

Finite-element applications to the nonlinear mechanics of solids

To cite this article: D Peric and D R J Owen 1998 *Rep. Prog. Phys.* **61** 1495

View the [article online](#) for updates and enhancements.

You may also like

- [Giant fluctuations in the flow of fluidised soft glassy materials: an elasto-plastic modelling approach](#)
Magali Le Goff, Eric Bertin and Kirsten Martens
- [Extension of the visco-plastic self-consistent model to account for elasto-visco-plastic behavior using a perturbed visco-plastic approach](#)
Youngung Jeong and Carlos N Tomé
- [A machine learning based prediction of elasto-plastic response of a short fiber reinforced polymer \(SFRP\) composite](#)
Subrat Kumar Maharana, Ganesh Soni and Mira Mitra

Finite-element applications to the nonlinear mechanics of solids

D Perić and D R J Owen

Department of Civil Engineering, University of Wales, Swansea SA2 8PP, UK

Received 18 August 1997

Abstract

The paper discusses some of the relevant computational advances which permit the simulation of large-scale problems involving nonlinear solids within realistic time frames and computational resources. The need for rigorous consideration of both theoretical and algorithmic issues is emphasized, particularly in relation to the computational treatment of finite-strain elasto-plastic (viscoplastic) deformation, the modelling of frictional contact conditions and element technology capable of dealing with material incompressibility. Practically important aspects such as adaptive mesh refinement procedures are discussed and attention is given to choice of appropriate error estimators for elasto-plastic materials and the transfer of solution parameters between successive meshes. The role of explicit solution techniques in the simulation of large-scale nonlinear problems is also discussed. The concept of discrete elements is briefly described and their applications to a wide range of solid mechanics problems illustrated. Some advances in the field of iterative equation solution methods are reviewed and their potential advantages in the simulation of large-scale nonlinear solid mechanics problems are demonstrated.

Contents

| | Page |
|---|------|
| 1. Introduction | 1497 |
| 2. Continuum constitutive modelling | 1499 |
| 2.1. Basics of geometry and kinematics | 1500 |
| 2.2. General principles of continuum thermodynamics | 1501 |
| 2.3. Constitutive principles | 1503 |
| 2.4. Thermodynamics with internal variables | 1505 |
| 3. Implicit finite-element solution strategy | 1507 |
| 3.1. Numerical integration algorithm. The incremental constitutive law | 1508 |
| 3.2. The incremental boundary value problem | 1509 |
| 3.3. Finite-element discretization | 1509 |
| 3.4. The Newton–Raphson scheme. Linearization | 1510 |
| 4. Finite-strain elasto-plasticity | 1511 |
| 4.1. Multiplicative decomposition | 1511 |
| 4.2. General hyperelastic-based elasto-plastic constitutive model | 1513 |
| 4.3. General stress integration procedure. The exponential map algorithm | 1516 |
| 4.4. The spatial tangent modulus | 1518 |
| 5. Contact-friction modelling | 1520 |
| 5.1. Mechanical description of frictional contact | 1521 |
| 5.2. Numerical integration of the frictional contact model | 1523 |
| 5.3. Consistent linearization of the discrete frictional contact problem | 1523 |
| 6. Element technology | 1527 |
| 6.1. The low-order elements for finite-strain problems | 1528 |
| 6.2. Some practical experiences | 1530 |
| 7. Further aspects of the deformation of inelastic solids at finite strains | 1533 |
| 7.1. Yield surface representation | 1534 |
| 7.2. Elasto-plastic damaging solids | 1536 |
| 7.3. Elasto-viscoplastic solids | 1537 |
| 8. Thermomechanical coupling | 1538 |
| 9. Adaptive strategies for nonlinear problems | 1544 |
| 9.1. Basic error estimates and adaptive refinement | 1545 |
| 9.2. Mesh generation/adaption techniques | 1547 |
| 9.3. Transfer operators | 1548 |
| 10. Explicit solution strategies | 1557 |
| 10.1. The discretized dynamic equations | 1559 |
| 10.2. Element methodology | 1560 |
| 11. A combined finite-/discrete-element method | 1561 |
| 12. Advanced equation solution strategies for solid mechanics | 1566 |
| 12.1. CGS and Bi-CGStab methods | 1567 |
| 12.2. Galerkin multigrid method | 1567 |
| 13. Concluding remarks | 1569 |
| Acknowledgments | 1571 |
| References | 1571 |

1. Introduction

Computational procedures based on the finite-element method (FEM) are now well established in many branches of engineering and science, for linear and nonlinear applications. The increasing acceptance of such approaches within both research and industrial environments is due to improved awareness, enhanced maturity of computational models and associated algorithms and, more importantly, dramatic increases in computational power/cost ratios.

Although the roots of the finite-element process can be traced back to the early 1940s and beyond (Courant 1943, Argyris 1954, 1955), the first publication that closely resembles its present form appeared in 1956 (Turner *et al* 1956) and the word '*finite element*' was coined in 1960 (Clough 1960). Early research was exclusively related to structural problems, with application to other fields, for example heat and fluid flow, only emerging in the late 1960s and 70s. Procedures for the treatment of a wide range of problems are now well established and can be found in standard texts such as Bathe (1996), Crisfield (1991, 1997), Hughes (1987), Oden (1972) and Zienkiewicz and Taylor (1989, 1991).

In contrast to the more traditional discretization methods, such as the finite-difference method (FDM), which are based on a 'strong' formulation whereby direct approximation of the governing differential equations is performed, the FEM is based on a weak, or variational (integral), formulation of the boundary or initial value problem. Consequently, the solution may be obtained by summing up the integral contributions over parts of the domain (finite elements) with much weaker conditions required in terms of the regularity of the prescribed data and solution than would be required in the FDM. The solution domain can be decomposed into a sum of contributions coming from an arbitrary number of finite elements with, significantly, the basic variables of the problem described in terms of simple polynomial approximations over a local finite-element subdomain that may have arbitrary geometry.

The principal advantages of the FEM, that have made it the most popular method for computer simulation of solid mechanics problems, may be summarized as follows: (i) arbitrary geometries may be easily accommodated within general unstructured meshes, (ii) the FEM is amenable to efficient computer implementations since individual contributions may be evaluated at the element level and summed up on an element-by-element basis, (iii) conservation properties of the governing physical laws (such as momentum balance and energy balance) are intrinsically built into the finite-element formulation, (iv) a solid mathematical foundation, although requiring complex functional analysis formalism and still under development (particularly for nonlinear problems), provides estimates in terms of the solution error and contributes significantly to the robustness and reliability of the FEM.

Over the last decade, there has been considerable advancement in several areas of finite-element research, including the treatment of incompressible and high-speed fluid flows, the modelling of granular and progressively fracturing solids by discrete-element methods and the behaviour of solids undergoing large-strain elasto-plastic deformations. This publication concentrates on some principal developments that have recently taken place in the application of FEM to problems in the field of nonlinear solid mechanics and highlights issues which must be successfully resolved for the solution of complex large-scale industrial problems. In particular, attention is focused on the computational treatment of

solids subjected to elasto-(visco)plastic deformations involving finite strains, together with ancillary topics such as frictional contact modelling, mesh adaptivity strategies, discrete-element procedures and iterative solution schemes for large-scale equation systems.

Probably the most striking example of developments in the field of nonlinear solid mechanics is the strides recently made in the numerical solution of finite-strain plasticity problems. The formulation of rigorous solution procedures has been the subject of intense debate over the last decade and only recently has some consensus been reached on an appropriate constitutive theory based on tensorial state variables to provide a theoretical framework for the macroscopic description of a general elasto-plastic material at finite strains. In computational circles, effort has been directed at the formulation of algorithms for integration of the constitutive equations relying on operator split methodology. The concept of consistent linearization has been introduced to provide quadratically convergent solution procedures. By employing logarithmic stretches as strain measures a particularly simple model for large inelastic deformations at finite strains is recovered. In particular, the effects of finite strains appear only at the kinematic level and the integration algorithms and corresponding consistent tangent operators for small-strain situations can be directly employed.

A further class of nonlinear problems for which considerable advances in numerical modelling have been made in recent years is that of contact-friction behaviour. Contact-friction phenomena arise in many important areas, for instance metal forming, and numerical treatment in the past has, of necessity, relied on temperamental and poorly convergent algorithms. This situation has changed markedly with recognition of the complete analogy that exists between contact-friction behaviour and the classical theory of elasto-plasticity. Hence, the operator split algorithms and consistent linearization procedures developed for the latter case translate directly to contact-friction models to provide robust and rapidly convergent numerical solutions.

A general feature encountered in the finite-element simulation of finite-strain plasticity problems is that the optimal mesh configuration changes continually throughout the deformation process requiring mesh derefinement as well as mesh refinement during use of any adaptive remeshing process. Considerable benefits may accrue by implementation of such strategies in terms of robustness and efficiency, realizing that the requirements of computational efficiency are ever increasing. At the same time, error estimation procedures will play a crucial role in quality assurance by providing reliable finite-element solutions.

A current area of crucial debate in the computational modelling of large-scale nonlinear problems in solid mechanics is the relative merits of explicit and implicit solution strategies. Issues of particular concern are the accuracy of explicit approaches in relation to implicit solutions and the relative computational efficiency of both approaches.

The ever increasing need to solve large-scale industrial problems by an implicit approach demands that advances be made in equation solution strategies. As equation systems extend beyond a certain size (of the order of 20 000 d.o.f.) direct solvers become increasingly inefficient and iterative solvers offer the most natural approach to solution. For problems with unsymmetric stiffness matrices, as caused by the presence of frictional contact, the conjugate gradient squared (CGS) and bi-stabilized conjugate gradient (Bi-CGStab) methods provide efficient solution. Considerable promise is also offered by multigrid methods, in which the problem is first solved on a coarse mesh to pre-condition the equation system for subsequent iterative solution on a finer mesh.

The remainder of this paper discusses the above issues in more detail and provides numerical examples illustrating the recent advances made in the finite-element analysis of nonlinear industrial problems. Although the text is based mostly on the authors' own research, and inevitably emphasizes the authors' point of view, we will attempt to offer a

balanced review of the main advances in the field of nonlinear solid mechanics that have taken place within the last decade or so.

2. Continuum constitutive modelling

A concept of a continuum itself is a mathematical idealization of the real material which is applicable when the fine structure of the matter can be ignored in relation to the geometric scale of the physical problem under consideration. A typical example of such an idealization is presented in the constitutive modelling of metals used in engineering practice which are on a scale of approximately $1\ \mu\text{m}$ composed of crystals with a uniform packing of atoms in a crystalline lattice. Since crystals are, in typical engineering materials, oriented in a more or less random manner, on a scale $1\text{--}100\ \mu\text{m}$, metals are considered as polycrystalline materials. On an even larger scale, the so-called macroscale, it will not be possible to distinguish any fine structure, and the material may be considered as a continuous medium. However, the observation (through experiments) of the phenomena at macroscale may not be sufficient for successful constitutive modelling of the material—the knowledge of underlying physical mechanisms at the microscale often provides the essential information required in description of the material behaviour.

Mathematical analysis of a continuous medium that includes mechanical and thermal phenomena is known as continuum thermomechanics. It has traditionally been divided into two areas: (i) general principles that on the one side include geometry and kinematics of continuum, and on the other balance laws, which govern conservation of mass, entropy, momentum and energy, (ii) constitutive description which is guided by a set of principles such as frame invariance, material symmetries, etc. General principles are applicable to the continuum irrespective of the material under consideration, i.e. the underlying constitutive equations. It is a fundamental requirement that numerical procedures based on the FEM or other techniques also preserve the properties of the continuum stated by the general principles and constitutive equations. It is worth pointing out that this has often proved to be a non-trivial task.

In this section some basic concepts of the thermomechanics of continuous media are briefly reviewed. The material presented here is standard and well established in the continuum mechanics literature (Truesdell and Noll 1965, Fung 1965, 1977, Gurtin 1981, Lubliner 1990). Nevertheless, its inclusion at this point is convenient for later discussion. First the general principles of continuum mechanics, such as conservation of mass, balance of linear and angular momentum, and the first and second principles of thermodynamics, are stated in their differential forms. In addition, some general principles that guide the constitutive description of materials, such as locality hypothesis, frame invariance and material symmetries, are briefly stated. Finally, a general framework for the phenomenological description of dissipative materials is presented, which is based on thermodynamics with internal variables. By emphasizing the link between micromechanical processes and their mathematical representation within the framework of continuum thermodynamics with internal variables, the purpose of this section is to establish fundamental principles in a clear logical sequence that guides the development of continuum constitutive models for general dissipative materials.

At this point, it is worth while emphasizing that in the numerical modelling of complex engineering phenomena, intuition is of little help in detecting false answers (and a physical check is often too late). Thus, it is important to adopt a high standard of mathematical rigour from the outset, in contrast to some approaches in theoretical physics where a pragmatic view of such rigour is allowable.

2.1. Basics of geometry and kinematics

2.1.1. *Deformation gradient.* Consider a generic continuum body \mathcal{B} which occupies a region of the three-dimensional (3D) Euclidean space \mathbb{R}^3 in its reference configuration, with boundary $\partial\mathcal{B}$. Let \mathcal{B} be subjected to a *motion* φ so that for each time t , the *deformation*

$$\varphi(\cdot, t) : \mathcal{B} \rightarrow \mathbb{R}^3 \quad (2.1)$$

maps each material particle \mathbf{X} of \mathcal{B} into the position \mathbf{x} it occupies at time t . The set

$$\mathbf{x} = \varphi(\mathbf{X}, t) \quad (2.2)$$

is called the *current* or *deformed* configuration. The two-point tensor \mathbf{F} defined by†

$$\mathbf{F}(\mathbf{X}, t) = \mathbf{D}\varphi(\mathbf{X}, t) = \frac{\partial \varphi_i}{\partial X_I} \mathbf{e}_i \otimes \mathbf{E}_I \quad (2.3)$$

is termed the *deformation gradient*, where $\{\mathbf{E}_I\}_{I=1,2,3}$ and $\{\mathbf{e}_i\}_{i=1,2,3}$ are fixed orthonormal bases in the reference and deformed configuration, respectively, typically chosen to be coincident with the standard basis in \mathbb{R}^3 . The Jacobian of the mapping (2.1) can be represented as

$$J = \det \mathbf{F}. \quad (2.4)$$

2.1.2. *Strain tensors.* By imposing the standard condition on the deformation gradient $\det \mathbf{F} > 0$ ensuring that (2.1) is a one-to-one mapping, the polar decomposition theorem admits the unique representation of the deformation gradient (2.3) in the form

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R} \quad (2.5)$$

where \mathbf{U} , \mathbf{V} are positive definite symmetric tensors and \mathbf{R} is an orthogonal tensor. Tensors \mathbf{U} and \mathbf{V} are known as the *right* and *left stretch tensors*, respectively, while \mathbf{R} is the *rotation tensor*. Tensors $\mathbf{U}(\mathbf{X}, t)$ and $\mathbf{V}(\mathbf{X}, t)$ measure local stretching near \mathbf{X} , while $\mathbf{R}(\mathbf{X}, t)$ measures the local rigid rotation of points close to \mathbf{X} (we refer to chapter 3 of Gurtin (1981) for detailed, yet introductory, analysis of the geometry and kinematics of a continuum).

The fact that \mathbf{U} and \mathbf{V} are associated with the stretching of material points suggests that these tensors should be used when measuring the straining of the continuum. The choice of a particular strain tensor has been, however, based on convenience and has resulted in a variety of strain measures that is nowadays used in the continuum mechanics literature (Ogden 1984). For later reference we introduce the *right* and *left Cauchy–Green tensors*, \mathbf{C} and \mathbf{B} , defined by

$$\mathbf{C} = \mathbf{F}^T \mathbf{F} = \mathbf{U}^2 \quad \mathbf{B} = \mathbf{F} \mathbf{F}^T = \mathbf{V}^2 \quad (2.6)$$

and the *logarithmic strain tensor* defined by

$$\boldsymbol{\varepsilon} = \ln \mathbf{V}. \quad (2.7)$$

It should be emphasized that, in order to preserve the invariance of the stress work per unit mass, a particular choice of strain measure necessitates usage of the associated dual stress measure within the constitutive relations.

When linearized about the reference state all strain measures give the standard small strain tensor

$$\boldsymbol{\epsilon} = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T) \quad (2.8)$$

where \mathbf{u} is the displacement vector given by

$$\mathbf{u}(\mathbf{X}, t) = \varphi(\mathbf{X}, t) - \mathbf{X}. \quad (2.9)$$

† Unless stated otherwise, the summation over twice-repeated indices is implied in the text.

2.2. General principles of continuum thermodynamics

In order to state the laws of thermodynamics, it is convenient to introduce the scalar fields $\theta(\mathbf{x}, t)$, $e(\mathbf{x}, t)$, $s(\mathbf{x}, t)$ and $r(\mathbf{x}, t)$ defined over the deformed configuration $\varphi(\mathcal{B}, t)$ of the body \mathcal{B} which represent, respectively, the *temperature*, *specific internal energy*, *specific entropy* and the *density of heat production*. In addition, the tensor field $\boldsymbol{\sigma}(\mathbf{x}, t)$ will denote the *Cauchy stress* and the vector fields $\mathbf{b}(\mathbf{x}, t)$ and $\mathbf{q}(\mathbf{x}, t)$ will denote, respectively, the *body force* and *heat flux*.

2.2.1. Conservation of mass. Given a part \mathcal{P} of the body \mathcal{B} , the postulate of conservation of mass requires that

$$\int_{\mathcal{P}} \rho_0(\mathbf{X}) \, dv_0 = \int_{\varphi(\mathcal{P})} \rho(\mathbf{x}, t) \, dv \quad (2.10)$$

for any part \mathcal{P} and time t , where $\rho_0 = \rho(\mathbf{X}, 0)$, v_0 and ρ, v denote the mass density and element volume in the reference and deformed configurations, respectively. By the change of variables on the right-hand side from \mathbf{x} to \mathbf{X} it follows that

$$\int_{\mathcal{P}} \rho_0(\mathbf{X}) \, dv_0 = \int_{\mathcal{P}} \rho(\varphi(\mathbf{X}, t)) J(\mathbf{X}, t) \, dv_0 \quad (2.11)$$

which must be valid for any subdomain \mathcal{P} of the body \mathcal{B} . Thus the local form of the conservation of mass may be represented as

$$\rho(\mathbf{x}, t) J(\mathbf{X}, t) = \rho_0(\mathbf{X}) \quad (2.12)$$

provided that $\mathbf{x} = \varphi(\mathbf{X})$.

If the postulate of the conservation of mass is expressed as

$$\frac{d}{dt} \int_{\varphi(\mathcal{P})} \rho(\mathbf{x}, t) \, dv = 0 \quad (2.13)$$

then, realizing that $\dot{J} = \text{div} \, \mathbf{v} J$ and after some simple transformations, the conservation of mass in the local form can be identified with the well known equations of continuity

$$\dot{\rho} + \rho \, \text{div}[\mathbf{v}] = 0 \quad (2.14)$$

where \mathbf{v} is the spatial *velocity* and $\text{div}[\cdot]$ denotes the *spatial divergence* of $[\cdot]$.

2.2.2. Momentum balance. The (linear) momentum balance, considered to be a continuum analogue to Newton's second law, states that for any region \mathcal{P} of the body \mathcal{B} and for any time t

$$\frac{d}{dt} \int_{\varphi(\mathcal{P})} \rho \mathbf{v} \, dv = \int_{\partial\varphi(\mathcal{P})} \mathbf{t}(\mathbf{x}, t, \mathbf{n}) \, da + \int_{\varphi(\mathcal{P})} \rho \mathbf{b} \, dv. \quad (2.15)$$

where $\mathbf{t}(\mathbf{x}, t, \mathbf{n})$ is the surface traction vector which gives the force per unit area at the position \mathbf{x} and time t over the surface element with a normal \mathbf{n} . By invoking Cauchy's theorem, the linear relationship between the traction vector $\mathbf{t}(\mathbf{x}, t, \mathbf{n})$ and the normal \mathbf{n} can be proved, which leads to the existence of the Cauchy stress tensor $\boldsymbol{\sigma}(\mathbf{x}, t)$ such that (for a formal proof of Cauchy's theorem we refer to Marsden and Hughes (1983) or Ciarlet (1988))

$$\mathbf{t}(\mathbf{x}, t, \mathbf{n}) = \boldsymbol{\sigma} \mathbf{n}. \quad (2.16)$$

By substituting expression (2.16) into (2.15), and by employing the divergence theorem it follows that

$$\frac{d}{dt} \int_{\varphi_t(\mathcal{P})} \rho \mathbf{v} \, dv = \int_{\varphi_t(\mathcal{P})} (\operatorname{div}[\boldsymbol{\sigma}] + \rho \mathbf{b}) \, dv. \quad (2.17)$$

By employing the change of variables theorem and invoking the conservation of mass (2.14), the left-hand side may be written as

$$\frac{d}{dt} \int_{\varphi_t(\mathcal{P})} \rho \mathbf{v} \, dv = \int_{\varphi_t(\mathcal{P})} \rho \dot{\mathbf{v}} \, dv \quad (2.18)$$

which results in the expression for the momentum balance in the form

$$\int_{\varphi_t(\mathcal{P})} (\operatorname{div}[\boldsymbol{\sigma}] + \rho \mathbf{b} - \rho \dot{\mathbf{v}}) \, dv = 0. \quad (2.19)$$

Since this expression must be valid for any subdomain \mathcal{P} of the body \mathcal{B} , including differential regions, the local form of the linear momentum balance may be recovered as

$$\operatorname{div}[\boldsymbol{\sigma}] + \rho \mathbf{b} = \rho \dot{\mathbf{v}}. \quad (2.20)$$

where $\dot{\mathbf{v}}$ stands for the *acceleration* field.

In a similar fashion the angular momentum balance

$$\frac{d}{dt} \int_{\varphi_t(\mathcal{P})} \mathbf{r} \times \rho \mathbf{v} \, dv = \int_{\partial\varphi_t(\mathcal{P})} \mathbf{r} \times \mathbf{t}(\mathbf{x}, t, \mathbf{n}) \, da + \int_{\varphi_t(\mathcal{P})} \mathbf{r} \times \rho \mathbf{b} \, dv \quad (2.21)$$

where \mathbf{r} denotes the vector from the origin to the point \mathbf{x} , leads to the local equation

$$\boldsymbol{\sigma}^T = \boldsymbol{\sigma} \quad (2.22)$$

which expresses the symmetry of the stress tensor. This equation is restricted to *non-polar* media, i.e. for continuum media in which stress couples are assumed absent.

2.2.3. The first principle. The first principle of thermodynamics is based on the postulate that the balance of energy holds for each subdomain \mathcal{P} of the body \mathcal{B} and any time t in the form

$$\frac{d}{dt} \int_{\varphi_t(\mathcal{P})} \rho (e + \frac{1}{2} \mathbf{v} \cdot \mathbf{v}) \, dv = \int_{\varphi_t(\mathcal{P})} \rho (\mathbf{b} \cdot \mathbf{v} + r) \, dv + \int_{\partial\varphi_t(\mathcal{P})} \rho (\mathbf{t} \cdot \mathbf{v} + h) \, dv. \quad (2.23)$$

This statement ensures that the change in internal energy e and kinetic energy of the body equals the work expended by the body force \mathbf{b} and surface force \mathbf{t} acting on the same portion of the body plus the rate of change of heat energy which comprises the heat production r and the flow through the boundary h .

After some manipulations, similar to those provided in section 2.2.2, the local form of the energy balance may be explicitly expressed by the equation

$$\rho \dot{e} = \boldsymbol{\sigma} : \mathbf{D} + \rho r - \operatorname{div}[\mathbf{q}] \quad (2.24)$$

where

$$\mathbf{D} = \frac{1}{2} (\nabla \mathbf{v} + \nabla \mathbf{v}^T) \quad (2.25)$$

is the *rate of deformation* or *stretching* tensor, with $\nabla(\cdot)$ denoting the spatial gradient of (\cdot) .

2.2.4. *The second principle.* Starting from the global form of the entropy production inequality

$$\frac{d}{dt} \int_{\varphi_t(\mathcal{P})} \rho s \, dv \geq \int_{\varphi_t(\mathcal{P})} \frac{\rho r}{\theta} \, dv + \int_{\partial \varphi_t(\mathcal{P})} \frac{h}{\theta} \, da \quad (2.26)$$

and following similar arguments exploited in sections 2.2.2 and 2.2.3 the local form of the entropy production inequality, also known as the second principle of thermodynamics, may be expressed as

$$\rho \dot{s} + \operatorname{div} \left[\frac{\mathbf{q}}{\theta} \right] - \frac{\rho r}{\theta} \geq 0. \quad (2.27)$$

The above inequality expresses the irreversibility of entropy production.

2.2.5. *The Clausius–Duhem inequality.* By combination of the first and second principles stated above, one easily obtains the inequality

$$\rho \dot{s} + \operatorname{div} \left[\frac{\mathbf{q}}{\theta} \right] - \frac{1}{\theta} (\rho \dot{e} - \boldsymbol{\sigma} : \mathbf{D} + \operatorname{div}[\mathbf{q}]) \geq 0. \quad (2.28)$$

The introduction of the *specific free energy* ψ (also known as the *Helmholtz free energy per unit mass*) defined by

$$\psi := e - \theta s \quad (2.29)$$

along with the identity

$$\operatorname{div} \left[\frac{\mathbf{q}}{\theta} \right] = \frac{1}{\theta} \operatorname{div}[\mathbf{q}] - \frac{1}{\theta^2} \mathbf{q} \cdot \nabla \theta \quad (2.30)$$

into equation (2.7) results in the *Clausius–Duhem inequality*:

$$\boldsymbol{\sigma} : \mathbf{D} - \rho(\dot{\psi} + s\dot{\theta}) - \frac{1}{\theta} \mathbf{q} \cdot \mathbf{g} \geq 0 \quad (2.31)$$

where $\mathbf{g} := \nabla \theta$ is the temperature gradient.

2.3. Constitutive principles

The balance principles presented so far are *a priori* valid for any continuum body. In order to distinguish between different types of material, a constitutive model must be introduced. This section presents three principles which form the basis for the development of a rather general class of constitutive models of continua. In the present context, the guidelines for constitutive modelling laid down by these principles should be followed regardless of the particular kind of material to be modelled. However, in order to give predictive capability to the constitutive relations an experimental program is required. Such a program should be designed to quantitatively describe the constitutive model which can then be used in modelling of generic processes with sufficient accuracy.

Before proceeding further, it is convenient to introduce the notion of a dynamic process (see Truesdell 1969) defined by the set

$$\{\boldsymbol{\sigma}(\mathbf{x}, t), e(\mathbf{x}, t), s(\mathbf{x}, t), r(\mathbf{x}, t), \mathbf{b}(\mathbf{x}, t), \mathbf{q}(\mathbf{x}, t)\} \quad (2.32)$$

of fields over \mathcal{B} such that the balance of momentum and the first and second principles of thermodynamics are satisfied.

2.3.1. Principle of locality. For a class of *simple materials*, which includes the majority of common engineering materials, the *local history* of \mathbf{F} , θ and \mathbf{g} suffices to determine the history of the dynamic process for constitutive purposes. In that case, regarding the body force \mathbf{b} and heat supply r as delivered, respectively, by the linear momentum balance (2.15) and conservation of energy (2.23), and with the specific free energy defined by (2.29), the principle of locality implies the existence of functions \mathfrak{F} , \mathfrak{G} , \mathfrak{h} and \mathfrak{J} such that, for a point \mathbf{X} ,

$$\boldsymbol{\sigma}(t) = \mathfrak{F}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) \quad (2.33a)$$

$$\psi(t) = \mathfrak{G}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) \quad (2.33b)$$

$$s(t) = \mathfrak{h}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) \quad (2.33c)$$

$$\mathbf{q}(t) = \mathfrak{J}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) \quad (2.33d)$$

and the Clausius–Duhem inequality (2.31) holds for every dynamic process of \mathcal{B} . A dependency on \mathbf{X} is understood on both sides of (2.33) and $(\cdot)^t$ on the right-hand sides denotes the *history* of (\cdot) at \mathbf{X} up to the present time t .

2.3.2. Material objectivity. Another important principle of the constitutive theory is the *principle of material objectivity* also known as the *material frame indifference*. It states that ‘the material response is independent of the observer’. The motion φ^* is related to the motion φ by a change in observer if

$$\varphi^*(\mathbf{X}, t) = \mathbf{y}(t) + \mathbf{Q}(t)\varphi(\mathbf{X}, t) \quad (2.34)$$

where $\mathbf{y}(t)$ is a point and $\mathbf{Q}(t)$ an orthogonal tensor. This relation corresponds to a rigid relative movement between the different observers and the deformation gradient corresponding to φ^* is given by

$$\mathbf{F}^* = \mathbf{Q}\mathbf{F}. \quad (2.35)$$

Scalar fields (such as θ , ψ and s) are unaffected by a change in observer but the Cauchy stress $\boldsymbol{\sigma}(t)$, heat flux $\mathbf{q}(t)$ and the temperature gradient $\mathbf{g}(t)$ transform according to the rules

$$\boldsymbol{\sigma} \rightarrow \boldsymbol{\sigma}^* = \mathbf{Q}\boldsymbol{\sigma}\mathbf{Q}^T \quad (2.36a)$$

$$\mathbf{q} \rightarrow \mathbf{q}^* = \mathbf{Q}\mathbf{q} \quad (2.36b)$$

$$\mathbf{g} \rightarrow \mathbf{g}^* = \mathbf{Q}\mathbf{g}. \quad (2.36c)$$

The principle of material objectivity places restrictions on the constitutive functionals (2.33). Formally, it requires that \mathfrak{F} , \mathfrak{G} , \mathfrak{h} and \mathfrak{J} satisfy

$$\boldsymbol{\sigma}^*(t) = \mathfrak{F}(\mathbf{F}^{t*}, \theta^t, \mathbf{g}^{t*}) \quad (2.37a)$$

$$\psi(t) = \mathfrak{G}(\mathbf{F}^{t*}, \theta^t, \mathbf{g}^{t*}) \quad (2.37b)$$

$$s(t) = \mathfrak{h}(\mathbf{F}^{t*}, \theta^t, \mathbf{g}^{t*}) \quad (2.37c)$$

$$\mathbf{q}^*(t) = \mathfrak{J}(\mathbf{F}^{t*}, \theta^t, \mathbf{g}^{t*}) \quad (2.37d)$$

for any transformation of the form (2.35), (2.36).

2.3.3. Material symmetry. The *symmetry group* of a material is the set of density preserving changes of reference configuration under which the material response functionals \mathfrak{F} , \mathfrak{G} , \mathfrak{h} and \mathfrak{J} are not affected. The symmetry group of a solid material is a subset of the orthogonal

group \mathcal{O} . A subgroup \mathcal{S} of \mathcal{O} is said to be the symmetry group of the material defined by the constitutive functionals \mathfrak{F} , \mathfrak{G} , \mathfrak{h} and \mathfrak{J} if the relations

$$\mathfrak{F}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) = \mathfrak{F}([\mathbf{F}\mathbf{Q}]^t, \theta^t, \mathbf{g}^t) \quad (2.38a)$$

$$\mathfrak{G}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) = \mathfrak{G}([\mathbf{F}\mathbf{Q}]^t, \theta^t, \mathbf{g}^t) \quad (2.38b)$$

$$\mathfrak{h}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) = \mathfrak{h}([\mathbf{F}\mathbf{Q}]^t, \theta^t, \mathbf{g}^t) \quad (2.38c)$$

$$\mathfrak{J}(\mathbf{F}^t, \theta^t, \mathbf{g}^t) = \mathfrak{J}([\mathbf{F}\mathbf{Q}]^t, \theta^t, \mathbf{g}^t) \quad (2.38d)$$

hold for any *time-independent* $\mathbf{Q} \in \mathcal{S}$. A solid is said to be *isotropic* if its symmetry group is the entire orthogonal group \mathcal{O} . In the development of any constitutive model, the constitutive functionals must comply with the restrictions imposed by the symmetries of the material in question.

2.4. Thermodynamics with internal variables

The constitutive equations (2.33) written in terms of functionals of the history of \mathbf{F} , θ and \mathbf{g} , in that format, are far too general to have practical applicability for modelling real materials undergoing a real thermomechanical process. This is especially true if one has in mind the experimental identification of the constitutive functions and the solution of the corresponding boundary value problems. Therefore, it is imperative that simplifying assumptions are added to the general forms of the constitutive relations stated above.

An effective alternative to the general description based on history functionals is the adoption of the so-called *thermodynamics with internal variables*. The starting point of thermodynamics with internal variables is the hypothesis that at any instant of a thermomechanical process, the thermodynamic state (defined by σ , ψ , s and \mathbf{q}) at a given point \mathbf{X} can be completely determined by the knowledge of a finite number of *state variables*. The thermodynamic state depends only on the instantaneous value of the state variables and not on their past history. This hypothesis is intimately connected with the assumption of existence of a (fictitious) state of thermodynamic equilibrium known as the *local accompanying state* (Kestin and Bataille 1977) described by the current value of the state variables. In other words, every process is considered to be a succession of equilibrium states[†].

From the mathematical point of view, the state variables can be seen as parametrizing the history of thermomechanical processes and replacing the complex constitutive description in terms of history functionals by an approximation involving a finite number of parameters. For the applications with which we are mostly concerned, it will be convenient to assume that at a certain time t , the thermodynamic state at a point is determined by the set

$$\{\mathbf{F}, \theta, \mathbf{g}, \boldsymbol{\alpha}\} \quad (2.39)$$

of state variables where \mathbf{F} , θ and \mathbf{g} are the *instantaneous* values of deformation gradient, temperature and the temperature gradient and $\boldsymbol{\alpha}$ is a set:

$$\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_k\} \quad (2.40)$$

of k *internal variables* associated with dissipative mechanisms. Each element $\alpha_i \in \boldsymbol{\alpha}$ may be, in general, an entity of scalar, vectorial or tensorial nature.

