_1 + \underset{\sim}{S}_2$ for brevity. Due to the presence of the vector $\underset{\sim}{F}(\underset{\sim}{\bar{\alpha}})$ this system of equations is nonlinear and must therefore be solved using a suitable iterative procedure, the discussion of which we defer until the next section. We shall in future refer to the assembled $\underset{\sim}{K}$, $\underset{\sim}{L}$, $\underset{\sim}{L}^T$ and $\underset{\sim}{S}$ matrices as $\underset{\sim}{K}^*$, where $\underset{\sim}{K}$ is the conventional elastic stiffness matrix. Note that $\underset{\sim}{S}$ is a diagonal matrix, and $\underset{\sim}{K}^*$ is constant in the sense of being independent of the solution.

It is convenient to rewrite (1.31) in the following abbreviated form :

$$\left[\underset{\sim}{K}^*\right] \begin{Bmatrix} \underset{\sim}{a} \\ \hline \underset{\sim}{\bar{\alpha}} \end{Bmatrix} + \begin{Bmatrix} \underset{\sim}{0} \\ \hline \varepsilon^{-1}\underset{\sim}{F}(\underset{\sim}{\bar{\alpha}}) \end{Bmatrix} = \begin{Bmatrix} \dot{\underset{\sim}{P}} \\ \hline \underset{\sim}{0} \end{Bmatrix} \ . \qquad (1.32)$$

Now, as we shall see in the following section, if the components of $\bar{\underset{\sim}{\alpha}}$ do not become negative during the solution of (1.32) then the term $\varepsilon^{-1}\underset{\sim}{F}(\bar{\underset{\sim}{\alpha}})$ is superfluous, with the result that the system of equations which is actually solved is

$$\left[\underset{\sim}{K}^*\right]\left\{\begin{array}{c}\underset{\sim}{a}\\ \hline \bar{\underset{\sim}{\alpha}}\end{array}\right\} = \left\{\begin{array}{c}\dot{\underset{\sim}{P}}\\ \hline \underset{\sim}{0}\end{array}\right\} \quad . \tag{1.33}$$

In such a situation the body is said to "continue loading". In fact, it is interesting to note that the system of equations (1.33) can quite easily be shown to be identical to that used in the conventional tangent stiffness approach (see, for example, ZIENKIEWICZ (1977), Section 18.4). Briefly, from the second matrix equation in (1.31), ignoring the term $\underset{\sim}{F}(\bar{\underset{\sim}{\alpha}})$, we have

$$\bar{\underset{\sim}{\alpha}} = \underset{\sim}{S}^{-1}\underset{\sim}{L}^T\underset{\sim}{a} \tag{1.34}$$

which, when substituted into the first matrix equation yields

$$[\underset{\sim}{K} - \underset{\sim}{L}\underset{\sim}{S}^{-1}\underset{\sim}{L}^T]\underset{\sim}{a} = \dot{\underset{\sim}{P}}$$

$$\Rightarrow \qquad \underset{\sim}{K}_T\,\underset{\sim}{a} = \dot{\underset{\sim}{P}} \quad . \tag{1.35}$$

Here, $\underset{\sim}{K}_T$ is the conventional tangent stiffness matrix which is constructed by modifying the elastic stiffness matrix $\underset{\sim}{K}$ with contributions which are due to the current state of plasticity in the body. Formal proof of the above has been given by DITTMER (1978) to which the reader may refer for further details.

We emphasise that the above equivalence holds only for the condition of "continued loading" and that the penalty-rate formulation and conventional tangent modulus formulation have different approaches to the condition of "elastic unloading". We will see in the next section how the latter condition is handled using the penalty-rate formulation.

Calculation of Element Matrices and Vectors*

Let us assume that the body occupies a domain Ω in R^2. The position vector of each point in Ω is $\underset{\sim}{x} = (x,y)$ relative to cartesian axes X,Y, as illustrated previously in Fig. 4.3

We adopt the notion of a master element $\hat{\Omega}$ having a natural coordinate system (ξ,η), and an invertible coordinate map T_e from $\hat{\Omega}$ to Ω_e (see Fig. 4.3). We assume that T_e is an isoparametric map and we let $|\underset{\sim}{J}(\xi,\eta)|$ denote the Jacobian of the transformation T_e. We restrict our attention to a family of quadrilaterial, conforming Lagrangian elements having N_e = 4,8 or 9, where N_e is the number of nodes defining the element.

---

* The calculations which we shall describe here follow closely those described in Chapter 5 of BECKER, CAREY and ODEN (1981), to which the reader is referred for further details.

Let $\{\hat{\phi}_i(\xi,\eta)\}_{i=1}^{N_e}$ be a family of basis functions defined on $\hat{\Omega}$ (see, for example, BECKER, CAREY and ODEN, page 198). Then the family of basis functions $\{\phi_i(x,y)\}_{i=1}^{N_e}$ are obtained from

$$\phi_i(x,y) = \hat{\phi}_i(\xi(x,y),\eta(x,y)) \quad , \quad 1 \leq i \leq N_e \quad . \tag{1.36}$$

The matrix $\underset{\sim}{B}(x,y)$, which relates strain rates to discrete velocities (eqn (1.9)), is transformed into $\hat{\underset{\sim}{B}}(\xi,\eta)$ in the usual way using the Jacobian $|\underset{\sim}{J}|$, (see BECKER, CAREY and ODEN, page 189).

We assume that, given a Gaussian quadrature rule of order N_G on $\hat{\Omega}$, we can construct a suitable family of basis functions $\{\hat{\phi}_i(\xi,\eta)\}_{i=1}^{N_G}$, which are defined at the Gauss points, as shown in Fig. 6.4. Then the family $\{\phi_i(x,y)\}_{i=1}^{N_G}$ are obtained in the manner indicated in (1.36), and similarly, the matrix $\hat{\underset{\sim}{B}}_p(\xi,\eta)$ may be obtained from $\underset{\sim}{B}_p(x,y)$ in the same way as $\hat{\underset{\sim}{B}}$ is obtained from $\underset{\sim}{B}$. It is worthwhile noting, however, that if we numerically integrate functions containing $\hat{\phi}_i(\xi,\eta)$ using the same quadrature rule as was used to define the $\hat{\phi}_i$, then the shape functions $\hat{\phi}_i$ need not be explicitly defined.

The matrix $\bar{\underset{\sim}{M}}$ of normalised yield function derivatives, being a function of position on Ω_e , is transformed to a function of position on $\hat{\Omega}$ via

$$\bar{\underset{\sim}{M}}(\underset{\sim}{x}(\xi,\eta),y(\xi,\eta)) = \hat{\underset{\sim}{M}}(\xi,\eta) \quad . \tag{1.37}$$

Thus, all the constituent matrices required for the element computations are available as functions on $\hat{\Omega}$.

Let $\underset{\sim}{f}(x,y)$ be any matrix of functions defined on $\Omega_e \subseteq R^2$ and $\hat{\underset{\sim}{f}}(\xi,\eta)$ the transformation of this matrix of functions to $\hat{\Omega}$ under the inverse map T_e^{-1} . Then, writing $d\underset{\sim}{x} = dxdy$, it is clear that

$$\int_{\Omega_e} \underset{\sim}{f}(x,y)\, d\underset{\sim}{x} = \int_{\Omega_e} \hat{\underset{\sim}{f}}(\xi,\eta) \left|\underset{\sim}{J}(\xi,\eta)\right| d\xi d\eta \quad . \tag{1.38}$$

Now let $I(\cdot)$ denote the numerical quadrature formula which is to be used to evaluate the right-hand-side of (1.38), given by

$$I(\hat{\underset{\sim}{f}}) = \sum_{i=1}^{N_G} \hat{\underset{\sim}{f}}(\xi_i,\eta_i) \left|J(\xi_i,\eta_i)\right| w_i \tag{1.39}$$

where N_G is the chosen number of quadrature points, (ξ_i,η_i) are the coordinates of the i-th quadrature point, and $w_i > 0$ is the quadrature weight associated with the i-th point. We use the above formula to integrate each of the element matrices and vectors for which the relevant function $\hat{\underset{\sim}{f}}$ is given in Table 6.1.

As far as the orders of integration are concerned we always integrate $\underset{\sim}{K}^{(e)}$ exactly (see BATHE (1982), Table 5.5), but allow the option of a different order of integration for $\underset{\sim}{L}^{(e)}$, $\underset{\sim}{S}_1^{(e)}$, $\underset{\sim}{S}_2^{(e)}$ and $\underset{\sim}{F}^{(e)}$. However, the computations are considerably simplified if we use the same quadrature rule to integrate the latter matrices as was used to define the basis functions $\phi_j(\underset{\sim}{x}_i)$ in (1.11). By doing so the matrix $\hat{B}_p(\underset{\sim}{x}_i)$ takes the trivial form of eqn (1.16), with the result that the matrices $\underset{\sim}{S}_1$ and $\underset{\sim}{S}_2$ are diagonal.

TABLE 6.1

Functions $\hat{\underset{\sim}{f}}$ for integration of element matrices and vectors using (1.39).

Matrix/vector	$\hat{\underset{\sim}{f}}$
$\underset{\sim}{K}^{(e)}$	$[\hat{\underset{\sim}{B}}(\xi_i,\eta_i)]^T \underset{\sim}{C}\ \hat{\underset{\sim}{B}}(\xi_i,\eta_i)$
$\underset{\sim}{L}^{(e)}$	$[\hat{\underset{\sim}{B}}(\xi_i,\eta_i)]^T \underset{\sim}{C}\ \hat{\bar{\underset{\sim}{M}}}(\xi_i,\eta_i)\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)$
$\underset{\sim}{S}_1^{(e)}$	$[\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)]^T [\hat{\bar{\underset{\sim}{M}}}(\xi_i,\eta_i)]^T \underset{\sim}{C}\ \hat{\bar{\underset{\sim}{M}}}(\xi_i,\eta_i)\ \hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)$
$\underset{\sim}{S}_2^{(e)}$	$E_p[\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)]^T\ \hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)$
$\underset{\sim}{F}^{(e)}$	$\varepsilon^{-1}[\hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)]^T\ \hat{\underset{\sim}{B}}_p(\xi_i,\eta_i)\bar{\alpha}_{i-}$
$\dot{\underset{\sim}{P}}^{(e)}$	$\Psi^T(\xi_i,\eta_i)\dot{\underset{\sim}{f}}$ and $\Psi^T(\xi_i,\eta_i)\dot{\underset{\sim}{t}}_e$

Note: $\underset{\sim}{C}$ is a matrix of elastic constants; see, for example, Appendix A, eqn A.2.

6.2 SOLUTION OF THE PENALTY-RATE PROBLEM AND INCREMENTAL PROBLEM

In the preceding section we developed a system of nonlinear algebraic equations representing the discrete approximation of the penalty-rate problem. We now wish to consider the solution of these equations, and subsequently, the solution of the incremental problem.

Solution of the Penalty-rate Problem

We assume that at the start of any increment we know the states of stress $\underset{\sim}{\sigma}^\circ$ and plastic strain $\underset{\sim}{p}^\circ$ at each Gauss point in the model, this information being gained from the solution for the preceding increment. Let us also assume that the solution for the preceding increment included N_p positive plastic multipliers $\underset{\sim}{\bar{\alpha}}^\circ = \{\bar{\alpha}_j^\circ : 1 \leqslant j \leqslant N_p\}$. Thus, with the load rates $\underset{\sim}{\dot{P}}$ being chosen, we may construct the system of equations* (1.32), with the exception of the term $\underset{\sim}{F}(\underset{\sim}{\bar{\alpha}})$ for which the following iterative scheme is proposed.

Let $F_j^{(r)}(\bar{\alpha}_j^{(r)})$ be the j-th component of $\underset{\sim}{F}^{(r)}$ in the r-th iteration : then, for $1 \leqslant j \leqslant N_p$, and following (1.29), we set

$$\begin{aligned} F_j^{(r)}(\bar{\alpha}_j^{(r)}) &= (\bar{w}_j\bar{\alpha}_j)^{(r)} \quad , \quad \text{if } \bar{\alpha}_j^{(r-1)} < 0 \\ F_j^{(r)}(\bar{\alpha}_j^{(r)}) &= \quad 0 \qquad\quad , \quad \text{if } \bar{\alpha}_j^{(r-1)} \geqslant 0 \end{aligned} \tag{2.1}$$

---

* We include only those plastic multipliers which are either known, or assumed, to be positive. We clarify what we mean by "assumed" later.

where $\bar{w}_j = |\underset{\sim}{J}(\xi_j,\eta_j)|w_j$ from (1.39).

The system of equations (1.32) is then solved by the successive iteration of

$$\left[\underset{\sim}{K}^*\right] \begin{Bmatrix} \underset{\sim}{a}^{(r)} \\ \hline \bar{\underset{\sim}{\alpha}}^{(r)} \end{Bmatrix} + \begin{Bmatrix} \underset{\sim}{0} \\ \hline \varepsilon^{-1}\underset{\sim}{F}^{(r)}(\bar{\underset{\sim}{\alpha}}^{(r)}) \end{Bmatrix} = \begin{Bmatrix} \dot{\underset{\sim}{P}} \\ \hline \underset{\sim}{0} \end{Bmatrix} \tag{2.2}$$

for r = 1,2, ... The procedure is terminated when, at the end of the r-th iteration, the following condition is satisfied :

$$|\bar{\alpha}_j^{(r)}| \leqslant c \quad , \quad \text{for all } \bar{\alpha}_j^{(r)} < 0 \quad . \tag{2.3}$$

where $c > 0$ is some predetermined tolerance of the order of ε . At this stage each $\bar{\alpha}_j^{(r)} < 0$ is assumed to be zero (and will be discarded from the system of equations which is formed at the start of the next increment).

It has been observed that if any $\bar{\alpha}_j$ is to become negative during the iterative procedure it will do so on the first iterations, $r = 1$. This is easily explained on physical grounds since $\bar{\alpha}_j < 0$ indicates that the corresponding Gauss point is "unloading elastically". At this point a second iteration is begun during which $\bar{\alpha}_j$ is penalised, and providing ε is chosen small enough this second iteration should be sufficient to satisfy eqn (2.3). Thus, it is seldom necessary to proceed beyond two iterations of the penalty algorithm.

It is also important to note that since the terms $\varepsilon^{-1}\bar{w}_j$ are added to the diagonal elements of $\underset{\sim}{S}$, which include constants of the order of magnitude of Young's modulus E , it is crucial to the success of the penalisation procedure to use an effective penalty parameter $\bar{\varepsilon}$, which may be taken as

$$\bar{\varepsilon} = \frac{\varepsilon}{E} \quad . \tag{2.4}$$

That is, in practice ε^{-1} in (2.2) is replaced by $\bar{\varepsilon}^{-1}$.

Solution of the Incremental Problem

Recall that in the solution of the penalty-rate problem the absolute magnitude of the load rates $\dot{\underset{\sim}{P}}$ is unimportant provided the direction is correctly chosen (see Fig. 6.1). Consequently, the solution vector $\{\underset{\sim}{a} : \underset{\sim}{\bar{\alpha}}\}^T$ remains a relative quantity until such time as it is scaled by the chosen interval Δt , representing the "size of the increment".

To emphasise the fact that the solution variable $\underset{\sim}{a}$ is a displacement rate we introduce the alternative notation $\dot{\underset{\sim}{b}} \equiv \underset{\sim}{a}$, and write the solution vector as $\{\dot{\underset{\sim}{b}} : \underset{\sim}{\bar{\alpha}}\}^T$. Then, letting $\dot{\underset{\sim}{b}}_e$ and $\underset{\sim}{\bar{\alpha}}_e$ denote the restrictions to Ω_e of $\dot{\underset{\sim}{b}}$ and $\underset{\sim}{\bar{\alpha}}$ we may compute the following additional rates, each at the j-th Gauss point on Ω_e : Using (1.9) we obtain the strain rates

$$\dot{\underset{\sim}{\varepsilon}}_j = \underset{\sim}{B}(\underset{\sim}{x}_j)\dot{\underset{\sim}{b}}_e \quad ; \tag{2.5}$$

using eqn (2.22) of Section 2.2 and eqns (1.17) and (1.20) we obtain the plastic strain rates

$$\dot{\underset{\sim}{p}}_j = \bar{\alpha}_j \, \bar{\underset{\sim}{M}}_j \tag{2.6}$$

where $\bar{\alpha}_j$ is the j-th component of $\bar{\underset{\sim}{\alpha}}_e$, and $\bar{\underset{\sim}{M}}_j$ is the vector of normalised yield function derivatives at the j-th Gauss point. Finally, from eqn (1.2) of Section 5.1 we obtain the stress rates

$$\dot{\underset{\sim}{\sigma}}_j = \underset{\sim}{C}[\dot{\underset{\sim}{\varepsilon}}_j - \dot{\underset{\sim}{p}}_j] \tag{2.7}$$

where $\underset{\sim}{C}$ is a matrix of elastic constants (see Appendix A, eqn (A.4)).

We may now choose Δt and compute the following solution increments :

$$\Delta\underset{\sim}{P} = \dot{\underset{\sim}{P}}\Delta t \quad , \quad \Delta\underset{\sim}{b} = \dot{\underset{\sim}{b}}\Delta t$$

$$\Delta\underset{\sim}{\varepsilon}_j = \dot{\underset{\sim}{\varepsilon}}_j\Delta t \quad , \quad \Delta\underset{\sim}{p}_j = \dot{\underset{\sim}{p}}_j\Delta t \quad , \quad \Delta\underset{\sim}{\sigma}_j = \dot{\underset{\sim}{\sigma}}_j\Delta t \tag{2.8}$$

where $\Delta\underset{\sim}{b}$ is the displacement increment. The above calculations are performed for each Gauss point in the body and the increments are then added to the totals that existed at the start of the increment to produce updated totals.

At this stage it is necessary to check whether any of the updated stress states $\underset{\sim}{\sigma}_j$ violate the yield condition, since if this is so a smaller interval Δt will need to be chosen. This check need only be performed for those Gauss points whose stress state at the beginning of the increment lay inside the yield surface; those stress states which lay on the yield surface at the beginning of the increment will, by virtue of the inclusion in (1.32) of an associated plastic multiplier, continue to move with the yield surface, or will move back inside of it*.

