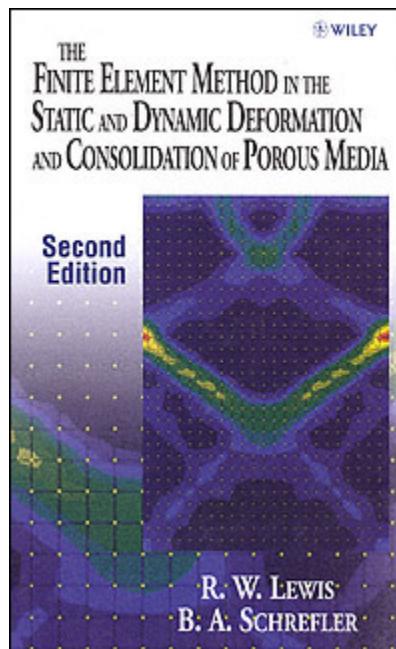


[cover](#)[next page >](#)

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

---

**title:** The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media  
**author:** Lewis, R. W.  
**publisher:**  
**isbn10 | asin:** 0471928097  
**print isbn13:** 9780471928096  
**ebook isbn13:** 9780585294810  
**language:**  
**subject:**  
**publication date:**  
**lcc:**  
**ddc:**  
**subject:**

[cover](#)[next page >](#)

[< previous page](#)[page\\_iii](#)[next page >](#)

Page iii

The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media

Second Edition

R W. Lewis

*University of Wales Swansea, UK*

B. A. Schrefler

*University of Padua, Italy*

**JOHN WILEY & SONS**

Chichester • Weinheim • New York • Brisbane • Singapore • Toronto

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[< previous page](#)[page\\_iii](#)[next page >](#)

[< previous page](#)[page\\_iv](#)[next page >](#)

Page iv

Copyright © 1998 by John Wiley & Sons Ltd,  
Baffins Lane, Chichester,  
West Sussex PO19 1UD, England

*National* 01243 779777

*International* (+44) 1243 779777

e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk

Visit our Home Page on <http://www.wiley.co.uk>

or

<http://www.wiley.com>

Reprinted January 2000

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency, 90 Tottenham Court Road. London WIP 9HE. UK, without the permission in writing of the Publisher.

*Other Wiley Editorial Offices*

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

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

Jacaranda Wiley Ltd, 33 Park Road, Milton,  
Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01.  
Jin Xing Distripark, Singapore 0512

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

*Library of Congress Cataloging-in-Publication Data*

Lewis, R. W. (Roland Wynne)

The finite element method in the static and dynamic deformation  
and consolidation of porous media / R. W. Lewis, B. A. Schrefler.  
2nd ed.

p. cm.

Rev. ed. of: The finite element in the deformation and  
consolidation of porous media / Roland W. Lewis, Bernard A.  
Schrefler. 1987.

Includes bibliographical references and index.

ISBN 0-471-92809-7

1. Porous materials-Mathematical models. 2. Finite element  
method. 3. Multiphase flow-Mathematical models. I. Schrefler, B. A.  
II. Lewis, R. W. (Roland Wynne). Finite element method in the  
deformation and consolidation of porous media. III. Title.

TA418.9.P6L49 1998

624.1 '5136-dc21

98-12080

CIP

*British Library Cataloguing in Publication Data*

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

ISBN 0 471 92809 7

Typeset on 10/12pt Times by Thomson Press (India) Ltd., New Delhi

Printed and bound in Great Britain by Bookcraft (Bath) Ltd

This book is printed on acid-free paper responsibly manufactured from sustainable

forestry, in which at least two trees are planted for each one used for paper production.

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[< previous page](#)

page\_iv

[next page >](#)

[\*\*< previous page\*\*](#)[\*\*page\\_v\*\*](#)[\*\*next page >\*\*](#)

Page v

*For Celia and Chantal*

Start of Citation[PU]John Wiley &amp; Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[\*\*< previous page\*\*](#)[\*\*page\\_v\*\*](#)[\*\*next page >\*\*](#)

[\*\*< previous page\*\*](#)**page\_vii**[\*\*next page >\*\*](#)**Page vii**

## Contents

Preface	xv
1	
Introduction	1
References	5
2	
Mechanics of Saturated and Partially Saturated Porous Media	9
2.1 Introduction	9
2.2 Averaging Principles	9
2.2.1 Averaging Process	11
2.2.2 Microscopic Balance Equations	13
2.2.3 Macroscopic Balance Equations	14
2.3 Macroscopic Balance Equations for a Non-Isothermal Partially Saturated Porous Material	18
2.3.1 Kinematic Equations	18
2.3.2 Mass Balance Equations	21
2.3.2.1 Solid Phase	21
2.3.2.2 Liquid Phase: Water	22
2.3.2.3 Gaseous Phases: Dry Air and Vapour	22
2.3.3 Linear Momentum Balance Equation	24
2.3.4 Angular Momentum Balance Equation	28
2.3.5 Balance of Energy Equation	29
2.3.6 Entropy Inequality	32
2.4 Constitutive Equations	34
2.4.1 Stress Tensor in the Fluid Phases	35
2.4.2 Gaseous Mixture of Dry Air and Water Vapour	35

2.4.3 Sorption Equilibrium	35
2.4.4 Clausius-Clapeyron Equation	36
2.4.5 Pore Size Distribution	37
2.4.6 Equation of State for Water	37
2.4.7 Darcy's Law	39
2.4.8 Fick's Law	40
2.4.9 Stress Tensor in the Solid Phase and Total Stress	41
2.4.10 Solid Density	43
2.4.11 Fourier's Law	44
2.5 General Field Equations	44
2.5.1 Mass Balance Equation	44
2.5.2 Linear Momentum Balance Equation	48
2.5.2.1 Fluids	48
2.5.2.2 Solid Phase	49
2.5.2.3 Multiphase Medium	49

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[< previous page](#)

page\_vii

[next page >](#)

[< previous page](#)

page\_viii

[next page >](#)

Page viii

2.5.3 Energy Balance Equation	50
2.5.3.1 Enthalpy Balance Equation for the Multiphase Medium	53
2.5.4 Summary of Governing Equations	54
2.5.4.1 Mass Balance Equations or Continuity Equations	54
2.5.4.2 Linear Momentum Balance Equations	56
2.5.4.3 Enthalpy Balance: Multiphase Medium	56
2.6 Physical Approach: Extended Biot Theory	56
2.6.1 The Physical Model	57
2.6.2 Constitutive Equations	61
2.6.3 Governing Equations	63
2.6.3.1 Linear Momentum Balance Equation: Multiphase Medium	63
2.6.3.2 Mass Balance Equations	64
2.6.3.3 Energy Balance Equation	66
References	68
Appendix 2A	71
Appendix 2B	72
Appendix 2C	72
 3 Numerical Solution for Isothermal Consolidation	75
3.1 Introduction	75
3.2 Coupled Solution: Saturated One-Phase Flow in a Deforming Porous Medium	75
3.2.1 Governing Equations	75
3.2.2 Initial and Boundary Conditions	77
3.3 Solution of the Boundary Value Problem	77
3.4 Application of the Finite Element Method	79
3.5 Choice of Elements	83
3.6 Discretisation in Time	84

3.7 Numerical Properties of the Time Discretisation	85
3.8 Saturated-Unsaturated Flow in a Deforming Porous Medium: One-Phase Flow	86
3.8.1 Governing Equations	87
3.8.2 Initial and Boundary Conditions	88
3.9 Discretisation of the Governing Equations for the Consolidation of Partially Saturated Soils	88
3.10 Stability, Convergence and Consistency in the Non-Linear Case	89
3.11 Airflow and Water Flow in a Deforming Porous Medium	93
3.11.1 Governing Equations	93
3.11.2 Initial and Boundary Conditions	94
3.12 Discretisation of the Governing Equations for Air and Water Flow in Deforming Porous Media	95
References	97

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[\*\*< previous page\*\*](#)

**page\_viii**

[\*\*next page >\*\*](#)

[\*\*< previous page\*\*](#)[\*\*page\\_ix\*\*](#)[\*\*next page >\*\*](#)[\*\*Page ix\*\*](#)

4

## Solid-Phase Constitutive Relationships, Variable Permeabilities and Solution Procedures

99

4.1 Introduction	99
4.2 Stress Invariants	100
4.3 Linear Elastic Analysis	102
4.4 Variable Elastic Analysis	103
4.4.1 Bilinear Models	104
4.4.2 Variable Elastic Model	105
4.4.2.1 Hyperbolic Model	105
4.4.2.2 $E$ - $v$ and $K$ - $G$ Variable Elastic Models	107
4.4.2.3 Spline functions	107
4.4.3 Thermo-Elastic Behaviour	107
4.4.4 Solution Procedures	109
4.5 Elastoplastic Models	112
4.5.1 Constitutive Law	112
4.5.2 Mohr-Coulomb Yield Surface	114
4.5.3 Critical State Model	118
4.5.3.1 Modified Cam Clay Model	118
4.5.3.2 $p$ - $q$ - $\theta$ Critical State Model	122
4.5.4 Corners of Yield and Potential Surfaces	124
4.5.5 Generalised Plasticity	124
4.5.6 Thermo-Plastic Behaviour	126
4.5.7 Solution Procedures	129
4.5.7.1 Explicit Algorithms	129
4.5.7.2 Implicit Algorithms	130
4.5.7.3 Consistent Stiffness Matrix	131

4.6 Partially Saturated Models	131
4.6.1 Elastic Behaviour	133
4.6.2 Plastic Behaviour	134
4.7 Variation of Permeability	136
4.8 Conclusions	139
References	139
5 Verification of Elastic and Elastoplastic Consolidation Programs	145
5.1 Introduction	145
5.2 Elastic Solutions for Drained and Undrained Conditions	146
5.2.1 Plane Strain, Uniform Loading	147
5.2.2 Radially Symmetric, Uniform Loading	149
5.3 Elastic Analysis of Consolidation under Strip and Circular Uniform Loading	152
5.4 Elastoplastic Solutions	154
5.4.1 Undrained Triaxial Tests On Normally Consolidated Soil	154
5.4.1.1 Mohr-Coulomb Analysis	154
5.4.1.2 Critical State Ellipse Analysis	155
5.4.2 Drained and Undrained Analyses of Strip Loading	156
5.4.2.1 Mohr-Coulomb Analysis	156
5.4.2.2 Critical State Ellipse Analysis	157

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[\*\*< previous page\*\*](#)

**page\_ix**

[\*\*next page >\*\*](#)

[\*\*< previous page\*\*](#)[\*\*page\\_x\*\*](#)[\*\*next page >\*\*](#)[\*\*Page x\*\*](#)

5.4.2.3 Analysis by Critical State Ellipse, with Mohr-Coulomb Cut-off ( $c = 0$ )	158
5.5 Elastoplastic Analysis of Consolidation under Uniform Loading and Strip Loading	158
5.5.1 One-Dimensional Consolidation	159
5.5.2 Two-Dimensional Consolidation	161
5.6 Linear and Non-Linear Elastic Consolidation with Variable Permeability	163
5.6.1 Consolidation of Swansea Blue Clay in a Rowe Consolidation Cell	163
5.6.2 Consolidation of Kaolin in a Rowe Consolidation Cell	165
5.7 Multiphase Flow in Porous Media: A Benchmark Problem for Non-Saturated Flow	167
5.8 Conclusions	174
References	175
 6 Modelling Subsidence: Numerical Aspects and Problems of Regional Scale	 177
6.1 Introduction	177
6.1.1 More about Coupling and Staggered Procedures	181
6.2 Problems of Regional Scale: Vertically Averaged Models	184
6.2.1 Spatially Averaged Quantities: The Megascopic Level	184
6.2.2 Macrolevel Governing Equations	186
6.2.2.1 Equilibrium Equation for the Two-Phase Medium	186
6.2.2.2 Fluid-Phase Behaviour	187
6.2.3 Implementation of the Numerical Model	188
6.3 Far-Field Boundary Conditions	192
6.3.1 Infinite Elements	193
6.4 A Coupled Solution for the Settlement above Gas Reservoirs	197
6.5 Single-Aquifer Withdrawal	201
6.5.1 Isolated Aquifer	202

6.5.2 Embedded Aquifer	206
6.6 Conclusions	208
References	209
7 Modelling Subsidence: Case Studies	213
7.1 Introduction	213
7.2 The Subsidence of Venice	214
7.2.1 Background	214
7.2.2 The Mathematical Model	220
7.2.3 Results	224
7.3 Subsidence in the Po Delta and the Polesine	230
7.3.1 Background	230
7.3.2 The Contarina Model	232
7.3.2.1 Available Data	232
7.3.2.2 Parametric Investigation	235
7.3.2.3 Subsidence Rebound	245

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[\*\*< previous page\*\*](#)

**page\_x**

[\*\*next page >\*\*](#)

[\*\*< previous page\*\*](#)[\*\*page\\_xi\*\*](#)[\*\*next page >\*\*](#)[Page xi](#)

7.4 Subsidence above Gas Reservoirs: The Ravenna Case	248
7.4.1 Background	248
7.4.2 Results	251
7.4.3 Comparison with a Volumetric Reservoir	257
7.4.4 New Results for the Ravenna Field	262
7.5 Subsidence of Abano Terme	268
7.5.1 Background	268
7.5.2 The Mathematical Model	273
7.6 Conclusions	276
References	277
 8 Modelling Three-Phase Flow in Deforming Saturated Oil Reservoirs	 281
8.1 Introduction	281
8.2 Development of the Governing Equations	282
8.2.1 The Equilibrium Equation for a Three-Phase System	282
8.2.2 Three-Phase Flow Equations	283
8.3 Application of the Finite Element Method	284
8.4 Numerical Procedures	287
8.4.1 Treatment of Fluid Non-Linear Terms	287
8.4.2 Stability Analysis	288
8.4.3 Mass Balance and Convergence Checks	289
8.4.4 Computational Procedures	290
8.5 Validation and Applications	290
8.5.1 Non-Linear Soil Column Analysis	291
8.5.2 Reservoir Compaction Problems	293
8.5.3 Surface Subsidence Analyses	298
8.5.3.1 Effect of Reservoir Parameters	300

8.5.3.2 Effect of Water Injection Schemes On Subsidence Analysis	302
8.6 Conclusions	304
References	305
9	
Fractured Reservoir Simulation	307
9.1 Introduction	307
9.2 Description of the Model	308
9.3 Development of the Governing Equations	310
9.4 A special Case: Single-Phase Flow in a Deforming Fractured Porous Medium	312
9.5 Discretisation in Space	313
9.6 Validation of the Model	322
References	338
10	
Heat and Fluid Flow in Deforming Porous Media	341
10.1 Introduction	341
10.2 Non-Isothermal Fully Saturated Consolidation	344
10.2.1 Governing Equations	344
10.2.2 Initial and Boundary Conditions	345

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[\*\*< previous page\*\*](#)

**page\_xi**

[\*\*next page >\*\*](#)

[\*\*< previous page\*\*](#)**page\_xii**[\*\*next page >\*\*](#)**Page xii**

10.3 Discretisation for Non-Isothermal Consolidation of Saturated Porous Media	346
10.4 Solution Procedures	348
10.4.1 Monolithic Augmentation Approach	349
10.4.2 Partitioned Solution Procedures	349
10.4.2.1 Numerical Properties of Partitioned Procedures	352
10.5 Non-Isothermal Airflow and Water Flow in a Deforming Porous Medium	354
10.5.1 Governing Equations	354
10.5.2 Initial and Boundary Conditions	356
10.6 Discretisation for Airflow and Water Flow in a Deforming Porous Medium	358
10.7 Numerical Examples	364
10.7.1 Thermo-Elastic Consolidation	364
10.7.2 Thermo-Elastoplastic Consolidation	368
10.7.3 Thermo-Elastic Consolidation around a Cylindrical Heat Source	370
10.7.4 Non-Isothermal Consolidation	376
10.7.5 Thermo-Elastic Consolidation of Partially Saturated Clay	379
10.8 Conclusions	386
References	386
Appendix 10A	389
Appendix 10B	394
11 Secondary Consolidation Creep in Solids	397
11.1 Introduction	397
11.2 Formulation of Secondary Consolidation	398
11.3 Application of the Creep Model	400
11.3.1 Background	400
11.3.2 Mathematical Modelling	403
11.3.3 Results	405