Following the above hypothesis, the specific free energy is assumed to have the form

$$\psi = \psi(\mathbf{F}, \theta, \boldsymbol{\alpha}) \quad (2.41)$$

[†] Despite the success of the internal variable approach in numerous fields of continuum physics, phenomena induced by very fast external actions (at timescales comparable to atomic vibrations) which involve states far from thermodynamic equilibrium are excluded from representation by internal variable theories.

so that its rate of change is given by

$$\dot{\psi} = \frac{\partial \psi}{\partial \mathbf{F}} : \dot{\mathbf{F}} + \frac{\partial \psi}{\partial \theta} \dot{\theta} + \frac{\partial \psi}{\partial \boldsymbol{\alpha}} \dot{\boldsymbol{\alpha}}. \quad (2.42)$$

In the last term on the right-hand side of the expression above, the following convention has been adopted:

$$\frac{\partial \psi}{\partial \boldsymbol{\alpha}} \dot{\boldsymbol{\alpha}} = \sum_{i=1}^k \frac{\partial \psi}{\partial \alpha_i} \dot{\alpha}_i \quad (2.43)$$

with the appropriate product implied. By introducing the connection

$$\boldsymbol{\sigma} : \mathbf{D} = \frac{\rho}{\rho_0} \mathbf{P} : \dot{\mathbf{F}} \quad (2.44)$$

where $\mathbf{P} := \det[\mathbf{F}] \boldsymbol{\sigma} \mathbf{F}^{-T}$ is the *first Piola–Kirchhoff* stress tensor and ρ_0 is the density in the reference configuration, one obtains for the Clausius–Duhem inequality

$$\left(\mathbf{P} - \rho_0 \frac{\partial \psi}{\partial \mathbf{F}} \right) : \dot{\mathbf{F}} - \rho_0 \left(s + \frac{\partial \psi}{\partial \theta} \right) \dot{\theta} - \rho_0 \frac{\partial \psi}{\partial \boldsymbol{\alpha}} \dot{\boldsymbol{\alpha}} - \frac{\rho_0}{\rho \theta} \mathbf{q} \cdot \mathbf{g} \geq 0. \quad (2.45)$$

Since this inequality must hold for any thermomechanical process, a standard argument leads to the well known expressions

$$\mathbf{P} = \rho_0 \frac{\partial \psi}{\partial \mathbf{F}} \quad s = - \frac{\partial \psi}{\partial \theta} \quad (2.46)$$

for the first Piola–Kirchhoff stress, \mathbf{P} , and entropy, s .

Then, by defining

$$A_i := \rho_0 \frac{\partial \psi}{\partial \alpha_i} \quad (2.47)$$

as the *thermodynamical force* conjugate to each internal variable $\alpha_i \in \boldsymbol{\alpha}$, the Clausius–Duhem inequality can be rewritten as

$$-A_i \dot{\alpha}_i - \frac{\rho_0}{\rho \theta} \mathbf{q} \cdot \mathbf{g} \geq 0. \quad (2.48)$$

For convenience, we shall define by \mathbf{A} the set

$$\mathbf{A} = \{A_1, A_2, \dots, A_k\} \quad (2.49)$$

of thermodynamical forces.

In order to completely characterize a constitutive model, complementary laws associated with the dissipative mechanisms are required. Namely, equations for the flux variables $\frac{1}{\theta} \mathbf{q}$ and $\dot{\boldsymbol{\alpha}}$ must be derived. Recalling the principle of thermodynamic compatible determinism, the Clausius–Duhem inequality, now expressed by (2.48), must hold and that will evidently place restrictions on the possible constitutive relations. An effective way of ensuring that (2.48) is satisfied is given by postulating the existence of a scalar valued *dissipation* (pseudo)potential of the form

$$\Psi = \Psi(\mathbf{A}, \mathbf{g}) \quad (2.50)$$

possibly having the state variables as parameters, which is assumed to be convex with respect to each A_i and \mathbf{g} and zero valued at the origin $\{\mathbf{A}, \mathbf{g}\} = \{\mathbf{0}, \mathbf{0}\}$. In addition, the hypothesis of *normal dissipativity* is introduced, i.e. the flux variables are assumed to be determined by the laws

$$\dot{\alpha}_i = - \frac{\partial \Psi}{\partial A_i} \quad \frac{1}{\theta} \mathbf{q} = - \frac{\partial \Psi}{\partial \mathbf{g}}. \quad (2.51)$$

It should be noted that the constitutive description by means of convex potentials as described above is *not* a consequence of thermodynamics but, rather, a tool for formulating constitutive equations without violating thermodynamics. Indeed, it is obvious that a constitutive model defined by (2.41), (2.46) and (2.51) satisfies *a priori* the dissipation inequality. Some examples of constitutive models supported by experimental evidence which do not admit representation by means of dissipation potentials are discussed by Onat and Leckie (1988).

2.4.1. The phenomenological approach. Undoubtedly, the success of a constitutive model intended to describe the behaviour of a particular material lies in the choice of an appropriate set on internal variables. Since no plausible model will be general enough to describe the response of a material under all conditions, the definition of the internal variables must be guided not only by the specific material in question but, rather, by the combined consideration of the material *and* the range of processes under which it will be analysed. In general, due to the difficulties involved in the identification of the underlying dissipative mechanisms, the choice of the appropriate set of internal variables is somewhat subtle and will obviously be biased by the preference of the investigator.

Basically, constitutive modelling by means of internal variables relies either on a micromechanical or on a phenomenological approach. The micromechanical approach involves the determination of mechanisms and related variables at the atomic, molecular or crystalline levels. In general, these variables are discrete quantities and their continuum (macroscopic) counterparts are determined by means of homogenization techniques. The phenomenological approach, on the other hand, bypasses the need for measurements of microscopic quantities. It is based on the study of the response of the *representative volume element*, i.e. the element of matter large enough to be regarded as a continuum. The internal variables in this case will be directly associated with the dissipative behaviour observed at the *macroscopic* level in terms of continuum quantities (such as strain, stress, temperature, etc). Despite the macroscopic nature of theories derived on the basis of the phenomenological methodology, it should be expected that ‘good’ phenomenological internal variables will be somehow related to the underlying microscopic dissipation mechanisms.

The phenomenological approach to irreversible thermodynamics has been particularly successful in the field of solid mechanics. Numerous well established models of solids, such as classical elasto-plasticity (Hill 1950), have been developed on a purely phenomenological basis providing evidence of how powerful such an approach to irreversible thermodynamics can be when the major concern is the description of the essentially macroscopic behaviour. Direct application of phenomenological thermodynamics with internal variables will be discussed in section 4 where the formulation of a continuum model for metals undergoing large elasto-plastic deformations at finite strains is addressed.

3. Implicit finite-element solution strategy

Let us assume that a particular material model has been defined within the framework of continuum thermodynamics with internal variables. The next step towards the prediction of the behaviour of this material in situations of practical interest is the establishment of the corresponding mathematical problem along with a numerical framework capable of producing accurate solutions over a wide range of conditions. In this section, a general framework for the efficient *implicit* finite-element simulation of large-strain problems involving dissipative materials is described. Its basis ingredients comprise:

- (i) an algorithm for numerical integration of the rate constitutive equations, leading to an incremental version of the original constitutive law;
- (ii) a finite-element discretization of the corresponding incremental (equilibrium) boundary value problem stated in the *spatial* configuration; and
- (iii) use of the full Newton–Raphson scheme for iterative solution of the resulting nonlinear algebraic systems of equations to be solved at each increment.

3.1. Numerical integration algorithm. The incremental constitutive law

Given a generic dissipative material model, the solution of the evolution problem defined by the corresponding rate constitutive equations and a set of initial conditions (initial values for the internal variables) is usually not known for complex deformation (and temperature) paths. Therefore, the use of an appropriate numerical algorithm for integration of the rate constitutive equations is an essential requirement in the numerical simulation of problems of interest. The choice of a particular technique for integration of a constitutive law will be obviously dependent on the characteristics of the model considered. In general, algorithms for integration of rate constitutive equations are obtained by adopting some kind of time (or pseudotime) discretization along with some hypothesis on the deformation path between adjacent time stations. Within the context of the purely mechanical theory, considering the time increment $[t_n, t_{n+1}]$ and given the set α_n of internal variables at t_n , the deformation gradient \mathbf{F}_{n+1} at time t_{n+1} must determine the stress σ_{n+1} uniquely through the integration algorithm. One may regard this requirement as the numerical counterpart of the principle of thermodynamic determinism stated in section 2. Such an algorithm defines an (approximate) incremental constitutive functional, $\hat{\sigma}$, for the stress tensor

$$\sigma_{n+1} = \hat{\sigma}(\alpha_n, \mathbf{F}_{n+1}) \quad (3.1)$$

which is path *independent* within one increment and whose outcome σ_{n+1} must tend to the exact solution to the actual evolution problem with vanishingly small deformation increments. Equivalently, an algorithmic functional, $\hat{\tau}$ for the Kirchhoff stress, τ , can be defined:

$$\tau_{n+1} = \hat{\tau}(\alpha_n, \mathbf{F}_{n+1}) = \det[\mathbf{F}_{n+1}] \hat{\sigma}(\alpha_n, \mathbf{F}_{n+1}). \quad (3.2)$$

Within the small-strain elasto-plasticity theory, procedures such as the classical return mappings (Ortiz and Popov 1985, Simo and Hughes 1987) provide concrete examples of numerical integration schemes for path-dependent constitutive laws.

Another important aspect concerning integration algorithms for general dissipative materials is the requirement of *incremental objectivity*. As a numerical version of the principle of material objectivity stated in section 2, incremental objectivity demands that the algorithmic constitutive law be invariant with respect to rigid-body rotations. If this principle is violated, an undesirable dependency of stresses on rotations exists and meaningless results may be obtained with the application of the integration algorithm. In cases such as hypo-elastic formulations (including hypo-elastic-based finite elasto-plasticity), incremental objectivity may not be easily imposed (Rubinstein and Atluri 1983) and, in some instances, its enforcement may result in rather cumbersome algorithms. We remark, however, that since the finite-strain damage models described in this paper are based on hyperelasticity, i.e. the stress tensor is the derivative of a (history dependent) free energy potential, incremental objectivity can be trivially ensured.

3.2. The incremental boundary value problem

The strong form of the momentum balance has been stated in section 2 by expression (2.20). Its weak counterpart, the *principle of virtual work*, is the starting point of displacement-based finite-element solution procedures (Bathe 1996, Hughes 1987, Oden 1972, Zienkiewicz and Taylor 1989, 1991). Consider the body \mathcal{B} subjected to body forces in its interior and surface tractions prescribed on the portion $\Gamma_\sigma \subset \partial\mathcal{B}$ of its boundary $\partial\mathcal{B}$. In addition, let the motion be prescribed by a given function $\bar{\varphi}$ on the remaining portion Γ_φ of $\partial\mathcal{B}$ †:

$$\varphi_t(\mathbf{X}) = \bar{\varphi}_t(\mathbf{X}) \quad \forall \mathbf{x} = \varphi_t(\mathbf{X}) \in \varphi_t(\Gamma_\varphi) \quad (3.3)$$

so that at a time t the set of *kinematically admissible* deformations of \mathcal{B} (often referred to as the *trial solution* set) is defined by

$$\mathcal{C} = \{\varphi_t(\cdot) \mid \varphi_t = \bar{\varphi}_t \text{ on } \Gamma_\varphi\} \quad (3.4)$$

where, for simplicity, the notation $\varphi_t(\bullet) \equiv \varphi(\bullet, t)$ has been used.

The principle of virtual work, in its *spatial* version, states that \mathcal{B} is in equilibrium at t if and only if the Cauchy stress field, $\boldsymbol{\sigma}$, satisfies the variational equation

$$G(\varphi, \boldsymbol{\eta}) := \int_{\varphi_t(\mathcal{B})} (\boldsymbol{\sigma} : \nabla \boldsymbol{\eta} - \mathbf{b} \cdot \boldsymbol{\eta}) \, dv - \int_{\varphi_t(\Gamma_\sigma)} \mathbf{t} \cdot \boldsymbol{\eta} \, da = 0 \quad \forall \boldsymbol{\eta} \in \mathcal{V} \quad (3.5)$$

where \mathbf{b} and \mathbf{t} are respectively the body force and surface traction fields referred to the current configuration and \mathcal{V} is the space of virtual displacements of \mathcal{B} :

$$\mathcal{V} = \{\boldsymbol{\eta} : \varphi_t(\mathcal{B}) \rightarrow \mathbb{R}^3 \mid \boldsymbol{\eta} = \mathbf{0} \text{ on } \varphi_t(\Gamma_\varphi)\}. \quad (3.6)$$

With the introduction of the algorithmic constitutive function $\hat{\boldsymbol{\sigma}}$ in the weak form of the equilibrium, the *incremental* boundary value problem can be stated as follows. *Given the set $\boldsymbol{\alpha}_n$ of internal variables at time t_n and given the body forces and surface traction fields at t_{n+1} , find a kinematically admissible configuration $\varphi_{n+1}(\mathcal{B})$ such that the following holds*

$$\int_{\varphi_{n+1}(\mathcal{B})} (\hat{\boldsymbol{\sigma}} : \nabla \boldsymbol{\eta} - \mathbf{b}_{n+1} \cdot \boldsymbol{\eta}) \, dv - \int_{\varphi_{n+1}(\Gamma_\sigma)} \mathbf{t}_{n+1} \cdot \boldsymbol{\eta} \, da = 0 \quad \forall \boldsymbol{\eta} \in \mathcal{V}. \quad (3.7)$$

Note that due to the introduction of $\hat{\boldsymbol{\sigma}}$, the constitutive relations are satisfied only approximately.

3.3. Finite-element discretization

In the conventional approach, approximations to the incremental boundary value problem above can be obtained by replacing the functional sets \mathcal{C} and \mathcal{V} with their discrete counterparts \mathcal{C}^h and \mathcal{V}^h , generated through a Galerkin finite-element discretization on the configuration $\varphi_{n+1}(\mathcal{B})$. A resulting system of ordinary differential equations, representing the time evolution, is subsequently integrated in time by employing a suitable time integration algorithm, such as, for instance, forward or backward Euler, midpoint rule, Newmark scheme, etc. The algebraic system of equations obtained in this way is then solved by an iterative method; typically the Newton–Raphson method is used within the implicit timestepping scheme. Standard textbooks that provide a detailed account of the FEM are Bathe (1996), Crisfield (1991, 1997), Hughes (1987), Oden (1972), Zienkiewicz and Taylor (1989, 1991). In this section we shall briefly review the main steps of the solution procedure based on the Galerkin finite-element discretization with the attention

† To simplify notation, the boundary regions Γ_σ and Γ_φ are assumed to be invariant with respect to time.

restricted to the quasistatic processes, i.e. the processes in which the inertial effects may be ignored.

Let $N_A : \mathcal{B} \rightarrow \mathbb{R}$, $A = 1, \dots, n_{\text{node}}$ denote the prescribed n_{node} global finite-element shape (interpolation) functions which, together with the associated nodal points $\mathbf{x}_A \in \varphi_t(\mathcal{B})$, discretize the deformed configuration $\varphi_t(\mathcal{B})$ of the body \mathcal{B} . The shape functions are subjected to the standard completeness condition $N_A(\mathbf{x}_B) = \delta_{AB}$, where δ_{AB} is the Kronecker symbol. In the Galerkin finite-element formulation the finite-dimensional subspaces \mathcal{C}^h and \mathcal{V}^h are defined, respectively, as

$$\varphi^h = \sum_{A=1}^{n_{\text{node}}} N_A(\mathbf{x}) \mathbf{u}_A \quad \text{and} \quad \boldsymbol{\eta}^h = \sum_{A=1}^{n_{\text{node}}} N_A(\mathbf{x}) \boldsymbol{\eta}_A \quad \forall \mathbf{x} \in \varphi_t(\mathcal{B}) \quad (3.8)$$

where \mathbf{u}_A is the nodal displacement. Inserting (3.8) into the weak form (3.7), and in view of the arbitrariness of the parameters $\boldsymbol{\eta}_A$ representing the virtual nodal displacements, after some rearrangements the discrete counterpart of (3.7) may be expressed, in the matrix form, as: find a vector \mathbf{U}_{n+1} of the global nodal displacements at t_{n+1} such that the following nonlinear algebraic system:

$$\mathbf{R}(\mathbf{U}_{n+1}) := \mathbf{F}_{n+1}^{\text{int}} - \mathbf{F}_{n+1}^{\text{ext}} = \mathbf{0} \quad (3.9)$$

is satisfied, where $\mathbf{F}_{n+1}^{\text{int}}$ and $\mathbf{F}_{n+1}^{\text{ext}}$ are, respectively, internal and external global force vectors resulting from the assemblage of the element vectors

$$\mathbf{F}^{(e)\text{int}} = \int_{\varphi(\mathcal{B}^{(e)})} \mathbf{B}^T \mathbf{s}_{n+1} \, dv \quad (3.10a)$$

$$\mathbf{F}^{(e)\text{ext}} = \int_{\varphi(\mathcal{B}^{(e)})} \mathbf{N}^T \mathbf{b}_{n+1} \, dv + \int_{\partial\varphi(\mathcal{B}^{(e)}) \cap \varphi(\Gamma_\sigma)} \mathbf{N}^T \mathbf{t}_{n+1} \, da \quad (3.10b)$$

with \mathbf{B} and \mathbf{N} being, respectively, the standard discrete symmetric gradient operator and the interpolation matrix of the element (e) in the configuration defined by displacement \mathbf{U}_{n+1} and \mathbf{s}_{n+1} is the vector containing the Cauchy stress components delivered by the algorithmic function (3.1).

3.4. The Newton–Raphson scheme. Linearization

An effective and efficient way to find a solution \mathbf{U}_{n+1} to the above nonlinear system is to use the standard Newton–Raphson iterative procedure, obtained from the exact linearization of (3.9). During a typical Newton–Raphson iteration (k), the following linear system is solved for the iterative displacement $\Delta\mathbf{U}^{(k)}$:

$$\mathbf{K}(\mathbf{U}_{n+1}^{(k)})[\Delta\mathbf{U}^{(k)}] = -\mathbf{R}(\mathbf{U}_{n+1}^{(k)}) \quad (3.11)$$

and the new guess for the solution \mathbf{U}_{n+1} is updated as

$$\mathbf{U}_{n+1}^{(k+1)} = \mathbf{U}_{n+1}^{(k)} + \Delta\mathbf{U}^{(k)}. \quad (3.12)$$

The *tangent stiffness* \mathbf{K} is defined by the directional derivative formula

$$\mathbf{K}(\mathbf{U})[\Delta\mathbf{U}] = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathbf{R}(\mathbf{U} + \varepsilon\Delta\mathbf{U}). \quad (3.13)$$

If the external loads are assumed independent of \mathbf{U} , then the element tangent stiffness is given by the formula

$$\mathbf{K}^{(e)} = \int_{\varphi(\mathcal{B}^{(e)})} \mathbf{G}^T \mathbf{aG} \, dv \quad (3.14)$$

where \mathbf{G} is the standard discrete spatial gradient operator and \mathbf{a} denotes the matrix form of the *spatial elasticity tensor* \mathbf{a} given, in Cartesian components, by

$$a_{ijkl} = \frac{1}{J} \frac{\partial \tau_{ij}}{\partial F_{kl}} F_{ll} - \sigma_{il} \delta_{jk}. \quad (3.15)$$

Note that, since the Kirchhoff stress tensor is the outcome of the algorithmic function (3.2), its derivative appearing in the expression above is, in fact, the derivative

$$\left. \frac{\partial \hat{\boldsymbol{\tau}}}{\partial \mathbf{F}} \right|_{(\boldsymbol{\alpha}_n, \mathbf{F}_{n+1}^{(k)})} \quad (3.16)$$

of the *incremental* (rather than the actual) constitutive functional. The need for such a consistency between the tangent stiffness and the local algorithm for integration of the rate constitutive equations was first addressed by Nagtegaal (1982), in the context of hypo-elastic-based finite-strain plasticity, and later formalized by Simo and Taylor (1985) who, within the context of infinitesimal von Mises elasto-plasticity, derived a closed formula for the so-called *consistent tangent operators* associated with classical return mapping schemes. It is worth mentioning here that whenever more complex integration algorithms and/or material models (particularly in the finite-strain range) are involved, consistent tangent operators may not be easily derived. Issues associated with consistent linearization aspects in finite multiplicative plasticity are discussed in detail by Simo (1992) and Cuitiño and Ortiz (1992). We remark that, within the present framework, consistent linearization is regarded as a crucial aspect of the formulations presented and will receive particular attention in section 4 where models for elastic and elasto-plastic materials are described. The asymptotically quadratic rates of convergence resulting from the exact linearization of the field equations more than justify the importance of this issue.

4. Finite-strain elasto-plasticity

4.1. Multiplicative decomposition

The main hypothesis underlying the present approach to finite-strain elasto-plasticity is the *multiplicative split* of the deformation gradient, \mathbf{F} , into elastic and plastic parts

$$\mathbf{F} := \mathbf{F}^e \mathbf{F}^p. \quad (4.1)$$

This assumption, first introduced by Lee (1969), admits the existence of a local unstressed *intermediate configuration* (see figure 1). Due to its suitability for the computational treatment of finite-strain elasto-plasticity, the hypothesis of multiplicative decomposition is currently widely employed in the computational mechanics literature (Cuitiño and Ortiz 1992, Eterovic and Bathe 1990, Moran *et al* 1990, Perić *et al* 1992, Perić and Owen 1992a, Simo 1992).

Following the multiplicative split of \mathbf{F} , the *velocity gradient*, $\mathbf{L} \equiv \dot{\mathbf{F}}\mathbf{F}^{-1}$, can be decomposed additively as

$$\mathbf{L} = \mathbf{L}^e + \mathbf{L}^p \quad (4.2)$$

where \mathbf{L}^e and \mathbf{L}^p are, respectively, the elastic and plastic contributions (Nemat-Nasser 1982) *defined* by

$$\mathbf{L}^e := \dot{\mathbf{F}}^e \mathbf{F}^{e-1} \quad \mathbf{L}^p := \mathbf{F}^e \dot{\mathbf{F}}^p \mathbf{F}^{p-1} \mathbf{F}^{e-1}. \quad (4.3)$$

Similarly, the *stretching* (or *rate of deformation*) tensor, $\mathbf{D} \equiv \text{sym}[\mathbf{L}]$, can be decomposed as

$$\mathbf{D} = \mathbf{D}^e + \mathbf{D}^p \quad (4.4)$$

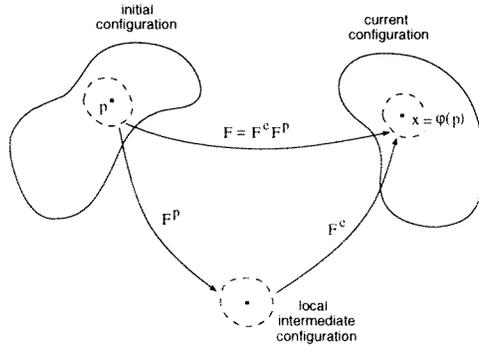


Figure 1. Multiplicative decomposition of the deformation gradient.

with the elastic and plastic stretchings given by

$$\mathbf{D}^e := \text{sym}[\mathbf{L}^e] \quad \mathbf{D}^p := \text{sym}[\mathbf{L}^p]. \quad (4.5)$$

It will be convenient to introduce the *modified* plastic contribution, to the velocity gradient†:

$$\bar{\mathbf{L}}^p := \dot{\mathbf{F}}^p \mathbf{F}^{p-1} \quad (4.6)$$

along with the *modified plastic stretching*:

$$\bar{\mathbf{D}}^p := \text{sym}[\bar{\mathbf{L}}^p]. \quad (4.7)$$

Note that $\bar{\mathbf{D}}^p$ measures the rate of plastic deformation on the *intermediate* configuration. Since the *spatial* configuration will be used to formulate constitutive equations in the following sections, the rotation of $\bar{\mathbf{D}}^p$, defined by

$$\tilde{\mathbf{D}}^p := \mathbf{R}^e \bar{\mathbf{D}}^p \mathbf{R}^{eT} = \mathbf{R}^e \text{sym}[\dot{\mathbf{F}}^p \mathbf{F}^{p-1}] \mathbf{R}^{eT} \quad (4.8)$$

will be adopted in the definition of the plastic flow rule. The *elastic rotation*, \mathbf{R}^e , results from the polar decomposition of \mathbf{F}^e :

$$\mathbf{F}^e = \mathbf{R}^e \mathbf{U}^e = \mathbf{V}^e \mathbf{R}^e \quad (4.9)$$

where \mathbf{U}^e and \mathbf{V}^e denote, respectively, the *right* and *left stretch tensors*.

4.1.1. The logarithmic strain measure. Eulerian (or spatial) elastic strain measures can be defined by using \mathbf{V}^e . Use of the *logarithmic* (or *natural*) strain measure is particularly convenient for the description of the elastic behaviour. In addition to its physical appeal, the use of logarithmic strains results, as we shall see in what follows, in substantial simplifications in the stress integration algorithm and allows a natural extension, to the finite-strain range of the now classical return mapping algorithms of infinitesimal elastoplasticity. The Eulerian logarithmic elastic strain is defined by

$$\boldsymbol{\varepsilon}^e := \ln[\mathbf{V}^e] = \frac{1}{2} \ln[\mathbf{B}^e] \quad (4.10)$$

where $\ln[\bullet]$ above denotes the *tensor logarithm* of (\bullet) and $\mathbf{B}^e = \mathbf{F}^e \mathbf{F}^{eT} = \mathbf{V}^{e2}$ is the elastic left Cauchy–Green strain tensor.

† Lee (1969) has regarded $\bar{\mathbf{L}}^p$ as the velocity gradient of the *purely plastic* deformation and concluded that the additive decomposition (4.2) was valid only if the elastic strains were infinitesimal. This conclusion has been later contested by Nemat-Nasser (1982) who showed that (4.3₂) is an equally acceptable definition for the plastic contribution to the velocity gradient.

The deviatoric/volumetric split of the elastic logarithmic strain gives

$$\boldsymbol{\varepsilon}^e = \boldsymbol{\varepsilon}_d^e + \varepsilon_v^e \mathbf{I} \quad (4.11)$$

where

$$\boldsymbol{\varepsilon}_d^e := \boldsymbol{\varepsilon}^e - \text{tr}[\boldsymbol{\varepsilon}^e] \mathbf{I} \quad (4.12)$$

and the volumetric elastic strain is given by

$$\varepsilon_v^e := \text{tr}[\boldsymbol{\varepsilon}^e] = \ln J^e \quad (4.13)$$

with

$$J^e := \det[\mathbf{F}^e]. \quad (4.14)$$

Note that, due to the properties of the logarithmic strain measure, as in the infinitesimal theory, a traceless $\boldsymbol{\varepsilon}^e$ corresponds to a volume-preserving elastic deformation.

4.2. General hyperelastic-based elasto-plastic constitutive model

Following the formalism of thermodynamics with internal variables described in section 2 and restricted to isothermal processes, a rather general class of isotropic hyperelastic-based finite-strain elasto-plastic constitutive models, formulated in the spatial configuration, can be defined by postulating:

(i) The existence of a *free energy potential*:

$$\psi(\boldsymbol{\varepsilon}^e, \boldsymbol{\alpha}) \quad (4.15)$$

expressed as a function of the elastic logarithmic strain and a generic set $\boldsymbol{\alpha} \equiv \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ of k internal variables.

(ii) A *yield function* $\Phi(\boldsymbol{\tau}, \mathbf{A}; \boldsymbol{\alpha})$ that, for fixed $\boldsymbol{\alpha}$, defines the *elastic domain*, where only reversible phenomena take place, as the set of all points $\{\boldsymbol{\tau}, \mathbf{A}\}$ in the space of thermodynamical forces for which

$$\Phi(\boldsymbol{\tau}, \mathbf{A}; \boldsymbol{\alpha}) \leq 0. \quad (4.16)$$

The yield surface is defined by $\Phi(\boldsymbol{\tau}, \mathbf{A}; \boldsymbol{\alpha}) = 0$.

(iii) A *dissipative potential* $\Psi(\boldsymbol{\tau}, \mathbf{A}; \boldsymbol{\alpha})$, from which the evolution laws for the plastic flow and internal variables are derived, respectively, as

$$\tilde{\mathbf{D}}^p = \dot{\gamma} \frac{\partial}{\partial \boldsymbol{\tau}} \Psi(\boldsymbol{\tau}, \mathbf{A}; \boldsymbol{\alpha}) \quad (4.17)$$

and

$$\dot{\alpha}_i = -\dot{\gamma} \frac{\partial}{\partial A_i} \Psi(\boldsymbol{\tau}, \mathbf{A}; \boldsymbol{\alpha}) \quad (i = 1, \dots, k) \quad (4.18)$$

where the *plastic multiplier* $\dot{\gamma}$ satisfies the *loading/unloading criterion*:

$$\Phi \leq 0 \quad \dot{\gamma} \geq 0 \quad \dot{\gamma} \Phi = 0. \quad (4.19)$$

4.2.1. *The dissipation inequality.* Following assumption (4.15), the time derivative of the free energy reads

$$\dot{\psi} = \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \dot{\boldsymbol{\varepsilon}}^e + \frac{1}{\rho_0} \mathbf{A} \dot{\boldsymbol{\alpha}} \quad (4.20)$$

where the notation: $\mathbf{A} \dot{\boldsymbol{\alpha}} = \sum_{i=1}^k A_i \dot{\alpha}_i$, with the appropriate product implied, has been adopted. Equivalently, by applying the chain rule to the definition (4.10) of $\boldsymbol{\varepsilon}^e$,

$$\dot{\psi} = \frac{1}{2} \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}} : \frac{\partial(\ln \mathbf{B}^e)}{\partial \mathbf{B}^e} : \dot{\mathbf{B}}^e + \frac{1}{\rho_0} \mathbf{A} \dot{\boldsymbol{\alpha}} = \frac{1}{2} \left[\frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \frac{\partial(\ln \mathbf{B}^e)}{\partial \mathbf{B}^e} \mathbf{B}^e \right] : \dot{\mathbf{B}}^e \mathbf{B}^{e-1} + \frac{1}{\rho_0} \mathbf{A} \dot{\boldsymbol{\alpha}}. \quad (4.21)$$

It should be noted that in the expression above, the tensors $\boldsymbol{\varepsilon}^e$, \mathbf{B}^e and $\partial \psi / \partial \boldsymbol{\varepsilon}^e$ share the same principal axes. Also, the tensor logarithm is a member of the class of isotropic tensor functions discussed in de Souza Neto and Perić (1996) and de Souza Neto *et al* (1998). These observations lead to the identity (see de Souza Neto *et al* (1998) for details):

$$\frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \frac{\partial(\ln \mathbf{B}^e)}{\partial \mathbf{B}^e} \mathbf{B}^e = \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} \quad (4.22)$$

and (4.21) can be rewritten as

$$\dot{\psi} = \frac{1}{2} \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \dot{\mathbf{B}}^e \mathbf{B}^{e-1} + \frac{1}{\rho_0} \mathbf{A} \dot{\boldsymbol{\alpha}}. \quad (4.23)$$

By definition, $\mathbf{B}^e := \mathbf{F}^e \mathbf{F}^{eT}$, or, in view of the multiplicative elasto-plastic decomposition assumption, $\mathbf{B}^e = \mathbf{F}(\mathbf{F}^p)^{-1}(\mathbf{F}^p)^{-T} \mathbf{F}^T$. Time differentiation of this last expression and substitution in (4.21) result, after some straightforward manipulations, in

$$\begin{aligned} \dot{\psi} &= \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \left\{ \mathbf{D} + \frac{1}{2} [\mathbf{F}^e \mathbf{F}^p (\mathbf{F}^{p-1}) \cdot \mathbf{F}^{eT} + \mathbf{F}^e (\mathbf{F}^{p-T}) \cdot \mathbf{F}^{pT} \mathbf{F}^{e-1}] \right\} + \frac{1}{\rho_0} \mathbf{A} \dot{\boldsymbol{\alpha}} \\ &= \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \left\{ \mathbf{D} - \frac{1}{2} \mathbf{V}^e \mathbf{R}^e [\bar{\mathbf{L}}^p + \bar{\mathbf{L}}^{pT}] \mathbf{R}^{eT} \mathbf{V}^{e-1} \right\} + \frac{1}{\rho_0} \mathbf{A} \dot{\boldsymbol{\alpha}} \end{aligned} \quad (4.24)$$

where use has been made of the relations: $\mathbf{F}^p (\mathbf{F}^{p-1}) \cdot = -\dot{\mathbf{F}}^p \mathbf{F}^{p-1} =: -\bar{\mathbf{L}}^p$ and $(\mathbf{F}^{p-T}) \cdot \mathbf{F}^{pT} = -\dot{\mathbf{F}}^{p-T} (\mathbf{F}^{pT}) \cdot =: -\bar{\mathbf{L}}^{pT}$, obtained, respectively, by time differentiation of the identities: $\mathbf{F}^p \mathbf{F}^{p-1} = \mathbf{I}$ and $\mathbf{F}^{p-T} \mathbf{F}^{pT} = \mathbf{I}$.