Consider a typical Gauss point in the model. Referring to Fig. 6.4, we assume that the stress state at the beginning of a particular increment is $\underset{\sim}{\sigma}^{\circ}$, with plastic strains $\underset{\sim}{p}^{\circ}$, and at the end of the increment it is $\underset{\sim}{\sigma} = \underset{\sim}{\sigma}^{\circ} + \Delta\underset{\sim}{\sigma}$. It is assumed further that

$$\phi(\underset{\sim}{\sigma},\underset{\sim}{p}^{\circ}) > 0 \tag{2.9}$$

where $\phi = 0$ defines the current yield surface. Our objective is to find the scale factor ρ such that the stress

$$\underset{\sim}{\sigma} = \underset{\sim}{\sigma}^{\circ} + \rho\Delta\underset{\sim}{\sigma} \tag{2.10}$$

satisfies the yield criterion

$$\phi(\underset{\sim}{\sigma},\underset{\sim}{p}^{\circ}) = 0 \quad . \tag{2.11}$$

The stress increment $\rho\Delta\underset{\sim}{\sigma}$ will, of course, be elastic by definition.

---

* We qualify this statement later for the case of elastic-perfectly plastic materials.

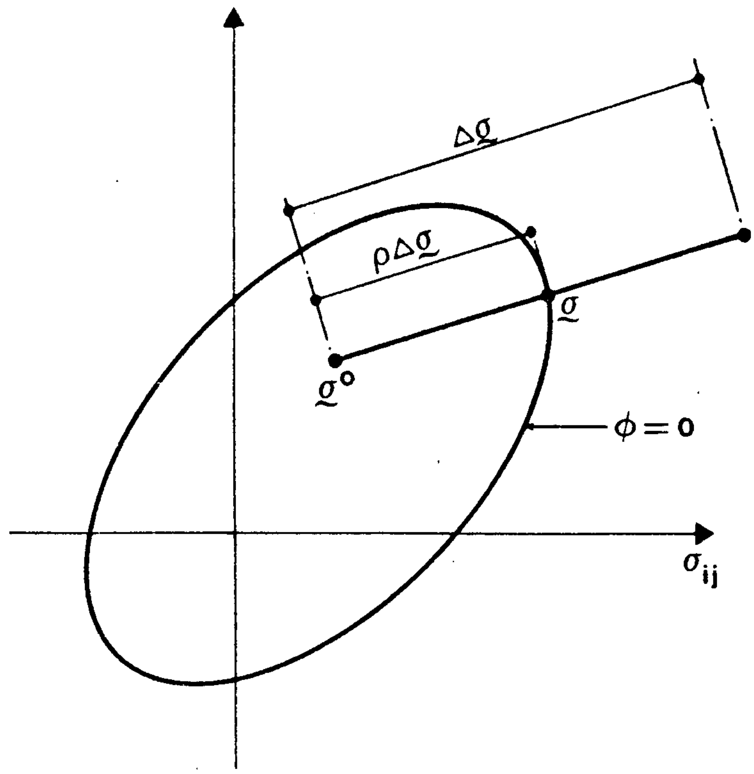


Figure 6.4 Definition of stress states for the scaling procedure.

We begin by writing down the expression for the effective stress $\bar{\sigma}$ corresponding to the stress state $\underset{\sim}{\sigma}$, using plane stress for the purposes of illustration. Thus, from Appendix A, eqn A.6, substituting from (2.11) above, we have

$$\bar{\sigma}^2 = (\sigma^o_{xx} + \rho\Delta\sigma_{xx} - E_p p^o_{xx})^2 + (\sigma^o_{yy} + \rho\Delta\sigma_{yy} - E_p p^o_{yy})^2$$

$$- (\sigma^o_{xx} + \rho\Delta\sigma_{xx} - E_p p^o_{xx})(\sigma^o_{yy} + \rho\Delta\sigma_{yy} - E_p p^o_{yy})$$

$$+ 3(\sigma^o_{xy} + \rho\Delta\sigma_{xy} - E_p p^o_{xy})^2 \quad . \qquad (2.12)$$

For convenience we define the quantities

$$\hat{\sigma}_{xx} = \sigma^o_{xx} - E_p p^o_{xx} \quad , \quad \Delta\hat{\sigma}_{xx} \equiv \Delta\sigma_{xx} \quad , \qquad (2.13)$$

and similarly for the other components. Then using (2.13) in (2.12) and simplifying, we get

$$\bar{\sigma}^2 = (\hat{\sigma}_{xx}^2 - \hat{\sigma}_{xx}\hat{\sigma}_{yy} + \hat{\sigma}_{yy}^2 + 3\hat{\sigma}_{xy}^2)$$

$$+ \rho(2\hat{\sigma}_{xx}\Delta\hat{\sigma}_{xx} - \hat{\sigma}_{xx}\Delta\hat{\sigma}_{yy} - \hat{\sigma}_{yy}\Delta\hat{\sigma}_{xx} + 2\hat{\sigma}_{yy}\Delta\hat{\sigma}_{yy}$$

$$+ 6\hat{\sigma}_{xy}\Delta\hat{\sigma}_{xy})$$

$$+ \rho^2(\Delta\hat{\sigma}_{xx}^2 - \Delta\hat{\sigma}_{xx}\Delta\hat{\sigma}_{yy} + \Delta\hat{\sigma}_{yy}^2 + 3\Delta\hat{\sigma}_{xy}^2)$$

$$= a\rho^2 + b\rho + c \quad . \tag{2.14}$$

Now (2.11) may be written as (Appendix A, eqn (A.7)),

$$\bar{\sigma}^2 - \sigma_o^2 = 0 \tag{2.15}$$

where σ_o is the unaxial yield stress in tension. Thus, combining (2.14) and (2.15) we have

$$a\rho^2 + b\rho + (c - \sigma_o^2) = 0 \tag{2.16}$$

from which the scale factor ρ is easily determined.

The above calculations are performed for every Gauss point for which the stress state $\underset{\sim}{\sigma}$ lies outside of the yield surface. From the set of values of ρ thus computed the minimum, ρ_{min} , is selected. The original choice of Δt is then scaled by ρ_{min} and the calculations in

(2.8) are repeated. The overall effect is that the analysis proceeds with the program automatically selecting the size of the interval Δt in such a way that the yield condition is never violated. Thus, if an initial choice of Δt yields a single stress path which violates the yield condition, such as that shown in Fig. 6.5(a) for example, then the scaling procedure must be invoked to compute a new value of Δt which causes the stress path at the offending point to terminate exactly at the yield surface, as shown in Fig. 6.5(b). Once a stress state lies on the yield surface it may load plastically* (Fig. 6.5(c)) or unload elastically (Fig. 6.5(d)) in subsequent increments without the need for further scaling.

The computations described above provide the only means by which a plastic multiplier can become active in the system of equations (2.2). For the first increment in a sequence of incremental problems none of the plastic multipliers are active and we solve an elastic system of equations. However, whenever a Gauss point is associated with the current ρ_{min} the plastic multiplier corresponding to that Gauss point is <u>assumed</u> to be active, that is, it will be assumed to be positive at the start of the next increment and will thereby be included in eqns (2.2).

---

* We refer to this as "plastic loading" to avoid confusion with Fig. 6.5(a). The situation shown in Fig. 6.5(c) will generally include both elastic and plastic strains.

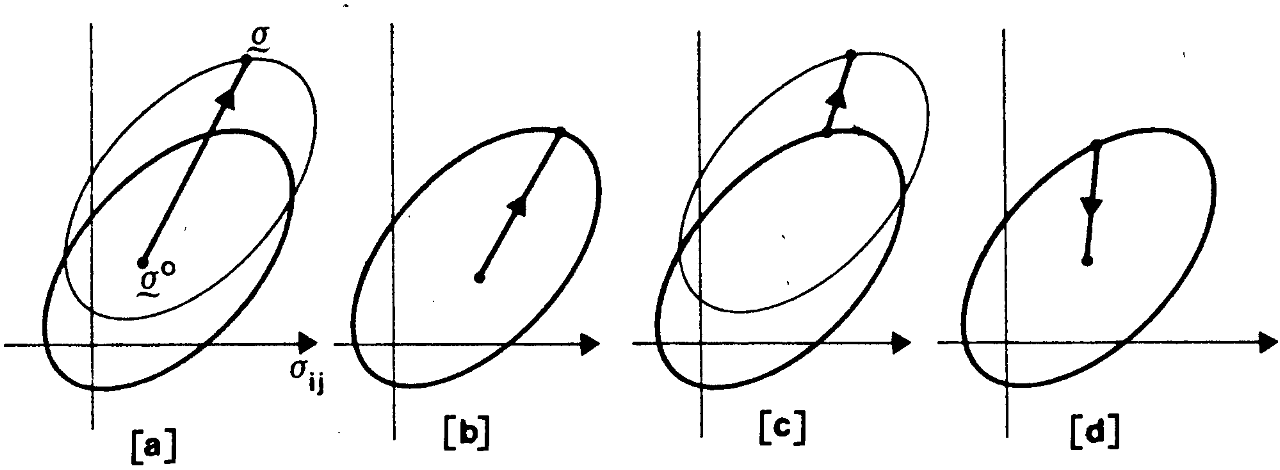


Figure 6.5 Alternative stress paths at a point in a hardening material during a single load increment. (a) inadmissible stress path (b) elastic loading (c) plastic loading (d) elastic unloading.

A final remark regarding elastic-perfectly plastic materials is required. Since the yield surface remains fixed for such materials it is not possible for plastic loading to take place as shown in Fig. 6.6(c); during plastic loading the stress point must remain on the initial yield surface. However, the numerical approximation may cause the stress point to move outside the initial yield surface so that to correct this we scale the updated total stress $\underset{\sim}{\sigma}_j$ as follows. Let $\bar{\sigma}_j$ be the effective stress and σ_o the initial yield stress; then we compute the factor $R_j = \sigma_o/\bar{\sigma}_j$ and multiply each component of $\underset{\sim}{\sigma}_j$ by R to obtain the final corrected stresses. Provided Δt is reasonably small this stress scaling procedure reduces the original error to negligible proportions.

This completes the description of the solution procedure. In the next section we briefly examine some computational aspects and compare the penalty-rate approach with a conventional rate approach to the solution of the incremental problem.

6.3 SUMMARY AND COMPARISON WITH A CONVENTIONAL APPROACH

The solution which we have presented here for the incremental elastic-plastic problem differs in principle from conventional solutions* only in the formulation of the rate problem. The equilibrium iterations and forward integration schemes normally used in conventional approaches may easily be incorporated, and whilst they would improve the accuracy of our stress calculations, they would not have any effect on the rate solution itself. In drawing comparisons with conventional methods it therefore suffices to examine in detail only the solution procedure for the rate problem itself.

The algorithm for the solution of the penalty-rate problem is given in Table 6.2. The first two steps are straightforward, although it is worth noting that $\dot{\underset{\sim}{P}}$, $\underset{\sim}{K}$, $\underset{\sim}{L}$ and $\underset{\sim}{S}$ all remain constant throughout the algorithm, and that $\underset{\sim}{S}$ is a diagonal matrix. In step 3 we compute the elements of the penalty vector according to (2.1); the non-zero elements are then added to the corresponding diagonal elements of $\underset{\sim}{S}$. In step 4 we use the Cholesky decomposition with forward/backward substitution to solve the equations (see BATHE (1982)). This is a significant choice since it allows us to assemble and triangularise the complete elastic

---

* For example, the tangent modulus approach as discussed by ZIENKIEWICZ (1977), and OWEN and HINTON (1980).

stiffness matrix $\underset{\sim}{K}$ prior to entering the iterative procedure. Thus, at step 4, all that remains is to triangularise the matrices $\underset{\sim}{L}$ and $\underset{\sim}{S}$, and perform the forward/backward substitution on $\underset{\sim}{K}^*$. Noting again that $\underset{\sim}{S}$ is a diagonal matrix the solution of the equations at step 4 obviously requires less effort than would be required if $\underset{\sim}{K}^*$ were a conventional banded matrix with no prior triangularisation. The convergence checks in step 5 are trivial, and as we have mentioned already, with a suitable choice of penalty parameter it should never take more than two iterations to obtain an acceptable solution. It is worth emphasising that the $\underset{\sim}{L}$ and $\underset{\sim}{S}$ matrices will generally change in size from one increment to the next since only plastic multipliers which are assumed or known to be positive are included in the system of equations in step 4.

We may now compare the penalty-rate formulation with the conventional tangent modulus approach by first investigating the effort required to obtain a single rate solution, and then looking at the information obtained from the respective solutions. For the penalty-rate formulation we use Table 6.2 to estimate the effort, and compare this with our estimate of the effort involved in assembling and solving eqn (1.35), which we will assume to be representative of the tangent modulus approach. The comparison is given in Table 6.3.

If we study first the effort involved in obtaining a penalty-rate solution it is evident that provided the number of plastic Gauss points in the model is relatively small (that is, $\underset{\sim}{K}^*$ is not significantly larger than $\underset{\sim}{K}_T$) , the penalty-rate formulation will involve less effort. For example, for step A.1 the efforts may be assumed roughly

TABLE 6.2

Algorithm for the solution of the penalty-rate problem.

1.	Compute and assemble the global load vector $\{\dot{\underset{\sim}{P}} : \underset{\sim}{0}\}$
2.	Compute and assemble into $\underset{\sim}{K}^*$ the matrices $\underset{\sim}{K}$, $\underset{\sim}{L}$ and $\underset{\sim}{S}$; $\underset{\sim}{L}$ and $\underset{\sim}{S}$ include only those parts corresponding to $\bar{\underset{\sim}{\alpha}}_j^{(r)}$ which are known or assumed to be positive
3.	Compute and assemble the vector $\varepsilon^{-1}\underset{\sim}{F}^{(r)}(\bar{\underset{\sim}{\alpha}}^{(r)})$
4.	Solve $\left[\underset{\sim}{K}^*\right]\begin{Bmatrix}\underset{\sim}{a}^{(r)}\\ \hline \bar{\underset{\sim}{\alpha}}^{(r)}\end{Bmatrix} + \begin{Bmatrix}\underset{\sim}{0}\\ \hline \varepsilon^{-1}\underset{\sim}{F}^{(r)}(\bar{\underset{\sim}{\alpha}}^{(r)})\end{Bmatrix} = \begin{Bmatrix}\dot{\underset{\sim}{P}}\\ \hline \underset{\sim}{0}\end{Bmatrix}$
5.	Check $\lvert\bar{\alpha}_j^{(r)}\rvert \leq c$, for all $\bar{\alpha}_j^{(r)} < 0$, where $c > 0$ is some tolerance
6.	If any check fails, return to step 3. Otherwise, the algorithm is complete and all $\bar{\alpha}_j^{(r)} < 0$ are assumed to be zero.

Notes: (i) It is assumed that $\underset{\sim}{K}^*$ and $\dot{\underset{\sim}{P}}$ are stored so that steps 1 and 2 need not be repeated at each iteration.

(ii) Step 3 is never performed on the first iteration ($r = 1$) since we always have $\bar{\underset{\sim}{\alpha}}^{\circ} \geq 0$ from the previous increment (see eqn (2.1)).

equal, bearing in mind that $\underset{\sim}{S}$ is diagonal and $\underset{\sim}{F}^{(r)}$ is a vector. At step A.3 the penalty-rate effort will always be slightly greater, but at step A.2 the penalty-rate effort will be significantly less. Conversely, as the number of plastic Gauss points increases a point will be reached at which the penalty-rate formulation becomes distinctly cumbersome, a drawback which we note does not effect the tangent modulus approach. On the other hand, turning to the solution we see that the penalty-rate formulation yields more information, in the form of the plastic multipliers, from which the plastic strains may be directly obtained. In the conventional tangent modulus approach plastic strains must be computed using a suitable state determination scheme. In fact, perhaps the most significant advantage of the penalty-rate formulation is that elastic unloading is detected and acted upon at the level of the rate problem and not at the level of the state determination, as is the case in the tangent modulus approach.

To summarise, it seems fair to conclude that provided the extent of plastic deformation in the body is relatively small, as occurs for example in a body subject only to localised stress concentrations, the penalty-rate formulation provides a more efficient and faster solution procedure than the conventional tangent modulus approach, and one which is, moreover, at least as simple to implement.

TABLE 6.3

Comparison of the penalty-rate and tangent modulus algorithms

Penalty-rate	Tangent modulus
A. Major areas of effort in solving the rate problem	
1. Compute $\underset{\sim}{L}$, $\underset{\sim}{S}$ and $\varepsilon^{-1}\underset{\sim}{F}^{(r)}$ and assemble into $\underset{\sim}{K}^{*}$	Compute plastic contributions "$\underset{\sim}{L}\underset{\sim}{S}^{-1}\underset{\sim}{L}^{T}$" and assemble into $\underset{\sim}{K}_{T}$
2. Triangularise $\underset{\sim}{L}$ and $\underset{\sim}{S}$	Triangularise $\underset{\sim}{K}_{T}$
3. FBS on $\underset{\sim}{K}^{*}$	FBS on $\underset{\sim}{K}_{T}$
B. Comparison of solution data	
Yields displacement rates and plastic multipliers, from which plastic strain rates are immediately obtainable	Yields displacement rates only; plastic strains must be computed using a suitable state determination scheme
$\bar{\alpha}_j^{(r)} < 0$ indicates "elastic unloading" at Gauss point j.	"Elastic unloading" can only be detected with additional computations

Notes: (i) FBS means forward/backward substitution

(ii) "elastic unloading" is illustrated in Fig. 6.5(d)

(iii) In part A we assume that the elastic stiffness matrix $\underset{\sim}{K}$ is already assembled into $\underset{\sim}{K}^{*}$ or $\underset{\sim}{K}_{T}$, and that in the penalty-rate case it is already triangularised ; $\underset{\sim}{K}_{T}$ cannot be triangularised prior to step A.1.