References	408
12	
Soil-Structure Interaction	409
12.1 Introduction	409
12.2 Governing Equations	410
12.3 Material Models	412
12.3.1 Interface Behaviour	412
12.3.2 Soil Behaviour	413
12.4 Applications	414
12.4.1 Test 1: Shallow Foundation	414
12.4.2 Test 2: Pile-Soil Interaction	418
12.4.3 Test 3: Frame on Soft Soil	421
12.4.3.1 Interaction of Two Adjacent Footings	424
12.4.3.2 Eccentric Loading	425
12.4.3.3 Horizontal Loading	425
12.5 Conclusions	425
References	425

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[\*\*< previous page\*\*](#)

**page\_xii**

[\*\*next page >\*\*](#)

[\*\*< previous page\*\*](#)**page\_xiii**[\*\*next page >\*\*](#)[Page xiii](#)

13		
Back Analysis in Consolidation		427
13.1 Introduction		427
13.2 Definition of Back Analysis		428
13.3 Methodology		429
13.3.1 Direct Method		429
13.3.2 Indirect Method		430
13.3.3 Probabilistic Approaches		431
13.3.4 Alternative Methods		431
13.4 Parameter Identification		432
13.4.1 Optimisation Methods		433
13.4.1.1 Simplex Method		433
13.4.1.2 Rosenbrock's Algorithm		434
13.4.1.3 Levenberg-Marquardt Method		434
13.4.2 Sensitivity Analysis		434
13.5 Case Study		436
13.5.1 Background		436
13.5.2 Hypothetical Case Study		437
13.6 Summary		442
References		443
14		
Large-Strain Quasi-Static and Dynamic Soil Behaviour		445
14.1 Introduction		445
14.2 Kinematic Equations		445
14.3 Constitutive Equations		448
14.4 Governing Equations and Their Weak Form		451
14.5 The Rate Form of Stress Power		453

14.6 Finite Element Discretisation	453
14.6.1 Spatial Discretisation	454
14.6.2 Discretisation in Time and Solution Procedure	456
14.7 Examples	459
14.7.1 Finite-Strain and Small-Strain Fully Saturated Consolidation	459
14.7.2 Finite-Strain and Small-Strain Partially Saturated Consolidation	461
14.7.3 Slope under Seismic Behaviour: Finite Strains	463
14.7.4 Dynamic Strain Localisation	464
14.8 Conclusions	473
References	474
Subject Index	477

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[< previous page](#)

page\_xiii

[next page >](#)

[< previous page](#)[page\\_xv](#)[next page >](#)

Page xv

## Preface

Our first text on this subject '*The Finite Element Method in the Deformation and Consolidation of Porous Media*', was published ten years ago and has been out of print for much of the past decade. It was the first book of its kind, despite the many available texts on groundwater flow through deforming porous media. The topic has been covered, albeit briefly, in many texts on geomechanics, petroleum engineering and finite element methods. However, there still exists no other book which covers all the mechanical and numerical aspects of flow in porous media in such detail.

In the intervening period there was a rapid expansion in the research and practical applications of these types of problem, which has prompted us to write this new and thoroughly updated version. It contains not only the results of research carried out at our two institutions but also reports on the work done under various European research programmes, e.g. Science (Greco Geomateriaux), TEMPUS PHARE (with the Technical University of Lodz and the Polish Academy of Sciences IPPT-PAN), and in particular Human Capital and Mobility, where an Alliance of Laboratories in Europe for Research and Technology (ALERT) was created, concentrating on research in geomaterials (soil, rock and concrete). Both our institutions were partners in this network, and the scientific exchanges proved to be extremely fruitful. Also, collaborative work carried out with the Norwegian Geotechnical Institute, under the BRINORD agreement, contributed to a better understanding of petroleum reservoir subsidence.

The chapters from the previous edition have been extensively updated and several new chapters have been added to give a much broader coverage of recent research interests. The theoretical part of the book is completely new: it now incorporates both phenomenological and averaging approaches.

We are indebted to many of our coworkers and in particular we thank Drs N. Abd. Rahman, P. Baggio, G. Bolzon, D. Gawin, H.R. Ghafouri, C.E. Majorana, E.A. Meroi, R.S. Ransing, V. Salomoni, L. Sanavia, L. Simoni, Y. Sukirman, D.V. Tran, E. Turska, X. Wang, X. Zhan, H.W. Zhang and Y. Zheng, who over the years have contributed to the work. Also, many thanks to Drs S.M. Hassanzadeh and D. Pigozzi for their advice on the theoretical chapter.

Finally, we would like to dedicate this book to the two ladies in our lives, Celia and Chantal, without whom it might never have been completed.

ROLAND W. LEWIS  
BERNARD A. SCHREFLER  
SWANSEA/ PADUA  
JULY 1997

Start of Citation[PU]John Wiley & Sons, Ltd. (UK)[/PU][DP]1998[/DP]End of Citation

[< previous page](#)[page\\_xv](#)[next page >](#)

# 1

## Introduction

Porous materials are classified as those with an internal structure. They comprise a solid phase and closed and open pores. Attention is focused in this book on the case where the open pores are filled with one or more fluids, i.e. a multiphase media. In the case of geomaterials, i.e. soil, rock, concrete, the fluids may be water, water vapour and dry air. Soil and rock may also contain gas, oil and/or water (reservoir engineering). The solid and the fluids usually have a relative velocity to each other, and because of this and the different material properties that exist, there is interaction between the constituents. Furthermore, the pore structure has, in general, an extremely complicated geometry which makes a mechanical description of the problem rather difficult. For engineering purposes a substitute model at the macroscopic scale is normally assumed where the interacting constituents are presumed to occupy the entire control space. This distribution is obtained by means of the volume fraction concept. Volume fractions are given by the ratio of the volume of the constituents to the volume of the control space. A consequence of the volume fraction concept is that the substitute constituents have reduced densities. These substitute continua may be treated via the methods of continuum mechanics.

Two strategies are generally used to arrive at the description of the behaviour of these substitute continua: one starts from a macromechanics viewpoint and the other from a micromechanics viewpoint. Phenomenological and mixture theory approaches, integrated by the concept of volume fractions, belong to the first strategy. Averaging theories, sometimes called hybrid mixture theories, belong to the second one.

Porous media theory has been of interest to research workers for a considerable time. An extensive review of the history of such porous media theories is given by de Boer [1]. This text mentions only a few important contributions. Woltman [2] introduced the concept of volume fractions and Delesse [3] dealt with surface fractions. Fick [4] studied the problem of diffusion of mixtures. Darcy [5] the motion of a liquid in a porous solid, and Stefan [6] the diffusion of gas through a porous diaphragm. Fillunger [7] introduced the concept of effective stresses and studied the problems of uplift, friction and capillarity in rigid, liquid-saturated porous solids. Terzaghi [8] investigated saturated deformable porous solids and also made use of the effective stress principle. In the years 1913 to 1934 both authors made many important contributions to modern soil mechanics [1]. Biot [9,10] developed the phenomenological approach of Terzaghi further and extended it to the three-dimensional case.

Modern mixture theories were developed by Morland [11], Goodman and Cowin [12], Sampaio and Williams [13] and Bowen [14,15]. Averaging theories were developed by

Hassanizadeh and Gray [16–18] and by Whitaker [19]. Finally a macroscopic thermodynamical approach to Biot's theory was used by Coussy [20].

This book is essentially devoted to consolidation and other slow (quasi-static) phenomena in porous media and their quantitative solution. Only one chapter deals with dynamic phenomena; they are dealt with much more extensively by Zienkiewicz *et al.* [21]. Consolidation plays an important role in many soil mechanics problems, which is evidenced by the vast amount of literature devoted to the solution of this problem since the pioneering work of Terzaghi [8].

A survey of the literature indicates two main areas where consolidation analysis is extensively applied. The first one is connected with the physical loading of soil layers and has probably received the greater attention. This aspect comprises the transient analysis of footings, pile foundations, soil–structure interaction, embankments, large fill dams, etc.

The second aspect of consolidation, connected with the change of hydraulic equilibrium in a system comprising aquifers and aquitards, has received attention only in recent times when the effects of extensive groundwater withdrawal for industrial and agricultural purpose, and the effects of extensive oil and/or gas pumpage became evident. Depending on the geologic environment of the exploited aquifers and/or hydrocarbon reservoirs, the heavy pumpage can result in large surface subsidence with possible damage to property [22].

The case histories of the San Joaquin Valley in California, Galveston, Mexico City, Tokyo and Venice, where groundwater was withdrawn, have been well documented. Extraction of oil or gas was the cause of subsidence in the Wilmington oilfield (California), Goose Creek (Texas), Lake Maracaibo (Venezuela), Po Delta and Ravenna (Italy), Niigata (Japan) and also in the North Sea. The list is by no means exhaustive. The problems are regional problems, where horizontal distances of interest are much larger than the thickness of the aquifer–aquitard systems.

The same type of problems exist, but on a local scale, where settlements are induced by drain wells or sand drains, by rock tunnelling beneath an aquifer, or by the change of the surface infiltration characteristics in an urban area. In all these cases the consequent decrease of pore pressure and resulting increase of effective (grain to grain) stress, produces compaction which continues until a new state of hydraulic equilibrium is achieved.

In engineering practice, settlements were calculated in most cases using Terzaghi's [8] one-dimensional consolidation theory. More recently Biots' [9–10], three-dimensional theory has been used, based on a linear stress-strain constitutive relationship and also a linear form of Darcy's flow rule. The extensive use of computers and the concomitant development of numerical techniques has made possible much more precise analyses. The non-linear behaviour of the skeleton and variation of permeability with strain can now be easily taken into account, if necessary. These possibilities shift attention from the question of problem solving to the problem of modelling observed phenomena. The eventual aim is to develop the capability for making predictive simulations.

This textbook is concerned with the development of mathematical models which are sufficiently general to study the phenomena related to heat and mass transfer in fully and partially saturated deforming porous media and for a numerical solution of the resulting governing equations. Also, the problem of modelling is addressed in detail.

The mathematical model for coupled thermo-hydro-mechanical processes is developed in Chapter 2. Two approaches have been used for this purpose. First come the averaging theories of Hassanzadeh and Gray [16–17], called ‘hybrid mixture theories’ by Achanta *et al.* [23]. The equations are then derived by extending Biot’s theory [9,10], i.e. using the classical phenomenological approach. Hybrid mixture theories yield a better understanding of the microscopic situation but the development of the necessary equations is rather lengthy. The reader can choose the theory he or she prefers.

Chapter 3 describes the finite element formulation of the governing equations for one-phase flow and saturated–unsaturated flow in deforming porous media. For the latter case, it looks at the situation where the air remains at atmospheric pressure and the situation where the airflow is taken into account. The numerical properties of the discretisation in time are extensively examined.

Chapter 4 considers the stiffness matrix and the flow matrix together with solution procedures. Both linear and non-linear elastic and elastoplastic constitutive relationships, which are suitable for applications in quasi-static and dynamic situations, are discussed. The plasticity theory comprises a Mohr–Coulomb and critical state model. Further generalised plasticity is addressed, particularly useful in soil mechanics, and models for fully and partially saturated porous media are shown within this framework. The influence of temperature on the constitutive relationships of the solid phase is also taken into account. Finally, the variations of permeability with strain are also considered.

Chapter 5 presents a series of test problems for the verification of elastic and elastoplastic consolidation programs involving both fully and partially saturated flow in deforming porous media. Limit situations, such as drained and undrained conditions, are considered together with the consolidation aspects. Examples include a variable permeability scheme. This chapter deals with the physical loading of soil layers.

Chapter 6 is the first of three chapters dealing with surface subsidence analysis. It addresses the numerical aspects of modelling of regional-scale problems, particularly coupling and uncoupling; staggered solution schemes are also presented; and a vertically averaged model is derived. The problem of far-field boundary conditions is dealt with, where finite and infinite elements are used for this purpose. A coupled solution for the settlement above gaseous reservoirs is given, where the full solution for a two-phase flow problem is avoided. Finally, single-aquifer withdrawal is studied in detail and numerical solutions are compared with analytical solutions.

Chapter 7 applies the theories outlined in the previous sections, along with the appropriate constitutive relationships, to the study of real case histories of surface subsidence and to parametric investigations. The various aspects of modelling are discussed in detail along with the difficulties encountered. For Venice it is shown that the proposed model is capable of making predictive simulations, since the predicted values match the measured values, known *a posteriori*. The applications in general indicate that a fully coupled consolidation model is a very useful and successful tool for simulating surface subsidence.

Chapter 8 develops a fully coupled system of governing equations for the case of immiscible and compressible three-phase fluid flow in a compacting three-dimensional saturated oil reservoir. The four simultaneous equations consist of the equilibrium and continuity equations, where displacement and fluid pressures are the primary unknowns. The effects of some non-linear dependent terms are taken into account when formulating

the continuity equation for each flowing phase. Chapter 8 also presents a detailed discretisation process of the final governing equations, along with the stability requirements. It concludes with a series of simulations for the purposes of validation and application.

Chapter 9 extends the formulation derived in the previous chapter for the case of naturally fractured reservoirs. The nine fully coupled governing equations are presented along with the boundary condition relationships for a rather general case where three different fluid phases exist and a double-porosity model is assumed. A three-dimensional field-scale example demonstrates the utility of the model and illustrates the nature of fluid flow in fractured reservoirs.

Chapter 10 takes into account the heat flow and the fluid flow through deforming porous media, both fully and partially saturated. This enables investigation of non-isothermal consolidation and drying phenomena, since phase change is included. Examples are given for thermoelastic and thermoelastoplastic consolidation in limited and unlimited domains.

Chapter 11 investigates the phenomenon of creep and proposes a secondary consolidation model, subsequently used to examine the behaviour of an experimental embankment test.

Chapter 12 covers the problems of interaction between structures, e.g. footings, pavements, roadways, and the underlying soil, with special reference to transient consolidation. Interfacial behaviour, structural stiffness and soil reaction are detailed and examples are solved for a shallow foundation, pile–soil interaction and a frame on soft soil.

Chapter 13 discusses the concepts of ‘back analysis’ for consolidation problems. The difficulties of determining accurate *in situ* measurements for these situations are mentioned, and the indirect method and probabilistic approaches are covered in some detail. The concluding case study epitomises the concepts involved in a satisfactory solution.

Chapter 14 studies isothermal dynamic and slow phenomena in fully and partially saturated soil, under the assumption of finite strains. In partially saturated conditions the air phase always remains at atmospheric pressure. The chapter concludes with examples of seismic analysis and dynamic strain localisation.

Having overviewed the contents, here is a brief summary of related work. Sandhu and Wilson [24], Schiffman *et al.* [25], Christian and Boehmer [26], Hwang *et al.* [27] and Yokoo *et al.* [28] have applied finite element techniques to the numerical analysis of soil consolidation along the lines of Biot’s self-consistent elastic theory. Compressibility of fluid was introduced by Ghaboussi and Wilson [29]. Lewis *et al.* [30] were the first to consider the important aspect of permeability changes during consolidation. Schrefler *et al.* [31] used Biot’s theory for surface subsidence modelling. Small *et al.* [32] introduced plasticity in consolidation analysis, using the Mohr–Coulomb model, whereas Runesson [33] and Desai and Siriwardane [34] used the critical state model. Zienkiewicz *et al.* [35] also dealt with elastoplasticity and further introduced compressibility of the solid grains. Runesson [33] and Prevost [36] derived a numerical model for consolidation analysis from the theory of mixtures. A finite strain formulation of consolidation was used by Carter *et al.* [37]. Norris [38] extensively studied the constitutive models by introducing kinematic hardening. All these authors, as well as this textbook, make use of displacements and pore pressures as basic variables. An equilibrium finite element

model for soil consolidation using total stresses and pore pressures as basic variables was proposed by Cividini and Gioda [39].