Finally, with the introduction of definition (4.8) of the spatial modified plastic stretching tensor, $\tilde{\mathbf{D}}^p$, and by taking into account the elastic isotropy, the rate of change of free energy can be expressed as

$$\dot{\psi} = \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : (\mathbf{D} - \tilde{\mathbf{D}}^p) + \frac{1}{\rho_0} \mathbf{A} \dot{\boldsymbol{\alpha}}. \quad (4.25)$$

Restricted to isothermal processes, the *Clausius–Duhem* inequality (2.31) can be expressed as

$$\boldsymbol{\tau} : \mathbf{D} - \rho_0 \dot{\psi} \geq 0 \quad (4.26)$$

so that introducing (4.25) one obtains

$$\left(\boldsymbol{\tau} - \rho_0 \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} \right) : \mathbf{D} + \rho_0 \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} : \tilde{\mathbf{D}}^p - \mathbf{A} \dot{\boldsymbol{\alpha}} \geq 0. \quad (4.27)$$

From a standard argument, the inequality above implies the following constitutive equation for the Kirchhoff stress:

$$\boldsymbol{\tau} = \rho_0 \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} \quad (4.28)$$

(i) Multiplicative decomposition of the deformation gradient

$$\mathbf{F} = \mathbf{F}^e \mathbf{F}^p$$

(ii) Hyperelastic law

$$\boldsymbol{\tau} = \rho_0 \frac{\partial \psi(\boldsymbol{\varepsilon}^e, \boldsymbol{\alpha})}{\partial \boldsymbol{\varepsilon}^e}$$

(iii) Evolution equations for \mathbf{F}^p and other internal variables $\boldsymbol{\alpha} \equiv \{\alpha_1, \dots, \alpha_k\}$

$$\begin{aligned} \tilde{\mathbf{D}}^p &:= \mathbf{R}^e \operatorname{sym} \left[\dot{\mathbf{F}}^p \mathbf{F}^{p-1} \right] \mathbf{R}^{eT} = \dot{\gamma} \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \\ \dot{\alpha}_i &= -\dot{\gamma} \frac{\partial \Psi}{\partial A_i} \quad (i = 1, \dots, k) \end{aligned}$$

(iv) Loading/unloading criterion

$$\Phi \leq 0 \quad \dot{\gamma} \geq 0 \quad \dot{\gamma} \Phi = 0$$

Box 4.1. General finite-strain elasto-plastic model.

and the non-negative dissipation requirement is reduced to

$$\boldsymbol{\tau} : \tilde{\mathbf{D}}^p - A \dot{\alpha} \geq 0. \quad (4.29)$$

The overall finite-strain elasto-plastic constitutive model is summarized in box 4.1.

Remark 4.1. Expressions (4.28) and (4.29), as well as the adopted plastic flow rule (4.17), are completely analogous to their small-strain counterparts. In the small-strain limit, $\boldsymbol{\varepsilon}^e$ and $\tilde{\mathbf{D}}^p$ correspond, respectively, to the standard infinitesimal elastic strain tensor and plastic strain rate respectively. Thus, the present approach allows a natural extension, to the finite-strain range, of general isotropic infinitesimal elasto-plastic constitutive models. A generic small-strain model defined by an elastic potential ψ_s , a yield function Φ_s and a dissipation potential Ψ_s can be extended to finite strains by adopting, in the constitutive equations above, ψ , Φ and Ψ with the same functional format as their respective small-strain counterparts.

Remark 4.2. With $J^p := \det[\mathbf{F}^p]$ we define the *plastic volumetric strain* as

$$\varepsilon_v^p := \ln J^p = \ln[\lambda_{(1)}^p \lambda_{(2)}^p \lambda_{(3)}^p] = \ln \lambda_1^p + \ln \lambda_2^p + \ln \lambda_3^p = \operatorname{tr}[\mathbf{V}^p] \quad (4.30)$$

where $\lambda_{(i)}^p$ are the principal *plastic* stretches, i.e. the eigenvalues of the plastic left Cauchy–Green strain tensor, $\mathbf{V}^p := \mathbf{F}^p \mathbf{F}^{pT}$. For volume-preserving \mathbf{F}^p ,

$$\det[\mathbf{F}^p] = 1 \iff \varepsilon_v^p = 0. \quad (4.31)$$

It can be easily shown that the plastic flow rule (4.17) implies that

$$\dot{\varepsilon}_v^p = \dot{\gamma} \operatorname{tr} \left[\frac{\partial \Psi}{\partial \boldsymbol{\tau}} \right] \quad (4.32)$$

so that, as in the infinitesimal theory, dissipation potentials whose derivatives with respect to τ are traceless (such as the classical von Mises and Tresca functions) produce *isochoric* plastic flow.

Remark 4.3. Analogously to the small-strain theory, if Φ is taken as the dissipation potential, then the well known *principle of maximum plastic dissipation* (Hill 1950) is extended to the finite-strain range. In that case, the loading/unloading criterion (4.19) corresponds to the Kuhn–Tucker optimality condition for the left-hand side of (4.29) to reach a maximum subject to the plastic admissibility constraint $\Phi \leq 0$.

4.3. General stress integration procedure. The exponential map algorithm

In the present context, knowing \mathbf{F}_n^p and the set α_n of internal variables at time t_n and given the deformation gradient \mathbf{F}_{n+1} at time t_{n+1} , the numerical integration of the constitutive equations of box 4.1 must determine τ_{n+1} , \mathbf{F}_{n+1}^p and the updated set α_{n+1} at the subsequent time t_{n+1} .

Due to the underlying additive structure of infinitesimal plasticity, *operator split* algorithms are especially suitable for the numerical integration of small-strain elasto-plastic constitutive equations and have been widely used in the computational literature (Crisfield 1991, 1997, Mitchell and Owen 1988, Owen and Hinton 1980, Simo and Hughes 1987, Simo 1992). These methods consist of splitting the problem into two parts: an *elastic predictor*, where the problem is assumed to be purely elastic (no plastic flow or internal variable evolution), and a *plastic corrector*, in which a discrete system of equations comprising the elasticity law, plastic consistency, plastic flow and internal variables evolution is solved, taking the results of the elastic predictor stage as initial conditions. In the present framework for multiplicative finite-strain plasticity an operator split algorithm will be adopted to integrate the constitutive equations of box 4.1. The general algorithm comprises the following steps.

(i) First, it is assumed that the pseudotime increment $[t_n, t_{n+1}]$ is purely elastic (no plastic yielding). The *elastic trial state* at t_{n+1} is then defined by the elastic trial deformation gradient:

$$\mathbf{F}_{n+1}^{e \text{ trial}} := \mathbf{F}_{n+1} \mathbf{F}_n^{p-1} \quad (4.33)$$

with \mathbf{F}^p and α frozen at t_n :

$$\mathbf{F}_{n+1}^{p \text{ trial}} = \mathbf{F}_n^p \quad (4.34)$$

and

$$\alpha_{n+1}^{\text{trial}} = \alpha_n. \quad (4.35)$$

The elastic trial Kirchhoff stress, corresponding to such an assumption, is given by

$$\tau_{n+1}^{\text{trial}} = \left. \frac{\partial \psi}{\partial \varepsilon^e} \right|_{(\varepsilon_{n+1}^{\text{trial}}, \alpha_{n+1}^{\text{trial}})}. \quad (4.36)$$

(ii) If $\Phi(\tau_{n+1}^{\text{trial}}, \alpha_{n+1}^{\text{trial}}) \leq 0$, then the increment is indeed purely elastic and we update:

$$(\bullet)_{n+1} := (\bullet)_{n+1}^{\text{trial}}. \quad (4.37)$$

(iii) Otherwise, plastic yielding occurs and the plastic flow rule (4.17) is discretized by means of a backward *exponential approximation* (see Weber and Anand 1990, Eterovic and

Bathe 1990, Cuitiño and Ortiz 1992), leading to the following discrete evolution law for \mathbf{F}^p :

$$\mathbf{F}_{n+1}^p = \exp \left[\Delta\gamma \mathbf{R}_{n+1}^{eT} \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \Big|_{n+1} \mathbf{R}_{n+1}^e \right] \mathbf{F}_n^p = \mathbf{R}_{n+1}^{eT} \exp \left[\Delta\gamma \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \Big|_{n+1} \right] \mathbf{R}_{n+1}^e \mathbf{F}_n^p. \quad (4.38)$$

In addition, a standard backward Euler scheme is used to integrate the rate evolution equation (4.18) for the internal variables:

$$\boldsymbol{\alpha}_{n+1} = \boldsymbol{\alpha}_n - \Delta\gamma \frac{\partial \Psi}{\partial \mathbf{A}} \Big|_{n+1}. \quad (4.39)$$

The incremental plastic multiplier, $\Delta\gamma$, satisfies the discrete counterpart of (4.19):

$$\Phi_{n+1} \leq 0 \quad \Delta\gamma \geq 0 \quad \Delta\gamma \Phi_{n+1} = 0. \quad (4.40)$$

4.3.1. Consequences of the exponential approximation. The small-strain return map. Some crucially important properties result from the use of the exponential map in the discretization of the plastic flow rule. First, the incompressibility of the plastic flow for pressure insensitive yield criteria is carried over *exactly* to the incremental rule (4.38). Indeed, for a traceless flow direction tensor, $\partial \Psi / \partial \boldsymbol{\tau}$, $\det[\exp[\Delta\gamma \partial \Psi / \partial \boldsymbol{\tau}]] = 1$ which ensures that the updating formula (4.38) is volume preserving. In addition, under isotropic conditions, the essential stress updating procedure can be written in the same format as the classical return mapping schemes of infinitesimal elasto-plasticity, with all large-strain related operations restricted to the kinematical level. This property is demonstrated in what follows.

Inversion of both sides of (4.38) followed by their pre-multiplication by \mathbf{F}_{n+1} and use of relation (4.1), gives

$$\mathbf{F}_{n+1}^e = \mathbf{F}_{n+1}^{\text{trial}} \mathbf{R}_{n+1}^{eT} \exp \left[-\Delta\gamma \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \Big|_{n+1} \right] \mathbf{R}_{n+1}^e. \quad (4.41)$$

Post-multiplication of both sides of (4.41) by \mathbf{R}_{n+1}^{eT} results in

$$\mathbf{V}_{n+1}^e = \mathbf{F}_{n+1}^{\text{trial}} \mathbf{R}_{n+1}^{eT} \exp \left[-\Delta\gamma \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \Big|_{n+1} \right] \quad (4.42)$$

or, equivalently,

$$\mathbf{V}_{n+1}^e \exp \left[\Delta\gamma \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \Big|_{n+1} \right] = \mathbf{F}_{n+1}^{\text{trial}} \mathbf{R}_{n+1}^{eT}. \quad (4.43)$$

Then, further post-multiplication of each side by its transpose gives

$$\mathbf{V}_{n+1}^e \exp \left[2\Delta\gamma \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \Big|_{n+1} \right] \mathbf{V}_{n+1}^e = (\mathbf{V}_{n+1}^{\text{trial}})^2. \quad (4.44)$$

Due to the assumed elastic isotropy, \mathbf{V}^e and $\boldsymbol{\tau}$ commute. If the potential Ψ is assumed to be an *isotropic* function of $\boldsymbol{\tau}$, then $\boldsymbol{\tau}$ and $\partial \Psi / \partial \boldsymbol{\tau}$ have the same principal directions so that all terms on the left-hand side of the above expression commute. Under such assumptions, expression (4.44) leads to the following simpler update formula in terms of the logarithmic Eulerian strain tensor:

$$\boldsymbol{\varepsilon}_{n+1}^e = \boldsymbol{\varepsilon}_{n+1}^{\text{trial}} - \Delta\gamma \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \Big|_{n+1} \quad (4.45)$$

which has the same format of the update formula for the elastic strains of the standard return mapping algorithms of the infinitesimal theory. For the elastic rotation, the following expression is obtained:

$$\mathbf{R}_{n+1}^e = \mathbf{R}_{n+1}^{\text{trial}}. \quad (4.46)$$

(i) Given incr. displ. $\Delta \mathbf{u}$, update the deformation gradient

$$\mathbf{F}_u := \mathbf{I} + \text{Grad}_{\varphi_n} [\Delta \mathbf{u}], \quad \mathbf{F}_{n+1} := \mathbf{F}_u \mathbf{F}_n$$

(ii) Compute elastic trial state

$$\begin{aligned} \mathbf{F}_{n+1}^{e \text{ trial}} &:= \mathbf{F}_{n+1} (\mathbf{F}_n^p)^{-1} \\ \mathbf{B}_{n+1}^{e \text{ trial}} &:= \mathbf{F}_{n+1}^{e \text{ trial}} (\mathbf{F}_{n+1}^{e \text{ trial}})^T \\ \mathbf{R}_{n+1}^{e \text{ trial}} &:= (\mathbf{V}_{n+1}^{e \text{ trial}})^{-1} \mathbf{F}_{n+1}^{e \text{ trial}} \\ \boldsymbol{\varepsilon}_{n+1}^{e \text{ trial}} &:= \ln [\mathbf{V}_{n+1}^{e \text{ trial}}] = \frac{1}{2} \ln [\mathbf{B}_{n+1}^{e \text{ trial}}] \\ \boldsymbol{\alpha}_{n+1}^{\text{trial}} &:= \boldsymbol{\alpha}_n \end{aligned}$$

(iii) **GOTO BOX 4.3** – small strain algorithm (update $\boldsymbol{\tau}$, $\boldsymbol{\varepsilon}^e$ and $\boldsymbol{\alpha}$)

(iv) Update \mathbf{F}^p and the Cauchy stress.

$$\begin{aligned} \mathbf{V}_{n+1}^e &:= \exp [\boldsymbol{\varepsilon}_{n+1}^{e \text{ trial}}] \\ \mathbf{R}_{n+1}^e &:= \mathbf{R}_{n+1}^{e \text{ trial}} \\ \mathbf{F}_{n+1}^p &:= (\mathbf{R}_{n+1}^e)^T (\mathbf{V}_{n+1}^e)^{-1} \mathbf{F}_{n+1} \\ \boldsymbol{\sigma}_{n+1} &:= \det [\mathbf{F}_{n+1}]^{-1} \boldsymbol{\tau}_{n+1} \end{aligned}$$

Box 4.2. General integration algorithm for finite multiplicative elasto-plasticity.

The resulting algorithm for integration of the large-strain elasto-plastic constitutive equations is summarized in boxes 4.2 and 4.3.

Remark 4.4. The operations carried out in box 4.2 are related exclusively to the kinematics of finite strains. Due to the use of logarithmic strains to describe elasticity along with the exponential approximation (4.38) to the plastic flow rule, the essential material related stress updating procedure, shown in box 4.3, preserves the small-strain format. It corresponds to the well established return mapping procedures of infinitesimal elasto-plasticity.

4.4. The spatial tangent modulus

The next step towards the complete incorporation of the present model into the numerical framework is the derivation of a closed formula for the spatial tangent modulus \mathbf{a} , whose general expression is given by (3.15), consistent with the integration algorithm described above.

In the small-strain return mapping of box 4.3, the updated stress $\boldsymbol{\tau}_{n+1}$ is obtained as a function of $\boldsymbol{\alpha}_n$ and the elastic trial logarithmic strain, so that this procedure can be regarded as an incremental constitutive functional of the form

$$\boldsymbol{\tau}_{n+1} = \tilde{\boldsymbol{\tau}}(\boldsymbol{\alpha}_n, \boldsymbol{\varepsilon}_{n+1}^{e \text{ trial}}). \quad (4.47)$$

In the general procedure of box 4.2, $\boldsymbol{\varepsilon}_{n+1}^{e \text{ trial}}$ is computed as a function of $\mathbf{B}_{n+1}^{e \text{ trial}}$ which, in turn,

(i) Elastic predictor

(i.1) Evaluate the trial elastic stress

$$\boldsymbol{\tau}_{n+1}^{\text{trial}} = \left. \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}^e} \right|_{(\boldsymbol{\varepsilon}_{n+1}^{\text{trial}}, \boldsymbol{\alpha}_{n+1}^{\text{trial}})}$$

(i.2) Check the plastic consistency condition

$$\text{IF } \Phi(\boldsymbol{\tau}_{n+1}^{\text{trial}}, \boldsymbol{\alpha}_{n+1}^{\text{trial}}) \leq 0$$

THEN

Set $(\bullet)_{n+1} = (\bullet)_{n+1}^{\text{trial}}$ and RETURN

ELSE go to (ii)

(ii) Plastic corrector (solve the algebraic system for $\Delta\gamma$, $\boldsymbol{\varepsilon}_{n+1}^e$ and $\boldsymbol{\alpha}_{n+1}$)

$$\left\{ \begin{array}{l} \Phi(\boldsymbol{\tau}_{n+1}, \boldsymbol{\alpha}_{n+1}) \\ \boldsymbol{\varepsilon}_{n+1}^e - \boldsymbol{\varepsilon}_{n+1}^{\text{trial}} + \Delta\gamma \left. \frac{\partial \Psi}{\partial \boldsymbol{\tau}} \right|_{n+1} \\ \boldsymbol{\alpha}_{n+1} - \boldsymbol{\alpha}_n + \Delta\gamma \left. \frac{\partial \Psi}{\partial \boldsymbol{A}} \right|_{n+1} \end{array} \right\} = \left\{ \begin{array}{l} 0 \\ 0 \\ 0 \end{array} \right\}$$

where

$$\boldsymbol{\tau}_{n+1} = \left. \frac{\partial \psi(\boldsymbol{\varepsilon}^e, \boldsymbol{\alpha})}{\partial \boldsymbol{\varepsilon}^e} \right|_{n+1}$$

(iii) RETURN

Box 4.3. General stress updating procedure—small strains.

is a function of \mathbf{F}_n^p and \mathbf{F}_{n+1} . With $\boldsymbol{\varepsilon}_{n+1}^{\text{trial}}$ at hand, the Kirchhoff stress is then updated by means of the incremental functional $\tilde{\boldsymbol{\tau}}$ (small-strain algorithm). Thus, the overall procedure defines a function $\hat{\boldsymbol{\tau}}$, for the Kirchhoff stress, that can be generally expressed as

$$\hat{\boldsymbol{\tau}}(\boldsymbol{\alpha}_n, \mathbf{F}_{n+1}) := \tilde{\boldsymbol{\tau}}(\boldsymbol{\alpha}_n, \boldsymbol{\varepsilon}_{n+1}^{\text{trial}}(\mathbf{B}_{n+1}^{\text{trial}}(\mathbf{F}_n^p, \mathbf{F}_{n+1}))). \quad (4.48)$$

Clearly, $\hat{\boldsymbol{\tau}}$ is a particular case of the general algorithmic constitutive functional (3.2).

Application of the chain rule to (4.48) gives

$$\frac{\partial \hat{\boldsymbol{\tau}}}{\partial \mathbf{F}_{n+1}} = \frac{\partial \tilde{\boldsymbol{\tau}}}{\partial \boldsymbol{\varepsilon}_{n+1}^{\text{trial}}} : \frac{\partial \boldsymbol{\varepsilon}_{n+1}^{\text{trial}}}{2\mathbf{B}_{n+1}^{\text{trial}}} : \frac{\partial \mathbf{B}_{n+1}^{\text{trial}}}{\partial \mathbf{F}_{n+1}}. \quad (4.49)$$

Substitution of this expression into (3.15) results, after straightforward manipulations, in the following closed formula for the components of the spatial tangent modulus consistent with the present integration algorithm:

$$\mathbf{a}_{ijkl} = \frac{1}{2J} [\tilde{\mathbf{h}} : \mathbf{n} : \mathbf{b}]_{ijkl} - \sigma_{il} \delta_{jk} \quad (4.50)$$

where $\tilde{\mathbf{h}}$ is the *small-strain elasto-plastic consistent tangent operator*, associated exclusively with the return map algorithm of box 4.3:

$$\tilde{\mathbf{h}} := \frac{\partial \tilde{\boldsymbol{\tau}}}{\partial \boldsymbol{\varepsilon}_{n+1}^{\text{trial}}}. \quad (4.51)$$

The tensor $\tilde{\mathbf{h}}$ is obtained from the linearization of the algorithm of box 4.3 by following the classical procedure introduced by Simo and Taylor (1985). The fourth-order tensor \mathbf{n} is defined as

$$\mathbf{n} := \frac{\partial \ln[\mathbf{B}_{n+1}^{\text{trial}}]}{\partial \mathbf{B}_{n+1}^{\text{trial}}} \quad (4.52)$$

i.e. it is the derivative of the tensor logarithm function at $\mathbf{B}_{n+1}^{\text{trial}}$. The tensor logarithm is a member of the class of isotropic tensor functions and its derivative can be easily obtained (de Souza Neto and Perić 1996). Finally, the Cartesian components of \mathbf{b} are defined by

$$\mathbf{b}_{ijkl} := \delta_{ik}(\mathbf{B}_{n+1}^{\text{trial}})_{jl} + \delta_{jk}(\mathbf{B}_{n+1}^{\text{trial}})_{il}. \quad (4.53)$$

Remark 4.5. Note that $\tilde{\mathbf{h}}$ is the *only* material related contribution to the spatial modulus \mathbf{a} . All other terms taking part in its assemblage in (4.50) are related to the geometry of finite deformations and do not depend on the particular material model adopted.

5. Contact-friction modelling

Over the last few years remarkable progress has been achieved in the field of computational contact mechanics. Undoubtedly, a better mathematical understanding has been one of the key factors in the development of techniques for numerical simulation of the contact problem with friction between deformable bodies. The formulation by means of variational inequalities (Duvaut and Lions 1976, Kikuchi and Oden 1988, Oden and Martins 1985) as well as the use of return mapping algorithms (Giannakopoulos 1989, Laursen and Simo 1993, Perić and Owen 1992b, Wriggers *et al* 1990, Wriggers 1995) have provided efficient frameworks for the numerical treatment of such problems.

Despite such advances, most numerical applications reported in the literature are still restricted to the standard Amontons–Coulomb law of perfect friction. As pointed out by Curnier (1984), such a simplified theory can represent only a limited range of tribological situations. In part, restriction to the standard Amontons–Coulomb law may be justified by the lack of more sophisticated (and well established) phenomenological models for friction. This is in clear contrast to other areas of continuum mechanics such as elasticity and plasticity.

The friction coefficient between two metallic bodies sliding relative to one another depends on the nature and topography of the surfaces in contact as well as on environmental factors (lubrication, presence of oxide films, etc) (Hutchings 1992). During continuous sliding, these conditions at the contact interface may be constantly changing as a consequence of complex phenomena such as the deformation of asperities, wear, internal straining, chemical reactions, etc (Oden and Martins 1985). In experiments with iron and carbon steel specimens, Suh and Sin (1981) observed variations of friction coefficient before steady state condition of the contact interfaces was reached. They attributed this phenomenon to the evolution of the surfaces' topography and consequent changes in the contributions of adhesion, ploughing and asperity deformation to the coefficient of friction at the early stages of sliding. Another example of deviations from the standard Amontons–Coulomb model was observed by Hashimoto *et al* (1990) in experiments with coated steel

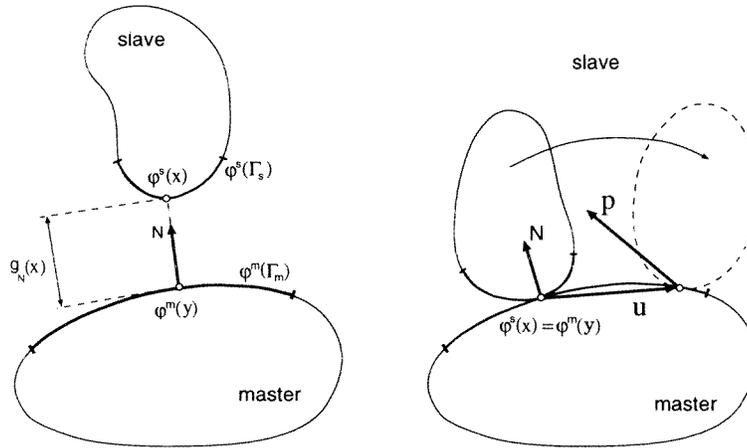


Figure 2. Evolution of frictional contact between two bodies: geometry and notation.

sheets. In situations involving high normal pressures and large sliding distances, typical in deep drawing operations, the surface coating present in these materials may be worn away exposing the bulk metal and causing a dramatic increase in the friction coefficient. In situations such as the ones described above, the evolution of surface wear is a crucial factor in the definition of the frictional behaviour and numerical predictions with reasonable accuracy may demand consideration of more realistic friction rules.

The purpose of this section is the description of a simple phenomenological model, along with an efficient computational framework, capable of simulating the frictional behaviour of contacting bodies subjected to large sliding distances.

5.1. Mechanical description of frictional contact

With reference to figure 2, at the stage of the deformation process corresponding to the deformation mappings φ^s and φ^m of the slave and master bodies respectively, the gap separating a material point \boldsymbol{x} on Γ_s from the master boundary Γ_m is defined by

$$g_N(\boldsymbol{x}) := [\varphi^s(\boldsymbol{x}) - \varphi^m(\boldsymbol{y})] \cdot \boldsymbol{N} \quad (5.1)$$

where \boldsymbol{y} is the material point on Γ_m currently defining the closest distance between \boldsymbol{x} and the master body.

Assume that contact has been established between \boldsymbol{x} and \boldsymbol{y} , i.e. $g_N = 0$. The subsequent relative displacement (under contact conditions) between the two material points will be denoted \boldsymbol{u} . For convenience, it will be decomposed as

$$\boldsymbol{u} = \boldsymbol{u}_T + u_N \boldsymbol{N} \quad (5.2)$$

where $\boldsymbol{u}_T := (\boldsymbol{I} - \boldsymbol{N} \otimes \boldsymbol{N})\boldsymbol{u}$ and $u_N := \boldsymbol{u} \cdot \boldsymbol{N}$ are, respectively, the tangential and normal components of \boldsymbol{u} .

The same additive decomposition will be considered for the contact pressure, \boldsymbol{p} , acting on the slave body, i.e.

$$\boldsymbol{p} = \boldsymbol{p}_T + p_N \boldsymbol{N}. \quad (5.3)$$

5.1.1. *Normal behaviour.* In the definition of the normal contact behaviour, it is assumed that penetration between the two bodies is admissible. In addition, a linear relation between the normal reaction and the penetration, $g_N < 0$, is postulated resulting in the following constitutive function for p_N :

$$p_N = \begin{cases} -k_N g_N & \text{if } g_N < 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

where k_N is the *normal stiffness* or *penalty factor*.

5.1.2. *Tangential behaviour.* Following standard arguments of the elasto-plasticity theory of friction (Curnier 1984, Michałowski and Mroz 1978, Oden and Martins 1985) the decomposition of the tangential relative velocity $\dot{\mathbf{u}}_T$ into an *adherence* and a *slip* component is adopted,

$$\dot{\mathbf{u}}_T = \dot{\mathbf{u}}_T^a + \dot{\mathbf{u}}_T^s \quad (5.5)$$

along with the following constitutive law for the tangent reaction on the slave boundary

$$\mathbf{p}_T = -k_T \mathbf{u}_T^a \quad (5.6)$$

where k_T is the *tangential contact stiffness*. The evolution law for the slip component is defined by

$$\dot{\mathbf{u}}_T^s = -\dot{\gamma} \frac{\partial \Psi}{\partial \mathbf{p}_T} \quad (5.7)$$

where the *slip potential* Ψ defines the direction of frictional sliding and $\dot{\gamma}$ is consistent with the loading/unloading condition:

$$\Phi(\mathbf{p}, A) \leq 0 \quad \dot{\gamma} \geq 0 \quad \dot{\gamma} \Phi(\mathbf{p}, A) = 0. \quad (5.8)$$

The *slip criterion* Φ above, is assumed to be a function of the contact reaction \mathbf{p} and (possibly) a set A of internal variables taking into account the history dependence of the friction phenomenon.

For *isotropic* frictional contact, frictional sliding may occur only in the direction opposite to the tangential reaction. Hence, the slip potential Ψ is given by

$$\Psi(\mathbf{p}_T) = \|\mathbf{p}_T\|. \quad (5.9)$$

As a central hypothesis of the present model for friction with hardening, the following particular form for the slip criterion is postulated:

$$\Phi(\mathbf{p}, w) = \|\mathbf{p}_T\| - \mu(w) p_N \quad (5.10)$$

where the *friction coefficient*, μ , is assumed to be a function of the single scalar internal variable w defined on the *slave* boundary.

In addition, the internal variable w is chosen as the density of *frictional work* expended on the point considered. Hence, its evolution equation is defined by

$$\dot{w} = -\mathbf{p}_T \cdot \dot{\mathbf{u}}_T^s. \quad (5.11)$$

The resulting constitutive model for frictional contact with hardening is analogous to classical work hardening plasticity.

5.2. Numerical integration of the frictional contact model

The similarity between rate-independent elasto-plasticity and the contact problem with friction makes techniques such as the operator split methodology, employed in section 4.3 in the context of large-strain plasticity, particularly suitable for the numerical integration of the constitutive equations for frictional contact presented above. Thus, an algorithm similar to the elastic predictor/plastic corrector scheme described in section 4.3 is used here to integrate the present frictional contact model.

First, at a given configuration φ_{n+1} , defined by an incremental displacement $\Delta \mathbf{u}$, the normal contact reaction is updated as

$$\mathbf{p}_{Nn+1} = g_N(\varphi_{n+1}). \quad (5.12)$$

Then, the corresponding tangential contact reaction, \mathbf{p}_{Tn+1} , is computed by means of the procedure described below.

5.2.1. Elastic predictor. With the fixed contact pressure determined above, the elastic trial state is evaluated:

$$\mathbf{p}_{Tn+1}^{\text{trial}} = \mathbf{p}_{Tn} - k_N \Delta \mathbf{u}_T. \quad (5.13)$$

If the consistency condition

$$\Phi_{n+1}^{\text{trial}} = \|\mathbf{p}_{Tn+1}^{\text{trial}}\| - \mu(w_n) p_{Nn+1} \leq 0 \quad (5.14)$$

is violated, then frictional sliding occurs and the corrector procedure must be employed to compute the frictional force.

5.2.2. Frictional slip corrector. In this stage, the rate evolution equation for w , discretized by a one-step backward Euler scheme, is solved in conjunction with the requirement that the slip function Φ vanishes during frictional sliding, i.e. the following system is solved for the variables $\Delta \gamma$ and w_{n+1} :

$$\|\mathbf{p}_{Tn+1}^{\text{trial}}\| - k_T \Delta \gamma - \mu(w_{n+1}) p_{Nn+1} = 0 \quad (5.15)$$

$$w_{n+1} - w_n - \mu(w_{n+1}) p_{Nn+1} \Delta \gamma = 0 \quad (5.16)$$

and the frictional force is updated as

$$\mathbf{p}_{Tn+1} = \mu(w_{n+1}) p_{Nn+1} \frac{\mathbf{p}_{Tn+1}^{\text{trial}}}{\|\mathbf{p}_{Tn+1}^{\text{trial}}\|}. \quad (5.17)$$

5.3. Consistent linearization of the discrete frictional contact problem

If contact occurs on a certain portion $\gamma_c = \varphi_t(\Gamma_c) \subset \partial \varphi_t(\mathcal{B})$ of the body analysed, then the contact contribution to the virtual work of the external forces:

$$G^c := \int_{\varphi_{n+1}(\Gamma_c)} \mathbf{p}_{n+1} \cdot \boldsymbol{\eta} \, da \quad (5.18)$$

must be included in the weak statement of the momentum balance (3.7). Thus, in order to preserve the quadratic rates of convergence of the Newton–Raphson scheme for solution of the incremental equilibrium problem, the consistent linearization of this extra term must be carefully addressed.

In this context, the appropriate linearization of the frictional contact algorithm plays an essential role. Since for given values of \mathbf{p}_n and w_n and an incremental displacement $\Delta \mathbf{u}$, the numerical integration algorithm described above determines uniquely the reaction \mathbf{p}_{n+1} ,

an *algorithmic* constitutive function $\hat{\mathbf{p}}$ can be defined such that the frictional force \mathbf{p}_{n+1} is delivered as

$$\mathbf{p}_{n+1} = \hat{\mathbf{p}}(\mathbf{p}_n, w_n, \Delta \mathbf{u}). \quad (5.19)$$

Following a standard procedure of classical elasto-plasticity, first given in a complete form by Simo and Taylor (1985), the differentiation of the algorithmic function at converged values for t_{n+1} provides the tangent relation

$$\hat{\mathbf{D}} := \frac{\partial \mathbf{p}_{n+1}}{\partial \Delta \mathbf{u}} \quad (5.20)$$

where the second-order tensor $\hat{\mathbf{D}}$ is the so-called *consistent tangent operator* associated with the present model (and integration algorithm). It can be easily shown that the closed form of $\hat{\mathbf{D}}$ is, for sticking and sliding conditions, respectively, given by

$$\hat{\mathbf{D}} = \begin{cases} -k_T(\mathbf{I} - \mathbf{N}_{n+1} \otimes \mathbf{N}_{n+1}) - k_N(\mathbf{N} \otimes \mathbf{N}) & \text{sticking} \\ -\xi k_T(\mathbf{I} - \mathbf{N}_{n+1} \otimes \mathbf{N}_{n+1} + \zeta k_T(\mathbf{T}_{n+1} \otimes \mathbf{T}_{n+1}) & \text{sliding} \\ -\beta \mu_{n+1} k_N(\mathbf{T}_{n+1} \otimes \mathbf{N}_{n+1}) - k_N(\mathbf{N}_{n+1} \otimes \mathbf{N}_{n+1}) & \end{cases} \quad (5.21)$$

where \mathbf{T}_{n+1} here is the unity vector in the direction of frictional sliding, i.e.

$$\mathbf{T}_{n+1} = \frac{\mathbf{p}_{Tn+1}}{\|\mathbf{p}_{Tn+1}\|} \quad (5.22)$$

and the scalar factors ξ , ζ and β are defined by

$$\xi = \frac{\mu_{n+1} p_{Nn+1}}{\mu_{n+1} p_{Nn+1} + k_T \Delta \gamma} = \frac{\|\mathbf{p}_{Tn+1}\|}{\|\mathbf{p}_{Tn+1}^{\text{trial}}\|} \quad (5.23a)$$

$$\zeta = \xi - \beta \frac{\mu_{n+1} p_{Nn+1}^2}{k_T} \frac{\partial \mu}{\partial w} \Big|_{n+1} \quad (5.23b)$$

$$\beta = \frac{1}{1 + \left(\frac{\mu_{n+1} p_{Nn+1}}{k_T} - \Delta \gamma \right) p_{Nn+1} \frac{\partial \mu}{\partial w} \Big|_{n+1}}. \quad (5.23c)$$

It is interesting to note that the consistent tangent matrix $\hat{\mathbf{D}}$ is symmetric under sticking conditions. However, due to the non-associative character of the frictional contact constitutive equations, its symmetry is lost during frictional sliding. Therefore, within a finite-element context, one should bear in mind the need for the use of unsymmetric solvers in the solution of this class of problems.