CHAPTER 7

NUMERICAL EXAMPLES

The solution methods described in Chapters 4 and 6 have been implemented in two separate computer programs called RATE and HOLO : the former uses the incremental penalty-rate formulation and the latter the incremental holonomic formulation. Both programs are limited to applications in the analysis of plane stress and to materials which obey the von Mises yield condition with either perfect plasticity or linear kinematic hardening. The element library for each program includes the 4, 8, and 9 node quadrilateral elements with conventional displacement interpolation, but each element may use either constant, linear, or quadratic interpolation for the plastic multipliers or strains, as the case may be.

The two programs must be regarded at this stage as pilot programs because of their somewhat limited scope and the fact that very little effort has been devoted to their efficient operation. Thus, in our view, exhaustive testing and comparison with other numerical solutions would be premature at this stage. What we need to do is determine the characteristic behaviour of the numerical solutions, especially in those areas where our theoretical developments have pointed to potential problems. For example, how well does the penalty-rate algorithm handle elastic unloading; what influence does the regularised stress-plastic strain curve for the holonomic formulation have on the convergence of the iterative solution, and so forth. These are the sorts of questions which we will address here and on the basis of the answers obtained we may judge whether our two methods are worthy of further development, and assuming this to be so, under what conditions.

We plan to discuss the following behavioural characteristics in this chapter :

(i) the role of the penalty and regularisation parameters ε in each solution, and the effect of varying their magnitudes;

(ii) the improvement of both solutions as the interval Δt is reduced, that is, as the number of subdivisions of the load path is increased;

(iii) the relationship between the RATE and HOLO solutions;

(iv) the behaviour of the solutions in limit load and cyclic loading/unloading analyses.

In the case of the RATE solution we also investigate its efficiency with respect to the extent of plastic deformation in the model.

7.1 THE EFFECT OF THE REGULARISATION PARAMETER ε IN THE INCREMENTAL HOLONOMIC SOLUTION

We consider here a square block of material subjected to a uniaxial tensile load; several loading/unloading analyses will be performed in which only the value of the regularisation parameter ε is changed. The single element model and the data for the analyses are given in Fig. 7.1. Since the problem is statically determinate the stresses are known a priori and are obviously $\sigma_x = p_1$, $\sigma_y = p_2$, and are constant throughout the element.

We apply the load $p_1 = 1.1$ ($p_2 = 0$) in a single increment, (that is, we treat the holonomic or deformation theory problem), this load being sufficient to cause significant plastic deformation, and monitor

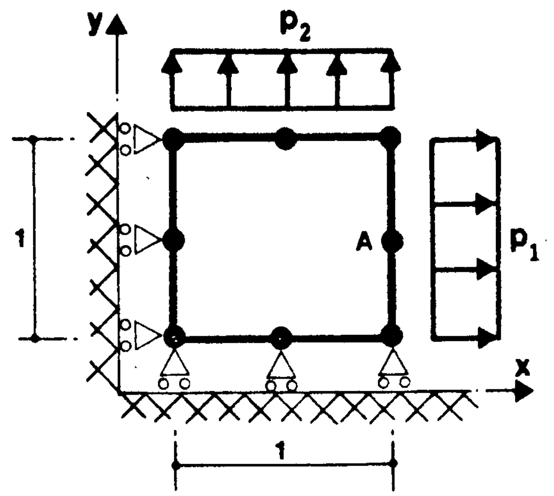
$E = 1.0$, $E_p = 0.25$

$\nu = 0$.

$\sigma_o = 1.0$

thickness = 1.0

ε = variable



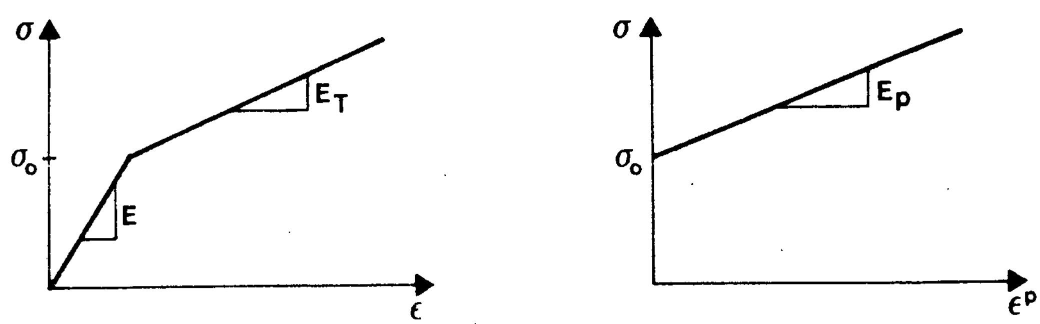


Figure 7.1 Single element model and corresponding data.

the x displacement at A , u_x , and the plastic strain ε_x^p as the solution iterates towards a convergent solution. These results are given in Fig. 7.2 for various values of ε . The initial conditions for the increment are the elastic displacement $u_x = 1.1$ and $\varepsilon_x^p = 0$ (see Section 4.2, eqn (2.12)); for $\varepsilon \leqslant 10^{-3}$ the solutions then proceed via different paths to the identical convergent solution which is reached after 9 iterations.

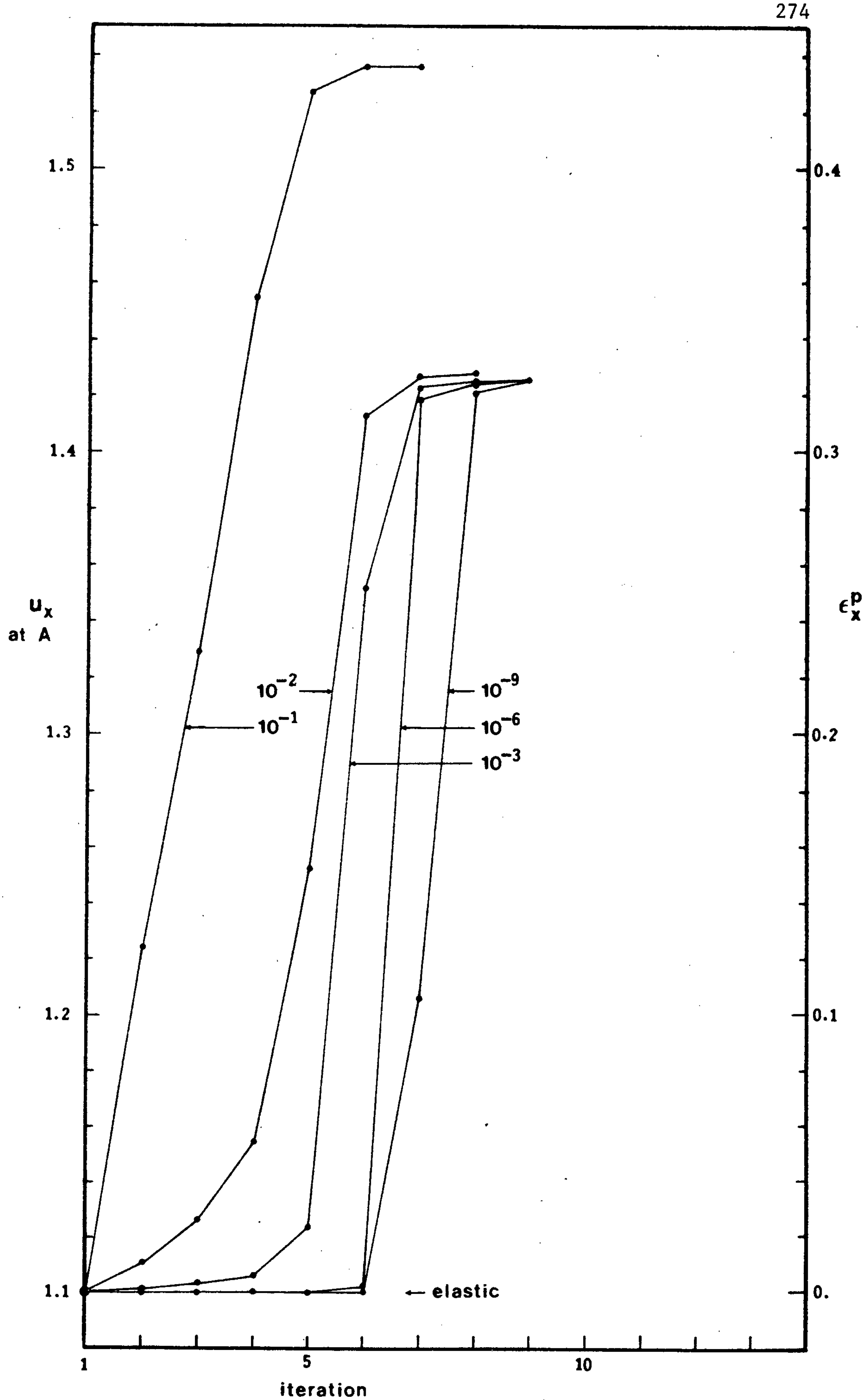


Figure 7.2 Solution behaviour over a single increment as a function of the regularisation prarmeter ε .

For $\varepsilon > 10^{-3}$ the solutions show a tendency to converge much faster but to a result which is greater than the correct one. These observations are summarised in Fig. 7.3 which shows the converged solution for u_x (for each value of ε) plotted against the corresponding value of $-\log \varepsilon$.

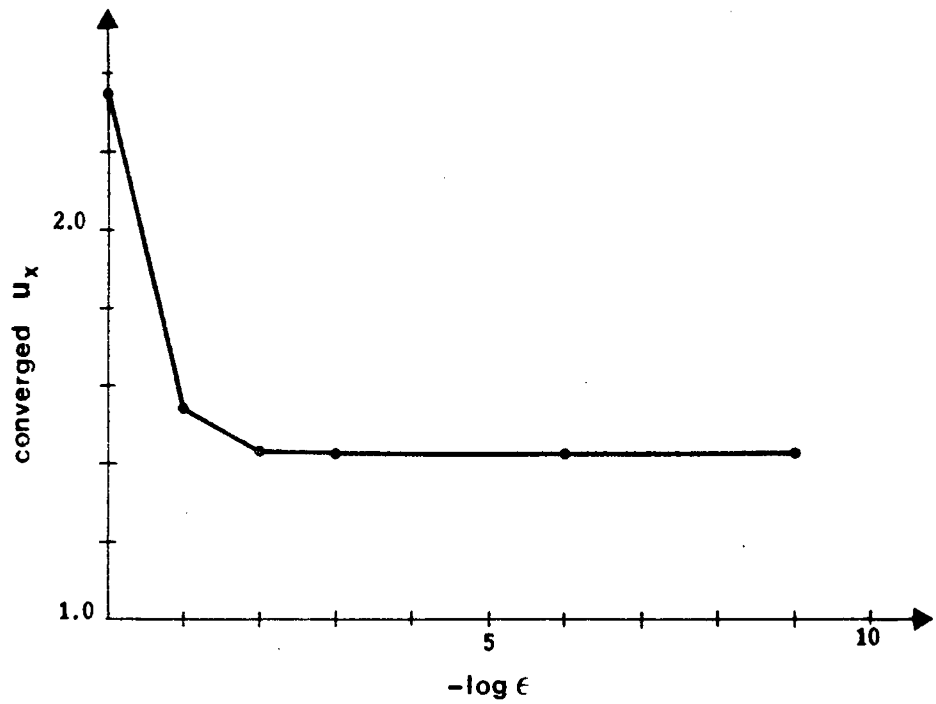


Figure 7.3 Convergent u_x for each value of ε .

We have already anticipated and attempted to explain this behaviour in Section 4.3, but a few additional comments may be useful. The important point to realise is that the regularised stress-plastic strain curve (see Fig. 4.4) approaches a step-function as $\varepsilon \to 0$, and it is the slope of this curve which is required in the Newton iterative procedure. Thus, for $\varepsilon = 10^{-9}$ the initial slope is very large and remains essentially constant until the yield stress is reached; at this point there is a very rapid change in slope towards the asymptotic value E_p . As ε becomes larger the stress-plastic strain curve becomes less "step-like" with the result that the changes in slope are smoother and more regular. This explanation is clearly evident in Fig. 7.2.

If we now remove the existing load completely in a single increment we find that, notwithstanding the assumption of elastic initial conditions, three iterations are required to obtain a convergent solution. However, this is somewhat misleading because in fact the correct elastic unloading solution is obtained in the first iteration, except that the stresses are not identically zero (they are of the order of 10^{-12}). Thus, a very small plastic strain increment must be found in order to satisfy the constitutive equations and this is why the additional iterations are required (the correct plastic strain increment should be zero). The additional two iterations contribute nothing meaningful to the solution so that it would be useful to find some formal means of dispensing with these. It is worth noting that the slopes of the unloading curves in the first iteration are exact for all ε , indicating that the unloading behaviour is independent of ε .

7.2 THE EFFECT OF THE PENALTY PARAMETER ε IN THE PENALTY-RATE SOLUTION

We again make use of the single element model and uniaxial loading shown in Fig. 7.1 and perform several analysis changing only the value of the penalty parameter ε .

Provided that during any loading increment each Gauss point in the model remains elastic or continues to load plastically (see Fig. 6.5(b), (c)) then no penalisation is required and the solution proceeds without any form of iteration. Thus, in considering the effect of ε we need only investigate elastic unloading conditions.

We begin by loading the model, using a sequence of increments, to the point $\sigma_x = p_1 = 1.1$ at which significant plastic deformation has taken place (Fig. 7.4(a)). We then remove the load in a single increment. During the first iteration of this unloading increment the plastic multipliers at each Gauss point will become negative; this indicates that the Gauss point is attempting to unload and the penalisation procedure must be initiated (see Section 6.2). Thus, a second iteration is begun in which the plastic multipliers at the unloading Gauss points are penalised. At the end of this iteration they will have negative values whose magnitudes are of the order of ε. Providing that ε is chosen small enough these values may be taken as being effectively zero and the objective of the penalisation is assumed to be accomplished. In practice, therefore, we never perform more than two iterations since we assume that ε will be chosen small enough. Nevertheless, it is of interest to investigate the effect of varying the magnitude of ε in order to determine how small this parameter should be.

The unloading lines for several values of ε are also shown in Fig. 7.4(a). The slopes of these lines vary between 0.9999 for $\varepsilon = 10^{-9}$ and 0.9260 for $\varepsilon = 10^{-1}$ (the exact slope is 1.0). The residual displacement at the end of the unloading increment is plotted against $-\log \varepsilon$ in Fig. 7.4(b), showing the relative insensitivity of the final displacement to ε, for $\varepsilon \leqslant 10^{-3}$. Since it is obvious that the plastic strain increment must be zero during unloading we take the liberty of enforcing this condition in practice, thus overriding the theoretical predictions given earlier in Fig. 5.3 (at the end of Section 5.5). Nevertheless, those predictions are clearly qualitatively true in the stress-strain case discussed here since it is clear from Fig. 7.4(a) that as $\varepsilon \to 0$ the slope of the unloading line clearly approaches the elastic slope E.

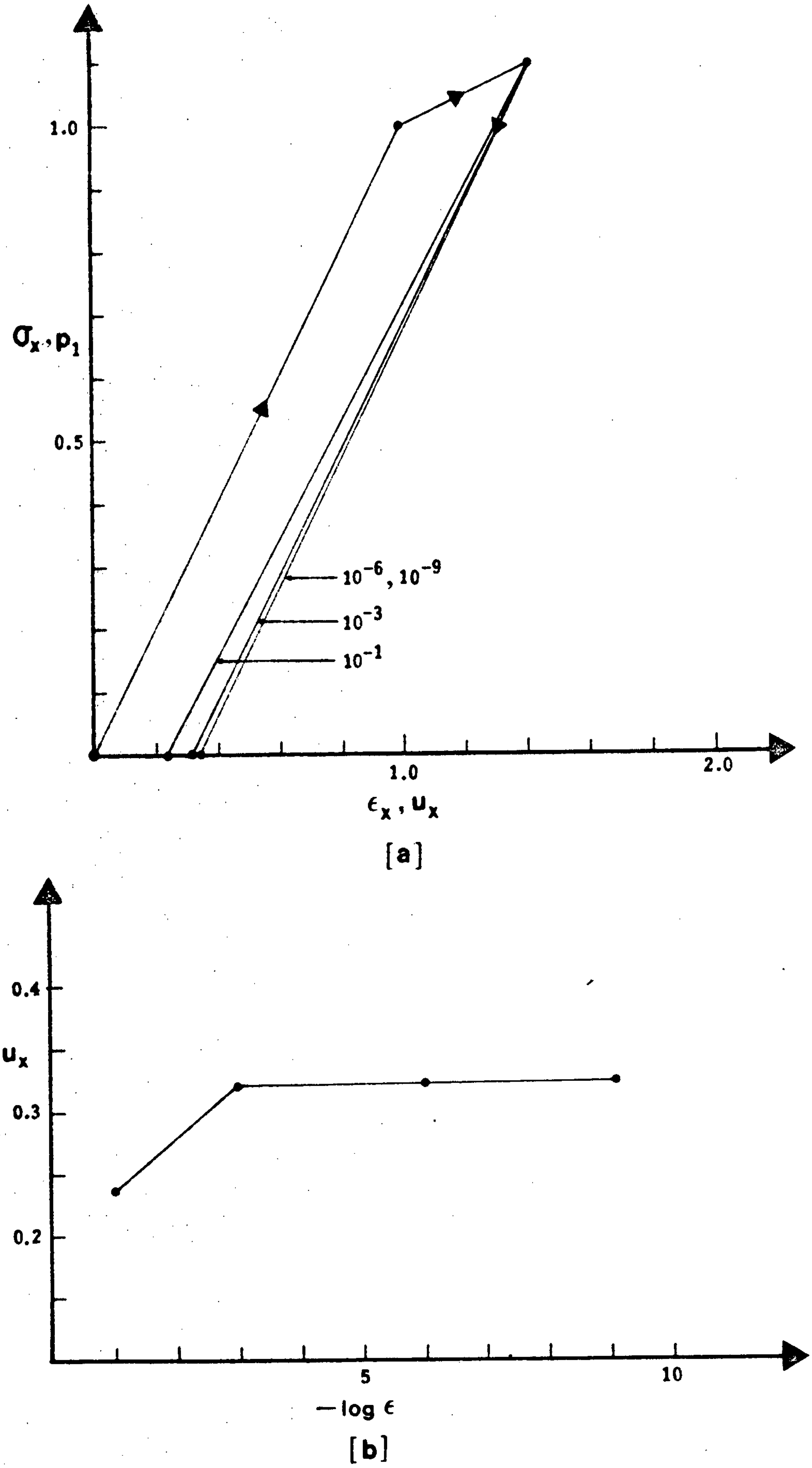


Figure 7.4 (a) stress-strain curve for unaxial loading/unloading
(b) residual displacement u_x after unloading, as a function of ε .