Fully saturated soil dynamics was studied by Zienkiewicz [40] and extended to finite strains by Zienkiewicz and Shiomi [41]. Models for fully saturated soil dynamics, based on the mixture theory, were also put forward by Prevost *et al.* [42] and Ehlers and Kubik [43]. Subsidence analysis of saturated–unsaturated soil was carried out by Safai and Pinder [44]. Several terms were neglected in that analysis; they were successively included in the model presented by Schrefler and Simoni [45]. The flow of only one fluid phase is considered in these models. A model for the quantitative study of static and dynamic isothermal behaviour of fully and partially saturated soils was put forward by Zienkiewicz and coworkers in two companion papers [46,47]. The model was based on the extended Biot theory, where for unsaturated conditions the air pressure was assumed to remain equal to the ambient air pressure. This model was then extended to finite strain assumptions by Meroi *et al.* [48].

Subsidence above compacting multiphase reservoirs was studied by Morgan *et al.* [49] with a two-step model. Isothermal small-strain multiphase flow solutions for slow phenomena involving a wetting and a non-wetting phase (water and bitumen) were developed by Li and Zienkiewicz [50], and for water and air by Schrefler and Zhan [51]. A variational property of the boundary value problem for saturated porous media with a poroplastic constitutive model was shown by Maier and Comi [62] and a finite element model in terms of generalised variables was obtained.

Non-isothermal consolidation was studied by Bear and Corapcioglu [52] and Aboustit *et al.* [53]. Non-isothermal elastoplastic consolidation was investigated by Lewis *et al.* [54]. Geraminegad and Saxena [55] put forward a thermoelastic model for heat and mass transfer in partially saturated soil, a modified version of Philip and de Vries' formulation [56] was applied to the heat transfer equation. Non-isothermal multiphase flow of brine and gas in a deformable medium was studied by Olivella *et al.* [57]. A non-isothermal three-phase model for slow phenomena with heat and mass transfer through conduction and convection and latent heat transfer, based on hybrid mixture theories, was presented by Gawin *et al.* [58]. A similar numerical model was also put forward by Thomas and He [59].

For an easier comprehension of this textbook, a basic knowledge of the finite element method is required. This knowledge is now quite common, since most of the engineering curricula include it. Readers who are not familiar with the finite element method are referred to basic textbooks [60,61].

No listing or diskette of code is included, but it is possible to obtain a general code for thermo-hydro-mechanical coupled problems in partially saturated porous media with full two-phase flow and phase change, written according to Chapter 10. It is accessible through the Italian Network for Computational Solid Mechanics (COMES) of the Italian Research Council.

## REFERENCES

1. de Boer, R. (1996) Highlights in the historical development of the porous media theory: toward a consistent macroscopic theory. *Appl. Mech. Rev.*, **49**(4), 201–62.

2. Woltman, R. (1794) *Beytraege zur hydraulischen Architektur*, Vol. 3, Johann Christian Dietrich, Göttingen.
3. Delesse, A. (1848) Pour determiner la composition des roches. *Ann. des Mines*, series 3, **13**, 379–88.
4. Fick, A. (1855) Über Diffusion. *Ann. der Physik und Chemie*, **94**, 59–86.
5. Darcy, H. (1856) *Les fontaines publiques de la ville de Dijon*, Dalmont, Paris.
6. Stefan, J. (1871) Über das Gleichgewicht und die Bewegung, insbesondere die Diffusion von Gasmengen. *Sitzungsber. Akad. Wiss., Math.-Naturwiss.* Section IIa, **63**, 63–124.
7. Fillunger, P. (1913) Der Auftrieb in Talsperren. *Österr. Wochenschrift für den öffentl. Bau Dienst*, **19**, 532–56, 567–70.
8. von Terzaghi, K. (1923) Die Berechnung der Durchlässigkeitsziffer des Tones aus dem Verlauf der hydrodynamischen Spannungerscheinungen. *Sitzungsber. Akad. Wiss., Math.-Naturwiss.* Section IIa, **132**, (3/4), 125–38.
9. Biot, M. A. (1941) General theory of three-dimensional consolidation. *J. Appl. Phys.*, **12**, 155–64.
10. Biot, M. A. (1956) General solution of the equation of elasticity and consolidation for a porous material. *J. Appl. Mech.*, **23**, 91–96.
11. Morland, L. W. (1972) A simple constitutive theory for fluid saturated porous solids. *J. Geophys. Res.*, **77**, 890–900.
12. Goodman, M. A. and Cowin, S. C. (1972) A continuum theory for granular materials. *Arch. Rational Mech. Anal.*, **44**, 249–66.
13. Sampaio, R. and Williams, W. O. (1979) Thermodynamics of diffusing mixtures. *J. de Mécanique*, **18**, 19–45.
14. Bowen, R. M. (1980) Incompressible porous media models by use of the theory of mixtures. *Int. J. Engng Sci.*, **18**, 1129–48.
15. Bowen, R. M. (1982) Compressible porous media models by use of theories of mixtures. *Int. J. Engng Sci.*, **20**, 697–735.
16. Hassanizadeh, M. and Gray, W. G. (1979) General conservation equations for multiphase systems: 1, averaging procedure. *Adv. Water Resources*, **2**, 131–44.
17. Hassanizadeh, M. and Gray, W. G. (1979) General conservation equations for multiphase systems: 2, mass, momenta, energy and entropy equations. *Adv. Water Resources*, **2**, 191–203.
18. Hassanizadeh, M. and Gray, W. G. (1980) General conservation equations for multiphase systems: 3, constitutive theory for porous media flow. *Adv. Water Resources*, **3**, 25–40.
19. Whitaker, S. (1977) Simultaneous heat mass and momentum transfer in porous media: a theory of drying. *Advances in Heat Transfer* 13, Academic Press, New York.
20. Coussy, O. (1995) *Mechanics of Porous Continua*, Wiley, Chichester.
21. Zienkiewicz, O. C., Chan, A., Pastor, M., Schrefler, B. A. and Shiomi, T. (1999) *Computational Soil Dynamics Applied to Earthquake Engineering*, Wiley, Chichester.
22. Poland, J. F. (1972) Subsidence and its control in underground waste management and environmental implications, in *AAPG Memoir 18*, American Association of Petroleum Geologists, Tulsa, pp. 50–71.
23. Achanta, S., Cushman, J. H. and Okos, M. R. (1994) On multicomponent, multiphase thermomechanics with interfaces. *Int. J. Engng Sci.*, **32**, 1717–38.
24. Sandhu, R. S. and Wilson, E. L. (1969) Finite element analysis of flow in saturated porous media. *J. Engng Mech. Div. ASCE*, **95**(EM3), 641–52.
25. Schiffman, R. L., Chen, A. T. F. and Jordan, J. C. (1969) An analysis of consolidation theories. *J. Soil Mech. Found. Div. ASCE*, **95**(SM1), 285–312.
26. Christian, J. T. and Boehmer, J. W. (1970) Plane strain consolidation by finite elements. *J. Soil Mech. Found. Div. ASCE*, **96**(SM4), 1435–57.
27. Hwang, C. T., Morgenstern, N. R. and Murray, D. W. (1971) On solution of plane strain consolidation problems by finite element methods. *Can. Geotech. J.*, **8**, 109–18.

28. Yokoo, Y., Yamagata, K. and Nagaoka, H. (1971) Finite element method applied to Biot's consolidation theory. *Soils and Foundation, Japanese Soc. Soil Mech. Found. Engng.*, **11**(10), 29–46.
29. Ghaboussi, J. and Wilson, E. L. (1973) Flow of compressible fluid in porous elastic media. *Int. J. Num. Meth. Engng.*, **5**, 419–42.
30. Lewis, R. W., Roberts, G. K. and Zienkiewicz, O. C. (1976) A non-linear flow and deformation analysis of consolidation problems, in *Numerical Methods in Geomechanics*, Vol. 2, C. S. Desai (ed.), ASCE, New York, pp. 1106–18.
31. Schrefler, B. A., Lewis, R. W. and Norris, V. A. (1977) A case study of surface subsidence of the Polesine area. *Int. J. Num. Anal. Meth. Geomech.*, **1**, 377–86.
32. Small, J. C., Booker, J. R. and Davis, E. H. (1976) Elastoplastic consolidation of soil. *Int. J. Solids Struct.*, **12**, 431–48.
33. Runesson, K. (1978) On nonlinear consolidation of soft clay. Thesis, Chalmers University of Technology, Gothenburg.
34. Desai, C. S. and Siriwardane, T. H. J. (1979) Subsidence due to consolidation including nonlinear behaviour, *Evaluation and Prediction of Subsidence*, S. K. Saxena (ed.), ASCE, New York, pp. 500–15.
35. Zienkiewicz, O. C., Humpheson, C. and Lewis, R. W. (1977) A unified approach to soil mechanics problems including plasticity and visco-plasticity, in *Finite Elements in Geomechanics*, G. Gudehus (ed.), Wiley, London, Ch. 4.
36. Prevost, J. H. (1981) Consolidation of anelastic porous media. *J. Engng. Mech. Div. ASCE*, **107**(EM1), 169–86.
37. Carter, J. P., Booker, J. R. and Small, J. C. (1979) The analysis of finite elasto-plastic consolidation. *Int. J. Num. Anal. Meth. Geomech.*, **2**, 107–29.
38. Norris, V. A. (1980) The elasto-plastic analysis of soil consolidation with special reference to kinematic hardening. PhD. thesis, University College of Swansea.
39. Cividini, A. and Gioda, G. (1982) Compatible and equilibrium FE for soil consolidation: a preliminary comparison, in *Numerical Methods in Geomechanics*, Z. Eisenstein (ed.), Balkema, Rotterdam, pp. 297–306.
40. Zienkiewicz, O. C. (1982) Basic formulation of static and dynamic behaviour of soil and other porous material, in *Numerical Methods in Geomechanics*, J. B. Martins (ed.), D. Reidl, Dordrecht, pp. 39–57.
41. Zienkiewicz, O. C. and Shiomi, T. (1984) Dynamic behaviour of saturated porous media: the generalised Biot formulation and its numerical solution. *Int. J. Num. Anal. Meth. Geomech.*, **8**, 71–96.
42. Prevost, J. H., Abdel-Ghaffar, A. M. and Lacy, S. J. (1985) Nonlinear dynamic analyses of an earth dam. *J. Geotech. Engng ASCE*, **111**(7), 882–97.
43. Ehlers, W. and Kubik, J. (1994) On finite dynamic equations for fluid-saturated soils. *Acta Meccanica*, **105**, 101–17.
44. Safai, N. M. and Pinder, G. F. (1979) Vertical and horizontal land deformation in a desaturating porous medium. *Adv. Water Resources*, **2**, 19–25.
45. Schrefler, B. A. and Simoni, L. (1988) A unified approach to the analysis of saturated-unsaturated elastoplastic porous media, in *Numerical Methods in Geomechanics*, G. Swoboda (ed.), Balkema, Rotterdam, pp. 205–12.
46. Zienkiewicz, O. C., Chan, A. H. C., Pastor, M., Paul, D. K. and Shiomi, T. (1990) Static and dynamic behaviour of soils: a rational approach to quantitative solutions. I. Fully saturated problems. *Proc. R. Soc. Lond. A*, **429**, 285–309.
47. Zienkiewicz, O. C., Xie, Y. M., Schrefler, B. A., Ledesma, A. and Bicanic, N. (1990) Static and dynamic behaviour of soils: a rational approach to quantitative solutions. II. Semi-saturated problems. *Proc. R. Soc. Lond. A*, **429**, 311–21.

48. Meroi, E., Schrefler, B. A. and Zienkiewicz, O. C. (1995) Large strain static and dynamic semi-saturated soil behaviour. *Int. J. Num. Anal. Meth. Geomech.*, **19**(2), 81–106.
49. Morgan, K., Lewis, R. W. and White, I. R. (1980) The mechanism of ground surface subsidence above compacting multiphase reservoirs and their analysis by the finite element method. *Appl. Math. Model.*, **4**, 215–24.
50. Li, X., Zienkiewicz, O. C. and Xie, Y. M. (1990) A numerical model for immiscible two-phase fluid flow in a porous medium and its time domain solution. *Int. J. Num. Meth. Engng.*, **30**, 1195–1212.
51. Schrefler, B. A. and Zhan, X. Y. (1993) A fully coupled model for water flow and airflow in deformable porous media. *Water Resources Research*, **29**, 155–67.
52. Bear, J. and Corapcioglu, M. Y. (1981) Mathematical model for regional land subsidence due to pumping. 2. Integrated aquifer subsidence equations for vertical and horizontal displacements. *Water Resources Research*, **17**, 947–58.
53. Aboustit, B. L., Advani, S. H. and Lee, J. K. (1985) Variational principles and finite element simulations for thermo-elastic consolidation. *Int. J. Num. Anal. Meth. Geomech.*, **9**, 49–69.
54. Lewis, R. W., Majorana, C. E. and Schrefler, B. A. (1986) A coupled finite element model for the consolidation of non-isothermal elastoplastic porous media. *Transport in Porous Media*, **1**, 155–78.
55. Geraminegad, M. and Saxena, S. K. (1986) A coupled thermoelastic model for saturated-unsaturated porous media. *Géotechnique*, **36**(4), 539–50.
56. Philip, J. R. and de Vries, D. A. (1957) Moisture movement in porous materials under temperature gradients. *Trans. Amer. Geophys. Union*, **38**, 222–32.
57. Olivella, S., Carrera, J., Gens, A. and Alonso, E. E. (1994) Non-isothermal multiphase flow of brine and gas through saline media. *Transport in Porous Media*, **15**, 271–93.
58. Gawin, D., Baggio, P. and Schrefler, B. A. (1995) Coupled heat, water and gas flow in deformable porous media. *Int. J. Num. Meth. Fluids*, **20**, 969–87.
59. Thomas, H. R. and He, Y. (1995) Analysis of coupled heat, moisture and air flow in a deformable unsaturated soil. *Géotechnique*, **45**(4), 677–89.
60. Zienkiewicz, O. C. and Taylor, R. (1991) *The Finite Element Method*, Vols. 1 and 2, McGraw-Hill, New York.
61. Segerlind, L. J. (1976) *Applied Finite Element Analysis*, Wiley, New York.
62. Maier, G. and Comi, C. (1997) Variational finite element modelling in poroplasticity, Recent Developments in *Computational and Applied Mechanics*, B.D. Reddy (ed), CIMNE, Barcelona, 180–199.

# 2

## Mechanics of Saturated and Partially Saturated Porous Media

### 2.1 INTRODUCTION

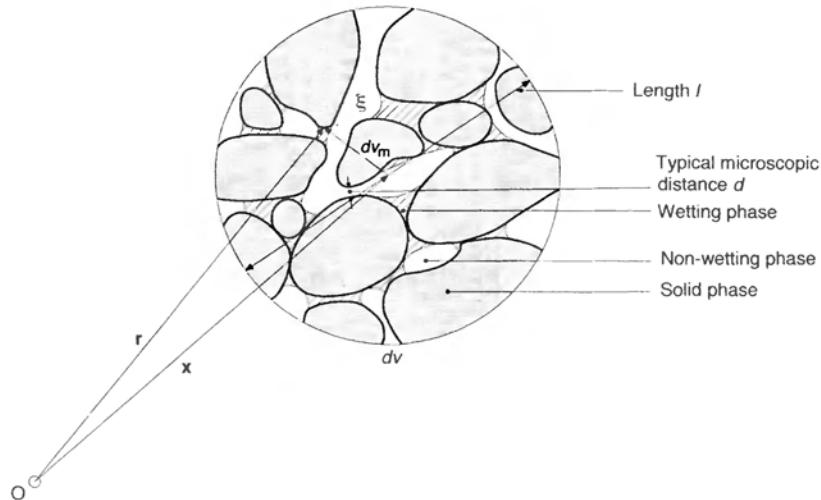
This chapter develops the governing equations for the full dynamic behaviour of a partially saturated porous medium. In particular, it considers the voids filled with water and air. The case of pores filled with one more fluid, i.e. oil will be considered in Chapters 8 and 9.