Example 5.1 (Friction modelling for coated steel sheet material). The experimental identification of the frictional hardening curve $\mu(w)$ was carried out by means of sliding tests of electrogalvanized (EG) steel sheets, typically employed in the manufacture of automotive body shells. The equipment used and the experimental procedure are schematically illustrated in figure 3. Once the normal force has been applied, the table slides 300 mm, driven by a hydraulic cylinder as represented in figure 3. After sliding, the normal force is released and the table returns to its initial position. The normal force is then reapplied and the cycle is repeated a number of times. During each sliding, the friction coefficient μ is determined by $\mu = F_T/F_N$ where F_T is the average tangential force measured by the load cell indicated in figure 3.

The experimental points used to determine the hardening curves were obtained with four different levels of *constant* normal force (always starting with a virgin specimen): $F_N = 0.98$ kN, $F_N = 1.96$ kN, $F_N = 2.94$ kN and $F_N = 3.92$ kN. Points obtained following more complex load paths, in which the normal force is variable, are also plotted in

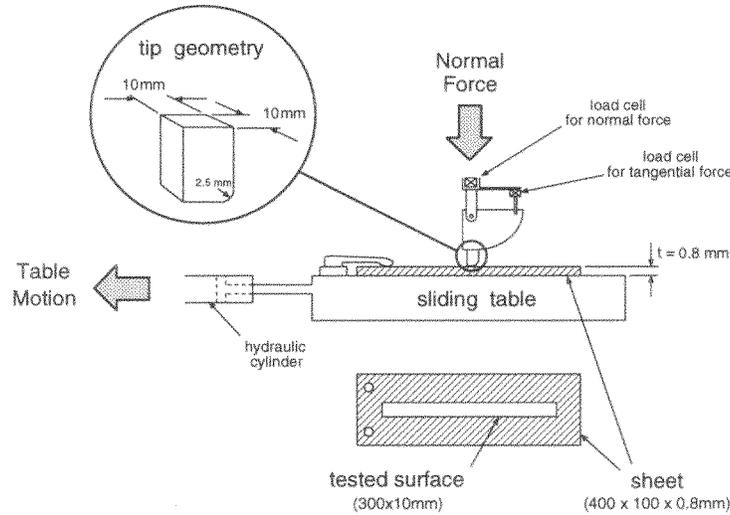


Figure 3. Experimental identification—sliding test.

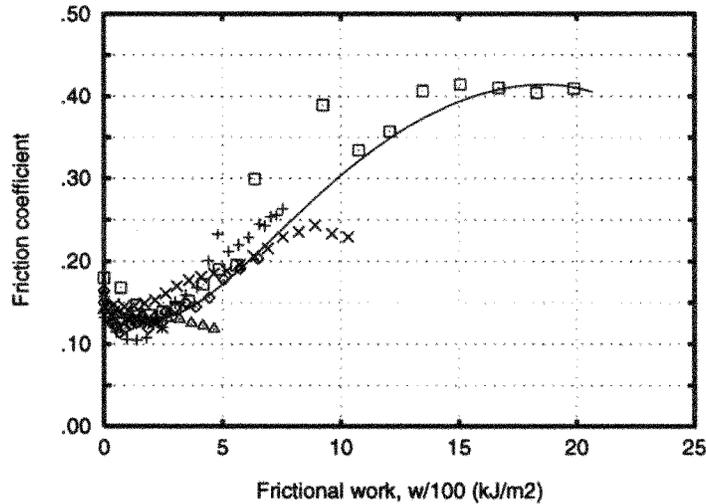


Figure 4. Frictional hardening curves. Experimental determination. EG steel sheet. (\diamond) increasing force, (+) decreasing force; (\square) 3.92 kN, (\times) 2.94 kN, (\triangle) 1.96 kN ($*$) 0.98 kN.

figure 4. In this case, two particular paths have been considered: *increasing* and *decreasing* normal pressure. For increasing normal load, starting with a virgin sheet, a sequence of four sets of five sliding cycles each is executed, respectively, with $F_N = 0.98$ kN, $F_N = 1.96$ kN, $F_N = 2.94$ kN and $F_N = 3.92$ kN (in this order). The procedure for decreasing normal load is the same but with the order of the forces reversed, i.e. it is started with $F_N = 3.92$ kN and, in the last five passes $F_N = 0.98$ kN.

Applying such a procedure for EG steel sheets (see de Souza Neto *et al* (1996a) for details), the resulting frictional hardening curve, shown in figure 4, has been represented by

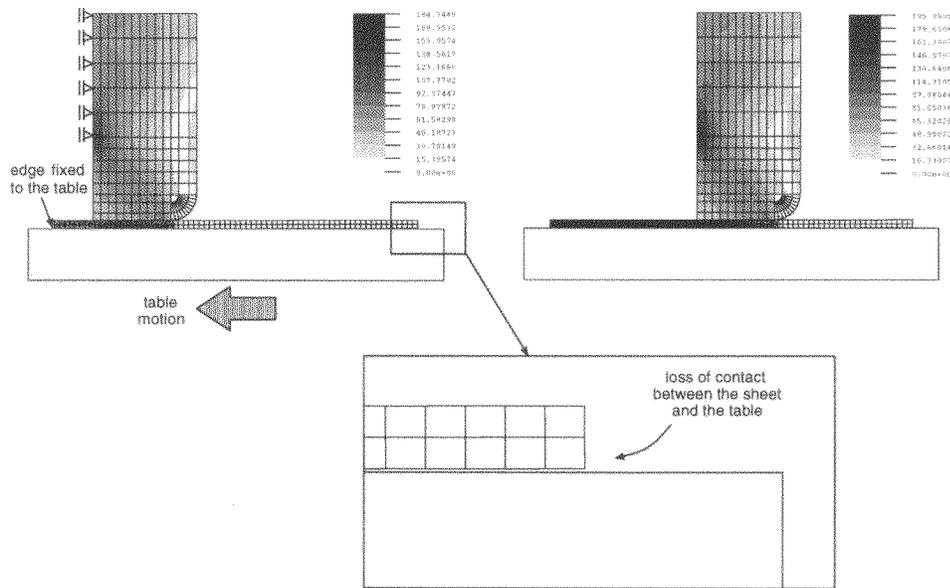


Figure 5. von Mises effective stress distribution (N mm^{-2}). EG steel sheet at constant 3.92 kN normal load—20th pass.

the following polynomial relationship:

$$\begin{aligned} \mu(w) = & -0.4096 \times 10^{-6} w^5 + 0.2890 \times 10^{-4} w^4 - 0.8212 \times 10^{-3} w^3 \\ & + 0.1035 \times 10^{-1} w^2 - 0.3148 \times 10^{-1} w + 0.1568 \end{aligned} \quad (5.24)$$

where w is expressed in $1/100$ of kJ m^{-2} .

To illustrate the predictive capability of the proposed model, the numerical simulation of a series of sliding tests was performed and the results are compared with the corresponding experiments (de Souza Neto *et al* 1996a). The simulation presented here relates to an EG steel sheet with a frictional hardening function given in (5.24). The above functional relationship is implemented as a sequence of linear segments with allows for a general frictional law of the type $\mu = \mu(w)$ to be easily included for simulation.

The finite-element model employed is illustrated in figure 5. The simulation of the sliding test is carried out under (i) an increasing, and (ii) a decreasing sequence of the normal pressures, as described above. For each level, the sliding cycle is repeated five times. The tip lies initially at 5 mm from the left edge of the sheet. Starting from this initial configuration and after the normal force is applied, a relative sliding $d = 25$ mm between the table and the tip is incrementally imposed. This ensures an approximately 10 mm long evenly worn region on the sheet surface (between 7.5 and 17.5 mm from the left edge). When d reaches 25 mm the normal force is released, the tip is lifted up and returned to its initial position.

The tangential reaction force on the tip obtained in the numerical simulation of the sliding tests with an EG steel sheet is plotted in figure 6 for each cycle. The results obtained experimentally are also plotted for comparison. For all levels of normal force a very good correspondence between numerical simulation and experiments is observed. This is an obvious consequence of the close proximity between the experimental points for variable force and the hardening curve observed in figure 4. Again, we remark that the

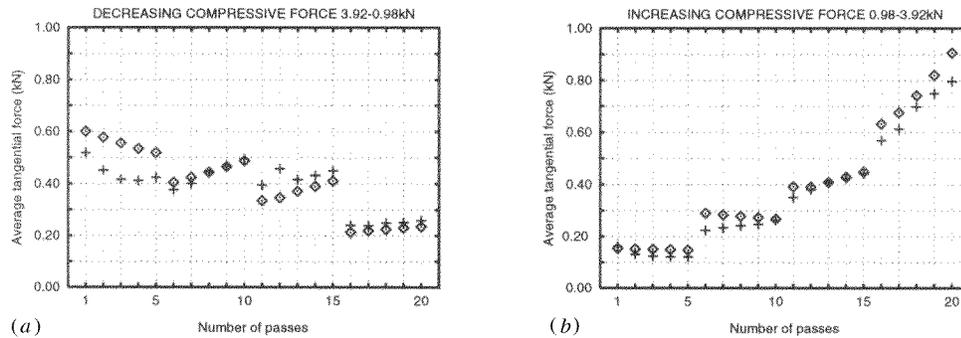


Figure 6. Sliding tests with EG steel sheet at constant normal forces. Finite-element simulation \diamond , and experiment (+). (a) Decreasing normal force: passes [1–5] $\rightarrow F_N = 3.92$ kN, passes [6–10] $\rightarrow F_N = 2.94$ kN, passes [11–15] $\rightarrow F_N = 1.96$ kN, passes [16–20] $\rightarrow F_N = 0.98$ kN; (b) increasing normal force: passes [1–5] $\rightarrow F_N = 0.98$ kN, passes [6–10] $\rightarrow F_N = 1.96$ kN, passes [11–15] $\rightarrow F_N = 2.94$ kN, passes [16–20] $\rightarrow F_N = 3.92$ kN.

agreement in such complex loading paths supports the use of the density of frictional work as the internal variable associated with variations of the friction coefficient.

6. Element technology

It is a well known fact that the performance of low-order kinematically based finite elements is extremely poor near the incompressible limit. Problems of practical engineering interest for which incompressibility plays a crucial role include the analysis of rubbery solids, typically modelled as incompressible hyperelastic materials, as well as elasto-plastic simulations under plastic dominant deformations and the assumption of isochoric plastic flow. In such situations, spurious *locking* frequently occurs as a consequence of the inability of low-order interpolation polynomials to adequately represent general volume-preserving displacement fields. However, due to their simplicity, low-order elements are often preferred in large-scale computations and several formulations have been proposed to allow their use near the incompressible limit. Within the context of the geometrically linear theory, the class of assumed enhanced strain methods described by Simo and Rifai (1990), which incorporates popular procedures such as the classical incompatible modes formulation (Taylor *et al* 1976) and B-bar methods (Hughes 1980) is well established. In the geometrically nonlinear regime, however, the enforcement of incompressibility is substantially more demanding and the development of robust and efficient low-order finite elements is by no means trivial. To tackle such a problem, different approaches have been proposed in the computational literature. Among others, the class of mixed variational methods developed by Simo *et al* (1985), the mixed u/p formulation proposed by Sussman and Bathe (1987), the nonlinear B-bar methodology adopted by Moran *et al* (1990) and the family of enhanced elements of Simo and Armero (1992) are particularly important. One aspect that should be observed here is that, in addition to handling incompressibility, robust formulations should also be able to cope with the extra requirements that different problems may present. For instance, in applications such as the prediction of failure in metal forming processes, the ability to capture strain localization phenomena becomes crucial; in problems involving extremely large strains, frequently encountered in the analysis of rubbery materials and metal forming simulations, it is not unusual that a solution can be obtained only if adaptive mesh refinement

is employed. Thus, since a single formulation is normally not sufficiently robust to produce an optimal performance under a very wide range of conditions, the design of low-order finite elements for large-strain analysis of quasi-incompressible materials remains an open issue.

This section describes the development of a simple four-node quadrilateral and a simple eight-node hexahedron for finite-strain analysis of nearly incompressible solids. The elements, termed F-bar elements, are based on the concept of multiplicative deviatoric/volumetric split in conjunction with the replacement of the compatible deformation gradient field with an assumed modified counterpart. The resulting formulation can be used regardless of the material model adopted. In addition, the strain-driven format of the algorithms for integration of the inelastic constitutive equations of the purely kinematic formulation is maintained. Despite some conceptual similarities, the present approach cannot be regarded as a geometrically nonlinear extension of the B-bar methodology. In order to preserve the asymptotically quadratic rates of convergence of the Newton–Raphson scheme in implicit finite-element computations, the closed form of the corresponding consistent tangent stiffnesses is derived. It has a particularly simple structure so that existing codes that support the conventional four-node displacement-based quadrilateral (or eight-node hexahedron, if 3D analysis is sought) can be easily adapted to incorporate these elements.

6.1. The low-order elements for finite-strain problems

Central to the developments presented below is the concept of multiplicative split of the deformation gradient, \mathbf{F} , into deviatoric (volume-preserving) and volumetric (dilatational) contributions. This multiplicative decomposition has been exploited by Simo *et al* (1985), Moran *et al* (1990) and Simo and Taylor (1991) in the treatment of the incompressibility constraint in finite-deformation problems. It consists of splitting the deformation gradient according to the expression

$$\mathbf{F} = \mathbf{F}_d \mathbf{F}_v \quad (6.1)$$

where \mathbf{F}_d and \mathbf{F}_v are, respectively, the *deviatoric* and *volumetric* components of \mathbf{F} , defined by

$$\mathbf{F}_d = (\det[\mathbf{F}])^{-1/3} \mathbf{F} \quad \mathbf{F}_v = (\det[\mathbf{F}])^{1/3} \mathbf{I}. \quad (6.2)$$

In the expressions above, \mathbf{I} denotes the identity tensor and $\det[\cdot]$ stands for the determinant of $[\cdot]$. Note that, by construction, \mathbf{F}_d and \mathbf{F}_v satisfy

$$\det[\mathbf{F}_d] = 1 \quad \det[\mathbf{F}_v] = \det[\mathbf{F}]. \quad (6.3)$$

6.1.1. Stress computation. The modified deformation gradient. Consider an ordinary displacement-based four-node quadrilateral and an eight-node hexahedron, with local coordinates denoted ξ , as illustrated in figure 7. Typically, the numerical integration of the element internal force vector requires the computation of the stresses at a prescribed number of Gauss points and, for geometrically nonlinear problems, the stresses are obtained from the deformation gradient by means of constitutive functionals. Let \mathbf{F} be the deformation gradient computed from the standard (bilinear for the quadrilateral and trilinear for the hexahedron) interpolation of the displacement field at a generic integration point i , with coordinate ξ_i , indicated in figure 7. For the conventional elements, the Cauchy stress tensor, $\boldsymbol{\sigma}$, in the case of elastic materials is determined at each integration point by

$$\boldsymbol{\sigma} = \hat{\boldsymbol{\sigma}}(\mathbf{F}) \quad (6.4)$$

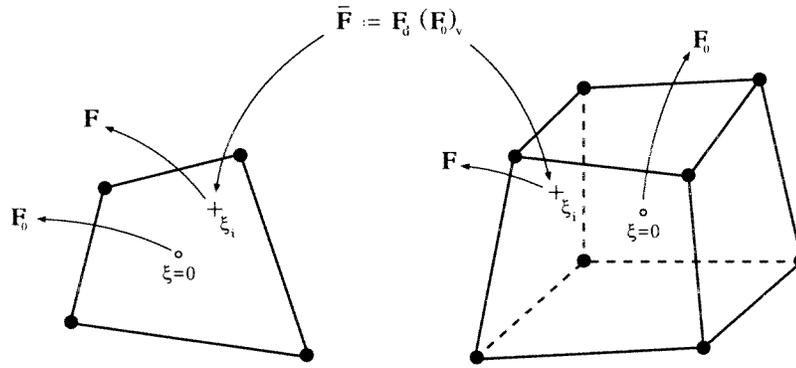


Figure 7. The four-node quadrilateral and eight-node hexahedron.

where $\hat{\sigma}$ is the corresponding elastic constitutive functional. For inelastic materials, incremental constitutive functionals are used instead. In these cases, considering a typical time interval $[t_n, t_{n+1}]$, the stress is evaluated according to the relation (3.1).

The basic objective here is to design a simple four-node quadrilateral and a simple eight-node hexahedron that circumvent, in the nonlinear range, the spurious *locking* exhibited by the standard bilinear and trilinear elements near the incompressible limit. The key idea underlying the present formulation is the use of an assumed *modified* deformation gradient to compute the stresses in (6.4). First, the volumetric/deviatoric split (6.1) is applied to the deformation gradient \mathbf{F} at the Gauss point of interest as well as to the deformation gradient \mathbf{F}_0 that results from the conventional displacement interpolation at the *centroid* of the element, $\boldsymbol{\xi} = \mathbf{0}$ (see figure 7):

$$\mathbf{F} = \mathbf{F}_d \mathbf{F}_v \quad \mathbf{F}_0 = (\mathbf{F}_0)_d (\mathbf{F}_0)_v. \quad (6.5)$$

The *modified deformation gradient*, $\bar{\mathbf{F}}$, is then defined as the composition of the deviatoric component of \mathbf{F} with the volumetric component of \mathbf{F}_0 , i.e.

$$\bar{\mathbf{F}} := \mathbf{F}_d (\mathbf{F}_0)_v = \left(\frac{\det[\mathbf{F}_0]}{\det[\mathbf{F}]} \right)^{1/3} \mathbf{F}. \quad (6.6)$$

Having defined the modified deformation gradient, the proposed four- and eight-node elements are obtained by replacing \mathbf{F} with $\bar{\mathbf{F}}$ in (6.4). Thus, for the present elements, the Cauchy stress at each Gauss point is computed as

$$\boldsymbol{\sigma} = \hat{\sigma}(\bar{\mathbf{F}}). \quad (6.7)$$

Remark 6.1. By construction of $\bar{\mathbf{F}}$, the deviatoric/volumetric split of the modified deformation gradient gives

$$\bar{\mathbf{F}}_d = (\det[\mathbf{F}])^{-1/3} \mathbf{F} = \mathbf{F}_d \quad (6.8a)$$

$$\bar{\mathbf{F}}_v = (\det[\mathbf{F}_0])^{1/3} \mathbf{I} = (\mathbf{F}_0)_v \quad (6.8b)$$

i.e. the deviatoric component of $\bar{\mathbf{F}}$ coincides with the current (integration point) deviatoric deformation gradient while its volumetric part corresponds to the dilatation at the centroid of the element. In view of (6.4), this implies that, for materials whose deviatoric and volumetric responses are *decoupled*, the present formulation results in constant pressure distributions throughout one element.

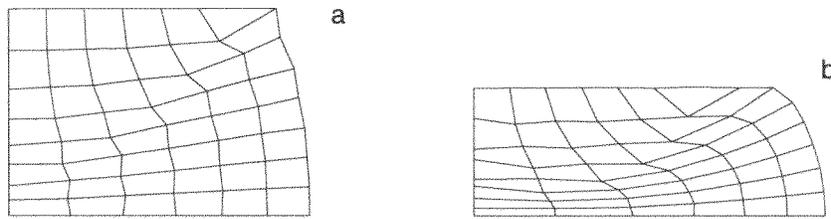


Figure 8. Upsetting of a cylindrical billet. Hourglass patterns. (a) Four-Gauss-point quadrature rule and (b) nine-Gauss-point quadrature rule.

6.2. Some practical experiences

Several numerical examples, involving hyperelastic as well as elasto-plastic simulations, are presented in this section. It provides an assessment of the performance of the proposed four-node quadrilateral and eight-node hexahedron over a wide range of circumstances. Some results obtained here with the four-node quadrilateral are compared with similar computations carried out using the geometrically nonlinear enhanced strain elements Q1/E4 and Q1/E5, respectively for plane-strain and axisymmetric problems, proposed by Simo and Armero (1992). We remark that the hyperelastic simulations are dealt with within the context of finite elasticity set on the spatial configuration as described by de Souza Neto *et al* (1994). In the elasto-plastic problems, the framework for treatment of finite multiplicative plasticity based on logarithmic strains described in section 4 is adopted.

Example 6.1 (Upsetting of an elasto-plastic cylindrical billet). As compared with the standard isoparametric formulation, a substantial diminution of incompressibility locking and accuracy gain in bending dominated problems is observed when enhanced assumed strain elements are used. Nevertheless, drawbacks still exist and need to be carefully addressed. Particularly in highly constrained compressive regimes, the presence of undesirable hourglass modes may be detected. This fact is well illustrated in figure 8 which shows deformed meshes during the simulation of the upsetting of an elastoplastic cylindrical billet with the assumed enhanced strain axisymmetric element named EAS-5 by Simo and Armero (1992). Figures 8(a), (b) correspond, respectively, to four- and nine-Gauss-point integration rules. Hourglass patterns are clearly visible and indicate that further research is required for the design of enhanced assumed strain finite elements with optimal performance in a wider range of situations.

Example 6.2 (Rubber cylinder pressed between two plates). The simulation of the compression of a rubber cylinder between two frictionless rigid plates using the F-bar element is carried out in this example. This problem has been considered by Sussman and Bathe (1987) in the context of the u/p formulation and by Simo and Taylor (1991), who employed a mixed formulation in conjunction with an augmented Lagrangian procedure. The geometry of the problem and the boundary conditions are illustrated in figure 9(a). Following Sussman and Bathe (1987), two different material models which fit the same experimental data are used: (i) a regularized Mooney–Rivlin material given by the strain energy function

$$\hat{\psi}(I_1, I_2) = C_1(I_1 - 3) + C_2(I_2 - 3) \quad (6.9)$$

where I_1 and I_2 are the first and second invariants of the left Cauchy–Green strain tensor $\mathbf{B} = \mathbf{F}\mathbf{F}^T$, respectively, with constants $C_1 = 0.293$ MPa, $C_2 = 0.177$ MPa and bulk

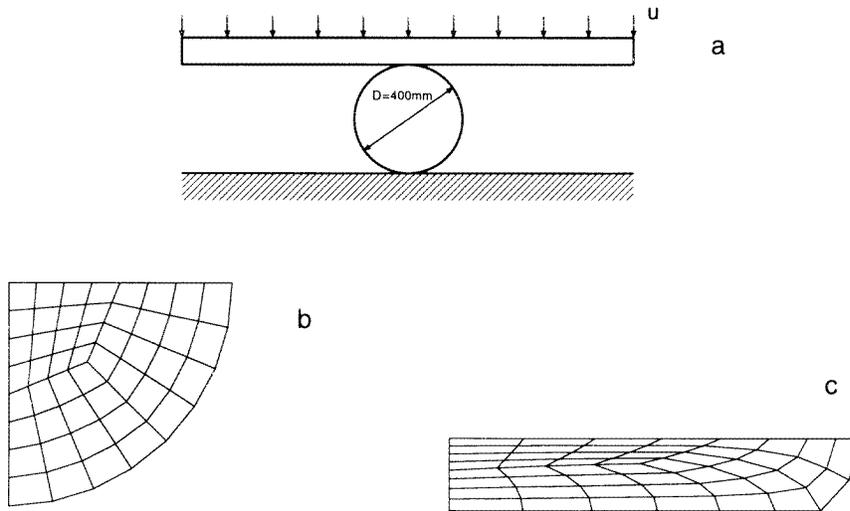


Figure 9. Rubber cylinder pressed between two plates. (a) Geometry and boundary conditions; (b) initial mesh; and (c) deformed mesh at $u = 250$ mm.

modulus $k = 1410$ MPa, and (ii) an Ogden material (Ogden 1984)

$$\bar{\psi}(\lambda_1, \lambda_2, \lambda_3) = \sum_{p=1}^3 \frac{\mu_p}{\alpha_p} (\lambda_1^{\alpha_p} + \lambda_2^{\alpha_p} + \lambda_3^{\alpha_p} - 3) \quad (6.10)$$

where λ_i , $i = 1, 2, 3$ are principal stretches while the material coefficients are given as $\mu_1 = 0.746$ MPa, $\mu_2 = -0.306$ MPa, $\mu_3 = 6.609 \times 10^{-5}$ MPa, $\alpha_1 = 1.748$, $\alpha_2 = -1.656$ and $\alpha_3 = 7.671$. A plane-strain state is assumed and, for symmetry reasons, only one quarter of the cylinder cross section is considered in the simulation. It is discretized with the mesh of 48 elements shown in figure 9(b). The deformed mesh obtained with a prescribed displacement $u = 250$ mm of the plate is depicted in figure 9(c). It is in good qualitative agreement with the deformed mesh shown by Simo and Taylor (1991). The reaction forces per unit thickness on the plate, obtained for the Mooney–Rivlin and Ogden models, are plotted in figure 10 against the plate deflection u . The curves are plotted up to $u = 200$ mm. For both materials the results agree with Sussman and Bathe (1987).

Example 6.3 (Elastomeric bead compression). The numerical simulation of the compression of an elastomeric axisymmetric bead is carried out in this example. The bead—a circular ring with trapezoidal cross section—is schematically represented in figure 11. Its function is to provide sealing when the plate, which contacts its top edge, is pressed downwards. In the finite-element simulation, the bottom edge of the ring seal is assumed clamped to a flat rigid base and both plate and base are idealized as rigid bodies with frictionless contact condition on the boundaries. The bead is modelled as a regularized neo-Hookean material

$$\tilde{\psi}(I_1) = C(I_1 - 3) \quad (6.11)$$

where $C = 2.5$ and with bulk modulus $k = 1000$. A mesh of 520 elements is used to discretize the bead. Figure 12(a) shows the mesh in its initial configuration. The compression of the bead is simulated here with both the axisymmetric version of the four-node quadrilateral presented in section 6.1 and the axisymmetric enhanced element Q1/E5. For the enhanced element, the five-point integration rule is employed in order to avoid

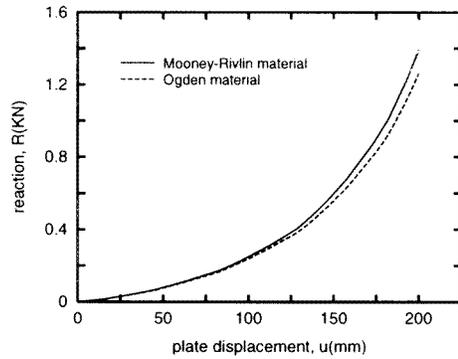


Figure 10. Rubber cylinder pressed between two plates. Load-deflection curves.

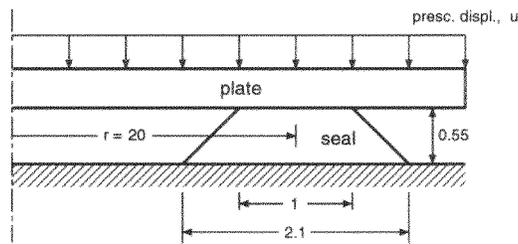


Figure 11. Elastomeric bead compression. Initial geometry and boundary conditions.

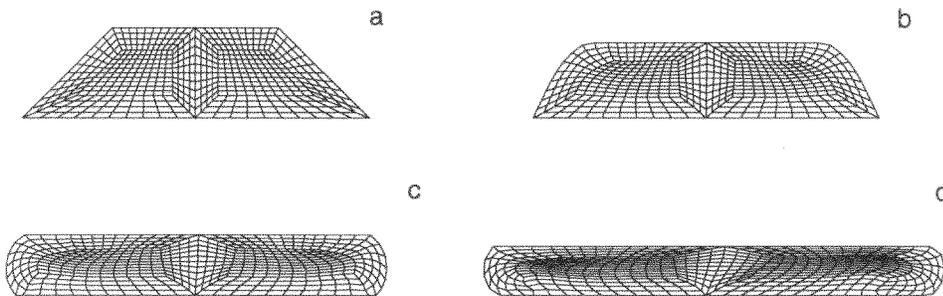


Figure 12. Elastomeric bead compression. Finite-element discretization on the initial configuration and deformed meshes obtained with the present element. (a) Initial configuration; (b) $u = 0.09$; (c) $u = 0.17$; and (d) final configuration, $u = 0.25$.

spurious mechanisms. With the present element, a total vertical displacement $u = 0.25$ is applied to the plate in 25 increments. Deformed meshes obtained at different stages of the compression process are depicted in figures 12(b)–(d). They correspond, respectively, to $u = 0.09$, 0.17 and 0.25. At the early stage shown in figure 12(b), the lateral surfaces of the seal make contact only with the top plate. At the later stages of figures 12(c), (d), contact also occurs with the rigid base. The reaction–displacement curve obtained is plotted in figure 13. The results of the simulation with the Q1/E5 element are shown in figure 14. In this case, due to the activation of non-physical hourglass deformation modes, the enhanced element fails to produce meaningful results. Although no severe spurious hourglass patterns can be observed in the deformed mesh of figure 14(a), obtained with $u = 0.09$, at the later

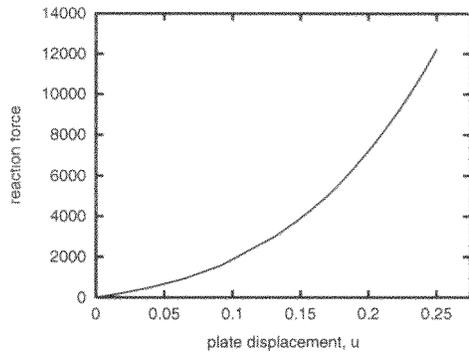


Figure 13. Elastomeric bead compression. Displacement–reaction diagram.

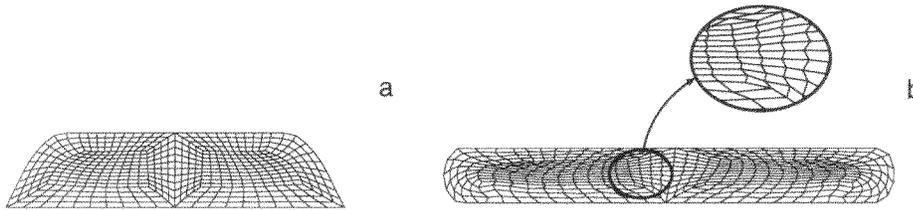


Figure 14. Elastomeric bead compression. Deformed meshes obtained with the enhanced element Q1/E5 (five-point integration rule). (a) $u = 0.09$; and (b) $u = 0.235$.

stage depicted in figure 14(b), with $u = 2.35$, extremely severe hourglassing has spread throughout the entire structure. The rank deficiency that triggers instability in this class of enhanced elements seems to have been first identified by Wriggers and Reese (1996) in the context of finite hyperelasticity. This fact was later confirmed by Crisfield *et al* (1995) and de Souza Neto *et al* (1995), for the elasto-plastic case and represents a serious limitation on the applicability of these elements in the finite-strain range, particularly under compressive dominant stresses. We remark that, so far, no pathological rank deficiency has been detected in the F-bar formulation.

7. Further aspects of the deformation of inelastic solids at finite strains

Previous sections provide a general continuum-based theoretical and computational framework for analysis of large deformations of inelastic solids at large strains. When considering a specific deformation process it is often necessary to resort to further constitutive descriptions that describe salient features of a particular material behaviour. Although phenomenological by construction, these models are based, directly or indirectly, on microscopic arguments and usually rely on a substantial body of physical evidence. A number of important classes of material description may be identified. Attention here is concentrated on three classes of materials that have proved to be useful in describing such specific features of material behaviour. The influence of the curvature of the elasto-plastic yield surface is examined utilizing the anisotropic yield criterion proposed by Barlat and Lian (1989). Additionally, by introducing damage parameters the effect of continuous deterioration of material properties is addressed. Finally, material rate dependence is introduced through a viscoplastic constitutive model. Rate sensitivity and strain rate effects

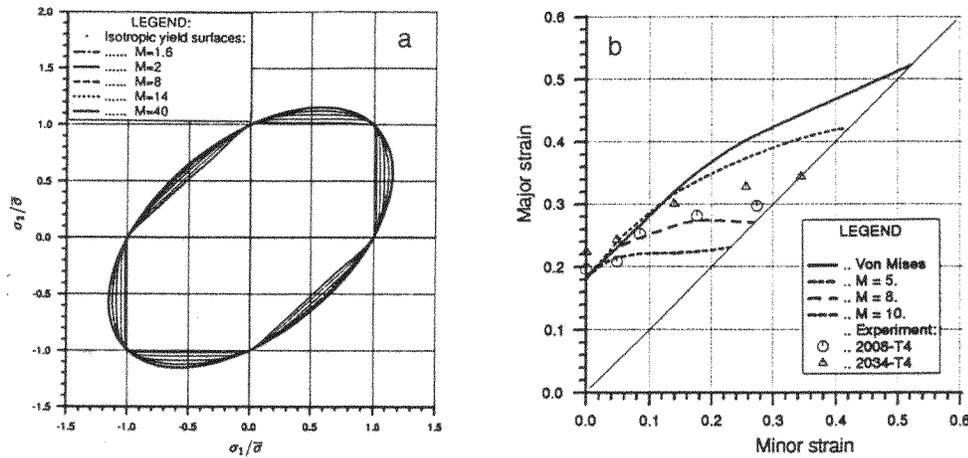


Figure 15. (a) Isotropic yield surfaces for several values of material constant M and for shear stress $\sigma_{xy} = 0$. (b) The effect of yield condition on limit forming predictions.

are known to have a significant role in constitutive description for a number of materials and processes.

7.1. Yield surface representation

For thin-sheet metal forming operations, due to the manufacturing process of the material, the plastic behaviour of textured polycrystalline sheet is predominantly anisotropic. Use of the Hill anisotropic criterion, which contains no shear stress, is restricted to a planar isotropy or for cases where the principal stress axes coincide with the anisotropy axes.

Full planar anisotropy is described by the yield function introduced by Barlat and Lian (1989) which for the plane-stress state is of the superquadric form (computational aspects of this criterion are discussed by Dutko *et al* (1993))

$$f(\boldsymbol{\sigma}, \bar{\sigma}) = a|K_1 + K_2|^M + a|K_1 - K_2|^M + (2 - a)|2K_2|^M - 2\bar{\sigma}^M \quad (7.1)$$

in which

$$K_1 = \frac{\sigma_{xx} + h\sigma_{yy}}{2} \quad \text{and} \quad K_2 = \sqrt{\left(\frac{\sigma_{xx} - h\sigma_{yy}}{2}\right)^2 + p^2\sigma_{xy}^2} \quad (7.2)$$

where a , h , p and M are material constants and $\bar{\sigma}$ is the yield stress from a uniaxial tension test. For a given value of the exponent M material constants a , h , p can be evaluated using R -values, i.e. plastic strain ratios of the in-plane strains to the thickness strain obtained from uniaxial tension tests in three different directions. The role of the material constant M is illustrated in figure 15(a) obtained by plotting the function (7.1) for various M values in the normalized σ_{xx} and σ_{yy} plane for the isotropic case (i.e. for $a = h = p = 1$) and taking $\sigma_{xy} = 0$. The resulting set of functions span the set of yield surfaces which include the standard von Mises and Tresca yield surfaces for $M = 2$ and $M \rightarrow \infty$, respectively.