On the basis of the above results and our general experience with the penalty algorithm we conclude that the penalty solutions are remarkably insensitive to the magnitude of ε . Moreover, the algorithm appears to be both robust and stable : for example, our general experience confirms that local unloading does not give rise to local numerical instability even in potentially unstable situations, as often occur, for example, in the vicinity of a limit load.

7.3 CONVERGENCE WITH RESPECT TO Δt OF THE INCREMENTAL PENALTY-RATE AND HOLONOMIC SOLUTIONS

In the introduction to Chapter 6 we stated that because of our use of the Euler forward method to advance the solution across an increment, the RATE solution depends for its accuracy on the size of the interval Δt , or in practice, the load increment $\Delta \underset{\sim}{P}$. Although no similar argument has been presented for the incremental holonomic problem we nevertheless assume, if only intuitively, that the solution must improve in some sense as the size of Δt is reduced. Our objective here is to demonstrate this.

First, however, we wish to show that under a certain assumption to be given below, the penalty-rate solution is exact. MARTIN (1975a), Section 5.5, has given the analytical solution for the plastic strain rates at a material point in a state of biaxial tension, obeying a von Mises yield condition with linear kinematic hardening. If we assume that all rate quantities in the analytical solution may be integrated using a simple Euler forward method then the analytical solution can be compared directly with the results of the penalty-rate formulation. Thus, we replace the rate quantities in the analytical solution with finite increments; for example $\dot{\sigma}_x$ is replaced with $\Delta\sigma_x$.

Let us assume that the current stress point is (1.0 , 0.) on the initial yield surface (Fig. 7.5(a)). We apply a stress increment $\Delta\sigma_x = 0.05$, and compute the following quantities (Martin, pages 252, 253) :

$$\frac{\partial\phi}{\partial\sigma_x} = 2(\sigma_x - E_p\ \varepsilon_x^p) - (\sigma_y - E_p\ \varepsilon_y^p) = 2.0$$

$$\frac{\partial\phi}{\partial\sigma_y} = 2(\sigma_y - E_p\ \varepsilon_y^p) - (\sigma_x - E_p\ \varepsilon_x^p) = -1.0$$

$$\beta = \frac{\frac{\partial\phi}{\partial\sigma_x}\Delta\sigma_x + \frac{\partial\phi}{\partial\sigma_y}\Delta\sigma_y}{(\frac{\partial\phi}{\partial\sigma_x})^2 + (\frac{\partial\phi}{\partial\sigma_y})^2} = 0.02$$

$$\Delta\hat{\sigma}_x = \beta\frac{\partial\phi}{\partial\sigma_x} = 0.04$$

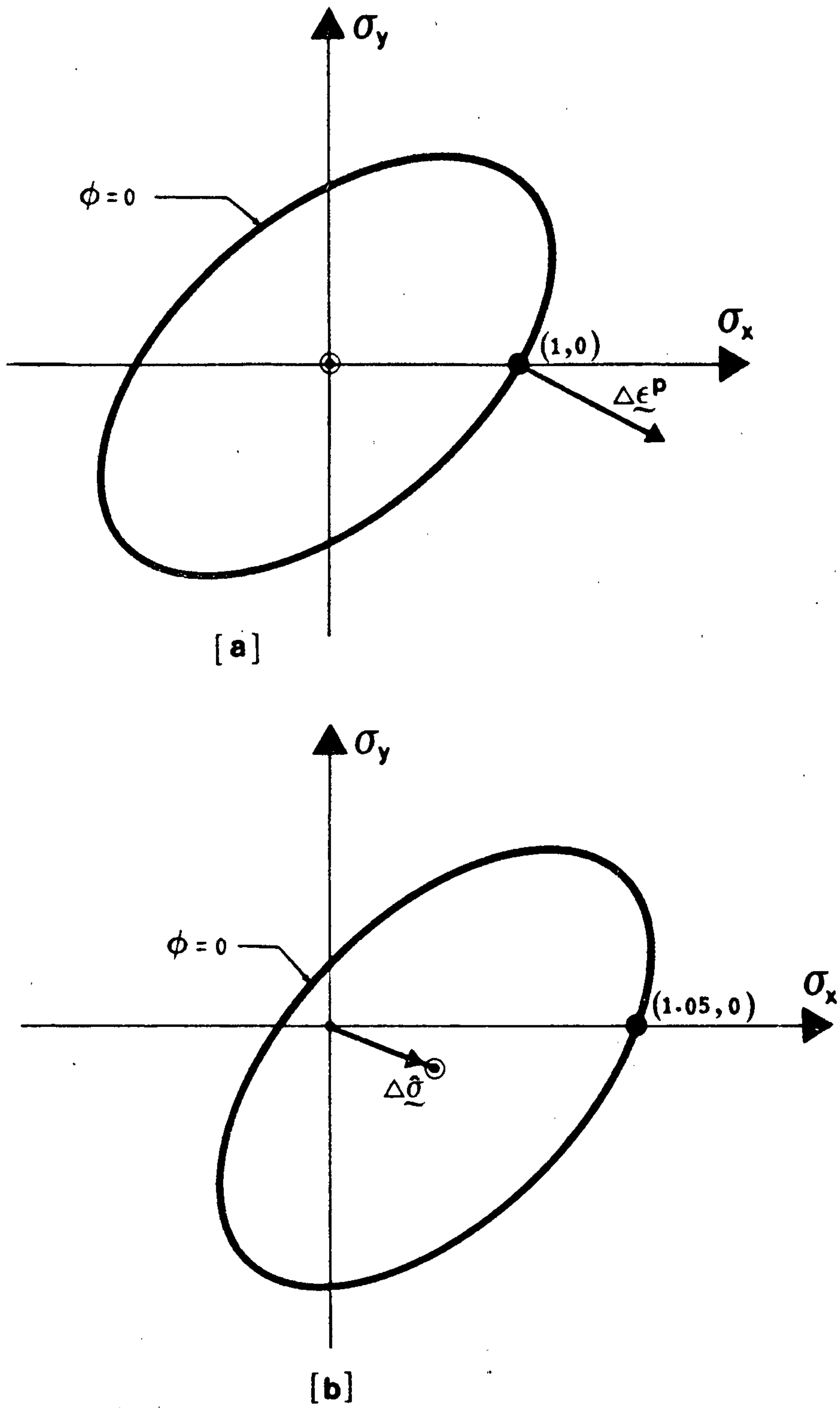
$$\Delta\hat{\sigma}_y = \beta\frac{\partial\phi}{\partial\sigma_y} = -0.02$$

$$\Delta\varepsilon_x^p = \frac{1}{E_p}\Delta\hat{\sigma}_x = 0.16$$

$$\Delta\varepsilon_y^p = \frac{1}{E_p}\Delta\hat{\sigma}_y = -0.08 \quad .$$

The centre of the von Mises ellipse now has coordinates (0.04, -0.02) , and the stress point is (1.05, 0.) as indicated in Fig. 7.5(b). We now apply a second stress increment $\Delta\sigma_x = 0.05$ and, following the same calculations as those above, we obtain

$$\Delta\varepsilon_x^p = 0.161914 \quad , \quad \Delta\varepsilon_y^p = -0.078528 \quad .$$



$$\phi = (\sigma_x - E_p\ \varepsilon_x^p)^2 - (\sigma_x - E_p\ \varepsilon_x^p)(\sigma_y - E_p\ \varepsilon_y^p) + (\sigma_y - E_p\ \varepsilon_y^p)^2 - \sigma_o^2$$

Figure 7.5 Translation of the von Mises ellipse at a material point in uniaxial tension. (a) initial ellipse (b) ellipse at the end of the first load increment.

We now perform the same analysis using the RATE program and the model given in Fig. 7.1 (which effectively represents a material point in a body), and we obtain precisely the same results, correct to 6 significant figures. Thus, we have confirmed the validity of our solution under the assumption of Euler forward integration and feel confident that our solution should approach the continuous elastic-plastic solution as $\Delta t \to 0$. This we demonstrate next.

We consider a monotonically increasing load p_1 , $0 \leqslant p_1 \leqslant 1.10$. For $p_1 \leqslant 1.0$ the behaviour is elastic and we may use a single increment to arrive at the stress point $(\sigma_x , \sigma_y) = (1.0, 0.)$ on the initial yield surface. Thereafter we increase the loading to $p_1 = 1.10$ using three different increment sizes : $\Delta p_1 = 0.1$, $\Delta p_1 = 0.05$, and $\Delta p_1 = 0.01$. The plastic strain components at the end of this program of loading are shown plotted against the interval* size Δt in Fig. 7.6. Also shown are the corresponding results for the incremental holonomic solutions. We observe that :

(i) both the penalty-rate and incremental holonomic solutions convergence linearly in Δt to the same solution, and

(ii) the penalty-rate solution converges from below, whereas the incremental holonomic solution converges from above; this is more than likely due to the use of forward difference

---

* To be consistent we used Δt to measure the size of increment but it is obviously synonymous with $\Delta\sigma_x$ in this particular case.

integration in the penalty-rate solution as opposed to backward difference integration in the case of the incremental holonomic solution.

Although we do not have an "exact" solution with which to compare, it still seems reasonable to conclude that the result towards which our two solutions converge as $\Delta t \to 0$ represents in some sense the "exact" or "continuous" solution. We should emphasise, in view of what we will discuss in the next section, that we do not lay claim to a general validity for the HOLO results in Fig. 7.6; we include them primarily out of interest and as an encouragement to further numerical experiment.

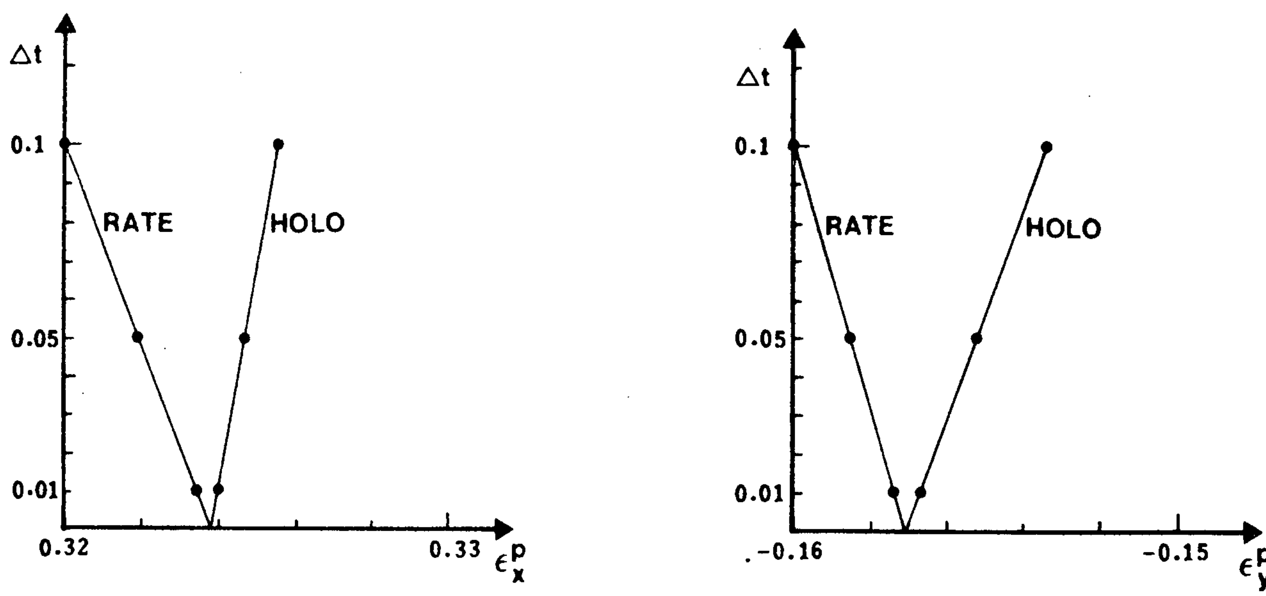


Figure 7.6 Convergence of the RATE and HOLO solutions with respect to Δt .

7.4 CONVERGENCE OF THE INCREMENTAL HOLONOMIC SOLUTION TO THE INCREMENTAL RATE SOLUTION

The incremental holonomic problem is formulated in terms of finite increments in the strains, stresses and so on. Thus, if we were to subject our single element model to a state of biaxial loading as defined by the load path in Fig. 7.7 we could find various ways of approaching the problem of determining the solution for the load point F. We might apply the loads (p_1 , p_2) corresponding to the point F directly in a single increment (that is, follow the load path AF), thereby ignoring the given load path OABCDEF. This would constitute the holonomic or deformation theory problem, but, assuming that the given load path OABCDEF gives rise to plastic deformation, we would not expect the holonomic solution to be completely accurate. Alternatively, we may attempt a sequence of incremental holonomic solutions by following the path OAF or OADF, for example; we would expect each of these sequences to yield a better estimate of the solution at F than the holonomic solution. In fact, as we follow the given load path OABCDEF with a sequence of incremental holonomic solutions we would expect the solution at F to approach that given by a corresponding sequence of incremental rate solutions.

Consider the following sequences of incremental holonomic problems (Fig. 7.7) : OF, OAF, OABF, OABCF, OABCDF, and OABCDEF. Each sequence includes one more increment than its predecessor, and in each case the additional increment is obtained by subdividing the last increment of preceding sequence; thus, to obtain the last sequence, for example, we divide the increment DF into two increments DE and EF. Now, we recall from Section 4.3 that the incremental holonomic constitutive equations are effectively updated at the beginning of every increment. Thus, we

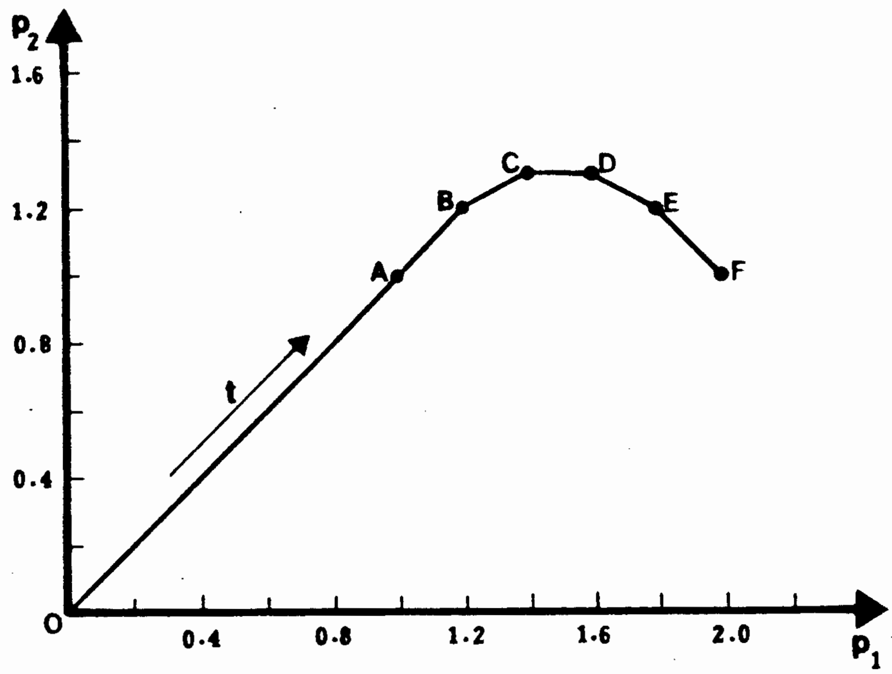


Figure 7.7 Piecewise-proportional load path.

may consider the sequence OF to embody 1 update (we consider the start of a sequence as an update), sequence OAF to embody 2 updates, and so forth. Clearly, we may choose as many points for updating as we like, so that as the number of updates $n \to \infty$ we expect to approach the continuous solution, as would be obtained, for example, by a sequence of incremental rate problems.

As before, let us parametrise the load path using the parameter t, $0 \leqslant t \leqslant \tau$, and let $[\Omega(t)]_n$ be the complementary work obtained by following any given path up to the point t, along which n updates to the constitutive equations have been made : the updates must be made at points on the actual load path, but the path followed between updates may be arbitrary. Now MARTIN (1986) has shown that for sequences of consistently formulated incremental holonomic problems

$$[\Omega(t)]_{n=1} \geq [\Omega(t)]_{n=2} \geq \text{------} \geq [\Omega(t)]_{n\to\infty} \quad . \qquad (4.1)$$

This continued inequality holds only if the sequence used to obtain $[\Omega(t)]_{n=i+1}$, is obtained by subdividing the last increment of the sequence used to obtain $[\Omega(t)]_{n=i}$, as we have already described above.

We have performed a numerical experiment to check the validity of eqn (4.1) using the single element model of Fig. 7.1 , and the load path shown in Fig. 7.7. Our results, given in Table 7.1, confirm the validity for $t = \tau$.

TABLE 7.1

Confirmation of Eqn (4.1) for $t = \tau$

Load Path	n	$[\Omega(\tau)]_n$	Remarks
OF	1	3.519	the holonomic solution
OAF	2	3.519	incremental holonomic solutions (n = 2 to 6)
OABF	3	3.417	
OABCF	4	3.334	
OABCDF	5	3.288	
OABCDEF	6	3.273	
OABCDEF	∞	3.213	penalty-rate solution for 41 increments. We take this solution to represent $n \to \infty$.

We have also compared the complementary work given by the incremental holonomic and penalty-rate solutions at each update point along the path OABCDEF. This comparison is shown in Table 7.2 and confirms that at each update point $\Omega(t)$ computed via the incremental holonomic sequence is greater than or equal to $\Omega(t)$ as computed via the incremental penalty-rate sequence.