As indicated in Chapter 1, the description of multiphase systems made of interpenetrating continuous bodies, such as porous media, is today based either on the mixture theory integrated by the concept of volume fractions, or on averaging theories and from a classical viewpoint on Biot's theory. Since the averaging theories offer possibilities for a better understanding of the microscopic situation and its relation to the macroscopic – the natural domain of all continuum mechanical models – we use the averaging theory based on spatial averaging operators. Within this theory we make use of macroscopic variables which correspond to real measurable quantities directly linked to laboratory practice, e.g. in soil mechanics. Note that, under appropriate assumptions, the averaging theory yields the same equations as the classical mixture theory, as shown by de Boer *et al.* [1]. Care has to be taken, however, in the linear momentum balance equation, as explained in Section 2.3.3.

For the reader mainly interested in the resulting governing equations and their numerical solution, we derive these equations again in Section 2.6 using Biot's theory. This also permits us to establish a link between the classical phenomenological approach and the description of the real microscopic composition of the multiphase system. Furthermore, it shows the essential correctness of Biot's findings. Tensorial notation is used throughout this chapter; it is defined in Appendix 2A.

### 2.2 AVERAGING PRINCIPLES

Here is a short summary of the principles necessary for the development of the governing equations. A full account of the averaging theories can be found in the literature [1–8]. Sections 2.2 and 2.3 follow, in particular, the work by Hassanizadeh and Gray [3–7] and de Boer *et al.* [1]. We introduce the following definitions:



**Figure 2.1** Typical averaging volume  $dv$  (two-phase flow) of a porous medium consisting of three constituents

*Microscopic level:* We consider the real non-homogeneous structure of the porous medium domain (Figure 2.1). The scale of the inhomogeneity is of the same order of magnitude as the dimensions of a pore or a grain, say  $d$ . Attention is focused on what happens at a mathematical point within a single phase and the field variables describing the status of a phase are defined only at the points occupied by that phase. For the practical description of the processes taking place in a porous medium, the level of detail given is not needed and sometimes even not desirable. Commonly instead we only need to know an average picture of subsurface processes.

*Macroscopic level:* The real multiphase system that occupies the porous medium domain is replaced by a model in which each phase is assumed to fill up the entire domain. This means that, at every point, all phases are supposed present at the same time (overlapping continua). This is the level of interest of continuum mechanics, where we investigate the continuous distribution of the constituents through a macroscopic control space. At this level, we usually deal with homogeneous media, but non-homogeneities may still be present, e.g. strata. Their scale is of an order of magnitude comparable with the order of magnitude of the entire domain, say  $L$ .

*Megascopic level:* At this level the conditions are similar to those of the previously defined level. The difference depends on the fact that some macroscopic inhomogeneities are eliminated by averaging and/or on the fact that the mathematical model is stated in a domain which has fewer dimensions than the real domain, e.g. a two-dimensional problem with field values averaged over the thickness [9,10]. Typical applications of this level are found in the simulation of land subsidence problems on a regional scale.

### 2.2.1 Averaging process

We consider here a multiphase system occupying a total volume  $V$ , and bounded by surface  $A$ . The constituents  $\pi$ ,  $\pi = 1, 2, \dots, \kappa$  have the partial volumes  $V^\pi$ . Each point of the total volume  $V$ , is considered to be the centroid of a representative elementary volume (REV) or average volume element  $dv$ . The position of the centre of a REV in a global coordinate system is described by position vector  $\mathbf{x}$  while  $\mathbf{r}$  indicates the position of a microscopic volume element  $dv_m$  (Figure 2.1). The volume of constituent  $\pi$  within a REV, called the average volume element  $dv^\pi$ , is obtained by defining a phase distribution function,

$$\gamma^\pi(\mathbf{r}, t) = \begin{cases} 1 & \text{for } \mathbf{r} \in dv^\pi \\ 0 & \text{for } \mathbf{r} \in dv^\alpha \quad \pi \neq \alpha \end{cases} \quad (2.1)$$

$$dv^\pi(\mathbf{x}, t) = \int_{dv} \gamma^\pi(\mathbf{r}, t) dv_m \quad (2.2)$$

where  $\mathbf{r} = \mathbf{x} + \xi$  and the integration refers to the microscopic local coordinate system  $\xi$  with its origin in  $\mathbf{x}$  (Figure 2.1). Similarly we write for the part  $da^\pi$  of area  $da$  of the REV, occupied by constituent  $\pi$

$$da^\pi(\mathbf{x}, t) = \int_{da} \gamma^\pi(\mathbf{r}, t) da_m \quad (2.3)$$

where  $da_m$  is the microscopic area element. Knowledge of  $dv^\pi$  allows us to introduce the concept of volume fraction,  $\eta^\pi$ , which is of paramount importance in multiphase systems:

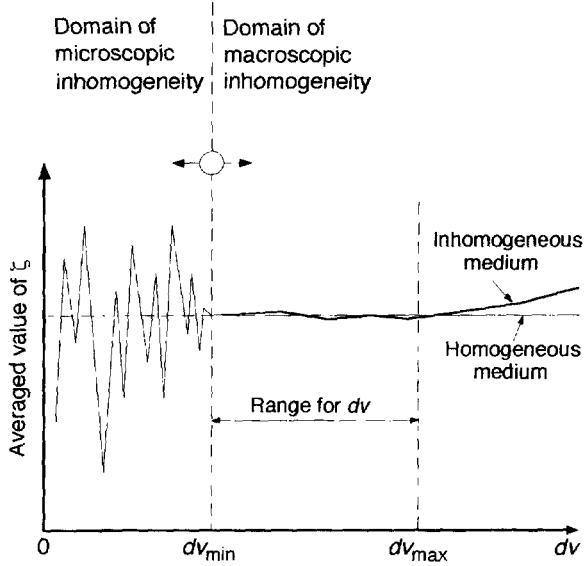
$$\eta^\pi(\mathbf{x}, t) = \frac{dv^\pi}{dv} = \frac{1}{dv} \int_{dv} \gamma^\pi(\mathbf{r}, t) dv_m \quad (2.4)$$

with

$$\sum_{\pi=1}^{\kappa} \eta^\pi = 1 \quad (2.5)$$

In fact, substitute continua fill the entire domain simultaneously, instead of the real fluids and the solid, which each fill only part of it. These substitute continua have a reduced density, which is obtained through the volume fractions. In the following, averaged quantities are obtained by integrating (averaging) a microscopic quantity over the volume  $dv$  or the area  $da$  of a REV. A field of macroscopic variables results from this, where the average volume  $dv$  and the average area  $da$  are associated with material points.

The size of a REV is an important choice. Average quantities have to be independent of the size of the average volume and continuous in space and time. Thus a REV has to fulfil the following requirements:  $dv$  has to be small enough to be considered as infinitesimal, i.e. the partial derivatives appearing in the governing equations must make sense; and  $dv$  must be large enough, with respect to the heterogeneities of the material, to give average quantities without fluctuations depending on the size of the REV (Figure 2.2).



**Figure 2.2** Averaged value  $\zeta$  versus size of the average volume  $dv$

To obtain meaningful average values, the characteristic length  $l$  of the average volume must satisfy the inequality

$$d \ll l \ll L$$

where  $l$  depends on the specific material which constitutes the medium. Some typical values of  $l$  are given by Lemaitre and Chaboche [11]:

Metals	0.5 mm
Plastics	1 mm
Wood	10 mm

Usually there are no boundary conditions associated with a REV, except in the case of the megascopic level. Attention is needed when defining the averaged values for a megascopic level [10]. Another exception is the case of a medium with periodic structure, where asymptotic analysis is used [12].

The following average operators are now defined and applied to a function,  $\zeta(\mathbf{r}, t)$ , which is a microscopic field variable. First come the *volume average* operators: the phase average is

$$\langle \zeta \rangle_{\pi}(\mathbf{x}, t) = \frac{1}{dv} \int_{dv} \zeta(\mathbf{r}, t) \gamma^{\pi}(\mathbf{r}, t) dv_m \quad (2.6)$$

and the intrinsic phase average is

$$\langle \zeta \rangle_{\pi}^{\pi}(\mathbf{x}, t) = \frac{1}{dv^{\pi}} \int_{dv} \zeta(\mathbf{r}, t) \gamma^{\pi}(\mathbf{r}, t) dv_m \quad (2.7)$$

where we have taken into account that

$$\int_{dv^\pi} \zeta(\mathbf{r}, t) dv_m = \int_{dv} \zeta(\mathbf{r}, t) \gamma^\pi(\mathbf{r}, t) dv_m \quad (2.8)$$

because of the form of  $\gamma^\pi$ .

From the definition of volume fraction (2.4) it follows that

$$\langle \zeta \rangle_\pi(\mathbf{x}, t) = \eta^\pi(\mathbf{x}, t) \langle \zeta \rangle_\pi^\pi(\mathbf{x}, t) \quad (2.9)$$

The *mass average* operator, with  $\rho(\mathbf{r}, t)$  microscopic mass density as weighting function, is

$$\bar{\zeta}^\pi(\mathbf{x}, t) = \frac{\int_{dv} \rho(\mathbf{r}, t) \zeta(\mathbf{r}, t) \gamma^\pi(\mathbf{r}, t) dv_m}{\int_{dv} \rho(\mathbf{r}, t) \gamma^\pi(\mathbf{r}, t) dv_m} \quad (2.10)$$

With constant microscopic mass density the following equation holds:

$$\frac{1}{\eta^\pi(\mathbf{x}, t)} \langle \zeta \rangle_\pi(\mathbf{x}, t) = \bar{\zeta}^\pi(\mathbf{x}, t) \quad (2.11)$$

The *area average* operator

$$\bar{\zeta}^\pi(\mathbf{x}, t) = \frac{1}{da} \int_{da} \zeta(\mathbf{r}, t) \cdot \mathbf{n} \gamma^\pi(\mathbf{r}, t) da_m \quad (2.12)$$

with  $\mathbf{n}$  the outward normal unit vector of an area element  $da_m$ . In (2.12)  $\zeta$  has a tensorial nature.

In the following, averages of velocity, external body force, internal energy, external supply of heat, internal entropy, external supply of entropy and total production of entropy are obtained through the mass average operator (2.10) [3]. These density functions are only additive in the form  $\rho \zeta dv$ . Volume-averaged quantities, through (2.7) and mass averaged quantities, through (2.10), are the same only for constant microscopic mass density (2.11). The area average operator (2.12) will be used to define the average of flux terms such as the stress tensor, heat flux and entropy flux.

For an isotropic distribution of phases, the volume and area average values differ only slightly from each other [3]. This follows from Delesse's law, which states that on each cut surface in an isotropic mixture, the surface ratio of each partial constituent must be equal to its volume ratio.

## 2.2.2 Microscopic balance equations

We now consider the classical balance equations of continuum mechanics which are used to describe the microscopic situation of any  $\pi$  phase. At the interfaces with other constituents, the material properties and thermodynamic quantities may present step discontinuities.

For a generic conserved variable  $\psi$ , (e.g. a scalar) the conservation equation within the  $\pi$  phase may be written as

$$\frac{\partial(\rho\psi)}{\partial t} + \operatorname{div}(\rho\psi\dot{\mathbf{r}}) - \operatorname{div}\mathbf{i} - \rho b = \rho G \quad (2.13)$$

where  $\dot{\mathbf{r}}$  is the local value of the velocity field of the  $\pi$  phase at a fixed point in space (Section 2.3.1),  $\mathbf{i}$  is the flux vector associated with  $\psi$ ,  $b$  is the external supply of  $\psi$  and  $G$  is the net production of  $\psi$ .

At the interface between two constituents  $\pi$  and  $\alpha$ , the jump condition holds

$$[\rho\psi(\mathbf{w} - \dot{\mathbf{r}}) + \mathbf{i}]|_{\pi} \cdot \mathbf{n}^{\pi\alpha} + [\rho\psi(\mathbf{w} - \dot{\mathbf{r}}) + \mathbf{i}]|_{\alpha} \cdot \mathbf{n}^{\alpha\pi} = 0 \quad (2.14)$$

where  $\mathbf{w}$  is the velocity of the interface,  $\mathbf{n}^{\alpha\pi}$  the unit normal vector pointing out of the  $\pi$  phase and into the  $\alpha$  phase, with

$$\mathbf{n}^{\pi\alpha} = -\mathbf{n}^{\alpha\pi} \quad (2.15)$$

and  $|_{\pi}$  indicates that the preceding term  $\{ \dots \}$  must be evaluated with respect to the  $\pi$  phase. No thermomechanical properties are attributed to these interfaces. This assumption does not exclude the possibility of exchange of mass, momentum or energy between the constituents. Moreover, the local thermodynamic equilibrium hypothesis is assumed to hold because the timescale of the modelled phenomenon is substantially larger than the relaxation time required to reach equilibrium locally.

### 2.2.3 Macroscopic balance equations

Instead of deriving the macroscopic balance equation separately for each quantity to which the conservation law applies, we derive it for the generic quantity  $\psi$  [1,3] then specialise the law for specific quantities: mass, linear momentum, angular momentum and energy. Note that the balance equations are written in a material-free manner. The constitutive equations are introduced successively.

A general average macroscopic balance equation is obtained from the microscopic balance equation (2.13) by multiplying (2.13) with the distribution function  $\gamma^{\pi}(\mathbf{r}, t)$  and by integrating this product over the volume element  $dv$  and over the total volume  $V$ . In this elaboration of the balance equations, macroscopic quantities are obtained through the previously defined averaging operators. Hence integral expressions of the following type appear in the equations:

$$\int_V \langle \zeta \rangle_{\pi}(\mathbf{x}, t) dV = \int_V \left[ \frac{1}{dv} \int_{dv} \zeta(\mathbf{r}, t) \gamma^{\pi}(\mathbf{r}, t) dv_m \right] dV \quad (2.16)$$

$$\int_A \bar{\zeta}^{\pi}(\mathbf{x}, t) dA = \int_A \left[ \frac{1}{da} \int_{da} \zeta(\mathbf{r}, t) \cdot \mathbf{n}(\mathbf{r}, t) \gamma^{\pi}(\mathbf{r}, t) da_m \right] dA \quad (2.17)$$

Furthermore, the following relations are needed to convert an integral over  $dv^{\pi}$ , including derivatives of  $\zeta$ , into the derivative of this integral over  $dv^{\pi}$ .

For the time derivative of  $\zeta(\mathbf{r}, t)$  and its integration over  $dV$  [1,3], where  $(\mathbf{r}, t)$  is omitted for the sake of brevity, we need

$$\begin{aligned} \int_V \left[ \frac{1}{dv} \int_{dv} \frac{\partial \zeta}{\partial t} \gamma^\pi dv_m \right] dV &= \int_V \frac{\partial}{\partial t} \left[ \frac{1}{dv} \int_{dv} \zeta \gamma^\pi dv_m \right] dV \\ &\quad - \int_V \left[ \frac{1}{dv} \sum_{\alpha \neq \pi}^K \int_{da^{\pi\alpha}} \zeta \mathbf{w} \cdot \mathbf{n}^{\pi\alpha} da_m \right] dV \end{aligned} \quad (2.18)$$

where  $da^{\pi\alpha}$  is the interface between the  $\pi$  phase and the other phases  $\alpha$  located inside  $dv$ .