An important property of the yield function described by equations (7.1), (7.2) is its convexity when constants a , h , p are positive and $M > 1$, as proved by Barlat and Lian (1989).

It has long been accepted that the von Mises criterion represents the yielding conditions in metals adequately. For example, even shear band formation in a thin sheet under uniaxial

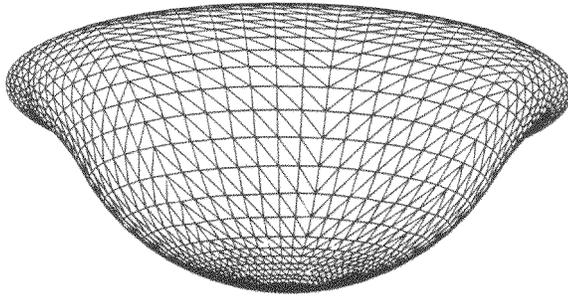


Figure 16. Stretching of a circular thin sheet by a hemispherical punch: deformed finite-element mesh at punch displacement $D_p = 40.0$ mm for the von Mises yield criterion $M = 2$.

tension may be well described utilizing simple von Mises flow theory of plasticity as shown by Perić *et al* (1991). However, under biaxial tension, prediction of the onset of localized necking requires relaxation of the normality condition through vertex formation on the yield surface, or increase of the curvature of the yield surface that is achieved by increasing the value of the M parameter in the Barlat model (7.1), (7.2). Figure 15(b) shows the use of such surfaces in simulating limit forming conditions in the biaxial stretching of an aluminium sheet. It is seen that the best correspondence with experiment is obtained for $M = 6-8$ and not the von Mises value of $M = 2$. Incidentally, for aluminium alloys which are composed of face centred cubic (FCC)-type crystalline structures the value of $M = 8$ has also been suggested by Barlat and Lian (1989) who, for this value of M , found a close correspondence with the yield surfaces obtained on the basis of microstructural arguments.

Example 7.1 (Stretching of a circular thin sheet by a hemispherical punch: elasto-plastic material). The geometry, material characteristics and other parameters are given by Owen and Perić (1994). The analysis is performed employing a 3D membrane formulation for a quarter of the model with appropriate boundary conditions. The material is assumed to follow Barlat's yield criterion (7.1), (7.2) where parameter M is varied. The solution for the standard von Mises material represented by $M = 2$ is also provided. To solve this problem the blank is discretized with 736 constant strain triangular finite elements, and 2145 and 612 triangular flat elements are used to discretize the surfaces of the punch and die, respectively. Results are obtained for a coefficient of friction between tools and blank of $\mu = 0.30$. Figure 16 gives the deformed mesh for punch displacement $D_p = 40$ mm and von Mises yield criterion $M = 2$.

The punch force versus punch displacement diagram, presented in figure 17(a), gives a comparison between results obtained for various M -values. The maximum punch force decreases with increase of the M -value, which indicates a strong influence of the curvature of the yield surface on the initiation of strain localization and its development.

The distribution of true strain in the radial direction is shown in figure 17(b) for various M -values and for punch displacements $D_p = 30$ mm. A typical localization behaviour may be observed where strain accumulates in a narrow zone, reaching high levels and leading to failure. The appearance of localization and associated failure is less pronounced with decrease of the curvature on the yield surface, specified by decrease of M -value, and is clearly delayed for a quadratic yield surface corresponding to the standard von Mises yield criterion.

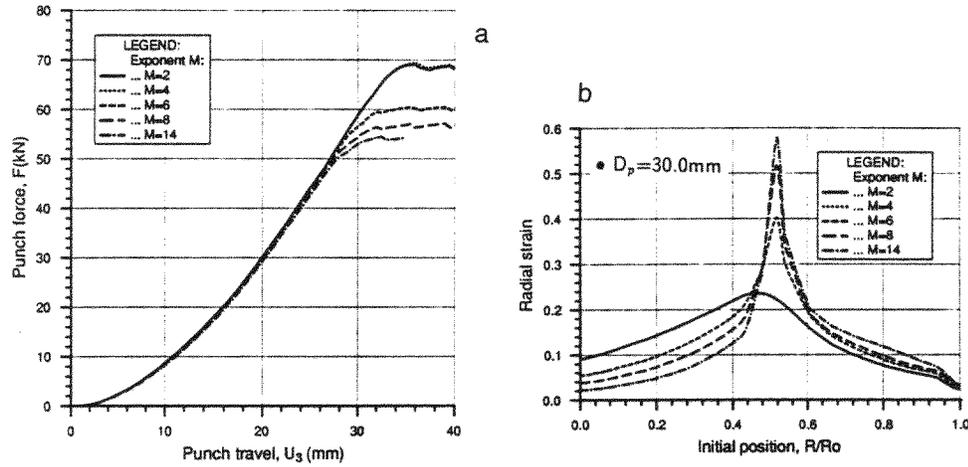


Figure 17. Stretching of a circular thin sheet by a hemispherical punch. (a) Punch force versus punch displacement curves for various M -values. (b) Distribution of true strain in the radial direction plotted over the initial configuration for various M -values at punch displacement $D_p = 30.0$ mm.

7.2. Elasto-plastic damaging solids

As experimentally verified for many materials (Lemaitre 1985, Lemaitre and Chaboche 1990), the nucleation and growth of voids and microcracks which accompany large plastic flow causes a reduction of the elastic modulus as well as material softening and can be strongly influenced by the triaxiality of the state of stress. In this case, the prediction of rupture with reasonable accuracy requires consideration of the coupling between elasto-plasticity and damage.

With attention restricted to isotropic solids, a scalar parameter D is introduced, which represents the density of microcracks and microcavities in the body. The elasticity law is then given by

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\varepsilon}^e (1 - D) \quad (7.3)$$

where \mathbf{C} is the elastic modulus which is assumed constant and isotropic.

Stresses are assumed to satisfy the von Mises yield criterion in the form

$$f(\boldsymbol{\sigma}, \bar{\sigma}, D) = \frac{\sqrt{3J_2(\boldsymbol{\sigma})}}{1 - D} - \bar{\sigma} \quad (7.4)$$

where $J_2(\boldsymbol{\sigma})$ is the second principal invariant of the deviatoric stress tensor $\text{dev}[\boldsymbol{\sigma}]$ and $\bar{\sigma}$ ($\bar{\varepsilon}^P$) is the uniaxial yield stress. The evolution laws for internal variables are obtained in a standard way employing the second law of thermodynamics in the form of the Clausius–Duhem inequality (2.31) and adopting the normal dissipation assumption which results in

$$\dot{\varepsilon}^p = \dot{\gamma} \frac{1}{(1 - D)} \frac{\text{dev}[\boldsymbol{\sigma}]}{\sqrt{3J_2(\boldsymbol{\sigma})}} \quad (7.5a)$$

$$\dot{D} = \dot{\gamma} \frac{1}{1 - D} \left(\frac{-Y}{S_0} \right)^{s_0} \quad (7.5b)$$

where S_0 and s_0 are material constants, and Y is the damage energy release rate expressible in terms of the stresses and accumulated damage, D . A computational model for the fully

coupled elasto-plastic damage model at finite strains is provided by de Souza Neto *et al* (1994).

7.3. Elasto-viscoplastic solids

The constitutive model for von Mises elasto-viscoplasticity is represented as

$$f(\boldsymbol{\sigma}, \bar{\sigma}) = F(\boldsymbol{\sigma}) - \bar{\sigma} \quad (7.6a)$$

$$F(\boldsymbol{\sigma}) = \sqrt{3J_2} \quad (7.6b)$$

$$\dot{\varepsilon}^{vp} = \gamma \langle \Phi(\bar{\sigma}) \rangle \sqrt{\frac{3}{2}} \frac{\text{dev}[\boldsymbol{\sigma}]}{\|\text{dev}[\boldsymbol{\sigma}]\|} \quad (7.6c)$$

in which Φ is the *viscoplastic flow potential*.

For metal forming operations under high-temperature conditions, the effective stress (usually termed the flow stress) is rate dependent, making viscoplastic approaches suitable for simulation. Several empirical relations exist for flow stress prediction, which are based on experimental tests, and are applicable to particular materials under specific conditions. Three of the most commonly employed expressions are summarized below.

(1) *Hajduk expression*. The flow stress $\bar{\sigma}$ is assumed to be expressed in the following form:

$$\bar{\sigma} = K_{f0} K_T K_{\bar{\varepsilon}} K_{\dot{\varepsilon}} \quad (7.7)$$

where the three coefficients are functions of the form

$$K_T = A_1 \exp[-m_1 T] \quad (7.8a)$$

$$K_{\bar{\varepsilon}} = A_2 \bar{\varepsilon}^{m_2} \quad (7.8b)$$

$$K_{\dot{\varepsilon}} = A_3 \dot{\varepsilon}^{m_3} \quad (7.8c)$$

where A_i and m_i are material constants determined from tests.

(2) *Sellars–Tegart expression*. This is based on the following interpolation equation:

$$Z = \dot{\varepsilon} \exp\left[\frac{Q}{R(T + 273)}\right] = C[\sinh[\alpha\bar{\sigma}]]^n \quad (7.9)$$

where Z is the Zener–Hollomon parameter, Q is an activation energy usually independent of temperature and in many cases also independent of strain, R is the gas constant $8.31 \text{ J mol}^{-1} \text{ K}^{-1}$, T is the temperature in degrees Celsius, while C , α and n are material constants.

(3) *ALSPEN expression* (Fjaer and Mo 1990). This expression is found to cover closely the properties of some aluminium alloys by fitting experimental curves in the form

$$\bar{\sigma} = c(T)(\alpha + \alpha_0)^{n(T)} \bar{\varepsilon}^{m(T)}. \quad (7.10)$$

Coefficients $c(T)$, $n(T)$ and $m(T)$ are described by Fjaer and Mo (1990) and Schönauer *et al* (1993), α_0 is a constant, and $d\alpha = d\bar{\varepsilon}^{vp}$ for temperatures below the onset limit $T_0 \approx 700 \text{ K}$, otherwise $d\alpha = 0$.

Since practically $\bar{\varepsilon} = \bar{\varepsilon}^{vp}$ and $\dot{\varepsilon}^{vp} = \gamma \Phi$ for J_2 plasticity, parameters γ and Φ can be obtained, as shown in box 7.1, for all three interpolation functions. In all three cases, the yield stress is taken to be zero, thus the assumption is made that some part of the strain is always inelastic.

(i) Additive strain rate decomposition

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^{vp}$$

(ii) Elastic response

$$\dot{\boldsymbol{\varepsilon}}^e = \mathbf{C}^{-1} : \dot{\boldsymbol{\sigma}}$$

(iii) Flow rule

$$\dot{\boldsymbol{\varepsilon}}^{vp} = \gamma \langle \Phi(\bar{\boldsymbol{\sigma}}) \rangle \sqrt{3/2} \frac{\text{dev}[\boldsymbol{\sigma}]}{\|\text{dev}[\boldsymbol{\sigma}]\|}$$

(iv) γ and Φ are given from interpolation functions:

(1) Hajduk

$$\gamma = \left[\frac{1}{A_1 A_2 A_3} \exp[m_1 T] \bar{\boldsymbol{\varepsilon}}^{vp(-m_2)} \right]^{(1/m_3)}$$

$$\Phi = \left(\frac{\bar{\boldsymbol{\sigma}}}{K_{fo}} \right)^{(1/m_3)}$$

(2) Sellars–Tegart

$$\gamma = C \exp \left[\frac{-Q}{R(T + 273)} \right]$$

$$\Phi = [\sinh[\alpha \bar{\boldsymbol{\sigma}}]]^n$$

(3) ALSPEN

$$\gamma = [c(T)]^{(-1/m(T))} [\alpha(\bar{\boldsymbol{\varepsilon}}^{vp}) + \alpha_0]^{(-n(T)/m(T))}$$

$$\Phi = (\bar{\boldsymbol{\sigma}})^{(1/m(T))}$$

Box 7.1. Elasto-viscoplastic material model.

8. Thermomechanical coupling

For some classes of forming problems consideration of the coupling between mechanical and thermal phenomena is essential for realistic simulation, since temperature changes can induce thermal stresses and change the material properties for mechanical analysis, and mechanical deformations can modify the thermal boundary conditions and generate heat by frictional sliding or dissipation of plastic work.

Following similar arguments as in the case for mechanical equilibrium (3.5)–(3.7) the incremental form of the heat balance can be expressed as

$$\int_{\varphi_{n+1}(\mathcal{B})} (\mathbf{q}_{kn+1} \cdot \nabla \eta + \rho c \dot{T}_{n+1} \eta - (q_i + q_t) \eta) \, dv - \int_{\varphi_{n+1}(\Gamma_q)} \mathbf{q}_N \cdot \mathbf{n} \eta \, da = 0 \quad \forall \eta \in \mathcal{V}_T \quad (8.1)$$

where c is the specific heat, T is the temperature, q_i and q_t are respectively the thermal effect of inelastic deformations and a heat source other than that caused by any mechanical effect, q_N represents the heat transfer at the boundary, q_k is the heat flux governed by Fourier's law of conductivity, $q_k = -k\nabla T$, in which k is the thermal conductivity tensor, and η is the admissible temperature field.

Finite-element discretization results in the system of semidiscrete equations

$$C_T \dot{T}_{n+1} + K_{n+1} T_{n+1} = Q_{n+1} \quad (8.2)$$

in which T_{n+1} are the nodal temperatures at the time instant t_{n+1} , C_T and K_{n+1} are, respectively, the heat capacity and heat conduction matrices, while Q_{n+1} is the heat flux vector. The system (8.2) is solved by employing certain numerical integration algorithms (for various choices we refer to Belytschko (1983), Hughes (1987) and Bathe (1996)) and is normally much simpler than the corresponding mechanical problem.

When the mechanical and thermal problems are coupled, a staggered solution approach is commonly adopted in which separate analyses are undertaken for each phenomenon with data exchange performed at the end of each timestep or increment. In particular, the nodal temperatures are transferred to the mechanical analysis, while the displacements, plastic work, frictional heat flux and contact data are communicated to the thermal solution. On the simplest level, identical meshes can be employed for analysis of the two phenomena. However, when mesh adaption procedures are to be introduced for both phenomena, different meshes will result, requiring data mapping between the two solution phases. The strategy described in section 9.3 for adaptive mesh refinement can also be employed in this context. Details of the formulation and numerical analysis of thermomechanical coupled processes may be found in a comprehensive review by Simo and Miehe (1992).

Example 8.1 (High-speed machining). This problem involves the removal of a continuous chip by a cutting tool and modelling requires, in addition to finite-strain elasto-plasticity and frictional contact representation, the introduction of an appropriate material separation criterion. Realistic modelling of this class of problem necessitates the introduction of thermomechanical coupling effects, since the generation of heat by, principally, frictional sliding between the chip and cutting tool and also the dissipation of plastic work can result in large temperature increases in the tool. Due to complex and interacting features of the problem, which, for realistic simulations, also necessitate the use of adaptive strategies (described later in section 9), some successes in computational modelling of general high-speed machining processes have only recently been accomplished (Marusich and Ortiz 1995, Owen and Vaz 1998, Sekhon and Chenot 1993, Vaz 1998).

Titanium alloys are particularly susceptible to adiabatic shear localization due to their very low thermal diffusivity and associated high sensitivity of the yield stress to temperature. High-speed machining, therefore, amplifies these effects due to the high rate of energy generation. The present example simulates machining of Ti-6Al-4V titanium alloy, for which the following aspects are addressed:

- (i) evaluation of the mesh refinement procedure;
- (ii) evaluation of the strain localization process;
- (iii) evaluation of the failure process.

The geometry of the problem and initial mesh for a rake angle of -3° are depicted in figure 18 whereas material data and other simulation parameters are presented in table 1. The yield stress adopted in the simulations is

$$\sigma_Y = \sigma_{Y0}(\varepsilon_0 + \varepsilon_p)^n + \frac{\partial \sigma_Y}{\partial T}(T - T_0) \quad (8.3)$$

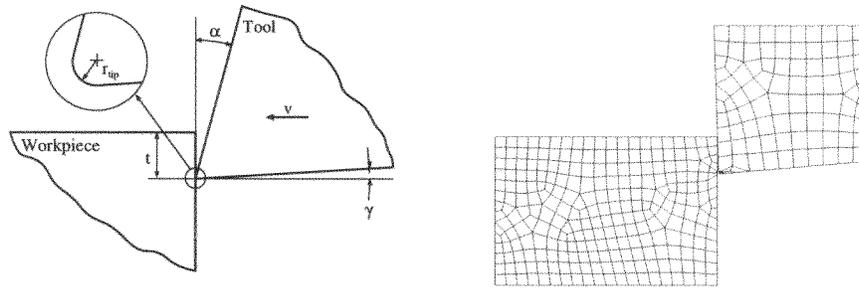


Figure 18. High-speed machining: Geometry and initial finite-element mesh.

Table 1. Material data for Ti-6Al-4V and other simulation parameters.

| Description | Symbol | Value |
|----------------------|--|---|
| Specific mass | ρ | 4420 kg m ⁻³ |
| Specific heat | c | 582.2 J kg ⁻¹ K ⁻¹ |
| Thermal conductivity | k | 5.86 J kg ⁻¹ m ⁻¹ K ⁻¹ |
| Young's modulus | E | 115.7 GPa |
| Poisson's ratio | ν | 0.321 |
| Yield stress data | σ_{Y0} | 1231.5 MPa |
| | ε_0 | 0.008 |
| | T_0 | 298 K |
| | n | 0.059 |
| | $\frac{\partial \sigma_Y}{\partial T}$ | -2.3 MPa K ⁻¹ |
| Damage data | s | 1 |
| Dissipation factor | ξ | 0.85 |
| Coupling interval | | Every timestep |
| Target error | η_{I_ϕ} | 8% |
| Maximum element size | h_{\max} | 0.1 mm |
| Minimum element size | h_{\min} | 0.007 mm |
| Cutting speed | v | 4–20 m s ⁻¹ |
| Cutting depth | t | 0.5 mm |
| Rake angle | α | -9° to +9° |
| Flank angle | γ | 5° |
| Coulomb friction | μ | 0.1 |
| Tool tip radius | r_{tip} | 25 μm |

in which the component $\partial \sigma_Y / \partial T$ represents the thermal softening and T and T_0 correspond to the current and initial temperature respectively. An enhanced four-node one-Gauss-point element is used in the simulations (Belytschko and Bindeman 1991).

The simulations employ the error indicator based on the *uncoupled integration of Lemaitre's damage model* and are performed for cutting speeds between 5 and 20 m s⁻¹ and rake angles between -9° and 9°. Separate tests are undertaken to evaluate the effect of the cutting speed and effect of the rake angle and to assess the capacity of the adaptive remeshing procedure to describe the process evaluation.

The material properties of titanium alloys create favourable conditions for strain localization at cutting speeds as low as 0.0051 m s⁻¹ (Recht 1964). Therefore, an efficient

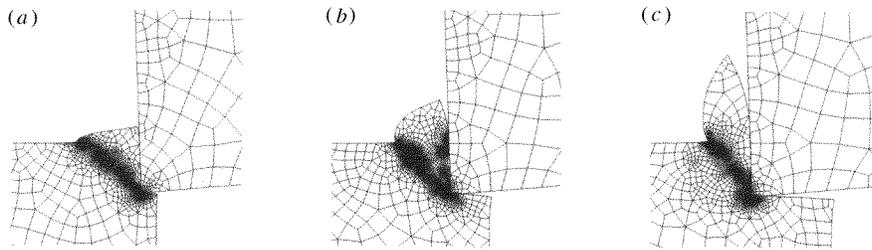


Figure 19. High-speed machining: typical finite-element meshes for negative rake angles. Cutting speed $v = 10 \text{ m s}^{-1}$ and rake angle $\alpha = -3^\circ$. (a) $U = 0.18 \text{ mm}$, (b) $U = 0.39 \text{ mm}$, (c) $U = 0.75 \text{ mm}$.

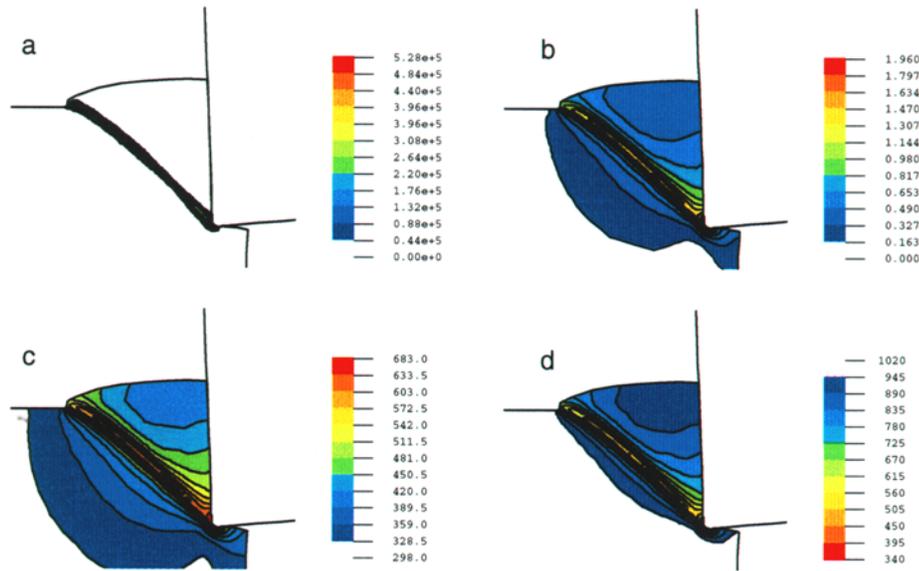


Figure 20. High-speed machining: thermoplastic strain localization for $\alpha = -3^\circ$, $v = 10 \text{ m s}^{-1}$ and $t = 0.5 \text{ mm}$. (a) Equivalent plastic strain rate, (b) equivalent plastic strain, (c) temperature, K, (d) yield stress, MPa.

error indicator should be able to translate the strain localization phenomena into refined meshes over the critical zones for a wide range of cutting conditions.

The error indicator based on the rate of fracture indicators has been employed (Vaz 1998, Perić *et al* 1998) based on the principle that mesh refinement should be able not only to capture the progression of the plastic deformation but also to efficiently produce refined meshes at regions of possible material failure. This concept is clearly illustrated in figures 19(a)–(c) which show the evolution of the mesh refinement process. It is worthy of note that, in this example, material failure is inhibited, otherwise the chip would have separated before reaching the stage shown in figure 19(c). It is also worth noting that the small tool tip radius requires a refined mesh to ensure a proper contact between the cutting edge of the tool and workpiece.

In the present case, localization was found to be confined, on average, to a $35 \mu\text{m}$ zone, corresponding to a region approximately five-elements wide, as depicted in figure 20(a).

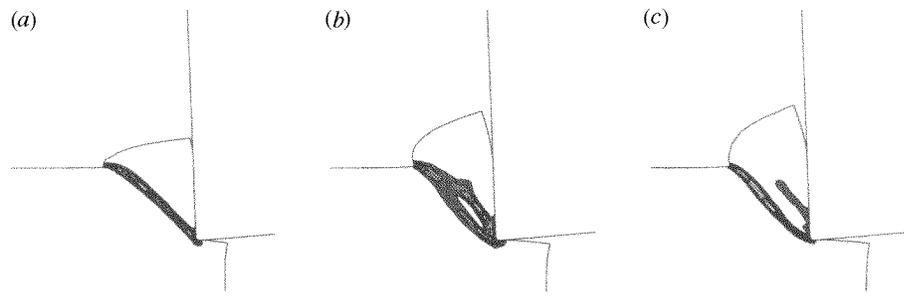


Figure 21. High-speed machining: progression of the shear band for a cutting speed of 10 m s^{-1} , a cutting depth of 0.5 mm and a rake angle of -3° . (a) $U = 0.23 \text{ mm}$, (b) $U = 0.37 \text{ mm}$, (c) $U = 0.40 \text{ mm}$.

The distribution pattern of the equivalent plastic strain rate can change during the process due to the cyclical character of the mechanisms of chip formation. However, even at very early stages, $\dot{\epsilon}_p$ was found to be significantly high, reaching values up to $1.2 \times 10^6 \text{ s}^{-1}$ for this particular example. A high strain rate causes localized plastic deformation which is reflected by the distribution of the equivalent plastic strain shown in figure 20(b). The shear band geometry in high-speed machining is not a straight line and, therefore, its characterization based on the *shear angle* is not particularly appropriate. On the other hand, the simulations show that initial inclination of the shear band increases with the rake angle and tool advance.

The very essence of thermoplastic shear localization is presented in figures 20(c) and (d), which show the distribution of the temperature and yield stress. Highly localized plastic deformation causes the temperature to rise sharply, well above the neighbouring regions, which in turn, causes the yield stress to decrease. Furthermore, the regions of low yield stress are more susceptible to larger plastic deformations. The cyclical response of *localized plastic deformation* \rightarrow *temperature rise* \rightarrow *yield stress reduction* \rightarrow *large and localized deformation* instigates a significant upward movement of the portion of the chip above the shear band, which can eventually lead to its separation.

Analysis of the progression of the strain localization process provides useful information on *when* a fractured chip is formed. Therefore, material failure is inhibited by de-activating the algorithm responsible for the failure process. Figures 21(a)–(c) show the shear band progression as a function of the tool advance, U , for a cutting speed of 10 m s^{-1} and a rake angle of -3° . When the tool tip touches the workpiece a well-defined shear band is rapidly developed, as shown in figure 21(a). A further tool progression gives rise to a second shear band, as shown in figures 21(b) and (c). The importance of the present assessment is highlighted by the fact that the chip breaks *before* the development of a new shear band. For further details related to numerical simulation of high-speed machining, including chip formation and breakage, we refer to Marusich and Ortiz (1995) and Owen and Vaz (1998).

Example 8.2 (A plane-strain bulk forming of a crane hook). The geometry and material characteristics for this example are shown in figure 22. The analysis is performed employing a plane-strain formulation for one half of the model with appropriate boundary conditions. The material is assumed to follow the Sellars–Tegart interpolation function (7.9) with material parameters given as: Young's modulus $E = 120\,000 \text{ N mm}^{-2}$, Poisson's ratio $\nu = 0.3$, $C = 1.2 \times 10^{12} \text{ s}^{-1}$, $\alpha = 1.12 \times 10^{-2} \text{ mm}^2 \text{ N}^{-1}$, $n = 3.5$ and $Q = 352.69 \text{ kJ mol}^{-1}$. The initial temperature of the workpiece is set at $T = 1100^\circ\text{C}$. Coulomb frictional law is assumed with a coefficient of friction 0.3. To solve this problem the bulk material is

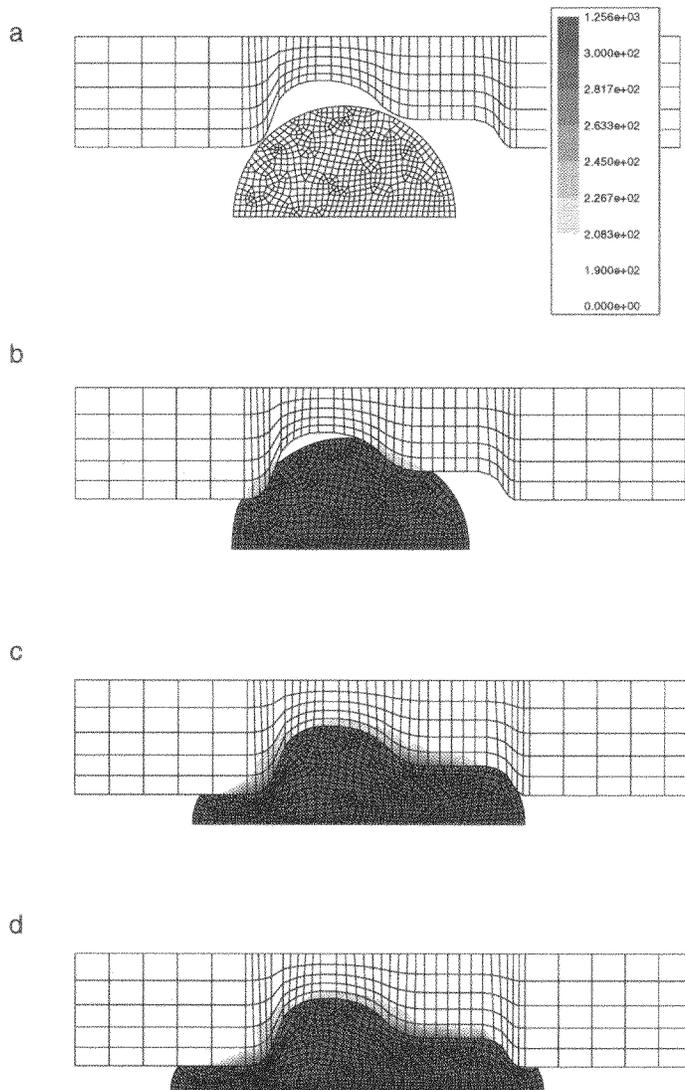


Figure 22. A plane strain bulk forming of a crane hook. (a) Initial finite-element meshes for tool and workpiece. (b)–(d) Deformed finite-element meshes with temperature distributions at various stages of punch displacement.

discretized with 600 quadrilateral finite elements while the punch is assumed rigid and described as a sequence of linear and circular segments. Figure 22(b)–(d) gives deformed meshes and temperature development in the tool over one forging cycle, which is primarily due to frictional sliding between workpiece and tool. It is seen that the areas of greatest temperature increase coincide with regions of high curvature of the tool profile, which results in large normal pressures and consequently relatively large amounts of frictional work dissipated as heat.

This may provide valuable information to industrial tool designers since high-temperature (low-hardness) and high-normal-pressure regions often coincide with regions

of high wear. A statistical investigation covering more than 100 forging geometries has categorized the principal causes of tool failure. Different types of wear (mainly at corners and roundings) is the cause for scrapping of some 60% of tools, various types of crack formation account for approximately 25% and local plastic deformation is responsible for 5%. Consequently, the prediction of tool wear is of primary concern for the implementation of a preventive maintenance strategy within the forging industry.

9. Adaptive strategies for nonlinear problems

The formal structure of adaptive finite-element methods for linear elliptic problems is now well understood, thus forming a solid foundation upon which effective and reliable techniques of error estimation and adaptive refinement may be established. Although certain issues still remain unresolved, it may be said that nowadays, adaptive strategies for linear problems can be routinely performed within finite-element computations. Among numerous contributions, the work by Babuška and Rheinboldt (1978), Bank and Weiser (1985), Demkowicz *et al* (1985), Eriksson and Johnson (1988), Zienkiewicz and Zhu (1987), and surveys by Johnson and Hansbo (1992), Oden *et al* (1989) and Zienkiewicz and Taylor (1989) illustrate the basic ideas and numerical strategies. In contrast, although some advances have been recorded for certain classes of nonlinear problems (see e.g. Babuška *et al* (1986) for some early contributions), only a limited amount of published work exists on a *posteriori* error estimates and adaptive approaches for history dependent nonlinear problems in solid mechanics. Notable exceptions are contributions by Ladeveze *et al* (1986), Jin *et al* (1989), Belytschko *et al* (1989), Ortiz and Quigley (1991), Johnson and Hansbo (1992), Lee and Bathe (1994), Perić *et al* (1994, 1996), Gallimard *et al* (1996) and Barthold *et al* (1997).

On the practical side, since for a large number of industrially relevant solid mechanics problems, the optimal mesh configuration changes continually throughout the deformation process, the introduction of adaptive mesh refinement processes is crucial for the solution of large-scale industrial problems. This necessitates several steps within an adaptive strategy: (i) a remeshing criterion, (ii) specification of an appropriate error estimation criterion, (iii) development of a strategy for adapting the mesh based on the error distribution and (iv) automatic mesh generation tools.

In this section, some basic concepts already in standard usage are generalized in order to describe an adaptive strategy for the elasto-plastic problem of evolution. In particular, a *posteriori* error estimates based on the Zienkiewicz–Zhu adaptive strategy and the energy norm (see Zienkiewicz and Zhu (1987), Zienkiewicz and Taylor (1989), Zienkiewicz *et al* (1990) and references therein) are appropriately modified to account for the elasto-plastic deformation at small and finite strains. We emphasize that, in this context, the generalized energy norm appears as a natural metric for this class of problems. In addition, the intrinsic dissipation functional, associated with the second law of thermodynamics is, by heuristic arguments, utilized as a basis for the development of a complementary *a posteriori* error indicator.

For most of the early practical simulations involving history dependent materials on evolving meshes work has been concentrated on adopting a suitable criterion to control the remeshing procedure. An adaptive remeshing scheme has been proposed by Ortiz and Quigley (1991), in which the resolution of the mesh is increased in certain regions so that the variation of the solution over each element is within a prescribed tolerance throughout the mesh. A similar procedure is adopted by Jin *et al* (1989) who use isolines of effective strain as a measure in their adaptive remeshing scheme. In each of the above schemes a

suitable transfer operator must be defined to ensure that the state variables are correctly distributed to the new mesh. Recently, Lee and Bathe (1994) and Perić *et al* (1994, 1996) have presented an adaptive strategy for large deformation processes at large inelastic strains which provides a discussion on important aspects of the adaptive procedure for history dependent materials on evolving finite-element meshes. Their discussion analyses various aspects of error indicators, remeshing algorithms and transfer operations, and some of these aspects are discussed in this section.