TABLE 7.2

Comparison of $\Omega(t)$ for $n = 6$ and $n = \infty$

Point	$[\Omega(t)]_n$	
	$n = \infty$	$n = 6$
A	0.700	0.700
B	1.168	1.168
C	1.771	1.772
D	2.337	2.344
E	2.803	2.828
F	3.213	3.273

The plastic strain paths corresponding to the load paths given in Table 7.1 are shown in Fig. 7.8. The paths indicate quite clearly the convergence of the incremental holonomic solution towards the penalty-rate solution. An additional incremental holonomic solution ($n = 11$) has been included here to show that further subdivision of each of the increments used in the $n = 6$ solution yields a solution which is still closer to the penalty-rate solution.

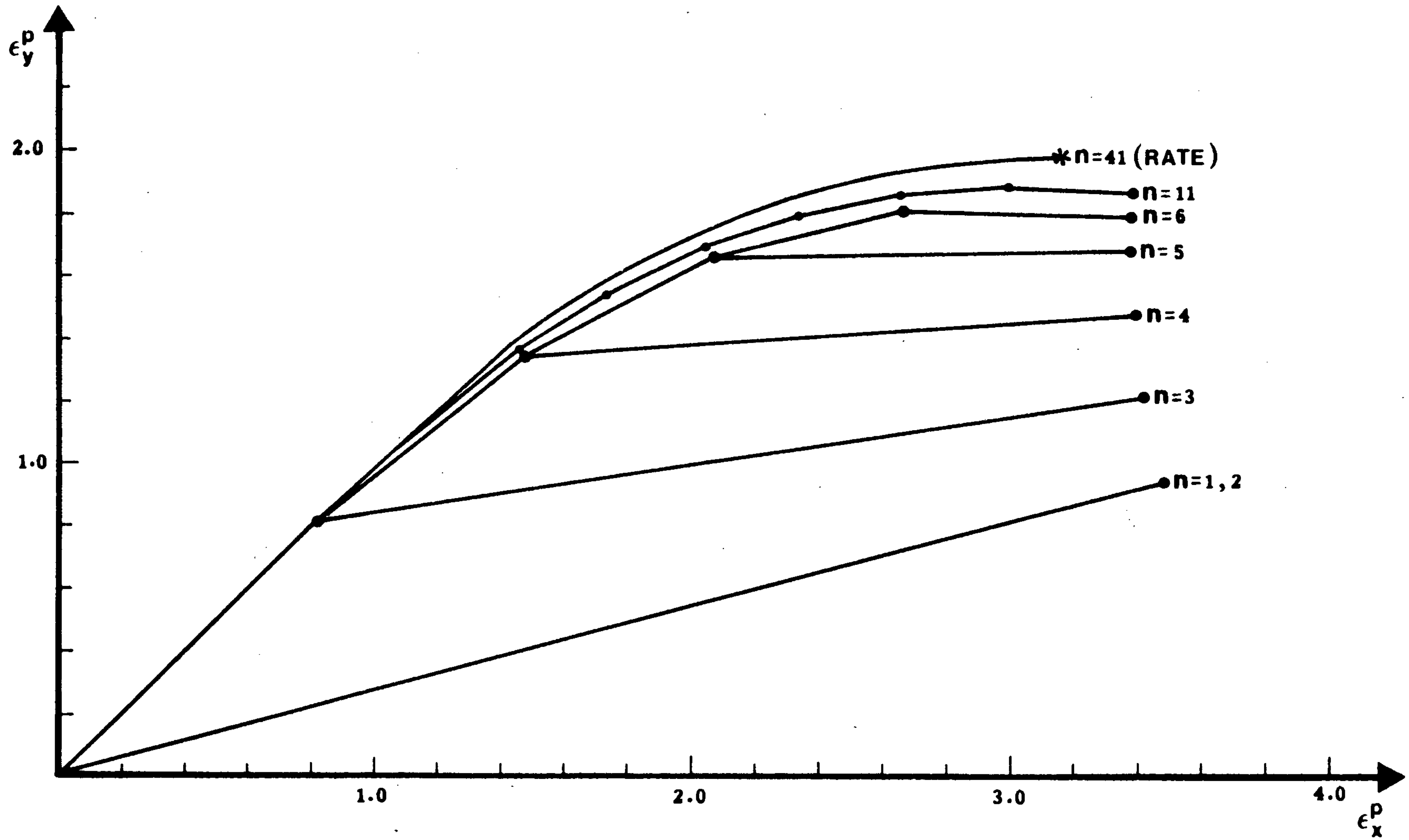


Figure 7.8 Showing convergence of the plastic strain solution as computed via HOLO to that computed via RATE.

The result of Martin's which we have briefly discussed here must be regarded as preliminary in the sense that there are rather stringent constraints imposed on the manner in which the load path may be subdivided. The ultimate objective is a principle which may in one sense be regarded as unifying the holonomic and rate problems. Certainly it seems intuitively obvious that the rate problem is a limiting case of the incremental holonomic problem as $\Delta t \rightarrow 0$ and we have demonstrated numerically that this is so at least for certain sequences of loading increments.

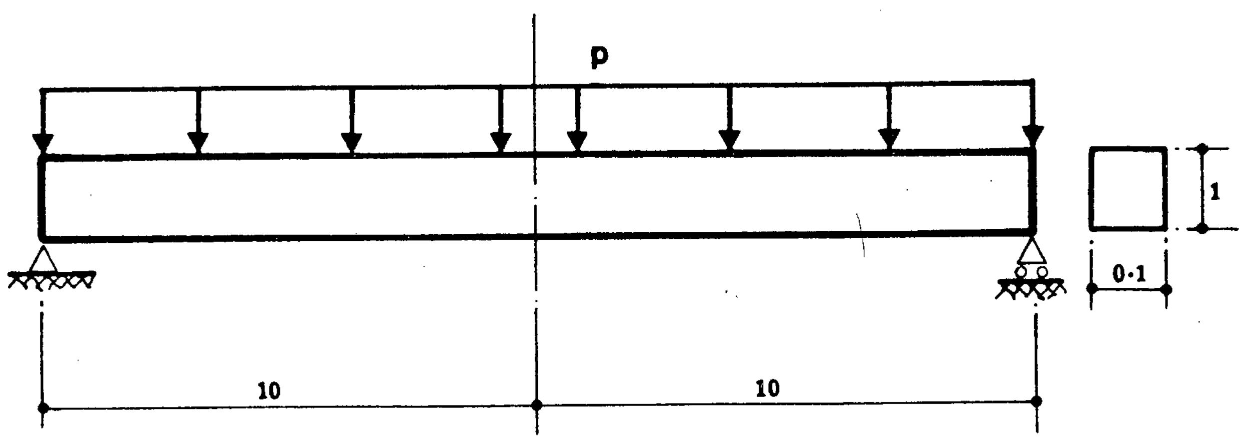
7.5 LIMIT LOAD ANALYSIS

The complete analytical solution for a simply supported beam under a uniformly distributed load has been given by PRAGER and HODGE (1951), where the material behaviour of the beam is assumed to be elastic-perfectly plastic. We have written a computer program which computes the continuous load-displacement curve for the centre of the beam and we will compare both the RATE and HOLO solutions with this. In addition we have used the program* ABAQUS to obtain an alternative finite element solution, based on the tangent stiffness approach. We discuss briefly the spread of plasticity through the beam up to the point of failure and comment on the efficiency of the RATE solution in particular. We also investigate the effect of different orders of integration for the plastic strains and stresses.

---

* ABAQUS is a proprietary finite element analysis package, developed and marketed by Hibbitt, Karlsson and Sorensen Inc., Providence, RI.

The beam and its properties are shown in Fig. 7.9. Due to symmetry we model only a quarter of the beam using ten quadrilateral 8-noded elements of equal size. The element stiffnesses are computed exactly using 3x3 Gaussian integration, but the interpolation of the plastic multipliers (RATE) and plastic strains (HOLO) may be based on either 3x3 or 2x2 Gaussian integration, i.e. either quadratic or linear (recall Figs. 4.2 and 6.3). The uniformly distributed load p is modelled by consistent nodal loads applied to the upper edge of the beam. Note that although we have proved existence of a solution to this problem only for $E_p > 0$, this does not imply non-existence of a solution for $E_p = 0$.



$E = 200\ 000$ $\nu = 0$ $\sigma_o = 250$ $E_p = 0$

Penalty (regularisation) parameter $\varepsilon = 10^{-9}$.

Figure 7.9 Simply supported beam and its material properties.

The six solutions which we propose to discuss here are summarised in Table 7.3. In the case of ABAQUS and RATE the size of the interval Δt , or the load increment Δp , is automatically chosen by the program. ABAQUS bases its choice on a user-specified equilibrium force tolerance, and includes Newton-Raphson iterations to ensure that equilibrium is maintained at the end of each increment. The number of increments used in the case of RATE solutions is governed by the stress scaling procedure described in Section 6.2. For the HOLO solutions we purposely chose few increments in order to demonstrate just how few are needed to yield good accuracy.

TABLE 7.3

Solution Statistics for the Limit Load Analysis

Name	Formulation	Plastic multiplier/ strain interpolation	No. of load increments	Average iterations per increment
Analytic	Analytical	-	-	-
ABAQUS	Tangent stiffness	-	15	3
RATE-2	Incremental penalty-rate	2x2	33	1
RATE-3		3x3	32	1
HOLO-2	Incremental holonomic	2x2	6	12
HOLO-3		3x3	6	13

The load-displacement curves and limit loads for the six solutions are shown in Fig. 7.10. The lower curve is the analytical solution for which the exact limit load is p = 0.125. The upper curve is the solution (to within the limits of the plotting scale) for ABAQUS, RATE-3 and HOLO-3; the RATE-2 and HOLO-2 solution lie between these two curves (and have been omitted for clarity), with the RATE-2 solution being the closer to the analytical solution.

In perfectly plastic analyses it is common for the solution to become unstable (or to exhibit "thrashing") in the vicinity of the limit load. Thus, the given limit loads are those for the solution just prior to the onset of what we have deemed to be unstable or meaningless behaviour. In all but the case of RATE-3 the limit loads occur at points well beyond the indicated limit of the displacement axis.

The general agreement between the various solutions is excellent, with the RATE-2 and HOLO-2 solutions being marginally closer to the analytical. The former solutions appear to be "less stiff" than the other numerical solutions, which is what we would expect from the use of a lower order of interpolation for the plastic multipliers/strains, and a lower order of integration to evaluate the plastic stiffness. This is well illustrated by the fact that although RATE-2 and RATE-3 have almost the same limit load, the RATE-2 displacement just prior to failure is much larger than that of RATE-3.

The plastic strain distibutions in the beam immediately prior to failure are shown for the five numerical solutions in Fig. 7.11. The difference between the RATE-3 and ABAQUS results, whose load-displacement

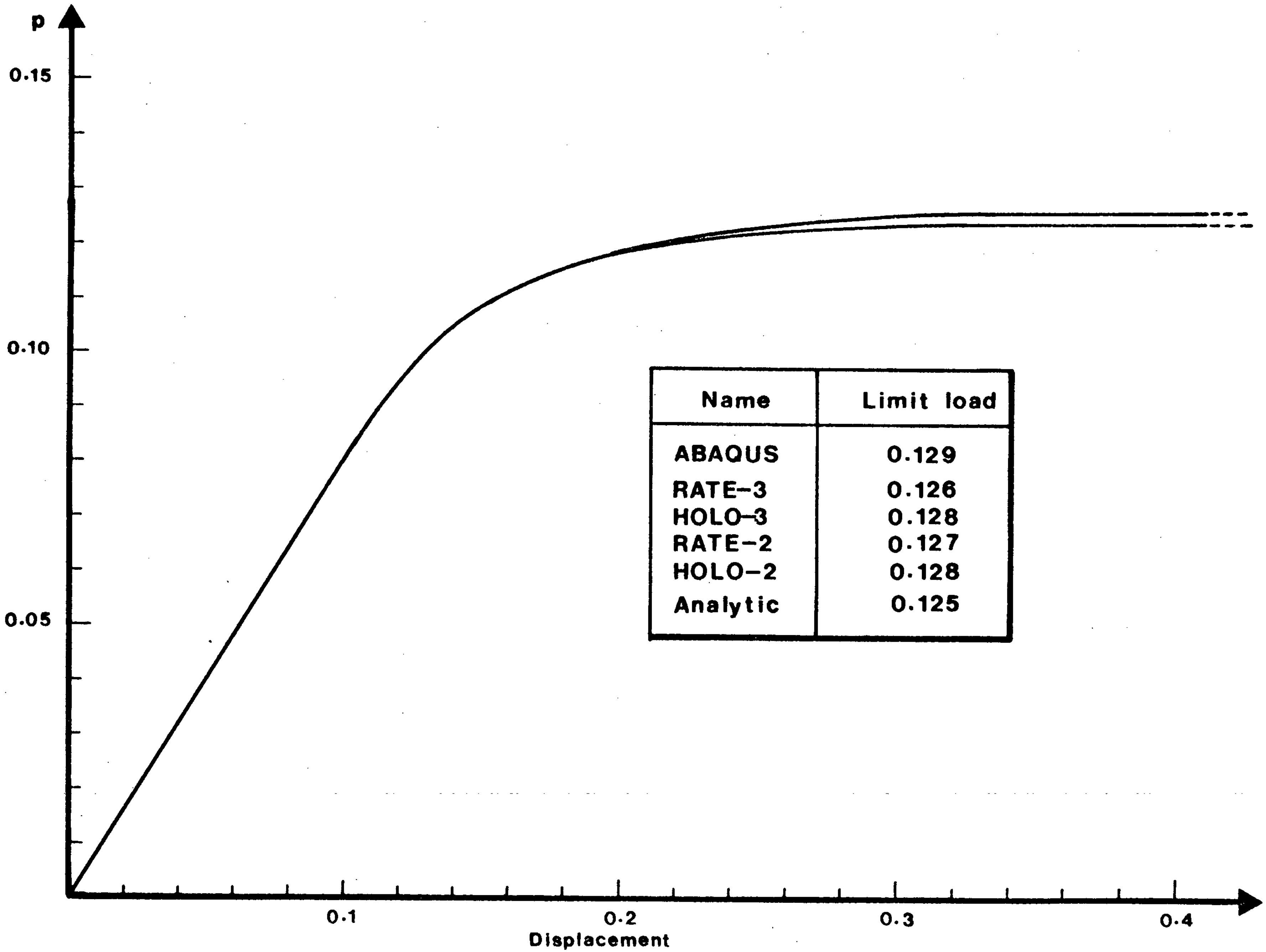


Figure 7.10 Load-displacement curves and limit loads for the simply supported beam.

solutions coincide, is due to the fact that the RATE-3 analysis fails much earlier than the ABAQUS analysis. For the HOLO analyses the user has direct control of the size of the load increments so that we have not been able to capture exactly the failure point and a larger spread of plasticity is clearly possible. It is also interesting to note that in the RATE-2 and ABAQUS cases, solutions have been obtained when the beam section has clearly failed in principle.

We turn now to the question of the relative efficiency of the incremental penalty-rate and holonomic solutions. If we study first the average statistics given in Table 7.4 we observe that, although there is a significant difference in the total CPU time for each solution, the CPU times per iteration show a different picture altogether. This observation is particularly encouraging with regard to the HOLO solutions for it seems to indicate that if we could drastically reduce the number of iterations per increment, we would have a solution procedure which is potentially very fast. For example, if we could cut down to three iterations per increment the HOLO-2 solution would take approximately 150 CPU seconds, which is competitive even by ABAQUS standards. The crucial advantage of the HOLO solutions is that they require fewer <u>increments</u> to obtain good solutions than do the rate approaches, RATE and ABAQUS.

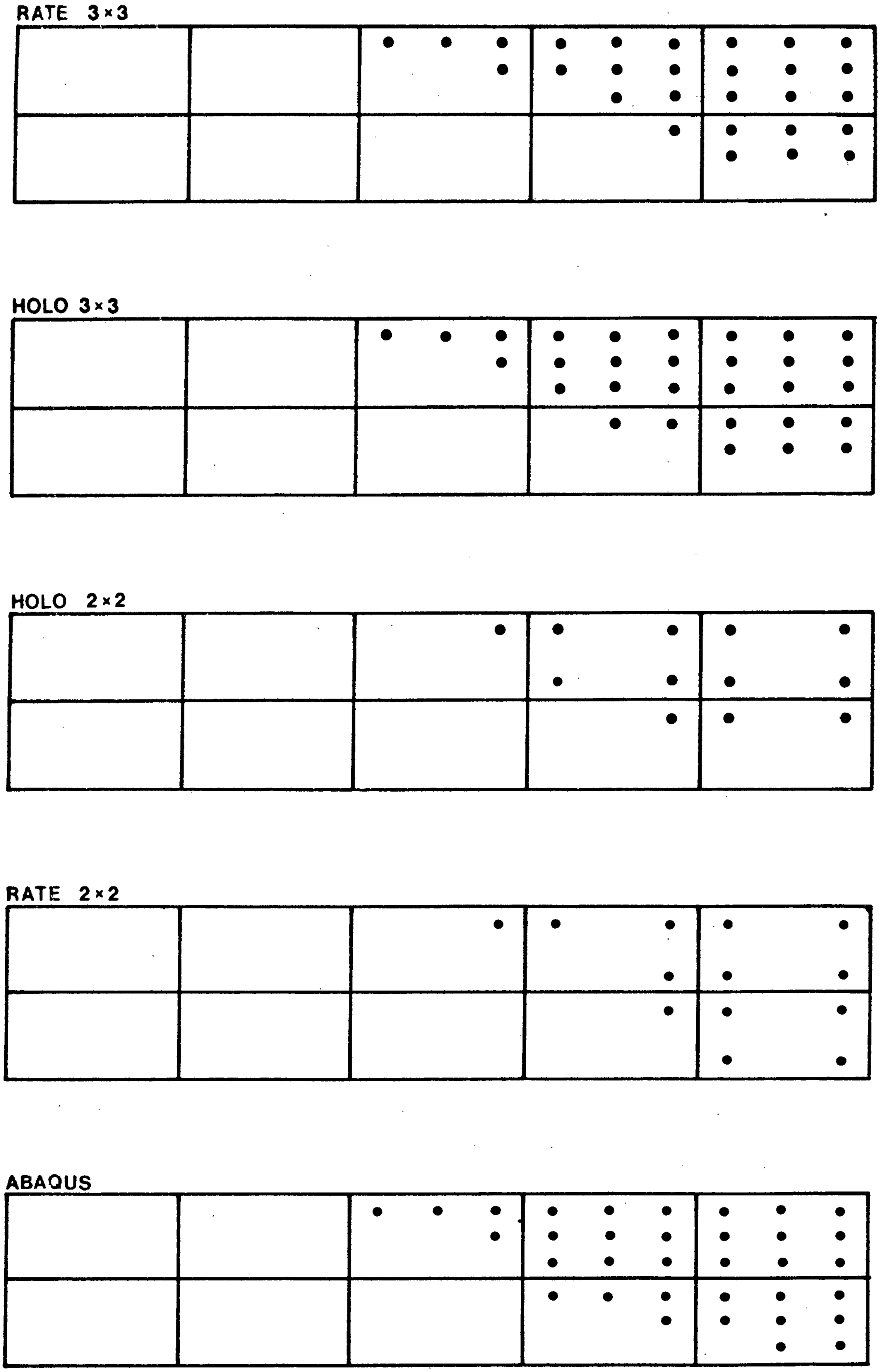


Figure 7.11 Extent of plastic deformation just prior to failure for the five numerical solutions.