For a spatial derivative of a vector or tensor  $\zeta$  and its subsequent integration over  $dv$ , we need

$$\begin{aligned} \int_V \left[ \frac{1}{dv} \int_{dv} \operatorname{div}(\zeta) \gamma^\pi dv_m \right] dV &= \int_V \left[ \frac{1}{dv} \int_{dv} \operatorname{div}(\zeta \gamma^\pi) dv_m \right] dV \\ &\quad + \int_V \left[ \frac{1}{dv} \sum_{\alpha \neq \pi}^K \int_{da^{\pi\alpha}} \zeta \cdot \mathbf{n}^{\pi\alpha} da_m \right] dV \\ &= \int_A \left[ \frac{1}{da} \int_{da} \zeta \gamma^\pi \cdot \mathbf{n} da_m \right] dA \\ &\quad + \int_V \left[ \frac{1}{dv} \sum_{\alpha \neq \pi}^K \int_{da^{\pi\alpha}} \zeta \cdot \mathbf{n}^{\pi\alpha} da_m \right] dV \end{aligned} \quad (2.19)$$

where the divergence theorem by Eringen and Suhubi [13] has been introduced; for a constant function, this allows us to write

$$\begin{aligned} \int_V \left[ \operatorname{div} \left( \frac{1}{dv} \int_{dv} \zeta dv_m \right) \right] dV &= \int_V \left[ \frac{1}{dv} \int_{dv} \operatorname{div} \zeta dv_m \right] dV \\ &= \int_A \left[ \frac{1}{da} \int_{da} \zeta \cdot \mathbf{n} da_m \right] dA \end{aligned} \quad (2.20)$$

We need, furthermore, to define the deviation of a microscopic function  $\zeta(\mathbf{r}, t)$  from the mass-averaged quantity  $\bar{\zeta}^\pi(\mathbf{x}, t)$  of constituent  $\pi$  as

$$\bar{\zeta}^\pi(\mathbf{x}, \zeta, t) = \zeta(\mathbf{r}, t) - \bar{\zeta}^\pi(\mathbf{x}, t) \quad (2.21)$$

Several identities useful for the application of the averaging procedure to equation (2.13) are listed in Appendix 2B. This averaging procedure yields [1,3] that

$$\begin{aligned} &\int_V \left[ \frac{1}{dv} \int_{dv} \frac{\partial (\rho(\mathbf{r}, t)\psi(\mathbf{r}, t))}{\partial t} \gamma^\pi(\mathbf{r}, t) dv_m \right] dV \\ &\quad + \int_V \left[ \frac{1}{dv} \int_{dv} \operatorname{div}(\rho(\mathbf{r}, t)\dot{\mathbf{r}}(\mathbf{r}, t)\psi(\mathbf{r}, t)) \gamma^\pi(\mathbf{r}, t) dv_m \right] dV \\ &\quad - \int_V \left[ \frac{1}{dv} \int_{dv} \operatorname{div} \mathbf{i}(\mathbf{r}, t) \gamma^\pi(\mathbf{r}, t) dv_m \right] dV \\ &\quad - \int_V \left[ \frac{1}{dv} \int_{dv} \rho(\mathbf{r}, t)b(\mathbf{r}, t) \gamma^\pi(\mathbf{r}, t) \right] dV \\ &= \int_V \left[ \frac{1}{dv} \int_{dv} \rho(\mathbf{r}, t)G(\mathbf{r}, t)\gamma^\pi(\mathbf{r}, t) dv_m \right] dV \end{aligned} \quad (2.22)$$

This equation is now modified using (2.18) and (2.19). For the integration over the volume element  $dv$ , note that averaged quantities are constant in  $dv$ . Furthermore, the mass averaged velocity  $\bar{\mathbf{v}}^\pi$  of the  $\pi$  phase is introduced (see also the kinematic equations in Section 2.3.1) as follows:

$$\bar{\mathbf{r}}^\pi(\mathbf{x}, t) = \dot{\mathbf{x}}^\pi(\mathbf{x}, t) = \bar{\mathbf{v}}^\pi = \frac{1}{dv \langle \rho \rangle_\pi} \int_{dv} \rho(\mathbf{r}, t) \dot{\mathbf{r}}(\mathbf{r}, t) \gamma^\pi(\mathbf{r}, t) dv_m \quad (2.23)$$

A detailed development of these modifications can be found in the literature [1,3,4,5]. They lead to the following form of the general balance equation for the macroscopic thermodynamic property  $\bar{\psi}^\pi$  associated with the  $\pi$  phase:

$$\begin{aligned} & \int_V \left\{ \frac{\partial}{\partial t} [\langle \rho \rangle_\pi(\mathbf{x}, t) \bar{\psi}^\pi(\mathbf{x}, t)] + \operatorname{div} [\langle \rho \rangle_\pi(\mathbf{x}, t) \bar{\mathbf{v}}^\pi(\mathbf{x}, t) \bar{\psi}^\pi(\mathbf{x}, t)] \right\} dV \\ & - \int_V \left\{ \frac{1}{dv} \sum_{\alpha \neq \pi}^K \int_{da^{\alpha\pi}} \rho(\mathbf{r}, t) \psi(\mathbf{r}, t) [\mathbf{w}(\mathbf{r}, t) - \dot{\mathbf{r}}(\mathbf{r}, t)] \cdot \mathbf{n}^{\alpha\pi}(\mathbf{r}, t) da_m \right\} dV \\ & - \int_V \left[ \frac{1}{dv} \sum_{\alpha \neq \pi}^K \int_{da^{\alpha\pi}} \mathbf{i}(\mathbf{r}, t) \cdot \mathbf{n}^{\alpha\pi}(\mathbf{r}, t) da_m \right] dV \\ & - \int_A \left\{ \frac{1}{da} \int_{da} [\mathbf{i}(\mathbf{r}, t) - \rho(\mathbf{r}, t) \tilde{\psi}^\pi(\mathbf{x}, \xi, t) \tilde{\mathbf{r}}^\pi] \cdot \mathbf{n}(\mathbf{r}, t) \gamma^\pi da_m \right\} dA \\ & - \int_V \langle \rho \rangle_\pi(\mathbf{x}, t) \bar{b}^\pi(\mathbf{x}, t) dV = \int_V \langle \rho \rangle_\pi(\mathbf{x}, t) \bar{G}^\pi(\mathbf{x}, t) dV \end{aligned} \quad (2.24)$$

or in more concise form

$$\begin{aligned} & \left[ \int_V \left[ \frac{\partial}{\partial t} (\langle \rho \rangle_\pi \bar{\psi}^\pi) + \operatorname{div} (\langle \rho \rangle_\pi \bar{\psi}^\pi \bar{\mathbf{v}}^\pi) - \operatorname{div} \mathbf{i}^\pi \right. \right. \\ & \left. \left. - \langle \rho \rangle_\pi [\bar{b}^\pi + e^\pi(\rho \psi) + I^\pi] \right] dV = \int_V \langle \rho \rangle_\pi \bar{G}^\pi dv \right] \end{aligned} \quad (2.25)$$

where

$\mathbf{i}^\pi$  = the flux vector associated with  $\bar{\psi}^\pi$

$\bar{b}^\pi$  = the external supply of  $\bar{\psi}^\pi$

$\langle \rho \rangle_\pi$  = the volume-averaged value of mass density

The balance equation (2.25) contains two further interaction terms, which describe chemical and physical exchanges. Exchange of  $\bar{\psi}^\pi$  due to mechanical interactions between the constituents is given by

$$I^\pi = \frac{1}{\langle \rho \rangle_\pi dv} \sum_{\alpha \neq \pi} \int_{da^{\alpha\pi}} \mathbf{n}^{\alpha\pi} \cdot \mathbf{i} da_m \quad (2.26)$$

whereas phase change of a constituent or possible mass exchange between the

constituent  $\pi$  and the other constituents  $\alpha$  is given by

$$e^\pi(\rho\psi) = \frac{1}{\langle\rho\rangle_\pi} \sum_{\alpha \neq \pi} \int_{da^{\pi\alpha}} \rho\psi(\mathbf{w} - \dot{\mathbf{r}}) \cdot \mathbf{n}^{\pi\alpha} da_m \quad (2.27)$$

The fourth term on the left-hand side of (2.24) is the macroscopic non-convective flux of  $\bar{\psi}^\pi$ , and an averaged flux density may be defined as

$$\bar{i}^\pi = \frac{1}{da} \int_{da} \left( \mathbf{i} - \rho\bar{\psi}^\pi \dot{\mathbf{r}}^\pi \right) \cdot \mathbf{n}^\pi da_m \quad (2.28)$$

With the averaged macroscopic flux there exists a flux vector  $\mathbf{i}^\pi$  such that  $\bar{i}^\pi$  can be defined as a linear function of the normal unit vector  $\mathbf{n}$  of  $da$  in  $\mathbf{x}$  [3]:

$$\bar{i}^\pi = \mathbf{n} \cdot \mathbf{i}^\pi \quad (2.29)$$

and the divergence theorem may be applied to obtain the term  $\operatorname{div} \mathbf{i}^\pi$  in (2.25), i.e.

$$\int_A \mathbf{n} \cdot \mathbf{i}^\pi dA = \int_V \operatorname{div} \mathbf{i}^\pi dV \quad (2.30)$$

For the proof of the existence of the average flux vector  $\mathbf{i}^\pi$ , the tetrahedron argument is invoked [3], as in stress analysis for Cauchy stresses.

A condition for all internal exchange processes is obtained by integrating equation (2.14) over all interfaces,  $da^{\pi\alpha}$  and  $da^{\alpha\pi}$ , and over the total volume  $V$ , then by summation over all constituents  $\pi$ , taking into account (2.26) and (2.27):

$$\int_V \sum_\pi \langle \rho \rangle_\pi [e^\pi(\rho\psi) + I^\pi] dV = 0 \quad \forall \mathbf{x} \in V \quad (2.31)$$

The differential form of the balance equation is derived under certain smoothness conditions by localisation at the macroscopic level:

$$\begin{aligned} & \forall \mathbf{x} \in V \\ & \frac{\partial}{\partial t} (\langle \rho \rangle_\pi \bar{\psi}^\pi) + \operatorname{div} (\langle \rho \rangle_\pi \bar{\psi}^\pi \bar{\mathbf{v}}^\pi) - \operatorname{div} \mathbf{i}^\pi \\ & - \langle \rho \rangle_\pi [\bar{b}^\pi + e^\pi(\rho\psi) + I^\pi] = \langle \rho \rangle_\pi \bar{G}^\pi \end{aligned} \quad (2.32)$$

subject to

$$\sum_\pi \langle \rho \rangle_\pi [e^\pi(\rho\psi) + I^\pi] = 0 \quad \forall \mathbf{x} \in V \quad (2.33)$$

If the body is separated by a discontinuity surface  $\Sigma$  at the macroscopic scale, which

moves with velocity  $\bar{\mathbf{w}}$ , the following additional relation must be fulfilled [3,4]:

$$[\langle \rho \rangle_\pi \bar{\Psi}^\pi (\bar{\mathbf{v}}^\pi - \bar{\mathbf{w}}) - \mathbf{i}^\pi] \cdot \mathbf{N} = 0 \quad \forall \mathbf{x} \in \Sigma \quad (2.34)$$

where  $\mathbf{N}$  is the normal unit vector of  $\Sigma$  in  $\mathbf{x}$ .

### 2.3 MACROSCOPIC BALANCE EQUATIONS FOR A NON-ISOTHERMAL PARTIALLY SATURATED POROUS MATERIAL

The macroscopic balance equations for mass, linear momentum, angular momentum and energy (enthalpy) are obtained then specialised for a deforming porous material where heat transfer and flow of water (liquid and vapour) and dry air are taking place. The starting-points are the microscopic balance equations, (2.13), where for each constituent the generic thermodynamic variable  $\psi$  is replaced by appropriate microscopic quantities suitable for a microscopic non-polar material.

For a proper description of the non-isothermal unsaturated porous medium, we need to take into account not only heat conduction and vapour diffusion, but also heat convection, liquid water flow due to pressure gradients or capillary effects, and latent heat transfer due to water phase change (evaporation and condensation) inside the pores. Furthermore, the solid is deformable, resulting in a coupling of the fluid, the solid and the thermal fields. All fluid phases are in contact with the solid phase.

The constituents are assumed to be immiscible except for dry air and vapour, and chemically non-reacting. Because of the local thermodynamic equilibrium hypothesis, the temperatures of all constituents at a point in the multiphase medium are taken to be equal. This does not mean that the temperature is uniform throughout the medium but only that at each point one temperature is sufficient to characterise the state. Momentum exchanges due to mechanical interaction are independent of the temperature gradient.

We use the following convention for stress. Tensile components are positive in the solid while the fluid in the state of compression has a positive pressure. Note that in this section the formulation is still material-free, i.e. no specific assumptions for the material behaviour have been introduced so far, except for the earlier general assumptions. Appropriate constitutive relationships will be introduced in section 2.4. For the development of macroscopic balance equations in the following sections, we still need to specify the kinematics.

#### 2.3.1 Kinematic Equations

As indicated in Section 2.2, a multiphase medium can be described as the superposition of all  $\pi$  phases, i.e. in the actual configuration each spatial point  $\mathbf{x}$  is simultaneously occupied by material points  $\mathbf{X}^\pi$  of all phases. The state of motion of each phase is, however, described independently. Based on these assumptions, the kinematics of a multiphase medium is dealt with next.

In a Lagrangian or material description of motion, the position of each material point,  $\mathbf{x}^\pi$  at time  $t$  is a function of its placement in a chosen reference configuration  $\mathbf{X}^\pi$  and of

the current time  $t$ :

$$\mathbf{x}^\pi = \mathbf{x}^\pi(\mathbf{X}^\pi, t) \quad (2.35)$$

To have this mapping continuous and bijective at all times, the Jacobian  $J$  of this transformation must be non-zero and strictly positive, since it is equal to the determinant of the deformation gradient tensor,  $\mathbf{F}^\pi$ :

$$\mathbf{F}^\pi = \text{Grad } \mathbf{x}^\pi \quad (\mathbf{F}^\pi)^{-1} = \text{Grad } \mathbf{X}^\pi \quad (2.36)$$

Because the Lagrangian relationship (2.35) is non-singular, it is possible to write its inverse and obtain the Eulerian or spatial description of motion:

$$\mathbf{X}^\pi = \mathbf{X}^\pi(\mathbf{x}^\pi, t) \quad (2.37)$$

It is also assumed that functions which describe the motion have continuous derivatives. If the path of the particle of the  $\pi$  phase (2.35) is known, its velocity and acceleration have the following material description:

$$\mathbf{v}^\pi = \frac{\partial \mathbf{x}^\pi(\mathbf{X}^\pi, t)}{\partial t} \quad (2.38)$$

$$\mathbf{a}^\pi = \frac{\partial^2 \mathbf{x}^\pi(\mathbf{X}^\pi, t)}{\partial t^2} \quad (2.39)$$

The corresponding spatial expression can be obtained by introducing (2.37) into (2.38) and (2.39). But, if only the spatial description is given for the velocity field in the form

$$\mathbf{v}^\pi = \mathbf{v}^\pi(\mathbf{x}^\pi, t) \quad (2.40)$$

to evaluate its time derivative with material coordinates held constant, we introduce the description of motion (2.35) into (2.40). By applying the chain rule of differentiation, it follows that

$$\mathbf{a}^\pi = \frac{\partial \mathbf{v}^\pi}{\partial t} + \text{grad } \mathbf{v}^\pi \cdot \mathbf{v}^\pi \quad (2.41)$$

The material time derivative of any differentiable function  $f^\pi(\mathbf{x}, t)$ , given in its spatial description and referring to a moving particle of the  $\pi$  phase, is

$$\frac{D^\pi f^\pi}{Dt} = \frac{\partial f^\pi}{\partial t} + \text{grad } f^\pi \cdot \mathbf{v}^\pi \quad (2.42)$$

If superscript  $\alpha$  is used for the operator  $D/Dt$ , we obtain

$$\frac{D^\alpha f^\pi}{Dt} = \frac{\partial f^\pi}{\partial t} + \text{grad } f^\pi \cdot \mathbf{v}^\alpha \quad (2.43)$$

the time derivative is taken moving with the  $\alpha$  phase.

Subtraction of (2.42) from (2.43) yields the following relation:

$$\frac{D^{\alpha}f^{\pi}}{Dt} = \frac{D^{\pi}f^{\pi}}{Dt} + \text{grad } f^{\pi} \cdot \mathbf{v}^{\alpha\pi} \quad (2.44)$$

where

$$\mathbf{v}^{\alpha\pi} = \mathbf{v}^{\alpha} - \mathbf{v}^{\pi} \quad (2.45)$$

is the velocity of the  $\alpha$  phase with respect to the  $\pi$  phase. This velocity is called the diffusion velocity [5].