9.1. Basic error estimates and adaptive refinement

9.1.1. *Energy norm.* It is convenient, at this stage, to introduce a unified notation associated with Halphen and Nguyen (1975), which allows an elegant presentation of the elasto-plastic problem of evolution. For that purpose denote by

$$\boldsymbol{\Sigma} := (\boldsymbol{\sigma}, \mathbf{A}) \quad (9.1a)$$

$$\mathbf{E} := (\boldsymbol{\varepsilon}, \mathbf{0}) \quad (9.1b)$$

$$\boldsymbol{\xi} := (\boldsymbol{\varepsilon}^p, \boldsymbol{\alpha}) \quad (9.1c)$$

and, in addition, introduce the unified modulus \mathbb{G} as

$$\mathbb{G} := \text{diag}[\mathbf{C}, \mathbf{D}] \quad (9.2)$$

where \mathbf{C} denotes the elasticity tensor and \mathbf{D} is the $\mathbb{R}^{n_{\text{int}}} \times \mathbb{R}^{n_{\text{int}}}$ matrix of the generalized hardening modulus where n_{int} is a number of internal variables. Denoting by $\boldsymbol{\Sigma}^h$ the approximate generalized stresses obtained as the finite-element solution, the corresponding error may be defined as

$$\mathbf{e}_{\boldsymbol{\Sigma}} := \boldsymbol{\Sigma} - \boldsymbol{\Sigma}^h. \quad (9.3)$$

Following standard usage for linear elliptic problems we introduce the generalized stress error in the energy norm for a generic element K , and a corresponding global error, respectively, as

$$\|\|\| \mathbf{e}_{\boldsymbol{\Sigma}} \|\|\|_K^2 = \int_K (\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^h) : \mathbb{G}^{-1} (\boldsymbol{\Sigma} - \boldsymbol{\Sigma}^h) \, dv \quad (9.4a)$$

$$\|\|\| \mathbf{e}_{\boldsymbol{\Sigma}} \|\|\|^2 = \sum_K \|\|\| \mathbf{e}_{\boldsymbol{\Sigma}} \|\|\|_K^2. \quad (9.4b)$$

An elementary procedure for the error estimation may be defined by the replacement of the exact values of variables and relevant derivatives of the problem by some post-processed values obtained from the available finite-element solution and the problem data. The post-processed solution is expected to have superior accuracy compared with the original finite-element solution. This characteristic of the post-processed solution is attributed to the so-called *superconvergence* properties, which are at present proved for certain regular meshes. In particular, the *a posteriori* error estimation procedure originally proposed and used by Zienkiewicz and Zhu (1987) for linear elliptic problems is based on the observation that exact stresses $\boldsymbol{\sigma}$ may be represented accurately by smoothed stresses $\boldsymbol{\sigma}^*$ obtained by a suitable projection of approximate stresses $\boldsymbol{\sigma}^h$ which satisfies

$$\int_B \boldsymbol{\Pi} (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}^h) \, dv = 0. \quad (9.5)$$

Some possible choices for the projection matrix $\boldsymbol{\Pi}$ are listed in Zienkiewicz and Taylor (1989).

Table 2. Error estimation procedure—energy norm error.

(i) Energy functional

$$\psi(\Sigma) = \frac{1}{2} \int_B \Sigma : \mathbb{G}^{-1} \Sigma \, dv = |||\Sigma|||^2.$$

(ii) Stress error in energy norm

$$|||e_\Sigma|||_K^2 = \int_K (\Sigma - \Sigma^h) : \mathbb{G}^{-1} (\Sigma - \Sigma^h) \, dv = |||\Sigma - \Sigma^h|||_K^2$$

$$|||e_\Sigma|||^2 = \sum_K |||e_\Sigma|||_K^2.$$

(iii) Stress projection

$$\int_B \Pi(\Sigma^* - \Sigma^h) \, dv = 0.$$

(iv) Error estimate

$$\begin{aligned} \varepsilon_{\Sigma,K}^2 &= \int_K (\Sigma^* - \Sigma^h)^T : \mathbb{G}^{-1} (\Sigma^* - \Sigma^h) \, dv \\ &= |||\Sigma^* - \Sigma^h|||_K^2 \, dv \approx |||e_\Sigma|||_K^2 \end{aligned}$$

$$|||\varepsilon_\Sigma|||^2 = \sum_K |||\varepsilon_\Sigma|||_K^2.$$

(v) Relative error

$$\eta_\Sigma = \frac{|||e_\Sigma|||}{\psi^{1/2}} \approx \frac{\varepsilon_\Sigma}{(\psi^h)^{1/2}}.$$

Extending these ideas to the elasto-plastic problem of evolution and generalized stress Σ the error estimator for an element K and its global counterparts, may be obtained as

$$\varepsilon_{\Sigma,K}^2 := \int_K (\Sigma^* - \Sigma^h) : \mathbb{G}^{-1} (\Sigma^* - \Sigma^h) \, dv = |||\Sigma^* - \Sigma^h|||_K^2 \approx |||e_\Sigma|||_K^2 \quad (9.6a)$$

$$\varepsilon_\Sigma^2 = \sum_K \varepsilon_{\Sigma,K}^2. \quad (9.6b)$$

In addition, the *relative energy norm error* is defined as

$$\eta_\Sigma := \frac{|||e_\Sigma|||}{\psi^{1/2}} \approx \frac{\varepsilon_\Sigma}{(\psi^h)^{1/2}}. \quad (9.7)$$

The above approximation is expected to hold asymptotically, i.e. for sufficiently small mesh size h .

For convenience, basic steps of the error estimation procedure based on the energy norm error are summarized in table 2.

9.1.2. Error indicator based on the dissipation functional. Recalling the expressions for the plastic dissipation (see Perić *et al* 1994), and setting $e_{\dot{\xi}} := \dot{\xi} - \dot{\xi}^h$ we observe that

$$|\langle e_\Sigma, e_{\dot{\Sigma}} \rangle| \leq \int_K \sum_{i,j=1}^{n_{\text{dim}}} |(\Sigma_{ij} - \Sigma_{ij}^h)(\dot{\xi}_{ij} - \dot{\xi}_{ij}^h)| \, dv. \quad (9.8)$$

Then, in the spirit of (9.1)–(9.5), the error based on the plastic dissipation and the associated *a posteriori* error indicator for an element K may be defined, respectively, as

$$|e_{\mathcal{D}}|_K^2 := \int_K \sum_{i,j=1}^{n_{\text{dim}}} |(\Sigma_{ij} - \Sigma_{ij}^h)(\dot{\xi}_{ij} - \dot{\xi}_{ij}^h)| \, dv \quad (9.9a)$$

$$\varepsilon_{\mathcal{D},K}^2 := \int_K \sum_{i,j=1}^{n_{\text{dim}}} |(\Sigma_{ij}^* - \Sigma_{ij}^h)(\dot{\xi}_{ij}^* - \dot{\xi}_{ij}^h)| \, dv \quad (9.9b)$$

where $|\cdot|$ is the standard Euclidean norm, and an adequate extension of the corresponding notation (9.1)–(9.5) is applied. The corresponding global quantities are obtained, in a standard way, as $|e_{\mathcal{D}}|^2 = \sum_K |e_{\mathcal{D}}|_K^2$ and $|\varepsilon_{\mathcal{D}}|^2 = \sum_K |\varepsilon_{\mathcal{D}}|_K^2$, respectively. In addition, the relative error $\eta_{\mathcal{D}}$ is defined as

$$\eta_{\mathcal{D}} = \frac{|e_{\mathcal{D}}|}{|\mathcal{D}|^{1/2}} \approx \frac{\varepsilon_{\mathcal{D}}}{(\mathcal{D}^h)^{1/2}}. \quad (9.10)$$

9.1.3. Error indicator based on the rate of plastic work. It may sometimes be useful to consider the rate of plastic work, as a basis for an error estimation procedure. Then, the error based on the rate of plastic work and the associated *a posteriori* error indicator for a generic element \mathbf{K} , may be expressed as

$$|e_{\mathcal{W}^p}|_K^2 := \int_K \sum_{i,j=1}^{n_{\text{dim}}} |(\sigma_{ij} - \sigma_{ij}^h)[\dot{\varepsilon}_{ij}^p - (\dot{\varepsilon}_{ij}^p)^h]| \, dv \quad (9.11a)$$

$$\varepsilon_{\mathcal{W}^p,K}^2 := \int_K \sum_{i,j=1}^{n_{\text{dim}}} |(\sigma_{ij}^* - \sigma_{ij}^h)[(\dot{\varepsilon}_{ij}^p)^* - (\dot{\varepsilon}_{ij}^p)^h]| \, dv. \quad (9.11b)$$

The corresponding global quantities are again obtained, in a standard way, as $|e_{\mathcal{W}^p}|^2 = \sum_K |e_{\mathcal{W}^p}|_K^2$ and $|\varepsilon_{\mathcal{W}^p}|^2 = \sum_K |\varepsilon_{\mathcal{W}^p}|_K^2$, respectively. In addition, the relative error $\eta_{\mathcal{W}^p}$ is defined as

$$\eta_{\mathcal{W}^p} = \frac{|e_{\mathcal{W}^p}|}{(\mathcal{W}^p)^{1/2}} \approx \frac{\varepsilon_{\mathcal{W}^p}}{(\mathcal{W}^{p,h})^{1/2}}. \quad (9.12)$$

Observe that if, for the present elasto-plastic model the equivalent expression for the rate of plastic work $\mathcal{W}^p = \bar{\sigma} \dot{\varepsilon}^p$ is also included in the error expression (9.9), the error indicator based on the plastic work $|\tilde{e}_{\mathcal{W}^p}|$ would be identical to the one based on the plastic dissipation (9.7b), i.e. $|\tilde{e}_{\mathcal{W}^p}| = |e_{\mathcal{D}}|$. The relative error thus would be given as $\tilde{\eta}_{\mathcal{W}^p} = \frac{|\tilde{e}_{\mathcal{W}^p}|}{(2\mathcal{W}^p)^{1/2}}$.

9.1.4. Refinement strategy. The mesh refinement procedure is constructed in a standard way (see e.g. Zienkiewicz and Zhu 1987, Zienkiewicz and Taylor 1989, Zienkiewicz *et al* 1990) with the objective of achieving a uniform distribution of local error. In addition, the relative error $\eta_{\Sigma}(\eta_{\mathcal{D}})$ is required to be within specified limits. In our numerical examples, the element size is predicted according to the asymptotic rate of convergence criteria of the linear model.

9.2. Mesh generation/adaption techniques

Structured and unstructured meshes (Weatherill 1990) can be used in the discretization of the domain. There exists little doubt that there is a role for both categories of meshes in the finite-element simulation of industrial applications.

Mesh adaption is generally required in the finite-element simulation of industrial problems such as forming operations where the geometry and material characteristics are evolving with deformation in the course of the process. Taking the above constraints into account, the unstructured approach in mesh generation/adaption is employed here. The unstructured mesh generation process implemented is a Delaunay-based two-stage scheme

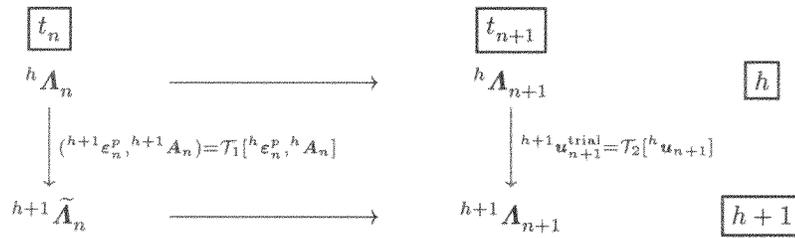


Figure 23. Transfer operator diagram.

(George 1996) whereas the mesh adaption scheme is designed upon the node-based h -adaption approach. The implemented mesh adaption scheme takes advantage of the data structure created in the mesh generation scheme. This mesh adaption scheme is particularly suited for local mesh refinement, which is more economical than the mesh regeneration approach.

The mesh adaption scheme implemented is capable of performing the mesh adaption according to the mesh prediction data. These mesh prediction data, usually in the form of mesh density variation or mesh refinement indices, are interpreted from the error data created by some error indicator, as was discussed in section 9.1.

9.3. Transfer operators

The issue of transfer of variables to new meshes must be properly addressed as it is expected to be critical in finite-element simulations based on the implicit approach. Important aspects which must be considered include:

- (i) consistency with the constitutive equations,
- (ii) requirements of equilibrium,
- (iii) compatibility of the state transfer with the displacement field on the new mesh,
- (iv) compatibility with evolving boundary conditions,
- (v) minimization of the numerical diffusion of the state fields.

In this section general aspects of the transfer operation for evolving finite-element meshes are provided for the case of a typical elasto-plastic material whose behaviour is described by a set of internal variables. First, some concepts related to the mapping of internal variables between two finite-element meshes, denoted by h and $h + 1$, are given. This mapping is the *essential* part of the adaptive strategy that is employed in simulation of history dependent material processes on evolving general unstructured meshes. The mapping of internal variables is formally denoted as the transfer operator \mathcal{T}_1 . In addition, a transfer operator is employed that transfers the displacement field from the old to a new mesh. In the context of the backward Euler scheme where solution is sought at time instant t_{n+1} , this transfer provides a trial solution. The mapping of displacements is formally denoted as the transfer operator \mathcal{T}_2 .

To describe the transfer operation, let us define a state array ${}^h\mathbf{\Lambda}_n := ({}^h\mathbf{u}_n, {}^h\boldsymbol{\varepsilon}_n, {}^h\boldsymbol{\varepsilon}_n^p, {}^h\boldsymbol{\sigma}_n, {}^h\mathbf{A}_n)$ where ${}^h\mathbf{u}_n, {}^h\boldsymbol{\varepsilon}_n, {}^h\boldsymbol{\varepsilon}_n^p, {}^h\boldsymbol{\sigma}_n, {}^h\mathbf{A}_n$ denote values of the displacement, strain tensor, plastic strain tensor, stress tensor and a vector of internal variables at time t_n for the mesh h (see figure 23). Assume, furthermore, that the estimated error of the solution ${}^h\boldsymbol{\Lambda}_p$ respects the prescribed criteria, while these are violated by the solution ${}^h\boldsymbol{\Lambda}_{n+1}$. In this case a new mesh $h + 1$ is generated and a new solution ${}^{h+1}\boldsymbol{\Lambda}_{n+1}$ needs to be computed.

As the backward Euler scheme is adopted, the plastic strain ${}^{h+1}\boldsymbol{\varepsilon}_n^p$ and the internal variables ${}^{h+1}\mathbf{A}_n$ for a new mesh $h + 1$ at time t_n need to be evaluated. In this way the state ${}^{h+1}\tilde{\mathbf{A}}_n = (\bullet, \bullet, {}^{h+1}\boldsymbol{\varepsilon}_n^p, \bullet, {}^{h+1}\mathbf{A}_n)$ is constructed, where a tilde is used to denote a reduced state array. It should be noted that this state characterizes the history of the material and, in the case of a fully implicit scheme, provides sufficient information for computation of a new solution ${}^{h+1}\mathbf{A}_{n+1}$.

Figure 23 summarizes, on a conceptual level, a typical transfer operation that includes both the mapping of the internal variables and mapping of the displacement field.

9.3.1. Mapping of internal variables—transfer operator \mathcal{T}_1 . Let \mathcal{T}_1 be the *transfer operator* between meshes h and $h + 1$ defined by

$$({}^{h+1}\boldsymbol{\varepsilon}_n^p, {}^{h+1}\mathbf{A}_n) = \mathcal{T}_1[{}^h\boldsymbol{\varepsilon}_n^p, {}^h\mathbf{A}_n]. \quad (9.13)$$

The variables $({}^h\boldsymbol{\varepsilon}_n^p, {}^h\mathbf{A}_n)$ specified at quadrature points of the mesh h , are transferred by the operator \mathcal{T}_1 to every point of the body \mathcal{B} , in order to specify the variables $({}^{h+1}\boldsymbol{\varepsilon}_n^p, {}^{h+1}\mathbf{A}_n)$ at the quadrature points of the new mesh $h + 1$. The operator \mathcal{T}_1 can be constructed in different ways.

(i) A simple version of the transfer operator may be constructed by taking constant values over the area associated with every quadrature point. Note that this construction is local.

(ii) Another possibility is to construct a solution which is continuous, for instance by a least-squares method, or by a suitable projection of ${}^h\boldsymbol{\varepsilon}_n^p$ which satisfies

$$\int_{\mathcal{B}} \boldsymbol{\Pi}[({}^h\boldsymbol{\varepsilon}_n^{p*}, {}^h\mathbf{A}_n^*) - ({}^h\boldsymbol{\varepsilon}_n^p, {}^h\mathbf{A}_n)] dv = 0 \quad (9.14)$$

where $\boldsymbol{\Pi}$ is the so-called projection matrix. Some possible choices for the projection matrix are listed in Zienkiewicz and Taylor (1989). This type of transfer operator can be global or local.

A local transfer of the second type (ii) is described here. Within this approach there exist various possibilities for the transfer. The choice for transfer operators that has been made was guided by the future application of the methodology to classes of problems where large deformations of inelastic materials at finite strains and complex boundary conditions with possible frictional contact will be standard operating conditions. Therefore *simple*, but *generally applicable*, transfer operators are adopted. The basic steps of the *implementation procedure* which is applicable for any type of finite element may be described as follows.

The continuous plastic strain tensor ${}^h\boldsymbol{\varepsilon}_n^{p*}$ and the internal variable vector ${}^h\mathbf{A}_n^*$ are obtained by projecting the Gauss-point (G.P.) components ${}^h\boldsymbol{\varepsilon}_{n,G}^p$ and ${}^h\mathbf{A}_{n,G}$ to the nodal points (N.P.) thus obtaining ${}^h\boldsymbol{\varepsilon}_{n,N}^p$ and ${}^h\mathbf{A}_{n,N}$. To project the Gauss-point components to the nodes, the shape functions of the element are used. The nodal-point averages of the projected values are then performed resulting in ${}^h\boldsymbol{\varepsilon}_{n,N}^{p*}$ and ${}^h\mathbf{A}_{n,N}^*$.

The nodal components of the plastic strain tensor ${}^h\boldsymbol{\varepsilon}_{n,N}^{p*}$ and the internal variable vector ${}^h\mathbf{A}_{n,N}^*$ for the mesh h are then transferred to the nodes of the new mesh $h + 1$ resulting in components ${}^{h+1}\boldsymbol{\varepsilon}_{n,N}^{p*}$ and the internal variable vector ${}^{h+1}\mathbf{A}_{n,N}^*$. The components at the Gauss points of the new mesh $h + 1$, ${}^{h+1}\boldsymbol{\varepsilon}_{n,G}^p$ and ${}^{h+1}\mathbf{A}_{n,G}$, are finally obtained by interpolation using the shape functions of the finite elements.

The transfer operation \mathcal{T}_1 is schematically presented in figure 24.

9.3.2. Mapping of the displacement field—transfer operator \mathcal{T}_2 . In the context of the backward Euler scheme with a *strain driven* format and Newton–Raphson iterative

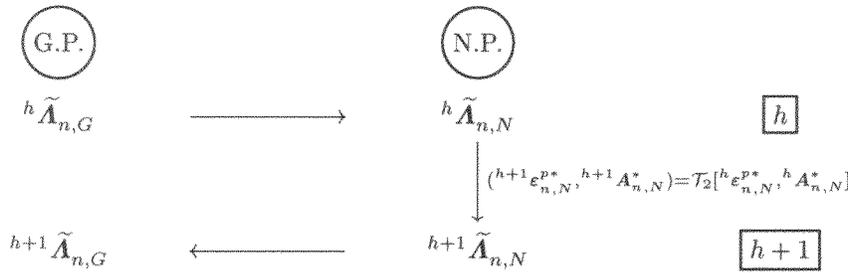


Figure 24. Schematic representation of the transfer operator $({}^{h+1}\epsilon_{n,G}^p, {}^{h+1}A_{n,G}) = \mathcal{T}_1[{}^h\epsilon_{n,G}^p, {}^hA_{n,G}]$.

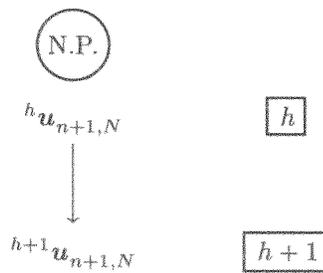


Figure 25. Schematic representation of the transfer operator ${}^{h+1}u_{n+1,N} = \mathcal{T}_2[{}^h u_{n+1,N}]$.

procedure, the transferred displacement field on a new mesh is used to provide an initial guess (trial solution) for the displacements in the first iteration of the Newton–Raphson scheme. Hence the trial displacement field is prescribed as

$${}^{h+1}u_{n+1}^{\text{trial}} = \mathcal{T}_2[{}^h u_{n+1}] \tag{9.15}$$

i.e. by the transfer of the displacement field ${}^h u_{n+1}$ obtained with the mesh h at time t_{n+1} . It is important to observe that the transfer operator \mathcal{T}_2 can be easily constructed since the nodal values of displacements and finite-element shape functions fully prescribe the displacement field.

For convenience, a schematic diagram of the transfer operator \mathcal{T}_2 is presented in figure 25.

Remark 9.1. Details of implementation of the general transfer operation given in sections 9.3.1 and 9.3.2 are described for the case of evolving finite-element meshes composed of constant strain triangles in Perić *et al* (1996). With minor modifications this procedure is applicable to other types of finite elements. For instance, in the case of a mesh composed of quadrilateral or higher order elements we proceed by simply subdividing these elements into a local mesh of constant strain triangles. Since, at present, our aim is to employ general unstructured meshes composed of low-order elements this procedure is considered as an appropriate choice. In a recently published article Lee and Bathe (1994) described an adaptive strategy based on quadrilateral element meshes that may be composed of higher order elements. The evaluation of local coordinates in element ${}^h\mathcal{B}^{(e)}$ of the background (old) mesh h corresponding to each node of the new mesh $h + 1$, for quadrilaterals and higher order finite elements results in a nonlinear problem. Lee and Bathe have solved this problem by employing the so-called ‘inverse isoparametric mapping technique’ which is based on a Newton–Raphson-type iterative procedure.

Remark 9.2. An adaptive mesh procedure for nonlinear problems has also been described by Belytschko *et al* (1989). The applications provided in this work relate to the explicit dynamic analysis of shells. In essence, their strategy consists ‘in splitting a single quadrilateral into four equal-sized smaller elements (fission), or four elements may be derefined or combined into a single element (fusion)’. In this way material history diffusion is minimized and transfer of the state variables is not a serious issue. In contrast, the transfer operation procedure described here is applicable to problems involving transfer between general *unstructured* meshes. This is of particular importance since our aim is to apply this procedure to the modelling of complex large deformation processes at large strains (using both implicit and explicit algorithms), where a complete regeneration of the finite-element meshes has often proved *essential* for a successful simulation.

Remark 9.3. In the context of the implicit timestepping scheme only internal variables [${}^h \boldsymbol{\varepsilon}_n^p$, ${}^h \mathbf{A}_n$] are transferred, while stresses and strains [${}^h \boldsymbol{\sigma}_n$, ${}^h \boldsymbol{\varepsilon}_n$] are not transferred to a new mesh $h+1$. As discussed earlier in section 9.3.1 (see also Lee and Bathe (1994), section 4.4 for a discussion) the transfer of the complete state array ${}^h \boldsymbol{\Lambda}_n = ({}^h \mathbf{u}_n, {}^h \boldsymbol{\varepsilon}_n, {}^h \boldsymbol{\varepsilon}_n^p, {}^h \boldsymbol{\sigma}_n, {}^h \mathbf{A}_n)$ would result in data that may not be self-consistent. For instance, the yield condition $\phi({}^{h+1} \boldsymbol{\sigma}_n, {}^{h+1} \mathbf{A}_n)$ may not be satisfied.

Remark 9.4. As has been observed earlier, the set of internal variables (${}^{h+1} \boldsymbol{\varepsilon}_n^p$, ${}^{h+1} \mathbf{A}_n$) prescribed at each quadrature point characterizes in full the history of the material on the new mesh $h+1$. In the case of a fully implicit strain driven computational scheme in which ${}^{h+1} \mathbf{u}_{n+1}$ (and ${}^{h+1} \boldsymbol{\varepsilon}_{n+1}$) are *given*, these internal variables provide sufficient information for computation of a new solution ${}^{h+1} \boldsymbol{\Lambda}_{n+1} = ({}^{h+1} \mathbf{u}_{n+1}, {}^{h+1} \boldsymbol{\varepsilon}_{n+1}, {}^{h+1} \boldsymbol{\varepsilon}_n^p, {}^{h+1} \boldsymbol{\sigma}_{n+1}, {}^{h+1} \mathbf{A}_{n+1})$. This may clearly be observed from box 4.3 in which the stress return procedure is presented.

Example 9.1 (Extension of a double-notched rubber sheet). In this example, the finite-element simulation of the plane-stress stretching of a notched rubber sheet is presented. The sheet, with initial mesh shown in figure 26(a) is assumed to be made of a *Mooney–Rivlin material* with constants $C_1 = 80.1938$ and $C_2 = 2.0$. The length of the sheet is $2L = 30.0$, width is $2B = 10.0$, and thickness is $t = 1.0$. Due to symmetry, only one quarter of the sheet is considered. As an error indicator the classical Zienkiewicz–Zhu L_2 stress norm is used. The maximum size of elements is restricted to $h_{\max} = 5.0$ and the minimum size is $h_{\min} = 0.3$.

A total of five increments is used to apply displacement $U = 3L$ and the convergence tolerance control is 1.0×10^{-6} on residual force and displacement norms.

Example 9.2 (Stretching of a strip with a hole). In this example, the finite-element simulation of a problem is presented with a localized shear band formation. The specimen is assumed to be made of an elasto-perfectly plastic material with Young’s modulus $E = 100$ GPa, Poisson’s ratio $\nu = 0.3$ and yield stress $\sigma = 100$ MPa. Due to symmetry, only one quarter of the strip is considered. In analysis an error indicator based on generalized energy norm (9.4) is used and the maximum size of elements is restricted to $h_{\max} = 0.7$ and the minimum size is $h_{\min} = 0.15$.

Meshes obtained after adaptive remeshing are shown in figures 27(b)–(d). Also the distribution of effective plastic strain is depicted in figure 27(d).

Example 9.3 (Axisymmetric piercing). The finite-element simulation of the axisymmetric piercing of a cylindrical workpiece is presented. The geometry of the problem is shown in figure 28(a) and the initial mesh in figure 28(b). The workpiece is assumed to be made of an elasto-plastic material with Young’s modulus $E = 210$ GPa, Poisson’s ratio $\nu = 0.3$,

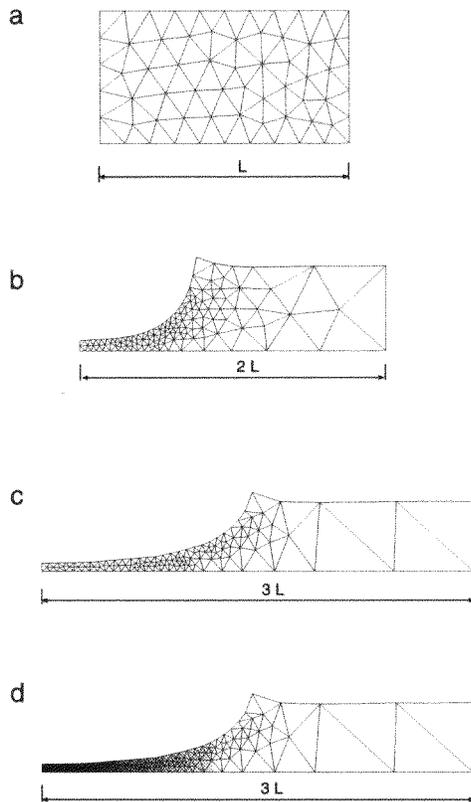


Figure 26. Extension of a double-notched rubber sheet: finite-element meshes. (a) Initial mesh, (b) $U_1 = 2L$, (c) $U_1 = 3L$, (d) distribution of effective stress for $U_1 = 3L$.

yield stress $\sigma_Y = 100$ MPa and linear hardening with hardening modulus $H = 900$ MPa, while the punch is assumed to be rigid. Frictional contact between workpiece and tool is defined by a Coulomb law with coefficient of friction $\mu = 0.1$.

In analysis an error indicator based on the rate of plastic work is used. The initial mesh consists of 101 quadrilateral elements and the final mesh contains 426 elements. Convergence of the finite-element solution is established on the basis of the standard Euclidean norm of the out-of-balance forces with a tolerance of 10^{-3} . No difficulties related to the convergence have been observed during the simulation despite frequent remeshings.

Distribution of effective plastic strain on deformed meshes at various stages of the process is shown in figure 29. The deformed meshes show no hourglassing patterns, which is in agreement with analyses of a similar class of problem carried out by de Souza Neto *et al* (1996b).

Example 9.4 (Closed-die rail forging). A simulation of plane-strain forging is presented to illustrate the application of the developed adaptive strategy to thermomechanically coupled problems. An *adiabatic* thermomechanical algorithm is adopted in the simulations, in which heat generation due to dissipation of plastic and frictional work is accounted for. An unstructured meshing approach based on Delaunay triangulation is adopted for both initial and subsequent mesh adaptations. Enhanced four-node one-Gauss-point elements are also employed to control hourglassing (Belytschko and Bindeman 1991).

The geometry of the problem is shown in figure 30 whereas the material data for mild steel and other simulation parameters are summarized in table 3.

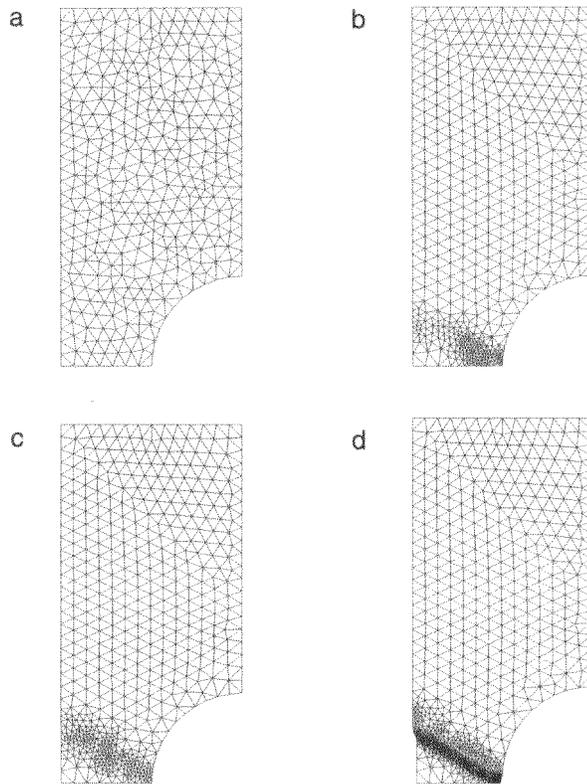


Figure 27. Stretching of an elastic-plastic strip with a hole. (a) Initial mesh, (b) $U_2 = 0.04$, (c) $U_1 = 0.06$, (d) $U_1 = 0.415$.

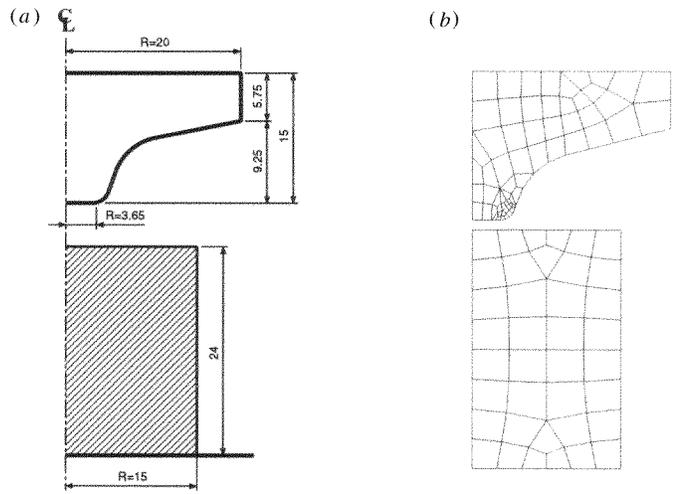


Figure 28. Axisymmetric piercing. (a) Geometry and (b) initial finite-element mesh.

The simulation is performed using a J_2 elasto-perfectly plastic material for the workpiece, rigid tools and a constant friction coefficient. The initial temperature is 300 K

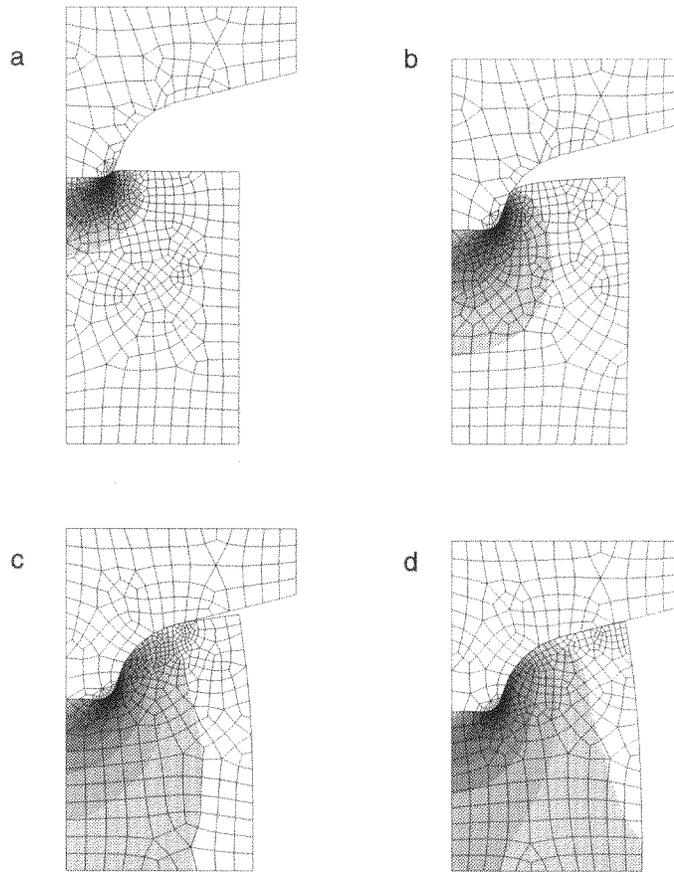


Figure 29. Axisymmetric piercing. Evolution of effective plastic strain. (a) $U = 0.62$, (b) $U = 5.66$, (c) $U = 9.83$, (d) $U = 11.00$.

for both die/punch and workpiece, which is updated every 200 timesteps. The dissipation factors for both plastic work and frictional work are assumed to be constant throughout the process.