TABLE 7.4

Average CPU statistics

	RATE-3	RATE-2	HOLO-3	HOLO-2	ABAQUS
CPU, total	351	236	660	316	175
CPU/increment	10.3	7.2	110.0	52.6	11.7
CPU/iteration	10.3	7.2	8.4	4.3	3.9

Note: CPU ≡ central processing unit : a measure of computer processing time.

We have already mentioned in the introduction to Chapter 6 that the Euler forward integration method is not the most efficient means of advancing the penalty-rate solution forward, and thus, we are not too concerned with the statistics given in Table 7.4 for the RATE solutions : these times can certainly be improved upon. However, what we are concerned with is the penalty-rate solution itself, particularly with regard to the effect which the number of plastic Gauss points in the model has on the time taken to complete the rate solution. (Recall that for each plastic Gauss point an additional equation must be added to the global $\underset{\sim}{K}^*$ matrix; see Section 6.2.) This effect is illustrated in Fig. 7.12. The time for zero plastic Gauss points is the time required to solve the elastic problem when the elastic stiffness matrix has already been triangularised. The line denoted "elastic" is the time required to complete a full elastic solution, and corresponds to just over six plastic Gauss points, or 8% of the total number of Gauss points in the model. Thus, for this example, with 8% of the model undergoing plastic deformation, the rate solution is completed in no more time than it takes to complete an elastic solution, which would be the time taken to complete a solution using the conventional tangent stiffness approach.

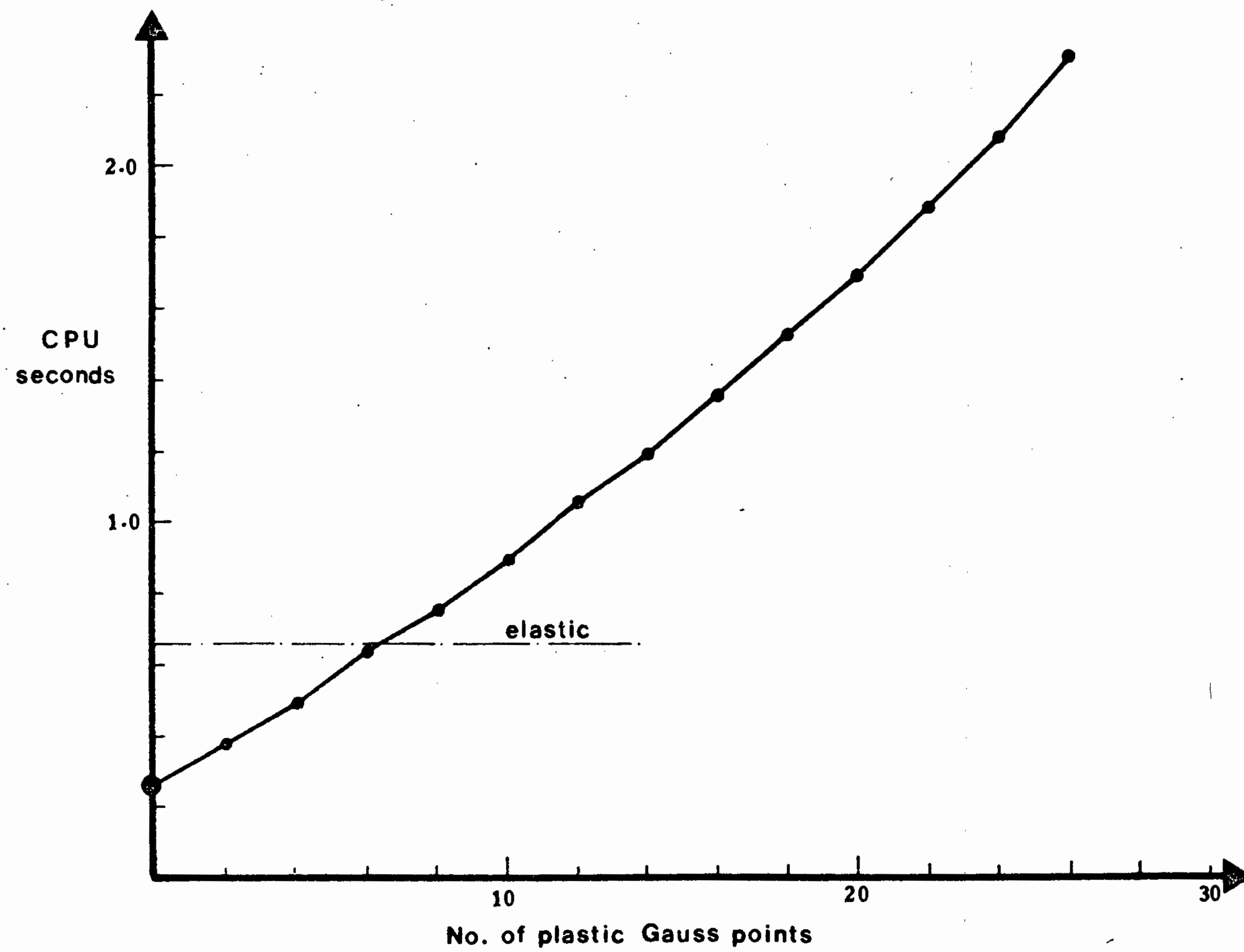


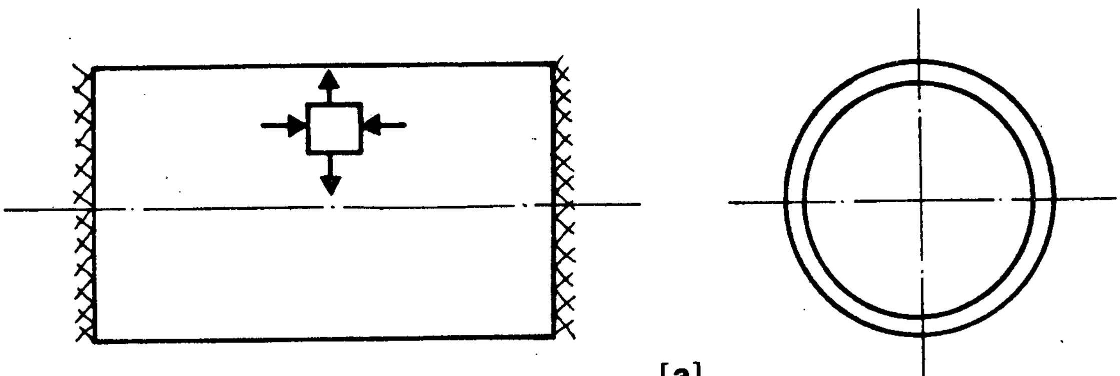
Figure 7.12 CPU time to complete the rate solution as a function of the number of plastic Gauss points in the model.

To summarise this example, we have shown that both the RATE and HOLO solutions agree closely with alternative solutions : as we would expect, the RATE-3 and HOLO-3 solutions agree with ABAQUS whereas the less stiff RATE-2 and HOLO-2 solutions agree more closely with the analytical solution. In particular we note that the HOLO solutions require many less increments to produce acceptable results than do the RATE and ABAQUS solutions, although the efficiency of the iterative procedure leaves something to be desired. Finally, we have shown that the penalty-rate solution is really only viable when the extent of plastic deformation in the model is relatively small.

7.6 CYCLIC LOADING AND UNLOADING

In recent years the problem of structures subjected to cyclic thermomechanical loading has become of increasing interest (PONTER and COCKS (1984a,b), COCKS and PONTER (1985)). In the nuclear industry, for example, components are often subjected to moderate dead weight loading together with large thermal fluctuations and in such situations it is important to be able to determine whether the component will shake down or whether the loads are such as might cause thermal ratchetting. Ratchetting is the phenomenon whereby an increment of plastic strain is produced during each load cycle; the component undergoes increasing deformation and if the number of load cycles increases indefinitely the component may fail via a process called incremental collapse. The ability to analyse ratchetting, whilst important in its own right, also provides an excellent benchmark for the assessment of the stability and accuracy of numerical algorithms under repeated cycles of loading.

We consider here a thin-walled tube which is restrained against axial deformation, subjected to a cyclicly varying temperature change ΔT , and internally pressurised to a hoop stress σ_p . The tube wall may be assumed to be in a state of uniform plane stress, so that it suffices to consider a small element of the wall which is then in a state of biaxial stress (Fig. 7.13(a)). We choose a single 8-node element to model the tube wall; the hoop stress σ_p is assumed to be uniformly applied and constant, and we replace the temperature change ΔT by an equivalent prescribed displacement change $\Delta u_x = \alpha L \Delta T$, where α is the coefficient of thermal expansion and L is the length of the element side parallel to the X axis (Fig. 7.13(b)).



[a]

$E = 1.0$

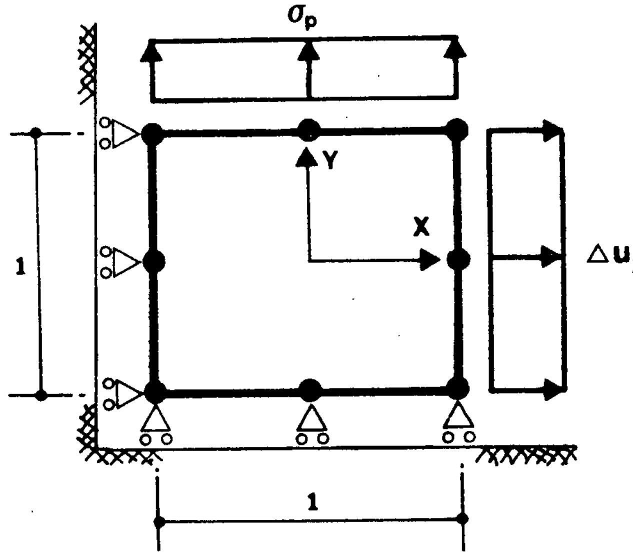
$E_p = 0.$ or 1.0

$\nu = 0.3$

$\sigma_o = 1.0$

thickness = 1.0

penalty parameter = 10^{-9}



[b]

Figure 7.13 (a) thin-walled tube subjected to internal pressure and temperature change (b) finite element model.

During the analysis the prescribed displacement Δu_x will be cycled as shown in Fig. 7.14(a). The interaction diagram showing the various regimes of behaviour which can be expected for any combination of σ_p and $\sigma_T = \Delta u_x/EL$ is given in Fig. 7.14(b). In order to demonstrate the ratchetting phenomenon we choose a point A which is just outside the shakedown regime, and which has the stress coordinates (0.4 , 1.996). Our objective now is to demonstrate the behaviour of our model for this combination of loads, and compare our solutions to both the analytical and ABAQUS solutions.

Consider first the elastic-perfectly plastic behaviour, as shown in Fig. 7.15. The inset shows the initial yield surface, which remains

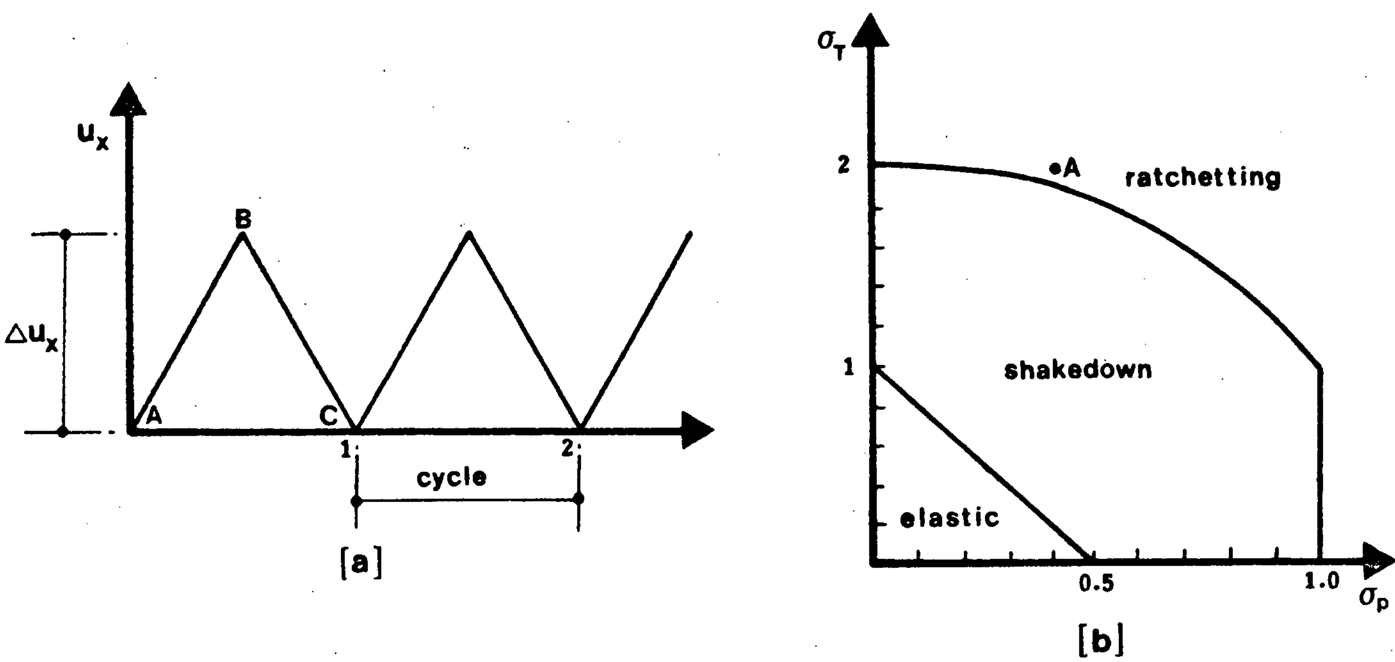


Figure 7.14 (a) prescribed displacement cycles (b) interaction diagram.

fixed, and the line AB at $\sigma_y = \sigma_p$ which the stress point traverses during the cycling; plastic straining occurs only at A or B. The plastic strain versus cycle solution shows a constant increase in plastic strain, $\Delta\varepsilon_y^p$, per cycle. The curve represents (to the scale of plotting) both the RATE and HOLO solutions, which are identical, and the ABAQUS solution, which actually lies just below the curve shown. The points marked A, B, C on the curve correspond to the points A, B, C in Fig. 7.14(a). Notice that the line joining the points corresponding to

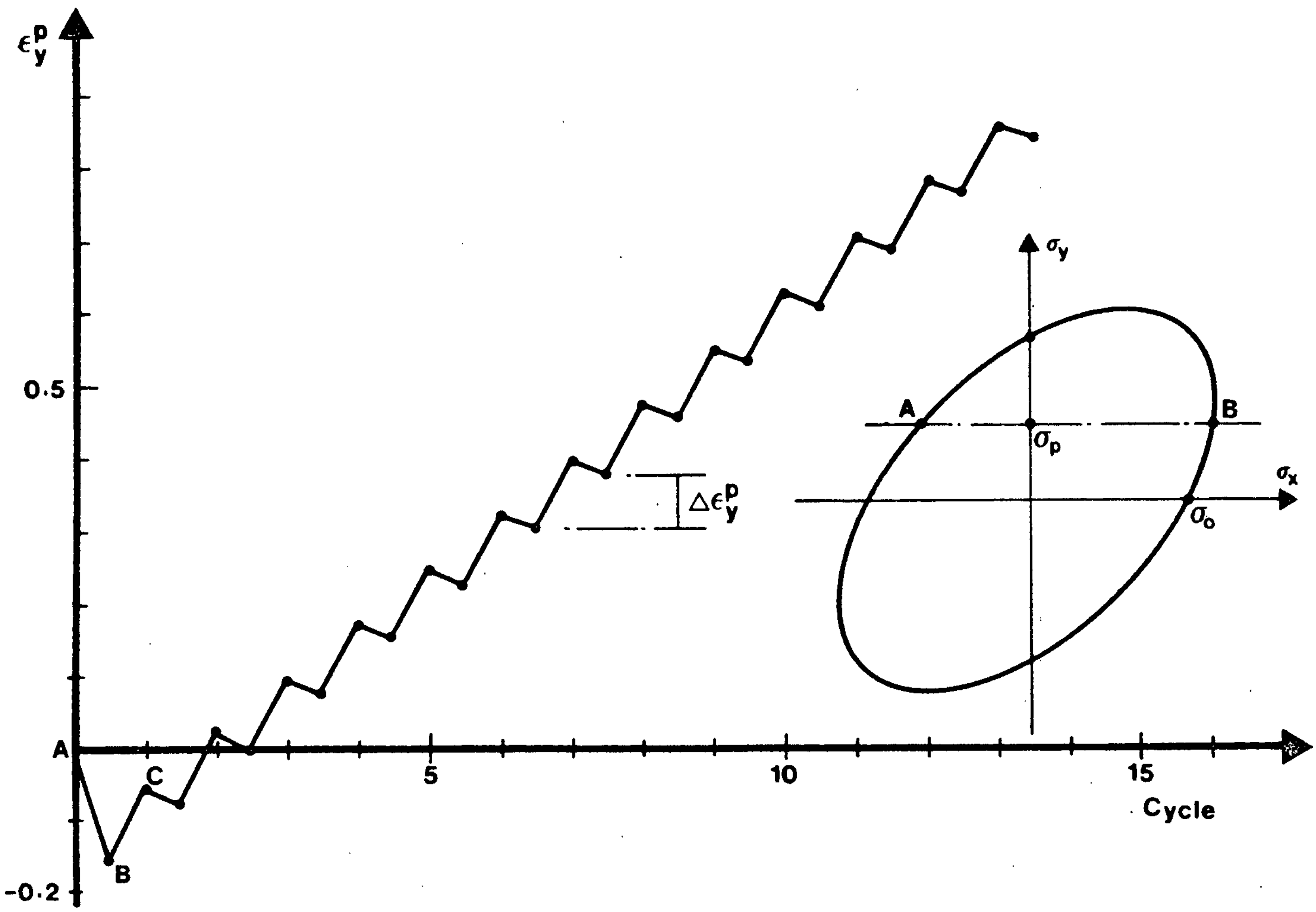


Figure 7.15 Ratchetting behaviour for the elastic-perfectly plastic case.

the beginning (or end) of each cycle is a straight line. Some statistics for the solutions are given in Table 7.5.