The operator  $D/Dt$  is a scalar operator and may be applied either to a vector quantity or a scalar quantity. If  $\mathbf{f}^{\pi}$  is a vector property per unit volume referring to the  $\pi$  phase, the total time derivative of its integral over a volume  $V$  is given by

$$\frac{d}{dt} \int_V \mathbf{f}^{\pi} dV = \int_V \left( \frac{\partial \mathbf{f}^{\pi}}{\partial t} + \text{grad } \mathbf{f}^{\pi} \cdot \mathbf{v}^{\pi} + \mathbf{f}^{\pi} \text{div } \mathbf{v}^{\pi} \right) dV = \int_V \left[ \frac{\partial \mathbf{f}^{\pi}}{\partial t} + \text{div}(\mathbf{f}^{\pi} \otimes \mathbf{v}^{\pi}) \right] dV \quad (2.46a)$$

For a scalar property  $f^{\pi}$ , we obtain

$$\frac{d}{dt} \int_V f^{\pi} dV = \int_V \left( \frac{\partial f^{\pi}}{\partial t} + \text{div}(f^{\pi} \mathbf{v}^{\pi}) \right) dV \quad (2.46b)$$

In these equations, velocities and accelerations of the  $\pi$  phase are considered as mass-averaged quantities – see (2.23) with the averaging symbol omitted – since they are the quantities usually measured in a field situation or in laboratory practice. In porous media theory it is customary to describe the motion of the fluid phases in terms of mass-averaged velocities relative to the moving solid. Their motion is described with reference to the actual configuration assumed by the solid skeleton. The velocities and accelerations of each fluid particle can then be written with reference to those of corresponding solid points, once the relative velocities are introduced. We specify the superscripts  $\pi$  now as  $s$  for soil,  $w$  for the liquid phase and  $g$  for the gas phase (dry air plus vapour) and write the following for the relative velocities of water and gas phase respectively (2.45):

$$\mathbf{v}^{ws} = \mathbf{v}^w - \mathbf{v}^s \quad \mathbf{v}^{gs} = \mathbf{v}^g - \mathbf{v}^s \quad (2.47)$$

Water and gas acceleration are given from (2.3.7), (2.3.9) and (2.3.13) as

$$\mathbf{a}^w = \mathbf{a}^s + \frac{D^s \mathbf{v}^{ws}}{Dt} + \text{grad}(\mathbf{v}^s + \mathbf{v}^{ws}) \cdot \mathbf{v}^{ws} \quad (2.48)$$

$$\mathbf{a}^g = \mathbf{a}^s + \frac{D^s \mathbf{v}^{gs}}{Dt} + \text{grad}(\mathbf{v}^s + \mathbf{v}^{gs}) \cdot \mathbf{v}^{gs} \quad (2.49)$$

The deformation process of the solid skeleton can be described by the velocity gradient tensor  $\mathbf{L}^s$ , which when referred to spatial coordinates is given by the

following equation [14,15]:

$$\mathbf{L}^s \equiv \text{grad } \mathbf{v}^s = \mathbf{D}^s + \mathbf{W}^s \quad (2.50)$$

Its symmetric part  $\mathbf{D}^s$  is called the Eulerian strain rate tensor, being related to pure straining; its skew-symmetric component  $\mathbf{W}^s$  is the spin tensor.

### 2.3.2 Mass balance equations

We identify the volume fractions,  $\eta^\pi$  of the constituents as follows:

*Solid phase:*  $\eta^s = 1 - n$  where  $n = (dv^w + dv^g)/dv$  is the porosity

*Water:*  $\eta^w = nS_w$  where  $S_w = dv^w/(dv^w + dv^g)$  is the degree of water saturation

*Gas:*  $\eta^g = nS_g$  with  $S_g = dv^g/(dv^w + dv^g)$  the degree of gas saturation

It follows immediately that

$$S_w + S_g = 1 \quad (2.51)$$

#### 2.3.2.1 Solid phase

In the microscopic situation, the variables for solid in (2.13) assume the following values:

$$\psi = 1 \quad \mathbf{i} = \mathbf{0} \quad b = 0 \quad G = 0 \quad (2.52)$$

and the microscopic mass balance equation results in

$$\frac{\partial \rho}{\partial t} + \text{div}(\rho \dot{\mathbf{r}}) = 0 \quad (2.53)$$

The averaged macroscopic solid mass balance equation is obtained from (2.32) by specifying

$$\begin{aligned} \bar{\psi}^s &= 1 & \mathbf{i}^s &= \mathbf{0} & \bar{b}^s &= 0 & I^s &= 0 & \bar{G}^s &= 0 \\ \frac{\partial \rho_s}{\partial t} + \text{div}(\rho_s \bar{\mathbf{v}}^s) &= \rho_s e^s(\rho) \end{aligned} \quad (2.54)$$

where  $\rho_s$  stands simply for  $\langle \rho \rangle_s$ , the phase-averaged solid density and  $\bar{\mathbf{v}}^s$  is the mass-averaged solid velocity. The same simplified notation will be used for the other constituents, once  $\pi$  is accordingly specified.

For the solid phase, the exchange term,  $\rho_s e^s(\rho)$ , i.e. the average microscopic exchange of density with the other phases is zero.

From (2.42) we have

$$\frac{D^s \rho_s}{Dt} = \frac{\partial \rho_s}{\partial t} + \text{grad } \rho_s \cdot \bar{\mathbf{v}}^s \quad (2.55)$$

By introducing (2.55) into (2.54) we obtain

$$\frac{D^s \rho_s}{Dt} + \rho_s \operatorname{div} \bar{\mathbf{v}}^s = 0 \quad (2.56)$$

because the following vector identity holds:

$$\operatorname{div}(\rho_s \bar{\mathbf{v}}^s) = \rho_s \operatorname{div} \bar{\mathbf{v}}^s + \operatorname{grad} \rho_s \cdot \bar{\mathbf{v}}^s \quad (2.57)$$

By introducing intrinsic phase-averaged densities through equation (2.9) we finally obtain

$$\frac{D^s(1-n)\rho^s}{Dt} + \rho^s(1-n) \operatorname{div} \bar{\mathbf{v}}^s = 0 \quad (2.58)$$

where the shorthand  $\rho^s = \langle \rho \rangle_s^s$  has been introduced for the intrinsic phase-averaged density.

### 2.3.2.2 Liquid phase: water

The microscopic mass balance equation is given by (2.53). The macroscopic mass balance equation is the same as (2.54) with the right-hand side being non-zero, because water may change into vapour or vice versa. Hence we have

$$\frac{\partial \rho_w}{\partial t} + \operatorname{div}(\rho_w \bar{\mathbf{v}}^w) = \rho_w e^w(\rho) \quad (2.59)$$

Transformation as for the solid phase yields

$$\frac{D^w \rho_w}{Dt} + \rho_w \operatorname{div} \bar{\mathbf{v}}^w = \rho_w e^w(\rho) \quad (2.60)$$

Finally

$$\rho_w e^w(\rho) = -\dot{m} \quad (2.61)$$

is the quantity of water lost through evaporation per unit time per unit volume.

### 2.3.2.3 Gaseous phases: dry air and vapour

The gaseous phase here is a multicomponent material composed of two different species, dry air and vapour. These species are miscible. We first write the mass balance equations for both species. Their microscopic mass balance equations are again given by (2.53) if we neglect net production of mass of each species, due to chemical reactions with the other species [6].

The macroscopic mass balance equation for dry air is given by (2.54) with appropriate superscripts and subscripts, and an exchange term equal to zero. We

introduce intrinsic phase-averaged densities and use super/subscript ga to indicate dry air. Because the two species, dry air and vapour, are miscible, they have the same volume fraction  $nS_g$ , so

$$\frac{\partial}{\partial t}(nS_g\rho^{ga}) + \operatorname{div}(nS_g\rho^{ga}\bar{\mathbf{v}}^{ga}) = 0 \quad (2.62)$$

Similarly, using super/subscript gw for vapour, we write

$$\frac{\partial}{\partial t}(nS_g\rho^{gw}) + \operatorname{div}(nS_g\rho^{gw}\bar{\mathbf{v}}^{gw}) = nS_g\rho^{gw}e^{gw}(\rho) = \dot{m} \quad (2.63)$$

We now derive the mass balance equation for the whole gaseous phase. This is obtained by summing the macroscopic balance equations of the two species and using appropriate definitions for bulk properties of the gaseous phase [6]:

$$\frac{\partial}{\partial t}(nS_g\rho^g) + \operatorname{div}(nS_g\rho^g\bar{\mathbf{v}}^g) = \dot{m} \quad (2.64)$$

with

$$\rho^g = \rho^{ga} + \rho^{gw} \quad (2.65)$$

and

$$\bar{\mathbf{v}}^g = \frac{1}{\rho^g}(\rho^{ga}\bar{\mathbf{v}}^{ga} + \rho^{gw}\bar{\mathbf{v}}^{gw}) = c^{ga}\bar{\mathbf{v}}^{ga} + c^{gw}\bar{\mathbf{v}}^{gw} \quad (2.66)$$

where

$$c^\pi = \rho^\pi/\rho^g \quad (2.67)$$

is the mass fraction of component  $\pi$ , subject to

$$\sum_\pi c^\pi = 1 \quad \pi = gw, ga \quad (2.68)$$

We introduce further the macroscopic diffusive dispersive velocity  $\mathbf{u}^\pi, \pi = ga, gw$  defined as follows [6]:

$$\mathbf{u}^\pi = \bar{\mathbf{v}}^{\pi g} = \bar{\mathbf{v}}^\pi - \bar{\mathbf{v}}^g \quad (2.69)$$

and subject to

$$\rho^{ga}\mathbf{u}^{ga} + \rho^{gw}\mathbf{u}^{gw} = \rho^g \sum_\pi c^\pi \mathbf{u}^\pi = 0 \quad (2.70)$$

Equation (2.64) for the gas phase can be transformed in a similar way to the equation for

the solid phase; this yields

$$\frac{D^g(nS_g\rho^g)}{Dt} + nS_g\rho^g \operatorname{div} \bar{\mathbf{v}}^g = \dot{m} \quad (2.71)$$

We proceed similarly for (2.63). The use of vector identity (2.57) for  $\operatorname{div}(nS_g\rho^{gw}\bar{\mathbf{v}}^{gw})$  and introduction of material time derivative (2.42) allows us to write (2.63) as

$$\frac{D^{gw}}{Dt}(nS_g\rho^{gw}) + nS_g\rho^{gw} \operatorname{div} \bar{\mathbf{v}}^{gw} = \dot{m} \quad (2.72)$$

Use of (2.44) for  $\alpha = g$ ,  $\pi = gw$  and introduction of (2.69) for  $\pi = gw$  yields

$$\frac{D^g}{Dt}(nS_g\rho^{gw}) + \mathbf{u}^{gw} \cdot \operatorname{grad}(nS_g\rho^{gw}) + nS_g\rho^{gw} \operatorname{div} \bar{\mathbf{v}}^{gw} = \dot{m} \quad (2.73)$$

Further use of vector identity (2.57) to obtain  $\operatorname{div}(nS_g\rho^{gw}\mathbf{u}^{gw})$  gives the following form of the mass balance equation for vapour:

$$\frac{D^g}{Dt}(nS_g\rho^{gw}) + \operatorname{div}(nS_g\rho^{gw}\mathbf{u}^{gw}) + nS_g\rho^{gw} \operatorname{div} \bar{\mathbf{v}}^g = \dot{m} \quad (2.74)$$

We introduce now the diffusive-dispersive mass flux of component gw as follows [13]:

$$\mathbf{J}_g^{gw} = nS_g\rho^{gw}\mathbf{u}^{gw} \quad (2.75)$$

which allows us to write (2.74) as

$$\frac{D^g}{Dt}(nS_g\rho^{gw}) + \operatorname{div} \mathbf{J}_g^{gw} + nS_g\rho^{gw} \operatorname{div} \bar{\mathbf{v}}^g = \dot{m} \quad (2.76)$$

Equations (2.71) and (2.76) will be used for obtaining the general field equations, or (2.71) together with (2.62), depending on the aim of the model.

### 2.3.3 Linear momentum balance equation

The microscopic variables for equation (2.13) are as follows:

$$\psi = \dot{\mathbf{r}} \quad \mathbf{i} = \mathbf{t}_m \quad \mathbf{b} = \mathbf{g} \quad \mathbf{G} = \mathbf{0} \quad (2.77)$$

where  $\mathbf{t}_m$  is the microscopic stress tensor and  $\mathbf{g}$  is the external momentum supply related to gravitational effects. The microscopic linear momentum balance equation is therefore

$$\frac{\partial}{\partial t}(\rho\dot{\mathbf{r}}) + \operatorname{div}(\rho\dot{\mathbf{r}} \otimes \dot{\mathbf{r}}) - \operatorname{div} \mathbf{t}_m - \rho\mathbf{g} = \mathbf{0} \quad (2.78)$$

Since

$$\operatorname{div}(\rho \dot{\mathbf{r}} \otimes \dot{\mathbf{r}}) = \dot{\mathbf{r}} \otimes \dot{\mathbf{r}} \cdot \operatorname{grad} \rho + \rho \operatorname{grad} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + \rho \dot{\mathbf{r}} \operatorname{div} \dot{\mathbf{r}} \quad (2.79)$$

and

$$\frac{\partial \dot{\mathbf{r}}}{\partial t} + \operatorname{grad} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} = \ddot{\mathbf{r}} \quad (2.80)$$

Equation (2.78) can be transformed by using the microscopic mass balance equation (2.53) to yield

$$\operatorname{div} \mathbf{t}_m + \rho(\mathbf{g} - \ddot{\mathbf{r}}) = \mathbf{0} \quad (2.81)$$

which is the usual form of the momentum balance equation for a single-phase material. For the averaged linear momentum balance equation obtained from (2.32) we have  $\bar{\psi}^\pi = \bar{\mathbf{v}}^\pi$ , the mass-averaged velocity of the constituent and

$$\tilde{\mathbf{t}}^\pi = \tilde{\mathbf{t}}^\pi = \frac{1}{da} \int_{da} (\mathbf{t}_m - \rho \tilde{\mathbf{r}}^\pi \otimes \tilde{\mathbf{r}}^\pi) \cdot \mathbf{n} \gamma^\pi da_m \quad (2.82)$$

According to (2.28) there exists a stress tensor  $\mathbf{t}^\pi$  such that

$$\tilde{\mathbf{t}}^\pi = \mathbf{t}^\pi \cdot \mathbf{n} \quad (2.83)$$

In mixture theories  $\mathbf{t}^\pi$  is called a partial stress tensor. The stress vector  $\tilde{\mathbf{t}}^\pi$  is composed of the sum of the averaged microscopic stress tensor in the  $\pi$  phase and an averaged stress influx, produced by the mass fluxes,  $\rho \tilde{\mathbf{r}}^\pi$ , which occur in the microscopic field relative to the macroscopic velocity of the  $\pi$  phase [1].

Note that only terms acting from the  $\pi$  phase on the  $\pi$  phase itself appear here. Hence  $\tilde{\mathbf{t}}^\pi$  is the surface density of an interaction force within the  $\pi$  phase which is balanced by the body force (2.84) and the inertia force (2.95), both additive and volume-bounded functions, and by diffusive interactions (2.85) and (2.88) which are also volume-bounded and additive. An important constitutive assumption which allows us to define  $\mathbf{t}^\pi$  as a stress tensor will be discussed below. Furthermore

$$\bar{\mathbf{g}}^\pi = \frac{1}{\rho_\pi dv} \int_{dv} \rho \mathbf{g} \gamma^\pi dv_m \quad (2.84)$$

is the external momentum supply, which we assume to be related to gravitational effects.  $\rho_\pi \bar{\mathbf{g}}^\pi$  is the density of an additive volume-bounded function.

$$\mathbf{e}^\pi(\rho \dot{\mathbf{r}}) = \mathbf{e}^\pi(\rho) \bar{\mathbf{v}}^\pi + \mathbf{e}^\pi(\rho \tilde{\mathbf{r}}^\pi) \quad (2.85)$$

where

$$\rho_\pi \mathbf{e}^\pi(\rho) = \frac{1}{dv} \sum_{\alpha \neq \pi}^K \int_{da_{\alpha\pi}} \rho [\bar{\mathbf{v}}^\pi \otimes (\mathbf{w} - \dot{\mathbf{r}})] \cdot \mathbf{n}^{\alpha\pi} da_m \quad (2.86)$$

is the momentum exchange due to averaged mass supply, whereas

$$\rho_\pi \mathbf{e}^\pi(\rho \tilde{\mathbf{r}}^\pi) = \frac{1}{dv} \sum_{\alpha \neq \pi}^K \int_{da^{\pi\alpha}} \rho [\tilde{\mathbf{r}}^\pi \otimes (\mathbf{w} - \dot{\mathbf{r}})] \cdot \mathbf{n}^{\pi\alpha} da_m \quad (2.87)$$

is the intrinsic momentum supply, with reference to the deviation  $\tilde{\mathbf{r}}^\pi$  of the velocity of constituent  $\pi$  from its mass-averaged velocity  $\bar{\mathbf{v}}^\pi$  due to a change of density.