Remeshing is an essential tool for a successful simulation of forging processes due to the large elasto-plastic deformation involved. The necessity of remeshing is assessed every 400 steps using criteria based on *element distortion*. Element distortion is evaluated by splitting the quadrilateral elements into four overlapping three-noded triangular elements at which a relative area change is computed as $\eta_e = \frac{1}{4} \sum_{i=1}^4 \frac{A_i - A_0}{A_0}$, where A_i and A_0 are respectively the current and initial (or immediately after remeshing) areas of the triangular elements. It is worth noting that, conceptually, *element distortion* is not an *error estimator* but an indicator of degradation of the element shape. In this simulation, the new mesh is created so that the element sizes over the die, punch and workpiece are as uniform as possible (the die and punch geometries require smaller elements near corners). Figure 31(a) shows a typical mesh, in which the elements over the die/punch and workpiece possess a characteristic length[†] around 20 and 5 mm respectively. The necessity of remeshing is clearly illustrated in figures 31 and 32(a) and (b), which show the finite-element mesh at

[†] In this example the maximum edge length is defined as the characteristic size of the elements.

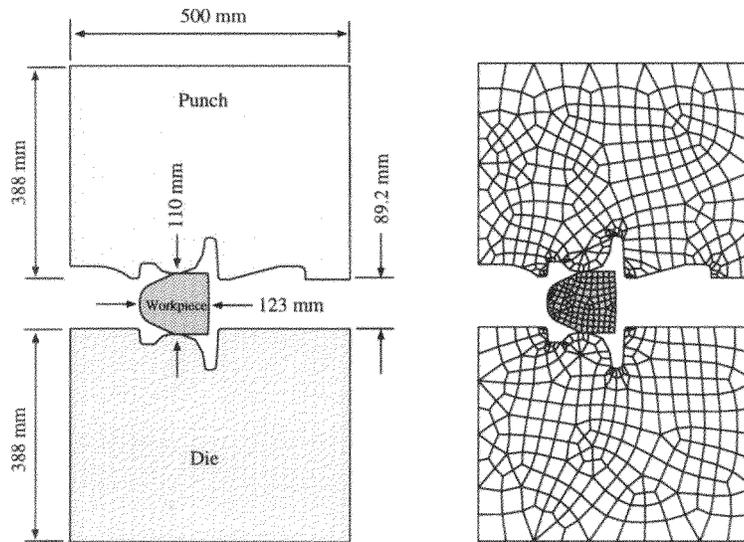


Figure 30. Closed-die rail forging: geometry and initial finite-element mesh.

Table 3. Material data and other simulation parameters.

| Description | Symbol | Value |
|--------------------------------|------------|---|
| Specific mass | ρ | 7800 kg m ⁻³ |
| Specific heat | c | 483.3 J kg ⁻¹ K ⁻¹ |
| Thermal conductivity | k | 48 J kg ⁻¹ m ⁻¹ K ⁻¹ |
| Young's modulus | E | 200.0 GPa |
| Poisson's ratio | ν | 0.3 |
| Coulomb's friction coefficient | μ | 0.1 |
| Yield stress | σ_Y | 275.0 MPa |
| Initial temperature | T_0 | 300 K |
| Plastic dissipation factor | ξ_p | 0.85 |
| Friction dissipation factor | ξ_f | 0.85 |
| Coupling interval | | 200 timesteps |
| Error checking | | 400 timesteps |
| Element distortion | η_e | 0.25% |
| Element size (workpiece) | h_w | 5.0 mm |
| Element size (die/punch) | | 20.0 mm |

the same compression stage for constant and adapted meshes respectively.

Figure 33 presents the distribution of the equivalent plastic strain for three different compression stages. In this example, due to the geometry of the die, the beginning of the process is characterized by a large plastic deformation near the centre of the workpiece and contact interface. As the punch advances, the excess material flows around the cavity edges to form the so-called *flash*, which, in conjunction with a narrowing gap, causes a greater increase of plastic deformation and friction.

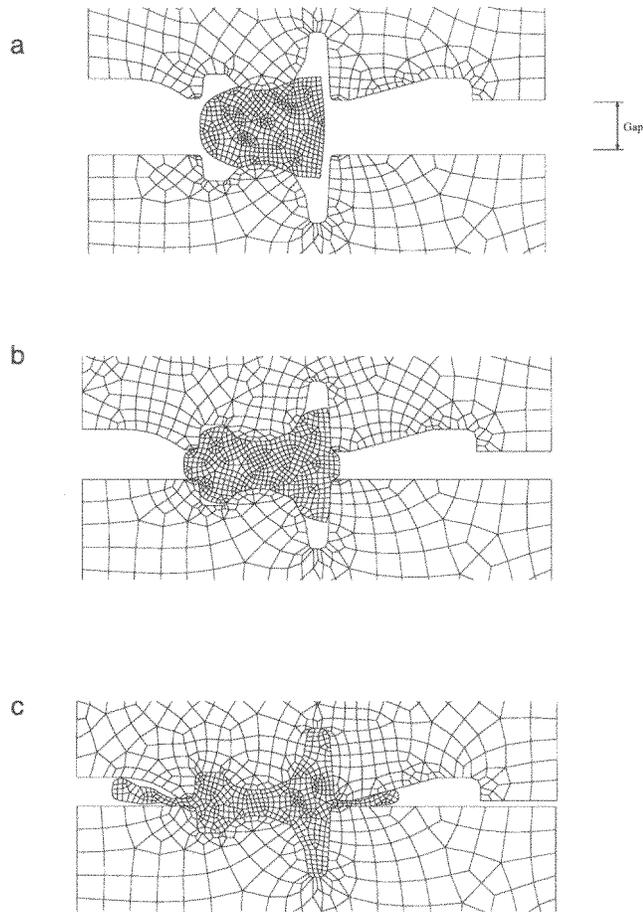


Figure 31. Closed-die rail forging: evolution of finite-element meshes. (a) Gap = 60 mm, (b) gap = 30 mm, (c) gap = 6 mm.

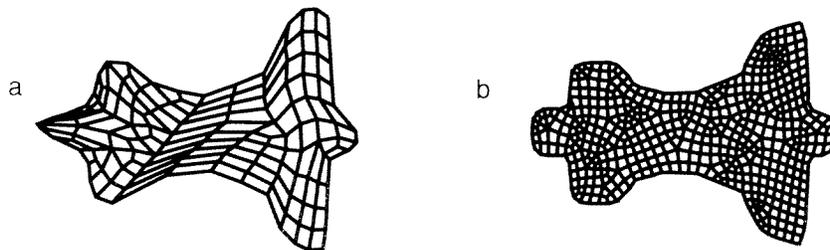


Figure 32. Closed-die rail forging: typical finite-element meshes. (a) Without remeshing, (b) with remeshing.

Heat generation due to dissipation of plastic work and frictional work are the main thermal components involved in the process. The importance of the individual effects at different compression stages can be evaluated from figure 34. At early stages, plastic deformation is the dominant factor, which is reflected as a similar distribution pattern for the

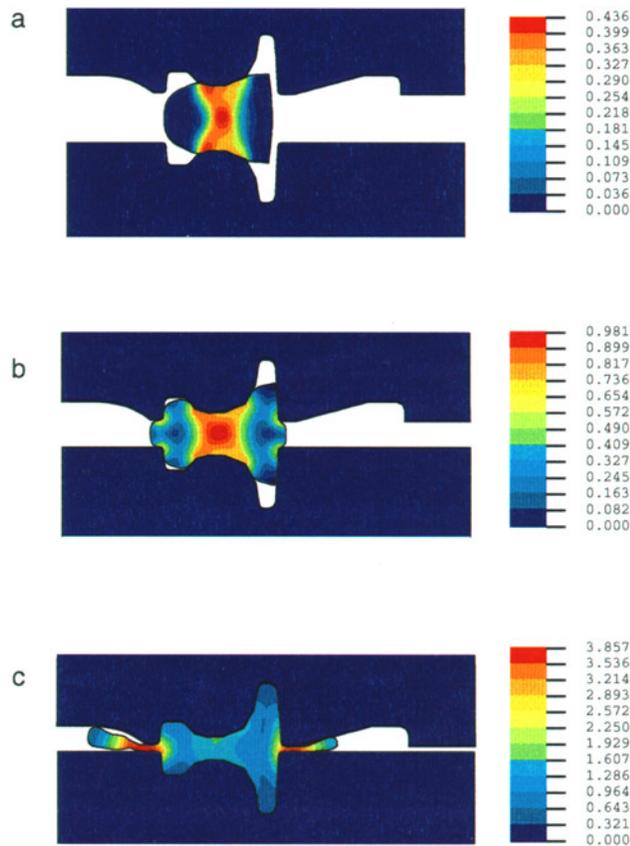


Figure 33. Closed-die rail forging: equivalent plastic strain evolution. (a) Gap = 60 mm, (b) gap = 30 mm, (c) gap = 6 mm.

equivalent plastic strain and temperature, as illustrated in figures 33 and 34 respectively. As the punch advances, the increasing friction effects cause localized heating near the contact zones. At the end of the process, the flow of the excess material through a narrow space causes a great increase of the frictional work and, consequently, the temperatures. In this example, the maximum temperatures are found near these regions on the surfaces of the die and punch.

10. Explicit solution strategies

The remarkable recent advances that have taken place in the development of implicit solution techniques for large-strain elasto-plastic problems with frictional contact conditions have been described in the preceding sections. In summary, consistent linearization has been achieved in all algorithmic aspects of the solution procedure and the element technology has been extended to deal with the incompressible nature of plastic deformation.

In contrast, the solution strategies for the explicit dynamic analysis of large-strain plasticity problems with contact are well established and there have been no significant developments in this direction over the last decade. The essential features of the explicit approach to finite-element analysis are summarized in section 10.1.

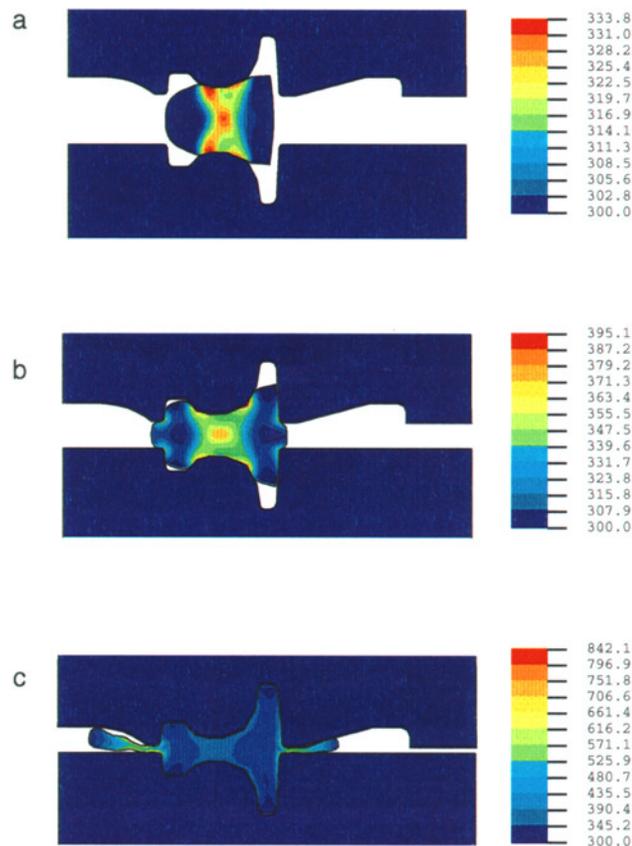


Figure 34. Closed-die rail forging: temperature evolution in K. (a) Gap = 60 mm, (b) gap = 30 mm, (c) gap = 6 mm.

In practice, the majority of industrial problems, such as forming operations, are sufficiently slow to be classified as quasistatic, with the material response being rate independent. Therefore, the sole justification of using explicit transient dynamic solution procedures for what are essentially quasistatic simulations is the much reduced computational times required, in comparison with quasistatic implicit analysis, for large-scale problems.

However, in order to achieve significant computational advantage several numerical artefacts have to be introduced into the explicit solution procedure. Whereas for quasistatic implicit analysis the process and material parameters utilized are the physical ones, some parameters are given artificial values in explicit analysis in order to provide acceptable CPU times. In particular, and with reference to metal forming operations, the following parameters are invariably modified.

(i) *Material density*. Since the maximum permissible timestep length, as defined by the Courant stability limit, is directly proportional to the square root of the material density, this parameter is increased; usually by one order of magnitude at least.

(ii) *Punch velocity*. In order to reduce the total number of timesteps necessary to model the forming process, the punch velocity is increased; again by at least an order of magnitude.

(iii) *Loading history*. Since increasing both the material density and punch velocity

results in increased inertia forces, the punch travel must be suitably controlled so as to minimize the inertia effects.

The amounts by which the material density and punch speed can be increased are limited, as the inertia forces may become unacceptably large—even with a judiciously designed loading history—resulting in erroneous solutions. For example, in deep drawing operations the predicted material stretching near the punch is often larger than seen in practice whilst the predicted draw is too small.

10.1. The discretized dynamic equations

Commencing from the linear momentum equation, the weak form of the equilibrium equations can be derived, which when discretized in the normal finite-element manner leads to

$$M\ddot{\mathbf{u}}_n + C\dot{\mathbf{u}}_n + \mathbf{P}(\mathbf{u}_n) = \mathbf{F}(t_n) \quad (10.1)$$

in which \mathbf{u}_n represents the displacement vector at time t_n , M and C are respectively the mass and damping matrices and $\mathbf{P}(\mathbf{u}_n)$ represents the internal force contribution from the element stress field which satisfies the (nonlinear) constitutive relations. The term $\mathbf{F}(t_n)$ represents the external forces arising from applied tractions and contact conditions.

The introduction of central difference approximations for the velocity and acceleration in terms of displacements gives a recurrence relation from which the displacements at time t_{n+1} can be evaluated in terms of quantities at time stations t_n and t_{n-1}

$$\mathbf{u}_{n+1} = (\mathbf{M} + C\Delta t/2)^{-1}[\Delta t^2(\mathbf{F}_n - \mathbf{P}_n) + 2M\mathbf{u}_n - (\mathbf{M} - C\Delta t/2)\mathbf{u}_{n-1}]. \quad (10.2)$$

The above time integration scheme is conditionally stable with the permissible timestep being governed by the *Courant stability limit*. The severe restrictions resulting from this condition make the use of mass lumping procedures possible which, with the added assumption of mass proportional damping, leads to the following uncoupled equation system

$$u_{A,n+1} = [M_A(1 + \alpha\Delta t/2)]^{-1}\{\Delta t^2(F_{A,n} - P_{A,n}) + 2M_A u_{A,n} - M_A(1 - \alpha\Delta t/2)u_{A,n-1}\} \quad (10.3)$$

in which α is the mass-proportional damping coefficient. This expression permits the evaluation of displacement on an individual nodal basis with internodal coupling occurring only through the calculation of the internal forces \mathbf{P}_n .

The kinematic equations must model finite-deformation effects and by using corotational measures of stress and strain the incremental large-strain constitutive relation can be written

$$\Delta\boldsymbol{\sigma}^c = \mathbf{C} : \Delta\boldsymbol{\varepsilon}^c. \quad (10.4)$$

Particular forms of the constitutive tensor \mathbf{C} follow from standard elasto-plastic or elasto-viscoplastic descriptions. Different strain measures may be employed, but considerable computational benefits arise from use of the logarithmic stretch so that

$$\Delta\boldsymbol{\varepsilon}^c = \ln[\mathbf{U}]. \quad (10.5)$$

The total stress at any time t_{n+1} is obtained from

$$\boldsymbol{\sigma}_{n+1}^c = \boldsymbol{\sigma}_n^c + \Delta\boldsymbol{\sigma}^c \quad (10.6)$$

and the Cauchy stress is then given by

$$\boldsymbol{\sigma}_{n+1} = \mathbf{R}_{n+1}\boldsymbol{\sigma}_{n+1}^c\mathbf{R}_{n+1}^T \quad (10.7)$$

where \mathbf{R}_{n+1} is an orthogonal rotation tensor defined by polar decomposition at time instant t_{n+1} .

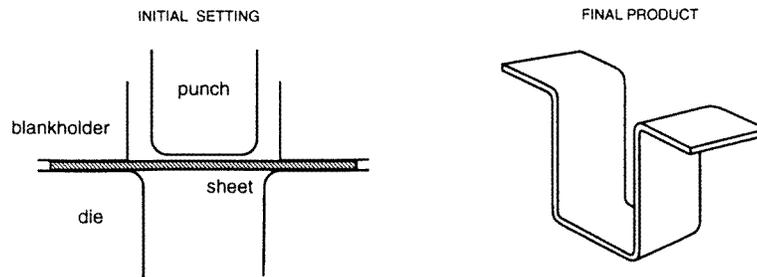


Figure 35. Drawing of a U-strip specimen: initial setting and final product.

10.2. Element methodology

The finite-element models employed for the simulation of industrial processes are invariably large and therefore in analysis a compromise between accuracy and computational efficiency has to be reached. This has led to the use of reduced integration procedures to primarily reduce computation time, but which also counter the overly stiff behaviour associated with full integration. For linear elements this implies single-point integration which, however, can result in spurious zero-energy (hourglass) modes of deformation (Belytschko *et al* 1984, Belytschko and Bindeman 1991, Belytschko and Leviathan 1994, Flanagan and Belytschko 1981, Hallquist 1991, Haug *et al* 1989, Schweizerhof *et al* 1992).

In order to obtain reliable results, various control methods have been proposed to eliminate hourglassing by providing restraint which the element lacks under single-point integration, but without stiffening the element's adequate response to other modes (Belytschko *et al* 1984, Flanagan and Belytschko 1981). Two principal ways of resisting hourglassing are with viscous damping and by introduction of artificial stiffness, both of which are capable of eliminating the spurious singular modes but have a negligible effect on the stable global modes. However, it should be stressed that hourglass control does not fully remove the kinematic modes and, in particular, coarse meshes and meshes loaded with large nodal forces, resulting either from boundary conditions or from contact, are susceptible to hourglassing despite the use of control techniques.

Example 10.1 (Drawing of a U-strip specimen). Figure 35 provides a schematic representation of a U-strip drawing operation, whereby a flat sheet of material is drawn into a die cavity by a much stiffer punch thus creating a U-strip. This problem is modelled both by a full 3D explicit analysis and also as a two-dimensional (2D) implicit quasistatic problem. For the 3D simulation four-node shell elements are employed with single-point integration and hourglass stabilization. The blankholder is modelled in two ways.

(i) *Deformable blankholder* in which the blankholder is represented by solid continuum elements with a pressure load applied to the upper surface.

(ii) *Rigid model* where a rigid contact surface is used in conjunction with a constraint equation to ensure that the total normal interface pressure is equal to the blankholder pressure.

For the 2D simulation continuum elements are used for both the strip and the blankholder.

Figure 36 illustrates the effect of punch speed on the prediction of the forming force. It is seen that the solution is significantly affected by the punch speed, with increasing velocities resulting in greater oscillations. The results of the explicit analysis could be considered acceptable up to a punch speed of 5 m s^{-1} provided that the loading is suitably

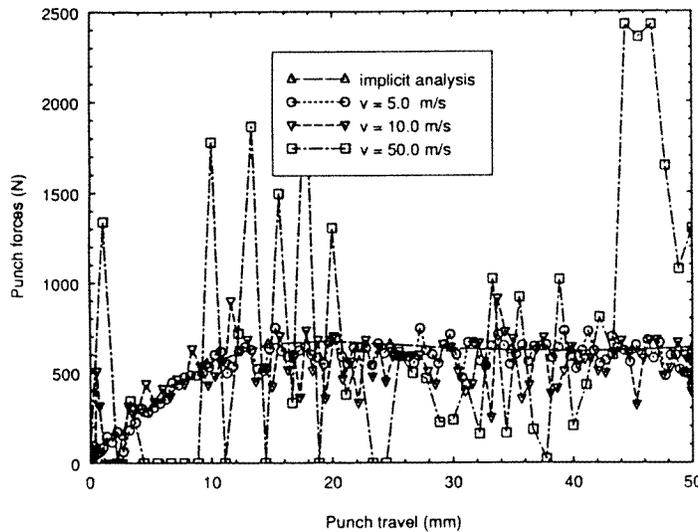


Figure 36. Drawing of a U-strip specimen: punch force with increasing punch velocity.

ramped. The material density has been scaled by 10 for solution.

The history of contact is shown in figure 37(a). The rigid contact surface algorithm provides an almost exact blankholder pressure, due to the constraint equation and consequently the total friction force on the blankholder is also accurate. The results for the deformable blankholder show large oscillations in the blankholder pressure, primarily due to points making and losing contact due to the inertia effects. This results in a friction force which is on average 25% lower than the expected value. The effect of this under-prediction of the blankholder friction force may be seen in figure 37(b). The punch force history for the rigid blankholder case agrees well with implicit solution, whereas the deformable blankholder prediction is some 30% too low.

Example 10.2 (Deep drawn automotive headlamp panel). Finally, the explicit simulation of an industrial component is illustrated in figure 38 which defines the finite-element description of the sheet and tooling for the pressing of a deep drawn automotive headlamp panel. Features of the problem include a non-planar blankholder, a deep draw resulting in relatively large sliding distances and wrinkling in certain regions of the finished part. The sheet is modelled by 5500 four-node shell elements and rigid tools are assumed, with the blankholder being pressure controlled. The final pressed shape is shown in figure 39 where it is seen that the wrinkling that occurs in the prototype is reproduced by the finite-element simulation. This wrinkling is caused by both lack of blankholder restraint at one end and a mismatch between the punch and die.

11. A combined finite-/discrete-element method

The finite-element approach is rooted in the concepts of continuum mechanics, and for physical situations exhibiting strong discontinuities in material and geometric behaviour, a finite-element description does not offer the most appropriate modelling approach. Examples of such problems include the behaviour of jointed rock, mining and rock blasting operations, ultimate load behaviour of masonry structures and various processes involving granular

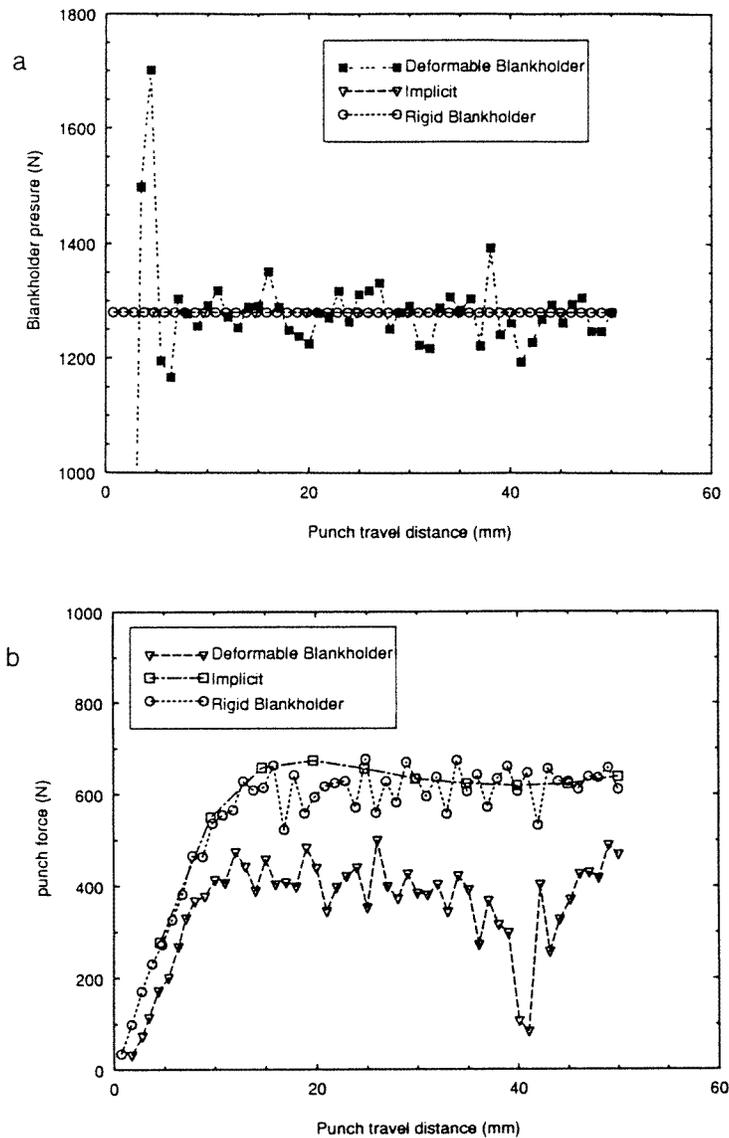


Figure 37. Drawing of a U-strip specimen. (a) Comparison of rigid and deformable blankholder algorithms. (b) Punch force against the punch travel.

materials, such as, material flow in silos and powder compaction processes. The treatment of these classes of problems is more naturally related in finite-/discrete-element concepts, in which individual material elements are considered to be separate and are (possibly) connected only along their boundaries by frictional contact.

The use of discrete elements was originally developed for geotechnical applications in the late 1960s. A model by Goodman *et al* (1968), for example, was aimed at the simulation of jointed rock and was accomplished by introducing discontinuities into an existing continuum. In the early 1970s models based on the *a priori* assumption of discontinuous behaviour were introduced, with continuum behaviour being treated as a

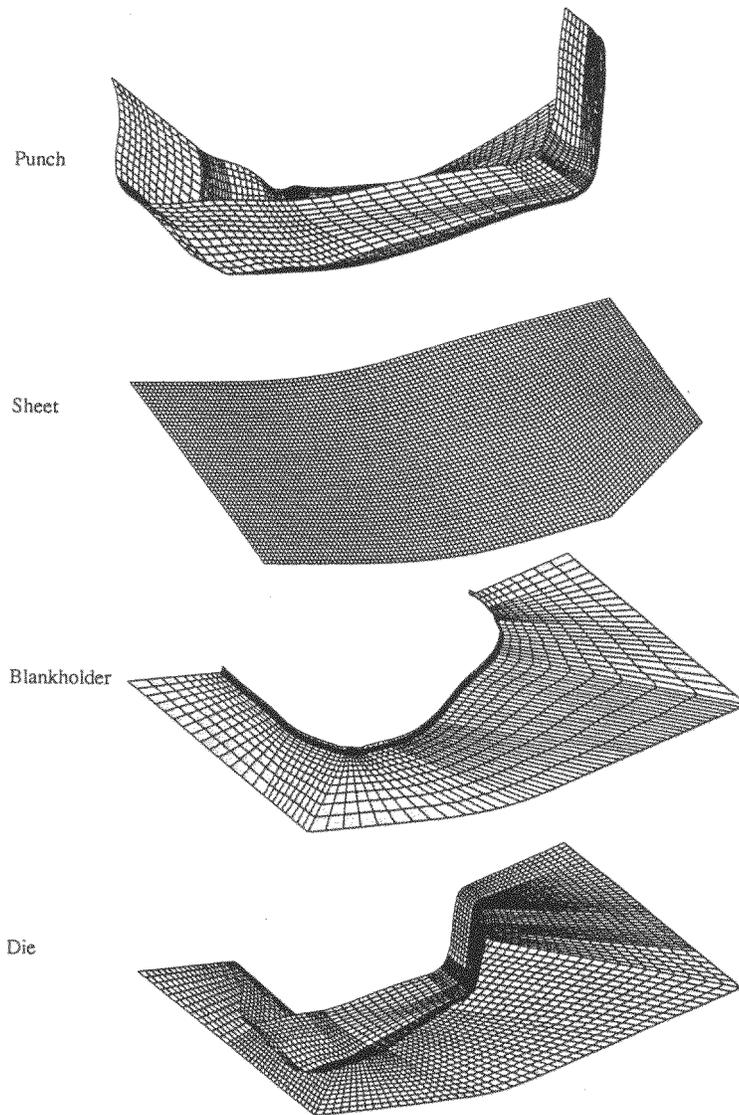


Figure 38. Deep drawn automotive headlamp panel: definition of finite-element model.

special case. The methodology was later applied in modelling various industrial processes incorporating granular media (Cundall and Strack 1979). The set of methods developed was generally termed the *distinct-element method*.

Further development took place in the 1980s with the incorporation of deformation kinematics for the more precise modelling of individual elements, and the term *discrete-element method* became widely accepted. The modelling of local deformation has permitted a more rigorous treatment of both the contact conditions and energy preservation requirements. This has also lead naturally to a combined finite-/discrete-element approach in which the problem is analysed by a combination of the two methods (Munjiza *et al* 1995). This solution method is particularly suited to problems in which progressive fracturing

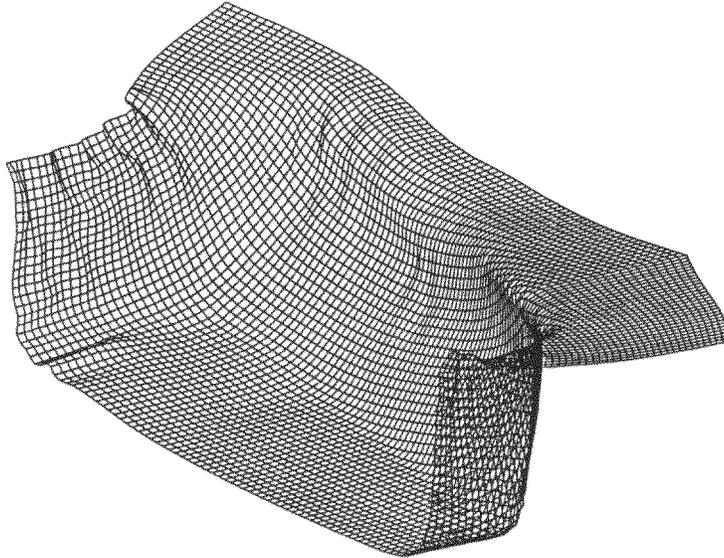


Figure 39. Deep drawn automotive headlamp panel: location of wrinkles in simulation.

takes place as is the case, for instance, in missile impact simulations and composite fracture/delamination. Conference proceedings edited by Jenkins and Satake (1983), Satake and Jenkins (1988), Mustoe (1992) and Williams (1995) offer a comprehensive review of the subject and areas of application.

In a finite-/discrete-element simulation, starting from a continuum representation by finite elements of the solid region in question, progressive fracturing within elements is allowed to take place according to a prescribed fracturing criterion, thereby forming discrete elements which may be composed of one or more deformable finite elements. Subsequent motion of these discrete elements and further fracturing of both the remaining continuum and previously created discrete elements is then modelled. This evolution process is continued until either the system comes to rest or up to the time of interest.

In essence, in a finite-/discrete-element analysis the main issues which require consideration, for both dynamic and quasistatic behaviour, are as follows.

(i) Appropriate element modelling of the continuum and discrete regions with a view to incorporating the deformation mechanisms necessary to model stress, strain and subsequent fracturing of the original continuum.

(ii) The development of appropriate material models, and particularly fracture criteria.

(iii) The development of remeshing algorithms to convert fractured zones into a discrete-element representation.

(iv) Detection procedures for monitoring contact between discrete-element regions.

(v) Representation of contact interaction conditions for contacting elements.

By far the most crucial operation which governs the efficiency of any discrete-element analysis is the multibody contact detection procedure. Algorithms which employ fast, bounding box search techniques, drastically reduce the number of local, geometrically based calculations required to define contact tractions between interacting bodies. In particular, the alternating digital tree (ADT) algorithm solves this problem with a computational expense proportional to $O = N \log_2(N)$, where N represents a number of contact interactions.

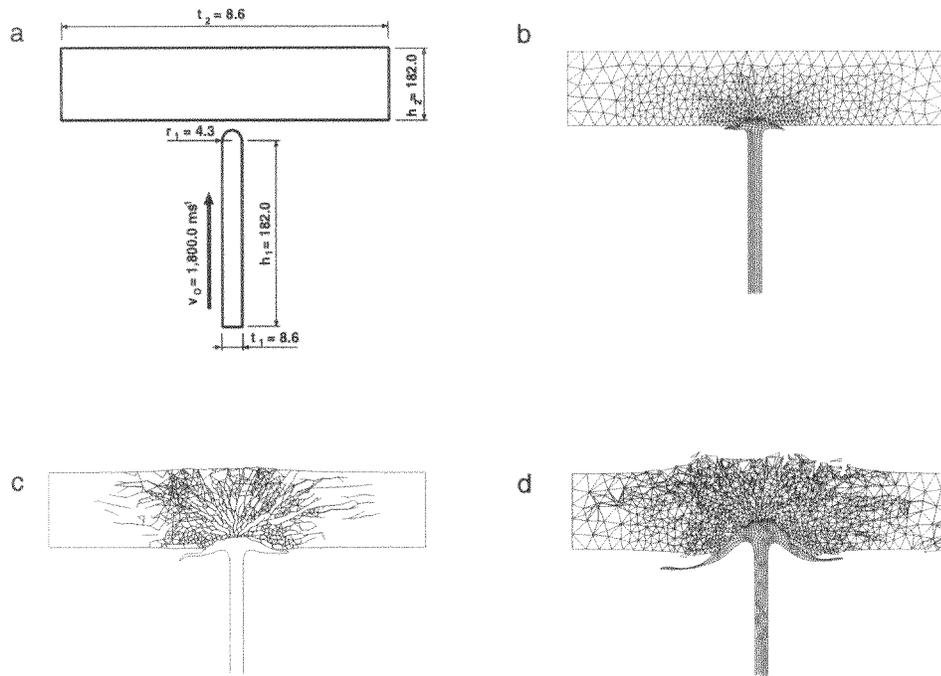


Figure 40. Impact of an elasto-plastic projectile on a ceramic block. (a) Geometry and loading conditions. Evolution of deformed mesh at various time instants: (b) $t = 0.013 \text{ ms}$, (c) $t = 0.026 \text{ ms}$, (d) $t = 0.050 \text{ ms}$.