TABLE 7.5

Solution Statistics for the Elastic-Perfectly Plastic Case

Solution	Load increments per cycle	Iterations per increment	Plastic strain interpolation
RATE	4	1	3x3 (quadratic)
HOLO	2	10	2x2 (linear)
ABAQUS	6	1 or 2	-

The analytical solution for the increment in plastic strain per cycle, $\Delta\varepsilon_y^p$, is easy to calculate (see ZARKA and CASIER (1977), page 134) and is 0.076646. Both the RATE and HOLO solutions yield exactly this result, but the ABAQUS result, which varies very slightly from one cycle to the next, has a mean value of 0.075790.

The RATE and HOLO solutions for linear kinematic hardening (E_p = 1.0) are shown in Fig. 7.16, and the solution statistics remain as given in Table 7.5. The inset shows the translation of the yield surface during cycling, with the stress point again remaining on the line $\sigma_y = \sigma_p$; the elastic region, A_oB_o or AB , obviously changes as the yield surface translates, and the plastic straining may occur only at A_o, A, B_o or B . We have not shown the full ratchetting effect, because of the large number of cycles, but have merely drawn a curve through the solution points corresponding to the end of each cycle. The RATE and HOLO solutions differ quite significantly for the first few cycles, but

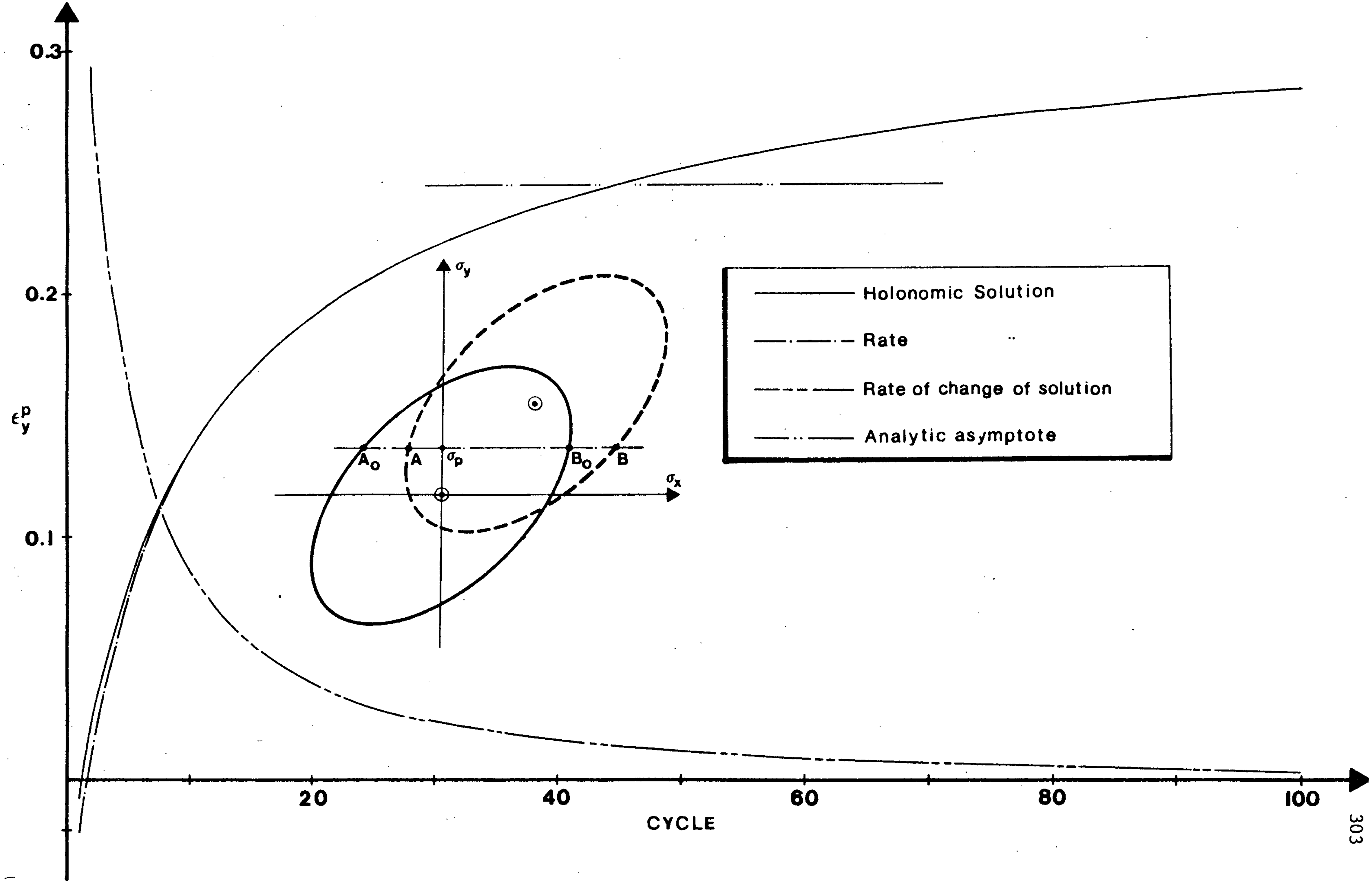


Figure 7.16 Ratchetting behaviour for the kinematic hardening case.

after 10 cycles this difference is negligible; we are unable to obtain an acceptable solution using ABAQUS.

Professor A R S Ponter has kindly supplied us with a simple first order analytic solution for the asymptotic plastic strain, which is also shown in Fig. 7.16 and has a value 0.2453. As can be seen, the numerical solutions exceed this asymptote quite significantly (the value at cycle 100 is 0.2847 , a difference of 16%). Nevertheless, the numerical solution does at least appear to be approaching an asymptote; this is more clearly seen from the "rate of change of solution" curve, which for all practical purposes may be considered zero at cycle 100 (the value is 0.0004). Thus, bearing in mind that the analytic solution depends on simplifying assumptions* which are not made in the numerical solution, and that the numerical solutions do appear to be approaching an asymptotic value, it is reasonable to conclude that our solutions are acceptable.

It is clear that both the RATE and HOLO formulations exhibit stable and accurate behaviour under repeated loading and unloading, which is certainly one of the most stringent tests to which a plasticity algorithm can be subjected. Moreover, the solutions which we have shown were obtained using the minimum number of increments consistent with the respective algorithms and with no special "tuning" of the computer

---

* For example, that the yield surface does not translate in the σ_y direction, but only along the σ_x axis; see the inset in Fig. 7.16 for the true behaviour.

programs : for example, it is not necessary to perform pilot analyses in order to determine a suitable penalty parameter or order of interpolation for the plastic strains. It is encouraging in this respect to note the success of the linear plastic strain interpolation in the HOLO solutions since this gives rise to very significant reductions in computational effort.

7.7 CONCLUSIONS REGARDING THE NUMERICAL EXPERIMENTS

As we have already said, it was never our intention to exhaustively test the RATE and HOLO programs here since they are not yet fully developed, especially with regard to efficiency. Nevertheless, we believe we have presented sufficient data on which to base our conclusions regarding their suitability for further development.

We summarise our conclusions as follows* :

1. Both methods have proved to be stable and robust in general loading/unloading conditions, and capable of excellent accuracy.
2. The methods are particularly fast in elastic unloading, where the elastic solution is obtained within a single iteration. This can be attributed to the internal variable formulation where plastic multipliers/strains are

---

* Our comments regarding the incremental penalty-rate formulation refer specifically to the penalty-rate solution and do not include the forward integration aspects of the incremental problem.

included amongst the solution variables, leaving the elastic stiffness matrix permanently unchanged. This approach differs from the conventional tangent stiffness method where the elastic stiffness matrix is continually modified to reflect the current state of plastic deformation.

3. Accurate solutions are obtained via the incremental holonomic formulation using many fewer load increments than are required using conventional methods. Unfortunately, this advantage is negated at present due to the large number of iterations required per increment, but we believe this situation can be improved upon. It is important to note in this respect that the number of iterations does <u>not</u> appear to be related to the size of the increment.

4. Under certain conditions the incremental holonomic solution converges to the continuous (rate) solution as the number of subdivisions of the loading path is increased.

5. The penalty-rate solution constitutes a viable alternative to the conventional tangent stiffness approach only when the extent of plastic deformation is limited to less than about 8% of the volume of the body.

6. Both methods are as simple to implement and use as any conventional method. From the user's point of view only one additional parameter, namely the penalty or regularisation parameter ε, is required; more important, in using the incremental holonomic method the user need

not be overly concerned with the size of load increments which he chooses since the solution is unconditionally stable in this respect.

It is clear that both the RATE and HOLO algorithms are stable and robust, and capable of providing acceptable solutions. They are also easy to implement and use. The question remains, however, as to whether the efficiency of the algorithms can be improved to the point where they are competitive with existing algorithms, for example, the tangent stiffness method. In the case of RATE we fear not, since the processing time for a solution depends critically on the extent of plastic deformation in the model. However, in the case of HOLO we believe that the efficiency can be dramatically improved : we note in this respect that the processing per iteration is already competitive with ABAQUS, and we regard this as encouraging enough to warrant further investigation of the HOLO algorithm.

CHAPTER 8

SUMMARY AND CONCLUSIONS

The behaviour of elastic-plastic materials is path-dependent in the sense that the history of behaviour must be taken into account when determining solutions. Thus, the constitutive equations for the material are written in rate form, leading directly to the rate boundary-value problems. Response rates obtained from solving this problem must then be integrated forward in order to advance the solution along the solution path corresponding to the given loading path. An alternative approach is to assume that the stress and strain paths corresponding to some finite interval along the loading path are extremal (in the sense of complementary work and work respectively). This allows the (forward) integration to be performed at the level of the constitutive equations, leading to constitutive equations which are written in terms of finite increments of stress and strain. These constitutive equations give us the incremental holonomic boundary-value problem. In the original formulation of this problem due to PONTER and MARTIN (1972) only the initial state (assumed to be the virgin state) and final state of the body were considered, thus giving a consistent formulation of the holonomic or deformation theory problem. We have extended their constitutive equations to a fully incremental form including initial stresses and plastic strains, and have subsequently defined an incremental holonomic boundary-value problem.

Having defined the two problems mentioned above, both of which include internal variables, we proceeded to show that both problems can be formulated as variational inequalities, from which minimisation problems follow in a natural way. We have proven existence and uniqueness of the solutions to the minimisation problems. Both problems involve certain difficulties which have been circumvented by the introduction of perturbed minimisation problems which depend on a positive parameter ε, referred to as a penalty or regularisation parameter. We show that the solutions to the perturbed problems converge to the original solutions as $\varepsilon \to 0$. In the case of the rate problem (or penalty-rate problem) we also discuss a saddle-point formulation. We use the perturbed problems as bases for finite element approximations and give estimates of the errors in these approximations with respect to both ε and the finite element mesh parameter h.

The finite element approximation of the perturbed minimisation problems gives rise to systems of algebraic equations relating discrete values of the displacements, plastic multipliers or strains, and loads, which are remarkably similar in structure. In the case of the penalty-rate problem these equations are linear but since the plastic multipliers are subject to inequality constraints the problem as a whole is nonlinear. Moreover, since these equations respresent only the rate problem itself, we use an Euler forward method to integrate the rate solution forward to obtain the incremental solution. In the case of the incremental holonomic problem the algebraic equations are nonlinear in the plastic strains and we use a Newton iterative method to solve them, obtaining the incremental solution directly.

We have presented numerical algorithms for the solution of the finite element approximations to both the incremental penalty-rate and holonomic problems, and these have been implemented in two computer programs RATE and HOLO. Since we regard these as pilot programs in their early stages of development we have only attempted to indicate the significant characteristics of each of the solution methods, and these have been summarised at the end of Chapter 7. In our view the incremental holonomic formulation has outstanding potential. Moreover, we believe that in the near future a true unification of the incremental holonomic and rate problems will emerge*. Assuming this to be so, and recognising that the incremental holonomic problem represents the fundamental approximation to the elastic-plastic problem, are we not looking at a future situation when the conventional tangent stiffness approaches are replaced by the potentially more efficient incremental holonomic formulations? We believe so, and we trust that the present work is a step in that direction.

---

* We recognise that this has already been shown by Professor Guilio Maier and his co-workers using quadratic programming methods within a finite-dimensional framework.

APPENDIX A

A REVIEW OF PLANE STRESS

Consider a body which occupies an open bounded domain Ω in R^2, with Lipschitz boundary $\Gamma = \Gamma_u \cup \Gamma_s$ (Fig. A.1). For convenience we let the body lie in the XY plane so that each material point on the body has a position vector $\underset{\sim}{x} = (x,y)$, where x and y are the cartesian coordinates of the point relative to global axes X and Y. Similarly, the displacement vector $\underset{\sim}{u}$ at any point has components u_x and u_y.

By definition the only non-zero components of the stress tensor $\underset{\sim}{\sigma}$ are σ_{xx}, σ_{yy} and $\sigma_{xy} = \sigma_{yx}$. Henceforth we redefine $\underset{\sim}{\sigma}$ to be the stress vector $\underset{\sim}{\sigma} = \{\sigma_{xx}, \sigma_{yy}, \sigma_{xy}\}^T$. Similarly, we refer to $\underset{\sim}{\varepsilon}$ as the total strain vector $\underset{\sim}{\varepsilon} = \{\varepsilon_{xx}, \varepsilon_{yy}, 2\varepsilon_{xy}\}^T$, with corresponding definitions for the elastic and plastic strain vectors $\underset{\sim}{e}$ and $\underset{\sim}{p}$ respectively.

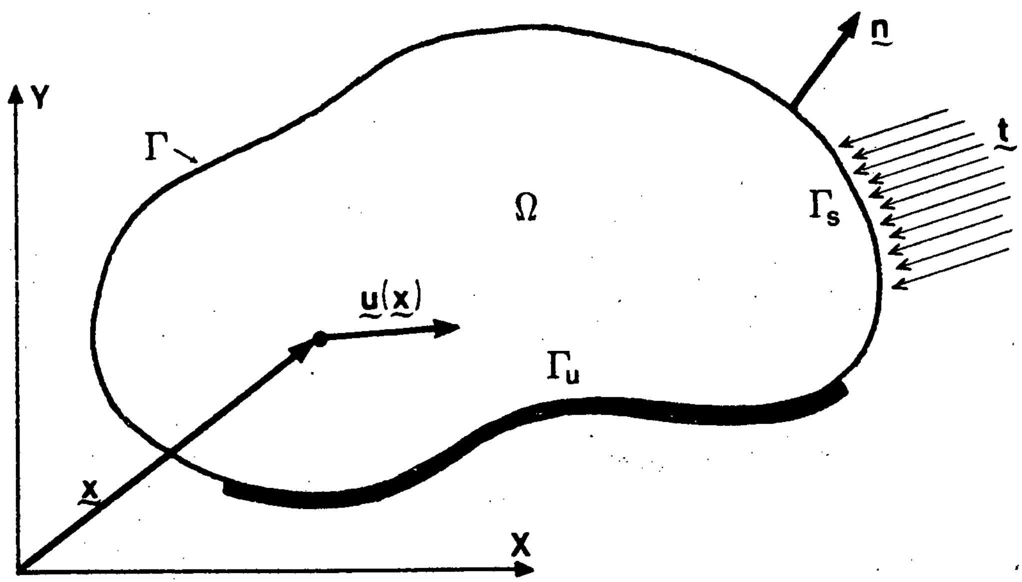


Figure A.1 Body Ω in a state of plane stress.

The elastic constitutive equations may be written in matrix form as

$$\underset{\sim}{\sigma} = \underset{\sim\sim}{Ce} \tag{A.1}$$

where the matrix $\underset{\sim}{C}$ is given by

$$\underset{\sim}{C} = \frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1-\nu}{2} \end{bmatrix} \tag{A.2}$$

and where E is Young's modulus and ν is Poisson's ratio.

The strains $\underset{\sim}{\varepsilon}$ are related to the displacements $\underset{\sim}{u}$ by

$$\underset{\sim}{\varepsilon} = \underset{\sim\sim}{Du} \tag{A.3}$$

where the operator $\underset{\sim}{D}$ is defined by

$$\underset{\sim}{D} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{bmatrix} . \tag{A.4}$$

We assume that on the portion Γ_s of the boundary a traction vector $\underset{\sim}{t}(s)$ is defined, $s \in \Gamma_s$, with components $\{t_x,t_y\}^T$; we assume that the displacements $\underset{\sim}{u}(s)$ on Γ_u , $s \in \Gamma_u$, are zero. Finally, a body force vector $\underset{\sim}{f} = \{f_x,f_y\}^T$ is assumed to be given on the domain Ω .