Note that in the sum on the right-hand side of (2.86) and (2.87) only the interactions exerted on the  $\pi$  phase from the  $\alpha$  phases appear, while those from the  $\pi$  phase on the  $\alpha$  phases are missing.

$$\mathbf{I}^\pi = \hat{\mathbf{t}}^\pi = \frac{1}{\rho_\pi dv} \sum_{\alpha \neq \pi}^K \int_{da^{\pi\alpha}} \mathbf{t}_m \cdot \mathbf{n}^{\pi\alpha} da_m \quad (2.88)$$

accounts for the exchange of momentum due to mechanical interaction with other phases. Again the sum on the right-hand side extends only to the interactions exerted from the  $\alpha$  phase on the  $\pi$  phase.

The sum of (2.85) and (2.88)

$$\rho_\pi \mathbf{I}^\pi + \rho_\pi \mathbf{e}^\pi(\rho \dot{\mathbf{r}}) \quad (2.89)$$

represents a supply quantity based on the volume  $dv$ . This sum may be interpreted as the volume density of a volume-bounded local diffusive interaction  $D$  between the  $\pi$  phase and the other constituents.

Finally

$$\tilde{\mathbf{G}}^\pi = \mathbf{0} \quad (2.90)$$

With the above definitions, the averaged linear momentum balance equation may be written as

$$\frac{\partial}{\partial t} (\rho_\pi \bar{\mathbf{v}}^\pi) + \operatorname{div}(\partial_\pi \bar{\mathbf{v}}^\pi \otimes \bar{\mathbf{v}}^\pi) - \operatorname{div} \mathbf{t}^\pi - \rho_\pi [\bar{\mathbf{g}}^\pi + \mathbf{e}^\pi(\rho \dot{\mathbf{r}}) + \hat{\mathbf{t}}^\pi] = \mathbf{0} \quad (2.91)$$

Recalling that averaged quantities are constant in  $dv$ , we make use of the relationship

$$\operatorname{div}(\rho_\pi \bar{\mathbf{v}}^\pi \otimes \bar{\mathbf{v}}^\pi) = \bar{\mathbf{v}}^\pi \otimes \bar{\mathbf{v}}^\pi \operatorname{grad} \rho_\pi + \rho_\pi \operatorname{grad} \bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi + \rho \bar{\mathbf{v}}^\pi \operatorname{div} \bar{\mathbf{v}}^\pi \quad (2.92)$$

of the definition of the  $\pi$  phase acceleration

$$\bar{\mathbf{a}}^\pi = \dot{\bar{\mathbf{v}}}^\pi = \frac{\partial \bar{\mathbf{v}}^\pi}{\partial t} + \operatorname{grad} \bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi \quad (2.93)$$

along with the averaged macroscopic mass balance equation (2.54).

It follows that

$$\operatorname{div} \mathbf{t}^\pi + \rho_\pi (\bar{\mathbf{g}}^\pi - \bar{\mathbf{a}}^\pi) + \rho_\pi [\mathbf{e}^\pi(\rho \dot{\mathbf{r}}) + \hat{\mathbf{t}}^\pi] = \mathbf{0} \quad (2.94)$$

The quantity

$$\rho_\pi \bar{\mathbf{a}}^\pi \quad (2.95)$$

is the volume density of the inertia force, which is also a volume-bounded quantity. The stress tensor  $\mathbf{t}^\pi$ , introduced in (2.83), needs some comments. Williams [16] has argued that  $\mathbf{t}^\pi$  in the form used above and in the traditional formulation of mixture theory should be the sum of the stress tensor in the  $\pi$  phase itself and of the cross stress tensors [16,17] of the type,  $1/2 \mathbf{t}^{\pi\alpha}$ , which describe the forces exerted by the other constituents  $\alpha \neq \pi$  on the  $\pi$  phase. If  $da$  is a surface element in the porous body then  $\int_{da} \mathbf{t}^{\pi\alpha} \cdot \mathbf{n} da_m = \mathbf{F}^{\pi\alpha}(da) + \mathbf{F}^{\alpha\pi}(da)$ , where  $\mathbf{F}^{\pi\alpha}$  is the contact force exerted by the  $\pi$  phase on the  $\alpha$  phase across  $da$  and  $\mathbf{F}^{\alpha\pi}$  is the force exerted by the  $\alpha$  phase on the  $\pi$  phase across  $da$ . Note that only their sum is balanced by a stress tensor because the interaction between the  $\pi$  phase interior to a diaphragm with the  $\alpha$  phase exterior to the same diaphragm is not the same as the interaction between the  $\alpha$  phase interior to the diaphragm with the  $\pi$  phase exterior to the diaphragm [17]. Williams and Noll consider additional identities of the type

$$\operatorname{div} \mathbf{t}^{\pi\alpha} - \mathbf{d}^{\alpha\pi} - \mathbf{d}^{\pi\alpha} = \mathbf{0} \quad (2.96)$$

where  $\mathbf{d}^{\alpha\pi}$ ,  $\mathbf{d}^{\pi\alpha}$  are the densities of the diffusive interactions  $D$ , i.e. of the volume-distributed force exerted by the  $\pi$  phase on the  $\alpha$  phase and vice versa. These identities come from the equations of balance of force and are regarded in the literature [16] as identities giving the constitutive descriptions of  $\mathbf{d}^{\alpha\pi} + \mathbf{d}^{\pi\alpha}$ . In accordance with the above definition [16] of the tensor  $\mathbf{t}^\pi$  the densities of the diffusive interaction  $D$  in (2.85) and (2.87) should be a combination of the type  $1/2(\mathbf{d}^{\pi\alpha} - \mathbf{d}^{\alpha\pi})$  for the  $\pi$  phase and each other constituent  $\alpha$ . Neither the stress tensor defined through (2.83) nor the diffusive interaction forces appear to take into account all the interaction terms from the  $\pi$  phase on the  $\alpha$  phases mentioned here. Therefore we must take  $\mathbf{t}^{\pi\alpha} = \mathbf{0}$  as a constitutive assumption, which means that the only contact forces applied to a portion of the  $\pi$  phase inside a diaphragm are those applied by the  $\pi$  phase external to the diaphragm [4].

With the assumption that  $\mathbf{t}^{\pi\alpha} = \mathbf{0}$  it follows that the interaction densities are equal and opposite, from (2.96). Hence,  $1/2(\mathbf{d}^{\pi\alpha} - \mathbf{d}^{\alpha\pi}) = \mathbf{d}^{\pi\alpha}$ , i.e. the supply terms considered (2.86), (2.87) and (2.88), express for  $\mathbf{t}^{\pi\alpha} = \mathbf{0}$  the totality of the interactions from the  $\alpha$  phases on the  $\pi$  phase and from the  $\pi$  phase on the  $\alpha$  phases. Sampaio and Williams [18] give examples with  $\mathbf{t}^{\pi\alpha} \neq \mathbf{0}$ . The assumption  $\mathbf{t}^{\pi\alpha} = \mathbf{0}$  is therefore a limiting assumption.

In the following, we make a further assumption that  $\rho_\pi \mathbf{e}^\pi(\rho \dot{\mathbf{r}}) \neq \mathbf{0}$  only for the fluid phases. Then the linear momentum balance equation for the solid becomes

$$\operatorname{div} \mathbf{t}^s + \rho_s (\bar{\mathbf{g}}^s - \bar{\mathbf{a}}^s) + \rho_s \hat{\mathbf{t}}^s = \mathbf{0} \quad (2.97)$$

and for the fluids it has the form of (2.94). The average linear momentum balance equations are subject to the constraint (2.33):

$$\sum_\pi \rho_\pi [\mathbf{e}^\pi(\rho \dot{\mathbf{r}}) + \hat{\mathbf{t}}^\pi] = \mathbf{0} \quad (2.98)$$

### 2.3.4 Angular momentum balance equation

As indicated in Section 2.3, all phases of the semisaturated porous medium are considered microscopically non-polar. The following microscopic variables are necessary for the balance equation (2.13) when angular momentum balance is considered:

$$\begin{aligned}\psi &= \mathbf{r} \times \dot{\mathbf{r}} \\ \mathbf{i} &= \mathbf{r} \times \mathbf{t}_m \\ \mathbf{b} &= \mathbf{r} \times \mathbf{g} \\ \mathbf{G} &= \mathbf{0}\end{aligned}\tag{2.99}$$

Hence we obtain the microscopic angular momentum balance equation in the form

$$\frac{\partial}{\partial t} [\rho(\mathbf{r} \times \dot{\mathbf{r}})] + \operatorname{div}[\rho(\mathbf{r} \times \dot{\mathbf{r}}) \otimes \dot{\mathbf{r}}] - \operatorname{div}(\mathbf{r} \times \mathbf{t}_m) - \mathbf{r} \times \rho \mathbf{g} = \mathbf{0}\tag{2.100}$$

Using (2.42) and (2.80), as well as taking into account that

$$\begin{aligned}\frac{\partial}{\partial t} [\rho(\mathbf{r} \times \dot{\mathbf{r}})] &= \frac{\partial \rho}{\partial t} (\mathbf{r} \times \dot{\mathbf{r}}) + \rho \left( \frac{\partial \mathbf{r}}{\partial t} \times \dot{\mathbf{r}} \right) + \rho \left( \mathbf{r} \times \frac{\partial \dot{\mathbf{r}}}{\partial t} \right) \\ \operatorname{div}[\rho(\mathbf{r} \times \dot{\mathbf{r}}) \otimes \dot{\mathbf{r}}] &= (\operatorname{grad} \rho \cdot \dot{\mathbf{r}})(\mathbf{r} \times \dot{\mathbf{r}}) \\ &\quad + (\rho \operatorname{div} \dot{\mathbf{r}})(\mathbf{r} \times \dot{\mathbf{r}}) + \rho \dot{\mathbf{r}} \cdot \operatorname{grad} \mathbf{r} \times \dot{\mathbf{r}} + \rho \mathbf{r} \times \dot{\mathbf{r}} \cdot \operatorname{grad} \mathbf{r}\end{aligned}\tag{2.101}$$

and

$$\operatorname{div}(\mathbf{r} \times \mathbf{t}_m) = \operatorname{grad} \mathbf{r} \times \mathbf{t}_m + \mathbf{r} \times \operatorname{div} \mathbf{t}_m$$

one obtains

$$\left[ \frac{D\rho}{Dt} + \rho \operatorname{div} \dot{\mathbf{r}} \right] (\mathbf{r} \times \dot{\mathbf{r}}) - \mathbf{r} \times [\operatorname{div} \mathbf{t}_m + \rho(\mathbf{b} - \ddot{\mathbf{r}})] - \operatorname{grad} \mathbf{r} \times \mathbf{t}_m = \mathbf{0}\tag{2.102}$$

After application of the mass balance equation (2.53) and the linear momentum balance equation (2.81), along with the relation

$$\operatorname{grad} \mathbf{r} = \mathbf{I}\tag{2.103}$$

where  $\mathbf{I}$  is the identity tensor, equation (2.102) reduces to

$$\mathbf{I} \times \mathbf{t}_m = \mathbf{0}\tag{2.104}$$

which implies the symmetry of the stress tensor

$$\mathbf{t}_m = (\mathbf{t}_m)^T\tag{2.105}$$

In a similar way, it can be shown, with an appropriate method chosen for the development of the average angular momentum equation [1,4], that for non-polar media, also at macroscopic level, the partial stress tensor is symmetric

$$\mathbf{t}^\pi = (\mathbf{t}^\pi)^T \quad (2.106)$$

and that the sum of the coupling vectors of angular momentum between the phases vanishes. The rather lengthy development is given in the literature [1,4].

### 2.3.5 Balance of energy equation

For the energy balance, the following components must be taken into account in the generic microscopic balance equation (2.13):

$$\begin{aligned} \psi &= E + \frac{1}{2} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \\ \mathbf{i} &= \mathbf{t}_m \dot{\mathbf{r}} - \mathbf{q} \\ b &= \mathbf{g} \cdot \dot{\mathbf{r}} + h \\ G &= 0 \end{aligned} \quad (2.107)$$

where  $E(\mathbf{r}, t)$  is the specific intrinsic energy,  $\mathbf{q}(\mathbf{r}, t)$  is the heat flux vector and  $h(\mathbf{r}, t)$  is the intrinsic heat source. Hence the microscopic energy balance equation has the form

$$\frac{\partial}{\partial t} \left( \rho E + \rho \frac{1}{2} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \right) + \operatorname{div} \left[ \rho \left( E + \frac{1}{2} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \right) \dot{\mathbf{r}} \right] - \operatorname{div}(\mathbf{t}_m \dot{\mathbf{r}} - \mathbf{q}) - \rho(\mathbf{g} \cdot \dot{\mathbf{r}} + h) = 0 \quad (2.108)$$

Taking into account the symmetry of the stress tensor (2.105),  $\mathbf{t}_m = \mathbf{t}_m^T$ , the definition of the velocity gradient tensor (2.50),  $\mathbf{L} = \operatorname{grad} \dot{\mathbf{r}}$ , and the following relation [1]:

$$\mathbf{t}_m : \mathbf{W} = 0 \quad (2.109)$$

we obtain

$$\begin{aligned} \frac{D}{Dt} (\rho E) + \rho E \operatorname{div} \dot{\mathbf{r}} + \frac{D}{Dt} \left( \rho \frac{1}{2} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \right) + \frac{1}{2} \rho \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \operatorname{div} \dot{\mathbf{r}} \\ = \rho \ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} + \mathbf{t}_m : \mathbf{D} - \operatorname{div} \mathbf{q} + \rho h \end{aligned} \quad (2.110)$$

Taking into account the microscopic mass balance equation (2.53) and the microscopic linear momentum balance equation (2.81), the final form of the conservation law of energy is obtained as

$$\rho \frac{DE}{Dt} = \mathbf{t}_m : \mathbf{D} - \operatorname{div} \mathbf{q} + \rho h \quad (2.111)$$

The average quantities necessary for the macroscopic balance equation (2.30) are the following [1,4]:

$$\bar{\psi}^\pi = \bar{E}^\pi + \frac{1}{2} \bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi \quad (2.112)$$

where  $\bar{E}^\pi$  accounts for the averaged specific energy of the volume element and for the averaged kinetic energy related to the deviation,  $\tilde{\mathbf{r}}^\pi$  from the mean velocity in  $dv$ . The flux term is obtained after substitution of the generic quantities in (2.28) using (2.107) and (2.112):

$$\tilde{i}^\pi = \frac{1}{da} \int_{da} \left[ \mathbf{t}_m \tilde{\mathbf{r}} - \mathbf{q} - \rho \left( E + \frac{1}{2} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} - \bar{E}^\pi - \frac{1}{2} \bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi \right) \tilde{\mathbf{r}}^\pi \right] \cdot \mathbf{n} \gamma^\pi da_m \quad (2.113)$$

After further transformation by use of (2.82) and the deviation definition (2.21) applied for the  $\pi$  phase velocity

$$\tilde{\mathbf{v}}^\pi = \tilde{\mathbf{r}}^\pi = \dot{\mathbf{r}}^\pi - \bar{\mathbf{r}}^\pi = \dot{\mathbf{r}}^\pi - \bar{\mathbf{v}}^\pi \quad (2.114)$$

the following form for  $\tilde{i}^\pi$  is obtained:

$$\tilde{i}^\pi = \tilde{\mathbf{t}}^\pi \cdot \bar{\mathbf{v}}^\pi + \frac{1}{da} \int_{da} \left[ \mathbf{t}_m \tilde{\mathbf{r}} - \mathbf{q} - \rho \left( E + \frac{1}{2} \dot{\mathbf{r}}^\pi \cdot \dot{\mathbf{r}}^\pi \right) \tilde{\mathbf{r}}^\pi \right] \cdot \mathbf{n} \gamma^\pi da_m \quad (2.115)$$

The integral expression represents the averaged internal heat flux  $\tilde{\mathbf{q}}^\pi$ . Following (2.28), a macroscopic heat flux vector can be defined such that

$$\tilde{\mathbf{q}}^\pi = \tilde{\mathbf{q}}^\pi \cdot \mathbf{n} \quad (2.116)$$

and (2.115) can be written in simplified form as

$$\tilde{i}^\pi = \tilde{\mathbf{t}}^\pi \cdot \bar{\mathbf{v}}^\pi - \tilde{\mathbf{q}}^\pi \quad (2.117)$$

where the negative sign takes into account the opposite directions of  $\mathbf{q}^\pi$  and  $\mathbf{n}$  [1].