The power of the method, and its potential applicability to the modelling of a wide range of practical problems is illustrated by the following examples.

Example 11.1 (Impact of an elasto-plastic projectile on a ceramic block). Geometry and loading for this problem are depicted in figure 40(a). The projectile is assumed to be made of an elasto-plastic copper material with Young's modulus $E = 117 \text{ GPa}$, Poisson's ratio $\nu = 0.35$ and yield stress $\sigma_Y = 400 \text{ MPa}$ with linear isotropic hardening modulus $H = 100 \text{ MPa}$. The target is made of an elasto-plastic brittle material with Young's modulus $E = 211.69 \text{ GPa}$, Poisson's ratio $\nu = 0.286$, yield stress $\sigma = 958 \text{ MPa}$ and fracture energy $D_f = 100 \text{ J m}^{-2}$.

A Coulomb law with coefficient of friction $\mu = 0.9$ defines frictional contact between the projectile and target and the initial velocity of the projectile is $v_0 = 1800 \text{ m s}^{-1}$. In analysis an error indicator based on L^2 stress norm is used. Analysis is performed by a transient dynamic explicit time integration approach in view of the high loading rates.

Continuous change in geometry of the projectile near the contact area necessitates frequent adaptive remeshings during the process. In addition, mesh adaption is required to provide for accurate description of multiple fracturing of the block. It should be emphasized that the present discrete-element model allows for crack propagation both along element boundaries, and through the finite elements. The deformation of the projectile, the adapted finite-element meshes and fracturing process of the block are depicted in figure 40(b)–(d).

Example 11.2 (Shot peening simulation). Shot peening operations, in which the surfaces of components are systematically impacted with steel or ceramic shot are extensively employed

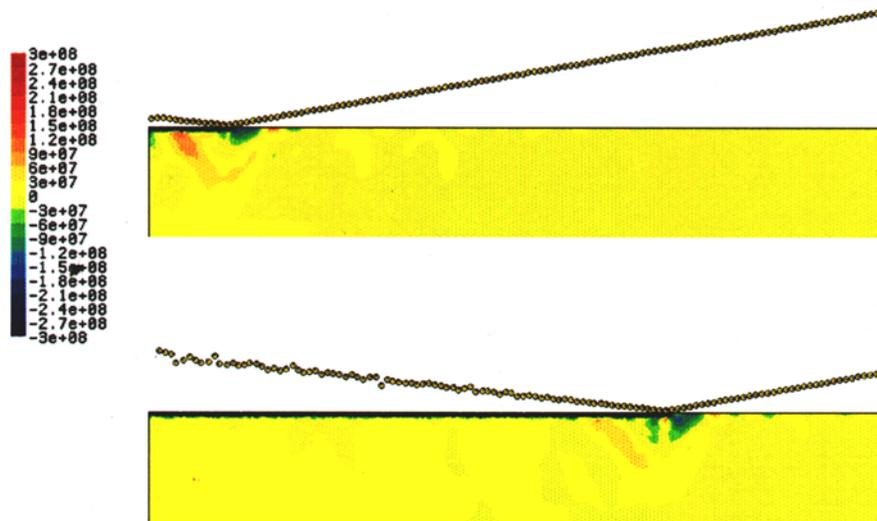


Figure 41. Shot peening simulation: evolution of the horizontal stress component (in N m^{-2}) at two stages of the process.

in the aerospace, automotive, off-shore and other industries, with three specific objectives: (i) the most common aim is to induce an initial state of compressive stress in the vicinity of the surface to inhibit fatigue crack growth under operational loading, (ii) to shape form components to produce a desired curvature and shape or to correct the shape of components, (iii) to relieve tensile stresses that contribute to stress-corrosion cracking. All these operations rely on plastically deforming the surface locally either to produce an initial compressive pre-stress or a local curvature condition.

Figure 41 illustrates the impacting of a sheet by some 150 shot. The problem is treated as a 2D dynamic transient analysis and the uniform compressive horizontal stress distribution produced in the vicinity of the surface is shown, which is in close agreement with experimental measurements in terms of both intensity and depth.

12. Advanced equation solution strategies for solid mechanics

The finite-element simulation of industrial-scale problems within an implicit scheme gives rise to large algebraic equation systems of the form

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (12.1)$$

where \mathbf{A} is an $n \times n$ non-singular matrix, \mathbf{b} is the known vector, and \mathbf{x} is the solution to be found.

In solution of the algebraic system (12.1), iterative methods offer compelling promise over direct methods with regard to the following aspects:

- (i) much easier to exploit system sparsity and thus the computer memory required may be substantially reduced, especially for large problems;
- (ii) relatively simple in implementation;
- (iii) accuracy may be more controllable and thus computing time may be saved in cases where only a lower level accuracy is required;

(iv) they may lead to a more synergistic incorporation in the solution of evolving nonlinear problems;

(v) these methods may prove more conducive to effective implementation in emerging computer systems with vector and parallel processing facilities.

Symmetric positive definite problems could arise if problems without friction are simulated. Otherwise, a set of non-symmetric algebraic equations can be expected if a process is analysed where frictional contact is essential, as is the case for most metal forming processes. In the former case, it seems quite clear that the *conjugate gradient (CG) method* is the most efficient iterative solver to date. In the past decade, a number of *Krylov-type methods* have been proposed that are applicable to non-symmetric cases, which include:

- (i) CGS, *conjugate gradient squared method* (Sonnenveld 1989)
- (ii) Bi-CGStab, *bi-conjugate gradient stabilized method* (van der Vorst 1992)
- (iii) GMRES, *generalized minimum residual method* (Saad 1986).

These methods require only three operations in the implementation: saxpy operations, inner products, and matrix–vector multiplications. As a result they can be efficiently implemented on scalar, vector and parallel computers.

12.1. CGS and Bi-CGStab methods

The CGS approach is an accelerated version of the *bi-conjugate gradient method*, which is an extension of CG to non-symmetric cases. CGS sometimes suffers from severe numerical instability, even though it often exhibits good convergence. Bi-CGStab is the stabilized version of CGS, whose superior performance can be demonstrated by numerical experiments. The Bi-CGStab algorithm involves two matrix–vector multiplications and two pre-conditioning operations at each iteration, and requires seven working vectors. For more details concerning the algorithm we refer to van der Vorst (1992).

12.2. Galerkin multigrid method

The essential multigrid principle is based on the observation that the smooth (or long-wavelength) part of the error, which may not be efficiently swept out by iterative methods such as CG and Bi-CGStab, can be substantially reduced by a coarse mesh correction. The success of multigrid strategies lies primarily in (i) their excellent convergence characteristics, which theoretically should not depend on the size of the finite-element mesh; (ii) their high efficiency whereby solutions of problems with n_{eq} unknowns are obtained with $O(n_{\text{eq}})$ in terms of work and storage for large classes of problems. Several different schemes of multigrid techniques have been put forward in the last decade (Hackbusch 1985). In this section we focus on one particular scheme termed the Galerkin multigrid (GMG) method proposed by Feng *et al* (1997).

To illustrate the basic idea of the GMG scheme we consider its two-grid form. Suppose that G_c and G are, respectively, coarse and fine grids which discretize the same geometrical domain \mathcal{B} , and that the fine grid is supposed to represent the current problem considered. We use subscript c to distinguish the quantities of the coarse grid from those of the fine grid. Let A_c be the coarse-grid matrix, and P and Q be, respectively, the matrix representations of the interpolation and projection operators. In the GMG method, the coarse-grid matrix A_c is constructed by direct projection of the fine-mesh matrix as

$$A_c = QAP. \quad (12.2)$$

Here projection operator Q is taken as $Q = P^T$, and therefore

$$A_c = P^T A P. \quad (12.3)$$

Efficient computation of A_c is crucial to achieve an overall high performance of the complete GMG method. Such an implementation can be found in Feng *et al* (1997). Let $S(x, \mu)$ denote the smoother with x as the initial guess and μ the maximum number of iterations, and μ_1 and μ_2 be the maximum iterations of the pre- and post-smoothing procedures performed respectively before and after the coarse-grid correction which is accomplished by a profile solver.

Obviously, the efficiency of GMG is dependent on the quality of the coarse grid and the appropriate selection of interpolation and projection operators. Once A_c , P and Q are determined, the performance of GMG will entirely depend on the smoother S and the numbers of iterations μ_1, μ_2 . The practical selection of smoothers can range from very simple Jacobi (or diagonal scaling), Gauss–Seidel, successive over-relaxation (SOR), symmetric SOR (SSOR) to incomplete decomposition, and even to any iterative algorithm. In our case, pre-conditioned Bi-CGStab is chosen as the smoother.

In order to enhance the performance of GMG, the outer loop of the multigrid iterations is further accelerated by GMRES and it is equivalent to nonlinear GMRES with multigrid as its pre-conditioning scheme. For more details regarding this enhancement, we refer to Feng *et al* (1998).

As the Galerkin strategy has been fully adopted for the generation of coarse-mesh equations and no material and loading information for coarse meshes is utilized, the GMG approach is relatively easy to incorporate into existing solution procedures, and is particularly suitable for implementation in material nonlinear cases, including frictional contact. For geometrically nonlinear cases, the approach uses a constant transfer operator throughout the whole solution process without significantly influencing the convergence property. Another important feature of the GMG method is that coarse and fine meshes can be non-nested and unstructured which not only allows for easy treatment of complex geometry problems, but also provides a possibility of easy combination with adaptive mesh refinement techniques.

Example 12.1 (Stretching of a circular thin sheet by a hemispherical punch). The geometry for this example is as follows: blank thickness is $t = 1$ mm, blank radius is $R = 59.18$ mm, while the punch and die radii are $R_p = 50.8$ mm and $R_d = 6.83$ mm respectively. It is assumed that the blank material is described by a large-strain elasto-plastic constitutive model with Young's modulus given by $E = 2.1 \times 10^5$ N mm⁻², Poisson's ratio $\nu = 0.3$ and isotropic hardening, which relates the equivalent stress $\bar{\sigma}$ and the equivalent plastic logarithmic strain $\bar{\epsilon}^p$ as $\bar{\sigma} = 589 \times (10^{-4} + \bar{\epsilon}^p)^{0.216}$ N mm⁻². The analysis is performed employing a membrane formulation for a quarter of the problem with appropriate boundary conditions. From a numerical point of view this problem is considered as a full 3D analysis with appropriate algorithmic treatment of the frictional contact problem. Results are obtained for a coefficient of friction between tools and blank of $\mu = 0.30$.

To solve this problem the blank is discretized with 17 666 constant-strain triangular finite elements resulting in 26 298 active d.o.f. The surfaces of the punch and die are respectively represented by 2145 and 612 triangular flat elements. Spatial discretization of the problem is depicted in figure 42(a), and figure 42(b) shows the deformed mesh at a punch displacement of $D_p = 30$ mm.

Within the multigrid strategy the above mesh represents the fine mesh, while the coarse mesh contains 841 constant-strain triangular finite elements with 1230 active d.o.f. It should

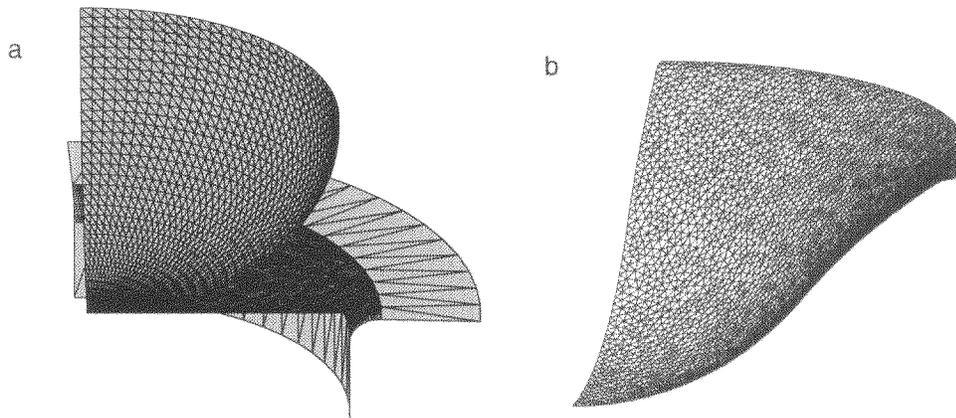


Figure 42. Stretching of a circular thin sheet by a hemispherical punch. (a) Spatial finite-element discretization of blank, punch and die surfaces. (b) Deformed finite-element mesh.

be emphasized that the fine and coarse meshes are fully unstructured meshes and have been obtained independently from each other.

A full Newton–Raphson method, with an unsymmetric tangent stiffness arising from the non-associated frictional contact law, is employed in all computations. Convergence of the finite-element solution is established on the basis of the standard Euclidean norm of the out-of-balance forces.

Three solution algorithms are employed in the solution of the algebraic system of equations: profile (direct) solver, Bi-CGStab iterative solution and multigrid algorithm. A comparison of performances of these algorithms, in terms of CPU time (seconds) and memory requirements (Mbytes), is presented in figure 43(a). It is clear that both the Bi-CGStab and multigrid solution provide a significant improvement over the standard direct solver. Memory requirements and CPU time have both been reduced by approximately 7–8 times in comparison with the direct solution. It should be emphasized that, in comparison with the Bi-CGStab solution, the multigrid strategy requires some small additional storage, but shows almost twofold reduction in the CPU time. In addition, the multigrid solution is expected to have a more stable behaviour in the solution of complex 3D problems that may arise in simulation of industrial processes.

The Bi-CGStab solution has also been implemented in parallel on a Silicon Graphics Power Challenge with eight R4400 processors. The results, in terms of speed-up and efficiency, are given in figure 43(b), for two, four, six and eight processors. Although very successful (94% efficiency) for two processors, the performance of the parallel version of the iterative solution slowly degrades with increase in the number of processors, and has an efficiency of 64% for eight processors. However, for larger problems, better efficiencies are to be expected.

13. Concluding remarks

Some recent advances in the finite-element analysis of nonlinear solid mechanics problems have been reviewed, indicating the progress that has been made both in the theoretical understanding of inelastic material behaviour under finite strains and the associated numerical implementation. Whilst the state of knowledge in some areas is relatively

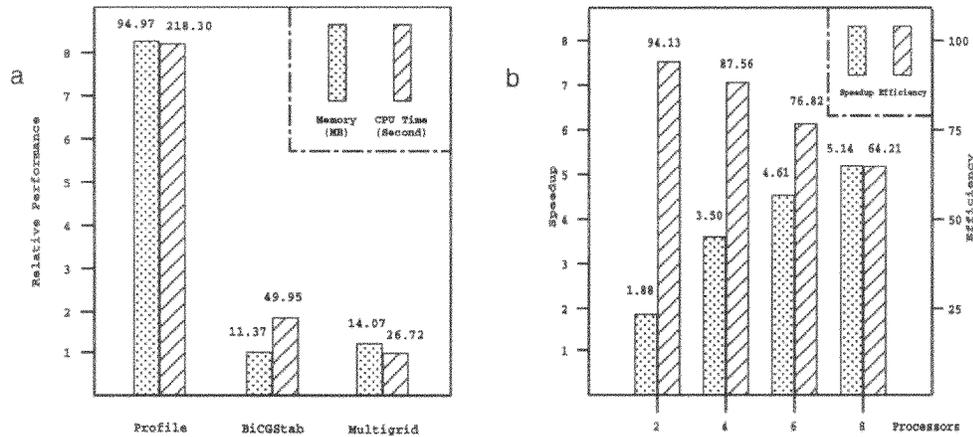


Figure 43. Stretching of a circular thin sheet by a hemispherical punch. (a) A comparison of performances, in terms of CPU time (seconds) and memory requirements (Mbytes), for three equation solution algorithms: profile (direct) solver, Bi-CGStab iterative solution and multigrid algorithm. (b) Performances, in terms of speed-up and efficiency, of a parallel version of the Bi-CGStab iterative solution on a shared memory parallel computer with two, four, six and eight processors.

mature, considerable further understanding and development is required in others. For example, issues related to the modelling of complex contact-friction phenomena are far from settled and a more comprehensive treatment of friction may necessitate the integration of micromechanical studies with computational approaches.

Although adaptive strategies are, at present, routinely performed for linear elliptic problems, their extension to nonlinear elliptic problems—in particular to forming problems where, typically, large inelastic deformations at finite strains are standard working conditions—is by no means trivial.

Apart from the issues briefly mentioned in section 9, several important aspects of adaptive strategies related to nonlinear industrial applications need further attention.

(i) Inclusion of various types of error estimators and their comparative analysis. In this respect, *a posteriori* error estimators of the residual type show particular promise.

(ii) The history dependent nature of elasto-plastic problems necessitates transfer of all relevant variables from the old mesh to the new one, as the successive mesh adaption is implemented during the process simulation. A careful further consideration must be given to: consistency with the constitutive equations and consequences for global equilibrium, compatibility of the state transfer with the displacement field on the new mesh.

The adequacy of explicit dynamic transient solutions for essentially quasistatic simulations is still a subject of debate. The primary justification of using such techniques is the much reduced computational effort involved, for large-scale industrial problems at least. However, the quality of the solutions obtained is invariably inferior to the corresponding implicit quasistatic solution, which is brought about by inertia effects and by the scaling of process and material parameters necessary to provide acceptable CPU times.

In future the competitiveness of implicit quasistatic solution methods may be improved by developments in iterative methods and sparse matrix techniques for equation solving which, when accompanied by increased computer memory availability, may make the solution of large-scale industrial problems by such approaches a realistic proposition.

Implementation of the procedures discussed in preceding sections will increasingly rely on parallel processing concepts in order to provide solutions to industrial problems within acceptable computing timescales. The use of parallel processing techniques and associated hardware for the solution of finite-element problems is currently an active research topic (Papadrakakis 1997). Parallel processing offers a natural approach to improving computational power and the next generation of workstations will increasingly employ multiprocessor architectures to achieve significantly advanced performance.

This review has concentrated on the field of solid mechanics, which, in the narrow sense, excludes problems of coupling of the traditional fields of solid and fluid mechanics, together with possible thermal field and phase changes, commonly known as *multiphysics* problems. At present the formulation and numerical solutions of multiphysics problems are being pursued for a large number of applications, ranging from defence to medicine.

The approach taken here to describe the solution of industrial problems at a macroscopic scale is rooted in the traditional view of solid mechanics based on the phenomenological approach. However, the last decade or so has witnessed establishment of the field of computational micromechanics (Ortiz (1996) provides a recent review of the field) whereby direct computational simulation at the microscale has been employed in generating information that is, directly or indirectly, used at the macroscopic level. This approach has been encouraged by considerable advances in computational resources, both in terms of hardware and software developments. Since computational micromechanics inevitably requires spanning over multiple length scales, the advances in this field and their practical applications will be closely linked with further advances in computational resources.

Acknowledgments

This work was partially supported by the EPSRC of UK through Grants GR/J13779, GR/J47008 and GR/K88965, awarded to the University of Wales, Swansea, and this support is gratefully acknowledged. A number of people have at various stages contributed to this work. In particular we acknowledge the contributions made by E A de Souza Neto, M Dutko, Y T Feng, M Vaz Jr, A J L Crook, J Yu, N Petrinić and Ch Hochard.

References

- Argyris J H 1954 *Aircr. Eng.* **26** 347–56
—1954 *Aircr. Eng.* **26** 383–7
—1954 *Aircr. Eng.* **26** 394
—1955 *Aircr. Eng.* **27** 42–58
—1955 *Aircr. Eng.* **27** 80–94
—1955 *Aircr. Eng.* **27** 125–34
- Babuška I and Rheinboldt W C 1978 *Int. J. Numer. Methods Eng.* **12** 1597–615
- Babuška I, Zienkiewicz O C, Gago J and de Oliveira E R (ed) 1986 *Accuracy Estimates and Adaptive Refinements in Finite Element Computations* (New York: Wiley)
- Bank R E and Weiser A 1985 *Math. Comput.* **44** 283–301
- Barlat F and Lian J 1989 *Int. J. Plast.* **5** 51–66
- Barthold F-J, Schmidt M and Stein E 1997 *Proc. 5th Int. Conf. on Computational Plasticity: Fundamentals and Applications* (Barcelona: CIMNE) pp 597–602
- Bathe K-J 1996 *Finite Element Procedures* 2nd edn (Englewood Cliffs, NJ: Prentice-Hall)
- Belytschko T 1983 *Computational Methods for Transient Analysis* (Amsterdam: North-Holland) pp 1–65
- Belytschko T and Bindeman L P 1991 *Comput. Methods Appl. Mech. Eng.* **88** 311–40
- Belytschko T and Leviathan I 1994 *Comput. Methods Appl. Mech. Eng.* **115** 277–86
- Belytschko T, Lin D and Tsay C S 1984 *Comput. Methods Appl. Mech. Eng.* **42** 225–51
- Belytschko T and Tsay C S 1983 *Int. J. Numer. Methods Eng.* **19** 405–19

- Belytschko T, Wong B L and Plakacz E J 1989 *Comput. Struct.* **33** 1307–23
- Ciarlet P G 1988 *Mathematical Elasticity. Volume 1: Three-Dimensional Elasticity* (Amsterdam: North-Holland)
- Clough R W 1960 *Proc. 2nd ASCE Conf. on Electronic Computation* (Pittsburgh, PA)
- Courant R 1943 *Bull. Am. Math. Soc.* **49** 1–23
- Crisfield M A 1991 *Non-linear Finite Element Analysis of Solids and Structures. Volume 1: Essentials* (Chichester: Wiley)
- 1997 *Non-linear Finite Element Analysis of Solids and Structures. Volume 2: Advanced Topics* (Chichester: Wiley)
- Crisfield M A, Moita G F, Jelenić G and Lyons L P R 1995 *Proc. 4th Int. Conf. on Computational Plasticity: Fundamentals and Applications* (Swansea: Pineridge) pp 293–320
- Cuitiño A and Ortiz M 1992 *Eng. Comput.* **9** 437–51
- Cundall P A and Strack O D L 1979 *Géotechnique* **29** 47–65
- Curnier A 1984 *Int. J. Solids Struct.* **20** 637–47
- Demkowicz L, Devloo Ph and Oden J T 1985 *Comput. Methods Appl. Mech. Eng.* **53** 67–89
- de Souza Neto E A, Hashimoto K, Perić D and Owen D R J 1996a *Phil. Trans. R. Soc. A* **354** 819–43
- de Souza Neto E A and Perić D 1996 *Comput. Methods Appl. Mech. Eng.* **130** 179–93
- de Souza Neto E A, Perić D, Dutko M and Owen D R J 1996b *Int. J. Solids Struct.* **33** 3277–96
- de Souza Neto E A, Perić D, Huang G C and Owen D R J 1995 *Commun. Numer. Methods Eng.* **11** 951–61
- de Souza Neto E A, Perić D and Owen D R J 1994 *Eng. Comput.* **11** 257–81
- 1998 *Arch. Comput. Methods Eng.* **5** 311–84
- Dutko M, Perić D and Owen D R J 1993 *Comput. Methods Appl. Mech. Eng.* **109** 73–93
- Duvaut G and Lions J-L 1976 *Inequalities in Mechanics and Physics* (Berlin: Springer)
- Eriksson K and Johnson C 1988 *Math. Comput.* **55** 361–83
- Eterović A L and Bathe K-J 1990 *Int. J. Numer. Methods Eng.* **30** 1099–114
- Eve R A, Reddy B D and Rockafellar R T 1990 *Q. Appl. Math.* **48** 59–83
- Feng Y T, Perić D and Owen D R J 1997 *Comput. Methods Appl. Mech. Eng.* **144** 307–25
- 1998 *Int. J. Numer. Methods Eng.* **42** 1441–62
- Fjaer H G and Mo A 1990 *Metall. Trans. B* **21** 1049–60
- Flanagan D P and Belytschko T 1981 *Int. J. Numer. Methods Eng.* **17** 679–706
- Fung Y C 1965 *Foundations of Solid Mechanics* (Englewood Cliffs, NJ: Prentice-Hall)
- 1977 *A First Course in Continuum Mechanics* (Englewood Cliffs, NJ: Prentice-Hall)
- Gallimard L, Ladeveze P and Pelle J-P 1996 *Int. J. Numer. Methods Eng.* **39** 189–217
- George P L 1996 *Handbook of Numerical Analysis* vol IV, ed P G Ciarlet and J L Lions (Amsterdam: North-Holland) pp 69–190
- Giannakopoulos A E 1989 *Comput. Struct.* **32** 157–67
- Goodman R E, Taylor R L and Brekke T 1968 *J. Soil Mech. Found. Div., Proc. ASCE* **94** SM3
- Gurtin M E 1981 *An Introduction to Continuum Mechanics* (Orlando, FL: Academic)
- Hackbusch W 1985 *Multi-grid Methods and Applications* (Berlin: Springer)
- Hallquist J O 1991 *LS-DYNA3D Theoretical Manual*
- Halphen B and Nguyen Q S 1975 *J. Mec.* **14** 39–63
- Hashimoto K, Ohwue T and Takita M 1990 *Proc. IDDRG 16th Congress*
- Haug E, Clinckemaillie J and Aberlenc F 1989 Post symposium course *2nd Int. Symp. on Plasticity (Nagoya)*
- Hill R 1950 *Mathematical Theory of Plasticity* (Oxford: Oxford University Press)
- Hughes T J R 1980 *Int. J. Numer. Methods Eng.* **15** 1413–18
- 1987 *The Finite Element Method: Linear Static and Dynamic Analysis* (Englewood Cliffs, NJ: Prentice-Hall)
- Hutchings I M 1992 *Tribology: Friction and Wear of Engineering Materials* (London: Edward Arnold)
- Jenkins J T and Satake M (ed) 1983 *Mechanics of Granular Materials: New Models and Constitutive Relations* (Amsterdam: Elsevier)
- Jin H, Bernspang L, Larsson R and Wiberg N E 1989 *Proc. 2nd Int. Conf. on Computational Plasticity: Models, Software and Applications* (Swansea: Pineridge) pp 347–57
- Johnson C and Hansbo P 1992 *Comput. Methods Appl. Mech. Eng.* **101** 143–81
- Kestin J and Bataille J 1977 *Proc. Int. Symp. on Continuum Models of Discrete Systems* (Waterloo: University of Waterloo Press) pp 39–67
- Kikuchi N and Oden J T 1988 *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods* (Philadelphia, PA: SIAM)
- Ladeveze P, Cofignal G and Pelle J-P 1986 *Accuracy Estimates and Adaptive Refinements in Finite Element Computations* ed Babuška *et al* (New York: Wiley) pp 181–203
- Laursen T and Simo J C 1993 *Int. J. Numer. Methods Eng.* **36** 3451–85

- Lee E H 1969 *J. Appl. Mech.* **36** 1–6
- Lee N-S and Bathe K-J 1994 *Finite Elem. Anal. Des.* **16** 99–139
- Lemaitre J 1985 *J. Eng. Mater. Technol.* **107** 83–9
- Lemaitre J and Chaboche J L 1990 *Mechanics of Solid Materials* (Cambridge: Cambridge University Press)
- Lubliner J 1990 *Plasticity Theory* (New York: Macmillan)
- Marusich T D and Ortiz M 1995 *Int. J. Numer. Methods Eng.* **38** 3675–94
- Michałowski R and Mroz Z 1978 *Arch. Mech.* **5** 259–76
- Mitchell G P and Owen D R J 1988 *Eng. Comput.* **5** 274–84
- Moran B, Ortiz M and Shih F 1990 *Int. J. Numer. Methods Eng.* **29** 483–514
- Munjiza A, Owen D R J and Bičanić N 1995 *Eng. Comput.* **12** 145–74
- Mustoe G W (ed) 1992 First US conference on discrete element methods (DEM) *Int. J. Eng. Comput.* **9** no 2 (special issue)
- Nagtegaal J C 1982 *Comput. Methods Appl. Mech. Eng.* **33** 469–84
- Nemat-Nasser S 1982 *Int. J. Solids Struct.* **18** 857–72
- Ogden R W 1984 *Non-linear Elastic Deformations* (Chichester: Ellis Horwood)
- Onat E T and Leckie F A 1988 *J. Appl. Mech.* **55** 1–10
- Oden J T 1972 *Finite Elements in Nonlinear Continua* (New York: McGraw-Hill)
- Oden J T, Demkowicz L, Rachowicz W and Westermann T A 1989 *Comput. Methods Appl. Mech. Eng.* **77** 113–80
- Oden J T and Martins J A C 1985 *Comput. Methods Appl. Mech. Eng.* **52** 527–634
- Ortiz M 1996 *Comput. Mech.* **18** 321–38
- Ortiz M and Popov E P 1985 *Int. J. Numer. Methods Eng.* **21** 1561–76
- Ortiz M and Quigley J J IV 1991 *Comput. Methods Appl. Mech. Eng.* **90** 781–804
- Owen D R J and Hinton E 1980 *Finite Elements in Plasticity: Theory and Practice* (Swansea: Pineridge)
- Owen D R J and Perić D 1994 *Finite Elem. Anal. Des.* **18** 1–15
- Owen D R J and Vaz M Jr 1998 *Comput. Methods Appl. Mech. Eng.* in press
- Papadrakakis M 1997 *Parallel Solution Methods in Computational Mechanics* (Chichester: Wiley)
- Paw G F and Owen D R J 1991 *Comput. Syst. Eng.* **2** 451–60
- Perić D 1993 *Int. J. Numer. Methods Eng.* **36** 1365–93
- Perić D, Hochard C, Dutko M and Owen D R J 1996 *Comput. Methods Appl. Mech. Eng.* **137** 331–44
- Perić D, Mitchell G P and Owen D R J 1991 *Int. J. Numer. Methods Eng.* **32** 79–101
- Perić D and Owen D R J 1992a *Finite Inelastic Deformations—Theory and Applications* ed D Besdo and E Stein (Berlin: Springer) pp 299–312
- 1992b *Int. J. Numer. Methods Eng.* **35** 1289–309
- Perić D, Owen D R J and Honnor M E 1992 *Comput. Methods Appl. Mech. Eng.* **94** 35–61
- Perić D, Yu J and Owen D R J 1994 *Int. J. Numer. Methods Eng.* **37** 1351–79
- Perić D, Vaz M Jr and Owen D R J 1998 *Comput. Methods Appl. Mech. Eng.* in press
- Recht R F 1964 *J. Appl. Mech. Trans. ASME* **31** 189–93
- Reddy B D and Simo J C 1995 *SIAM J. Numer. Anal.* **32** 1705–28
- Rubinstein R and Atluri S N 1983 *Comput. Methods Appl. Mech. Eng.* **36** 277–90
- Saad Y 1986 *SIAM J. Sci. Stat. Comput.* **7** 856–69
- Satake M and Jenkins J T (ed) 1988 *Micromechanics of Granular Materials* (Amsterdam: Elsevier)
- Schönauer M, Rodič and Owen D R J 1993 *J. Physique IV* **3** 1199–209
- Schweizerhof K, Nilsson L and Hallquist J O 1992 *Int. J. Comput. Appl. Technol.* **5** 134–56
- Sekhon G S and Chenot J L 1993 *Eng. Comput.* **10** 31–48
- Simo J C 1992 *Comput. Methods Appl. Mech. Eng.* **99** 61–112
- Simo J C and Armero F 1992 *Int. J. Numer. Methods Eng.* **33** 1413–49
- Simo J C and Hughes T J R 1987 *Constitutive Laws for Engineering Materials: Theory and Applications* ed C S Desai *et al* (New York: Elsevier) pp 221–31
- Simo J C and Miehe C 1992 *Comput. Methods Appl. Mech. Eng.* **98** 41–104
- Simo J C and Rifai S 1990 *Int. J. Numer. Methods Eng.* **29** 1595–1638
- Simo J C and Taylor R L 1985 *Comput. Methods Appl. Mech. Eng.* **48** 101–18
- 1991 *Comput. Methods Appl. Mech. Eng.* **85** 273–310
- Simo J C, Taylor R L and Pister K S 1985 *Comput. Methods Appl. Mech. Eng.* **51** 177–208
- Sonnenveld P 1989 *SIAM J. Sci. Stat. Comput.* **10** 36–52
- Suh N P and Sin H C 1981 *Wear* **69** 91–114
- Sussman T and Bathe K-J 1987 *Comput. Struct.* **26** 357–409
- Taylor R L, Beresford P J and Wilson E L 1976 *Int. J. Numer. Methods Eng.* **10** 1211–19
- Truesdell C 1969 *Rational Thermodynamics* (New York: McGraw-Hill)

- Truesdell C and Noll W 1965 The nonlinear field theories of mechanics *Handbuch der Physik* vol III/3, ed S Flügge (New York: Springer)
- Turner M J, Clough R W, Martin H C and Topp L J 1956 *J. Aero. Sci.* **23** 805–23
- van der Vorst H 1992 *SIAM J. Sci. Stat. Comput.* **13** 631–44
- Vaz M Jr 1998 Computational approaches to metal cutting processes *PhD Thesis* (Swansea: University of Wales)
- Weatherill N P 1990 *Grid Generation* Von-Karman Institute for Fluid Dynamics, Lecture series 1990-10
- Weber G and Anand L 1990 *Comput. Methods Appl. Mech. Eng.* **79** 173–202
- Williams J R (ed) 1995 2nd US Conf. on discrete element methods (DEM) *Int. J. Eng. Comput.* **12** no 2 (special issue)
- Wriggers P 1995 *Arch. Comput. Meth. Eng.* **2** 1–49
- Wriggers P and Reese S 1996 *Comput. Methods Appl. Mech. Eng.* **135** 201–9
- Wriggers P, Vu Van T and Stein E 1990 *Comput. Struct.* **37** 319–31
- Zienkiewicz O C, Huang G C and Liu Y C 1990 *Int. J. Numer. Methods Eng.* **30** 1527–53
- Zienkiewicz O C and Taylor R L 1989 *The Finite Element Method. Volume 1: Basic Formulation and Linear Problems* 4th edn (Maidenhead: McGraw-Hill)
- 1991 *The Finite Element Method. Volume 2: Solid and Fluid Mechanics, Dynamics and Non-linearity* 4th edn (Maidenhead: McGraw-Hill)
- Zienkiewicz O C and Zhu J Z 1987 *Int. J. Numer. Methods Eng.* **24** 337–57