As far as plastic behaviour is concerned we again confine our attention to materials which exhibit linear kinematic hardening and obey the von Mises yield criterion, which is written in terms of total stresses as

$$\phi = \frac{1}{3}\,[(\sigma_{xx} - E_p p_{xx})^2 - (\sigma_{xx} - E_p p_{xx})(\sigma_{yy} - E_p p_{yy}) + (\sigma_{yy} - E_p p_{yy})^2 + 3(\sigma_{xy} - E_p p_{xy})^2 - 3k^2] = 0 \tag{A.5}$$

where $E_p > 0$ is the plastic modulus (Section 2.3, eqn 3.13), $k = \sigma_o/\sqrt{3}$, and σ_o is the uniaxial yield stress in tension. From (A.5) we define the effective stress $\bar{\sigma}$ as

$$\bar{\sigma} = [(\sigma_{xx} - E_p p_{xx})^2 - (\sigma_{xx} - E_p p_{xx})(\sigma_{yy} - E_p p_{yy}) + (\sigma_{yy} - E_p p_{yy})^2 + 3(\sigma_{xy} - E_p p_{xy})^2]^{1/2} \tag{A.6}$$

so that (A.5) may be written as

$$\phi = \frac{1}{3}\,[\bar{\sigma}^2 - \sigma_o^2] = 0 \quad . \tag{A.7}$$

APPENDIX B

COMPLEMENTARY WORK IN PLANE STRESS

The elastic complementary work density along a path in stress space between the virgin state and some state $\underset{\sim}{\sigma}$ is, for the case of plane stress,

$$\bar{\Omega}^e = \frac{1}{2E}\,[\sigma_{xx}^2 - 2\nu\sigma_{xx}\sigma_{yy} + \sigma_{yy}^2] + \left(\frac{1+\nu}{E}\right)\sigma_{xy}^2 \quad . \tag{B.1}$$

Let $\bar{\Omega}_A$ and $\bar{\Omega}_B$ be the elastic complementary work densities corresponding to the start and end respectively of a loading increment. The the change in elastic complementary work over the increment is given by

$$\Delta\bar{\Omega}^e = \bar{\Omega}_B^e - \bar{\Omega}_A^e \quad . \tag{B.2}$$

In general, initial plastic strains $\underset{\sim}{p}^\circ$ may be present at the start of a load increment and these will also contribute the following complementary work :

$$\Delta\bar{\Omega}^\circ = \underset{\sim}{p}^\circ \cdot (\underset{\sim}{\sigma}_B - \underset{\sim}{\sigma}_A)$$

$$= p_{xx}^\circ\Delta\sigma_{xx} + p_{yy}^\circ\Delta\sigma_{yy} + 2p_{xy}^\circ\Delta\sigma_{xy} \tag{B.3}$$

where $\Delta\underset{\sim}{\sigma} = (\underset{\sim}{\sigma}_B - \underset{\sim}{\sigma}_A)$.

If a change in plastic strain $\Delta\underset{\sim}{p}$ takes place during the load increment then the complementary plastic work due to this change is (recall Section 2.5, eqn (5.19)),

$$\Delta\bar{\Omega}^p = \frac{1}{2}\,\Delta \underset{\sim}{p} \cdot (\underset{\sim}{\sigma}_B - \underset{\sim}{\tilde{\sigma}})$$

$$= \frac{1}{2}\,[\Delta \underset{\sim}{p}_{xx}\Delta\hat{\sigma}_{xx} + \Delta p_{yy}\Delta\hat{\sigma}_{yy} + 2\Delta p_{xy}\Delta\hat{\sigma}_{xy}] \tag{B.4}$$

where $\Delta\hat{\underset{\sim}{\sigma}} = (\underset{\sim}{\sigma}_B - \hat{\underset{\sim}{\sigma}})$. Thus, the total change in complementary work density during a given load increment is

$$\Delta\bar{\Omega} = \Delta\bar{\Omega}^e + \Delta\bar{\Omega}^o + \Delta\bar{\Omega}^p \quad . \tag{B.5}$$

The above computations are performed at the element level and then integrated over the volume of each element to give the complementary work

$$\Delta\Omega_e = \int_V \Delta\bar{\Omega} \quad dV \tag{B.6}$$

The total complementary work for the body is then the sum of the element contributions $\Delta\Omega_e$.

REFERENCES

ANZELLOTTI, G. 1983 On the existence of the rates of stress and displacement for Prandtl-Reuss plasticity. Q. appl. Math. 40, 181.

BASUSKA, I. 1973 The finite element method with Lagrange multipliers. Numer. Math. 20, 179.

BATHE, K-J. 1982 Finite Element Procedures in Engineering Analysis. New Jersey : Prentice-Hall.

BREZZI, F. 1974 On the existence, uniqueness and approximations of saddle-point problems arising from lagrangian multipliers. RAIRO, R.2, 129.

BECKER, E.B., CAREY, G.F., & ODEN, J.T. 1981 Finite Elements : An Introduction. Volume I. New Jersey : Prentice-Hall.

CARTER, P., & MARTIN, J.B. 1977 A note on the derivation of the kinematic rate theorem of plasticity from the free energy minimum principle. J. de Mecanique 16, 791.

CARTER, P., & MARTIN, J.B. 1979 The derivation of the static rate theorems plasticity from an internal variable model. J. Mech. Phys. Solids 26, 345.

CAPURSO, M. 1969 Prinicpi di minimo per la soluzione incrementale di problemi elasto-plastici. Accad. Lincei-Rend. Sc. Fis. Mat. Nat. 46, 417.

CERADINI, G. 1966 A maximum principle for the analysis of elastic-plastic systems. Meccanica 1, No. 3/4, 77.

CIARLET, P. 1978 The Finite Element Method for Elliptic Problems. Amsterdam : North-holland.

COCKS, A.C.F., & PONTER, A.R.S. 1985 The plastic behaviour of components subjected to constant primary stresses and cyclic secondary strains. J. of Strain Analysis 20, No. 1, 7.

COHN, M.Z., & MAIER, G. 1979 Engineering Plasticity by Mathematical Programming. Proceedings of the NATO Advanced Studies Institute, University of Waterloo, August 1977. Permagon Press.

COHN, M.Z., & FRANCHI, A.	1984	Structural plasticity computer system : STRUPL. Proc. ASCE, J. Struc. Div. 105, ST4.
COLONNETTI, G.	1955	L'equilibre des Corps Deformables. Paris : Dunod.
CHOU, S.I., & WANG, C.C.	1979	Error estimates of finite element approximations for problems in linear elasticity. Part 1. Problems in Elastoplastics. Archs ration. Mech. Analysis 72, 41.
DITTMER, C.T.	1978	A Programming Approach to the Numerical Analysis of Elasto-Plastic Continua. Ph.D. Thesis, University of Cape Town.
DE DONATO, O.	1968	Extensions to continua of some minimum theorems in elastoplastic theory. Meccanica 3, No. 4, 1.
DITTMER, C.T., GRIFFIN, T.B., & MARTIN, J.B.	1985	A programming approach to the numerical analysis of elastoplastic continua. Eng. Frac. Mech. 21, 817.
DRUCKER, D.C.	1951	A More Fundamental Approach to Plastic Stress-Strain Relations. Proc. 1st U.S. Nat. Congr. Appl. Mech., 487.
DUVAUT, G., & LIONS, J.L.	1972	Les Inequations en Mecanique et en Physique. Paris : Dunod.
DUVAUT, G., & LIONS, J.L.	1976	Inequalities in Mechanics and Physics. Berlin : Springer. [This is the English Translation of DUVAUT and LIONS (1972).]
EKELAND, I., & TEMAN, R.	1976	Convex Analysis and Variational Problems. Amsterdam : North Holland.
FINZI, L.	1955	Principio della minima energia elastica differenziale. Rc. Accad. naz. XL, serie VIII, XVIII-3.
FRANCHI, A., & GENNA, F.	1984	Minimum principles and initial stress method in elastic-plastic analysis. Eng. Struct. 6, 65.
FUCIK, S., & KUFNER, A.	1980	Nonlinear Differential Equations. Elsevier.
GLOWINSKY, R., LIONS, J.L., & TREMOLIERES, R.T.	1981	Numerical Analysis of Variational Inequalities. Amsterdam : North-Holland.
GREENBERG, H.J.	1949a	On the variational principles of plasticity. Tech. Rept. A11-54, Div. Appl. Math., Brown University, Providence, RI.

GREENBERG, H.J. 1949b Complementary minimum principles for an elastic-plastic material. Q. appl. Math. 7, 85.

HAAR, A., & VON KARMAN, T. 1909 Zur theorie der Spannungzustände in plastischen und sandartigen Medren. Nachr. der Wiss. zu Gottingen, Math.-phys. Klasse, 204.

HAVNER, K.S., & PATEL, H.P. 1976 On convergence of the finite element method for a class of elastic-plastic solids. Q. appl. Math. 34, 59.

HENCKY, H. 1924 Zur theorie plastischer Deformationen und der hierdurch im Material herrorgerufenen Nachspannungen. ZAMM 4, 323.

HILL, R. 1950 The Mathematical Theory of Plasticity. Oxford : Clarendon Press.

HLAVACEK, I. 1980 A finite element solution for plasticity with strain-hardening. RAIRO Analyse Num. 14, 347.

HLAVACEK, I., & NECAS, J. 1970a On inequalities of Korn's type : I Boundary-value problems of elliptic systems of partial differential equations. Archs ration. Mech. Analysis 36, 305.

HLAVACEK, I., & NECAS, J. 1970b On inequalities of Korn's type : II Applications to linear elasticity. Archs ration. Mech. Analysis 36, 312.

HODGE, P.G. 1966 A deformation bounding theorem for flow law plasticity. Q. appl. Math. 24, 171.

HODGE, P.G. 1968 Numerical applications of minimum principles in plasticity. Engineering Plasticity (Heyman, J. and Leckie, F.A., Eds). Cambridge University Press, 237.

JIANG, L.S. 1984 On an elastic-plastic problem. J. diff. Eqns. 51, 97.

JOHNSON, C. 1976a Existence theorems for plasticity problems. J. Math. Pures·Appl. 55, 431.

JOHNSON, C. 1976b On finite element methods for plasticity problems. Num. Math. 26, 79.

JOHNSON, C. 1977 A mixed finite element method for plasticity problems with hardening. SIAM J. Numer. Anal. 14, 575.

JOHNSON, C. 1978 On plasticity with hardening. J. Math. Analysis Applics. 62, 325.

KIKUCHI, N.	1981	Convergence of a penalty-finite element approximation for an obstacle problem. Numer. Math. 37, 105.
KOITER, W.T.	1953	Stress-strain relations, uniqueness and variational theorems for elastic-plastic materials with a singular yield surface. Q. appl. Math. 11, 350.
MAIER, G.	1968	A quadratic programming approach for certain classes of nonlinear structural problems. Meccanica 3, No. 2, 121.
MAIER, G.	1969a	Complementary plastic work theorems in piecewise-linear elastoplasticity. Int. J. Solids & Structures 5, 261.
MAIER, G.	1969b	Some theorems for plastic strain rates and plastic strains. J. de Mecanique 8, 5.
MAIER, G.	1970	A matrix structural theory of piecewise-linear elastoplasticity with interacting yield planes. Meccanica 5, No. 1, 54.
MAIER, G., & MUNRO, J.	1982	Mathematical programming applications to engineering plastic analysis. Appl. Mech. Reviews 35, No. 12, 1631.
MAIER, G., & NAPPI, A.	1984	On a unified framework provided by mathematical programming in plasticity. Mechanics of Material Behaviour (Dvorak, G.J., and Shield, R.T., Eds). Elsevier.
MARCAL, P.V., & KING, I.P.	1967	Elastic-plastic analysis of two-dimensional stress systems by the finite element method. Int. J. Mech. Sci. 9, 143.
MARQUES, J.M.M.C., & OWEN, D.R.J.	1984	Some reflections on elastoplastic stress calculation in finite element analysis. Comp. & Struct. 18, 1135.
MARSDEN, J.E., & HUGHES, T.J.R.	1983	Mathematical Foundations of Elasticity. New Jersey : Prentice-Hall.
MARTIN, J.B.	1966a	Extended displacement bound theorems for work hardening continua subjected to dynamic loading. Int. J. Solids & Structures 2, 9.
MARTIN, J.B.	1966b	The determination of upper bounds on displacement resulting from static and dynamic loading by the application of energy methods. Proc. 5th U.S. Nat. Congr. Appl. Mech., ASME (N.Y.), 221.
MARTIN, J.B.	1970	A complementary energy bounding theorem for time-independent materials. Developments in Theoretical and Applied Mechanics. 8 (D. Frederick and E.H. Harris, Eds). Permagon Press.

MARTIN, J.B., & PONTER, A.R.S. 1972 On dual energy theorems for a class of elastic-plastic problems due to G. Maier. J. Mech. Phys. Solids 20, 301.

MARTIN, J.B. 1975a Plasticity : Fundamentals and General Results. Boston : MIT Press.

MARTIN, J.B. 1975b On the kinematic minimum principle in classical plasticity. J. Mech. Phys. Solids 23, 123.

MARTIN, J.B. 1975c A note on the implications of thermodynamic stability in the internal variable theory of inelastic solids. Int. J. Solids & Structures 11, 247.

MARTIN, J.B., & REDDY, B.D. 1977 A programming approach to the solution of the rate problem in elastic, plastic solids. Mechanics in Engineering (Dubey, R.N., & Lind, N.C., Eds). University of Waterloo Press.

MARTIN, J.B., REDDY, B.D., GRIFFIN, T.B., & BIRD, W.W. 1984 Applications of mathematical programming concepts to incremental elastic-plastic analysis. Tech. Rept. 52, UCT/CSIR Applied Mechanics Research Unit, University of Cape Town.

MARTIN, J.B. 1986 A complementary work bounding principle for forward integration along the path of loading for elastoplastic bodies. Tech. Rept 63, UCT/CSIR Applied Mechanics Research Unit, University of Cape Town.

MATTHIES, H., STRANG, G., & CHRISTIANSEN, E. 1979 The saddle-point of a differential program. Energy Methods in Finite Element Analysis (Glowinski, R., Rodin, P., & Zienkiewicz, O.C., Eds). New York : Wiley.

NADAI, A. 1931 Plasticity. New York : McGraw-Hill.

ODEN, J.T., & KIKUCHI, N 1980 Theory of variational inequalities with applications to problems of flow through porous media. Int. J. Engng Sci. 18, 1173.

ODEN, J.T., & WHITEMAN, J.R. 1982 Analysis of some finite element methods for a class of problems in elasto-plasticity. Int. J. Engng Sci. 20, 977.

ODEN, J.T., & CAREY, G.F. 1983 Finite Elements : Mathematical Aspects. Volume IV. New Jersey : Prentice-Hall.

ODEN, J.T., & REDDY, J.N. 1976 Variational Methods in Theoretical Mechanics. Berlin : Springer.

ORTIZ, M., & POPOV, E.P. 1984 Accuracy and stability of integration algorithms for elastoplastic constitutive relations. Technical Report, Division of Engineering, Brown University, Providence, RI.

OWEN, D.R.J., & HINTON, E.	1980	Finite Elements in Plasticity : Theory and Practice. Swansea : Pineridge Press.
PONTER, A.R.S., & MARTIN, J.B.	1972	Some extremal properties and energy theorems for inelastic materials and their relationship to the deformation theory of plasticity. J. Mech. Phys. Solids 20, 281.
PONTER, A.R.S.	1968	Convexity conditions and energy theorems for time-independent materials. J. Mech. Phys. Solids 16, 283.
PONTER, A.R.S., & COCKS, A.C.F.	1984a	The incremental strain growth of an elastic-plastic body loaded in excess of the shakedown limit. ASME J. appl. Mech. 51, 465.
PONTER, A.R.S., & COCKS, A.C.F.	1984b	The incremental strain growth of elastic-plastic bodies subjected to high levels of cyclic thermal loading. ASME. J appl Mech. 51, 470.
PRAGER, W.	1942	Fundamental theorems of a new mathematical theory of plasticity. Duke Math. J. 9, 228.
PRAGER, W.	1946	Variational principles in the theory of plasticity. Proc. 6th Int. Congress. App. Mech., Paris.
PRAGER, W., & HODGE, P.G.	1951	Theory of Perfectly Plastic Solids. New York : Dover.
REDDY, B.D., & MITCHELL, G.P.	1983	The analysis of elastic-plastic plates : A quadratic programming problem and its solution by finite elements. Comp. Meth. appl. Mech. Engng 41, 237.
REDDY, B.D., GRIFFIN, T.B., & MARAIS, M.J.	1985	A penalty approach to the rate problem in small-strain plasticity. IMA J. appl. Maths 34, 303.
REDDY, B.D.	1986	Functional Analysis and Boundary-Value Problems. London : Pitman.
SAMUELSSON, A., & FROIER, M.	1978	Finite elements in plasticity - a variational inequality approach. MAFELAP III (J.R. Whiteman, Ed). Academic Press.
SIMO, J.C., & TAYLOR, R.L.	1985	Consistent tangent operators for rate-independent elastoplasticity. Comp. Meth. appl. Mech. Engng 48, 101.
SOECHTING, J.F., & LANCE, R.H.	1969	A bounding principle in the theory of work hardening plasticity. J. Appl. Mech. 36, 228.
SUQUET, P.M.	1978a	Existence et regularite des solutions des equations de la plasticite. CR Acad. Sci. Paris 286, 1129 and 1201.

SUQUET, P.M.	1978b	Existence and regularity of solutions for plasticity problems. Variational Methods in the Mechanics of Solids (S. Nemat-Nasser, Ed), Northwestern University, Permagon Press.
SUQUET, P.M.	1981	Evolution problems for a class of dissipative materials. Q. appl. Math. 38, 391.
TEMAM, R., & STRANG, G.	1980	Functions of bounded deformation. Archs ration. Mech. Analysis 75, 7.
TING, T.W.	1966	Elastic-plastic torsion of a square bar. Trans. Am. math. Soc. 123, 369.
VIANBERG, M.M.	1964	Variational Methods for the Study of Nonlinear Operators. Holden-Day.
ZARKA, J., & CASIER, J.	1977	Elastic-plastic response of a structure to cyclic loading : practical rules. Mechanics Today (S. Nemat-Nasser, Ed) 6, 93. Permagon Press.
ZIENKIEWICZ, O.C.	1977	The Finite Element Method. 3rd Edition. Maidenhead : McGraw-Hill.