The term for the external energy supply is

$$\bar{b}^\pi = h^\pi + \mathbf{g}^\pi \cdot \bar{\mathbf{v}}^\pi \quad (2.118)$$

where  $h^\pi$  results from the sum of the averaged heat sources,  $h$  and the averaged work done by external volume forces due to the velocity difference  $\tilde{\mathbf{r}}^\pi$ .

The supply terms (2.26) and (2.27) are dealt with next.

$$e^\pi(\rho\dot{\psi}) = e^\pi \left( \rho E + \frac{1}{2} \rho \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \right) = \frac{1}{\rho_\pi dv} \sum_{\alpha \neq \pi} \int_{da^{\alpha\pi}} \rho \left( E + \frac{1}{2} \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \right) (\mathbf{w} - \dot{\mathbf{r}}) \cdot \mathbf{n}^{\alpha\pi} da_m \quad (2.119)$$

Using the deviation term (2.21), It is possible to transform (2.119) as

$$e^\pi(\rho\psi) = \mathbf{e}^\pi(\rho\tilde{\mathbf{r}}^\pi) \cdot \bar{\mathbf{v}}^\pi + \frac{1}{2}e^\pi(\rho)\bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi + e^\pi(\rho\hat{E}) \quad (2.120)$$

where

$$e^\pi(\rho\hat{E}) = \frac{1}{\rho_\pi dv} \sum_{\gamma \neq \pi} \int_{da^{\pi\gamma}} \rho \left( E + \frac{1}{2} \tilde{\mathbf{r}} \cdot \tilde{\mathbf{r}} \right) (\mathbf{w} - \tilde{\mathbf{r}}) \cdot \mathbf{n}^{\pi\gamma} da_m \quad (2.121)$$

is the exchange term of internal energy due to phase change and possible mass exchange. Further more, according to (2.26), we have

$$I^\pi = \frac{1}{\rho_\pi dv} \sum_{\gamma \neq \pi} \int_{da^{\pi\gamma}} (\mathbf{t}_m \tilde{\mathbf{r}} - \mathbf{q}) \cdot \mathbf{n}^{\pi\gamma} da_m \quad (2.122)$$

Taking into account (2.88) and (2.21), we may write (2.122) as

$$I^\pi = \hat{\mathbf{t}}^\pi \cdot \bar{\mathbf{v}}^\pi + \frac{1}{\rho_\pi dv} \sum_{\gamma \neq \pi} \int_{da^{\pi\gamma}} (\mathbf{t}_m \tilde{\mathbf{r}} - \mathbf{q}) \cdot \mathbf{n}^{\pi\gamma} da_m = \hat{\mathbf{t}}^\pi \cdot \bar{\mathbf{v}}^\pi + Q^\pi \quad (2.123)$$

The integral expression indicated as  $Q^\pi$  represents the exchange of energy due to mechanical interaction.

Finally

$$\bar{G}^\pi = 0 \quad (2.124)$$

The macroscopic energy balance equation may now be written as

$$\begin{aligned} & \frac{\partial}{\partial t} \left[ \rho_\pi \left( \bar{E}^\pi + \frac{1}{2} \bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi \right) \right] + \operatorname{div} \left[ \rho_\pi \left( \bar{E}^\pi + \frac{1}{2} \bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi \right) \bar{\mathbf{v}}^\pi \right] \\ & - \operatorname{div}(\hat{\mathbf{t}}^\pi \bar{\mathbf{v}}^\pi - \tilde{\mathbf{q}}^\pi) - \rho_\pi(h^\pi + \mathbf{g}^\pi \cdot \bar{\mathbf{v}}^\pi) - \rho_\pi \mathbf{e}^\pi(\rho \tilde{\mathbf{r}}^\pi) \cdot \bar{\mathbf{v}}^\pi \\ & - \frac{1}{2} \rho_\pi e^\pi(\rho) \bar{\mathbf{v}}^\pi \cdot \bar{\mathbf{v}}^\pi - \rho_\pi e^\pi(\rho \hat{E}) - \rho_\pi \hat{\mathbf{t}}^\pi \cdot \bar{\mathbf{v}}^\pi - \rho_\pi Q^\pi = 0 \end{aligned} \quad (1.125)$$

Using the mass balance equation (2.54) for the generic  $\pi$  phase, the definition of material derivatives, the linear momentum balance equation (2.94) along with (2.50), the energy balance equation may be written as follows:

$$\rho_\pi \frac{D^\pi \bar{E}^\pi}{Dt} = \hat{\mathbf{t}}^\pi : \mathbf{D}^\pi + \rho_\pi h^\pi - \operatorname{div} \tilde{\mathbf{q}}^\pi + \rho_\pi R^\pi \quad (2.126)$$

where

$$\rho_\pi R^\pi = \rho_\pi [e^\pi(\rho \hat{E}) - e^\pi(\rho) \bar{E}^\pi + Q^\pi] \quad (2.127)$$

represents the exchange of energy between the  $\pi$  phase and other phases of the medium due to phase change and mechanical interaction. Equation (2.126) coincides with the energy balance equation in the mixture theory as shown in [1].

For energy balance, (2.33) has the form

$$\sum_{\pi} \rho_{\pi} \left[ e^{\pi}(\rho \hat{E}) + \mathbf{e}^{\pi}(\rho \bar{\mathbf{r}}) \cdot \bar{\mathbf{v}}^{\pi} + \frac{1}{2} e^{\pi}(\rho) \bar{\mathbf{v}}^{\pi} \cdot \bar{\mathbf{v}}^{\pi} + \hat{\mathbf{t}}^{\pi} \cdot \bar{\mathbf{v}}^{\pi} + Q^{\pi} \right] = 0 \quad (2.128)$$

and physically means that the total balance of energy exchange between all the phases is zero. Phase change and the corresponding supply terms will be considered in the following, only for the fluid phases.

### 2.3.6 Entropy inequality

Exploitation of entropy inequality is a tool for developing constitutive equations in a systematic manner, leading to a consistent thermodynamic description of the material behaviour at the macroscale. The use of entropy inequality further assures that the second law of thermodynamics is not violated. The procedure was proposed by Coleman and Noll [19]. It has been exploited by Sampaio and Williams [18], and recently by Gray and Hassanzadeh [20] for the development of constitutive equations for unsaturated flow in dry or partially saturated soil, including interfacial phenomena. Throughout this book we use existing constitutive relations which are closer to current laboratory practice, but we quote here the entropy inequalities from de Boer *et al.* [1] both for the microscale and for the macroscale.

The variables in the microscopic balance equation (2.13) are now

$$\psi = \lambda \quad \mathbf{i} = \Phi \quad b = s \quad G = \varphi \quad (2.129)$$

where  $\lambda$  is the specific entropy,  $\Phi$  the entropy flux vector and  $s$  an intrinsic entropy source. The net production  $\varphi$  denotes an increase of entropy. The balance equation then becomes

$$\frac{\partial}{\partial t} (\rho \lambda) + \operatorname{div}(\rho \lambda \dot{\mathbf{r}}) - \operatorname{div} \Phi - \rho s = \rho \varphi \quad (2.130)$$

This equation may be transformed in the usual manner, also making use of the mass balance equation, to obtain

$$\rho \frac{D\lambda}{Dt} - \rho s - \operatorname{div} \Phi = \varphi \quad (2.131)$$

By respectively identifying the entropy flux  $\Phi$  and the entropy source  $s$  with the energy flux vector  $\mathbf{q}$  and the energy source  $h$ , both divided by the absolute temperature,  $\theta$  [1], we may write (2.131) as

$$\rho \frac{D\lambda}{Dt} - \frac{1}{\theta} \rho h + \operatorname{div} \frac{1}{\theta} \mathbf{q} = \rho \varphi \quad (2.132)$$

According to the second law of thermodynamics, the entropy production is zero for reversible processes, whereas for irreversible processes

$$\rho\varphi \geq 0 \quad (2.133)$$

Equations (2.132) and (2.133) yield the entropy inequality for single-component media

$$\rho \frac{D\lambda}{Dt} - \frac{1}{\theta} \rho h + \operatorname{div} \frac{1}{\theta} \mathbf{q} \geq 0 \quad (2.134)$$

The macroscopic variables are identified as follows:

$$\bar{\psi}^\pi = \bar{\lambda}^\pi = \frac{1}{\rho_\pi dv} \int_{dv} \rho \lambda \gamma^\pi dv_m \quad (2.135)$$

is the averaged specific entropy of constituent  $\pi$ ,

$$\tilde{i}^\pi = \tilde{\Phi}^\pi = \frac{1}{da} \int_{da} (\Phi - \rho \bar{\lambda}^\pi \tilde{\mathbf{r}}^\pi) \cdot \mathbf{n} \gamma^\pi da_m \quad (2.136)$$

is the averaged entropy supply density, for which an entropy supply vector can be defined, such that

$$\tilde{\Phi}^\pi = \Phi^\pi \cdot \mathbf{n} \quad (2.137)$$

$$e^\pi(\rho\psi) = e^\pi(\rho\lambda) = \frac{1}{\rho_\pi dv} \sum_{\alpha \neq \pi} \int_{da^{\pi\alpha}} \rho \lambda (\mathbf{w} - \dot{\mathbf{r}}) \cdot \mathbf{n}^{\pi\alpha} da_m \quad (2.138)$$

is the entropy supply due to mass exchange.

$$I^\pi = \frac{1}{\rho_\pi dv} \sum_{\alpha \neq \pi} \int_{da^{\pi\alpha}} \Phi \cdot \mathbf{n}^{\pi\alpha} da_m \quad (2.139)$$

denotes the entropy supply due to mechanical interactions. Further

$$\bar{b}^\pi = \bar{s}^\pi \quad (2.140)$$

is the averaged entropy source and

$$\bar{G}^\pi = \bar{\varphi}^\pi \quad (2.141)$$

the averaged total entropy increase.

If the sum of both supply terms (2.138) and (2.139), is indicated as  $(1/\rho_\pi)\hat{\lambda}^\pi$  and the usual transformation is applied, we obtain

$$\rho_\pi \frac{D^\pi \bar{\lambda}^\pi}{Dt} + \rho_\pi e^\pi(\rho) \bar{\lambda}^\pi - \operatorname{div} \Phi^\pi - \rho_\pi \bar{s}^\pi - \hat{\lambda}^\pi = \rho_\pi \bar{\varphi}^\pi \quad (2.142)$$

Equation (2.33) is simply

$$\sum_{\pi} \hat{\lambda}^{\pi} = 0 \quad (2.143)$$

As for the microscopic situation, we now define

$$\sum_{\pi} \rho_{\pi} \dot{\phi}^{\pi} \geq 0 \quad (2.144)$$

By assuming that the absolute temperature  $\theta$  and the external acceleration  $g$  are microscopically constant, the entropy supply vector may be written as

$$\Phi^{\pi} = \frac{1}{\theta^{\pi}} \mathbf{q}^{\pi} \quad (2.145)$$

and equation (2.140) as

$$\tilde{s}^{\pi} = \frac{1}{\theta^{\pi}} h^{\pi} \quad (2.146)$$

The entropy inequality for the mixture, which is the quantity of interest here, taking into account (2.143) and (2.144), finally becomes

$$\sum_{\pi} \left[ \rho_{\pi} \frac{D^{\pi} \tilde{\lambda}^{\pi}}{Dt} + \rho_{\pi} e^{\pi}(\rho) \tilde{\lambda}^{\pi} + \text{div} \left( \frac{1}{\theta^{\pi}} \mathbf{q}^{\pi} \right) - \frac{1}{\theta^{\pi}} \rho_{\pi} h^{\pi} \right] \geq 0 \quad (2.147)$$

Again, this corresponds to the form used in the mixture theory [1]. Before making further transformations of the macroscopic balance equations, we introduce the constitutive equations for the constituents.

## 2.4 CONSTITUTIVE EQUATIONS

To complete the description of the mechanical behaviour, we now need to specify the constitutive equations. The balance equations developed in the previous sections allow for the introduction of quite elaborate constitutive theories, especially if the balance equations for the bulk material are extended to the interfaces, as done by Gray and Hassanizadeh [20–22] for the aspects concerning multiphase flow. For the solid phase, second-grade material theories are also possible, where the gradients of relevant thermodynamic properties, such as densities, are considered as independent variables [1]. However, since this book is application oriented, i.e. we aim for the quantitative solution of real engineering problems, we make a different choice.

We select constitutive models which are based on quantities currently measurable in laboratory or field experiments, and which have been extensively validated both with reference to known exact solutions and to experiments. Most of these constitutive models correspond to linearisations of more complex arguments and have been obtained

from the entropy inequality [20–22]. We deal first with the properties of the fluid phases, and only briefly mention the solid phase here, because the whole of Chapter 4 is devoted to that subject.

### 2.4.1 Stress tensor in the fluid phases

By applying entropy inequality (2.147) for the bulk material [5,20] the stress tensor in the fluid phases that appears in equation (2.94) may be written as

$$\mathbf{t}^\pi = -\eta^\pi p^\pi \mathbf{I} \quad (2.148)$$

where  $\mathbf{I}$  is the identity tensor. This operator is equivalent to the Kronecker symbol  $\delta$ , (Section 2.6.2). Quantity  $p^\pi$  is the macroscopic pressure of the  $\pi$  phase.

The volume fraction  $\eta^\pi$  appears in (2.148) because  $\mathbf{t}^\pi$  is the force exerted on the fluid phase per unit area of multiphase medium. Note that the stress vector in the fluid phase does not have any dissipating part. The macroscopic effects of deviatoric stress components will be accounted for in linear momentum balance equations through momentum exchange terms (2.164).

### 2.4.2 Gaseous mixture of dry air and water vapour

The moist air in the pore system is usually assumed to be a perfect mixture of two ideal gases, i.e. dry air and water vapour. Hence we use the ideal gas law, relating the partial pressure  $p_{g\pi}$  of species  $\pi$ , the mass concentration  $\rho^{g\pi}$  of species  $\pi$  in the gas phase and the absolute temperature  $\theta$ .

The equations of state of a perfect gas, applied to dry air (ga), vapour (gw) and moist air (g) are

$$\begin{aligned} p^{ga} &= \rho^{ga} \theta R / M_a \\ p^{gw} &= \rho^{gw} \theta R / M_w \end{aligned} \quad (2.149)$$

$$\begin{aligned} \rho^g &= \rho^{ga} + \rho^{gw} \\ p^g &= p^{ga} + p^{gw} \end{aligned} \quad (2.150)$$

$$M_g = \left( \frac{\rho^{gw}}{\rho^g} \frac{1}{M_w} + \frac{\rho^{ga}}{\rho^g} \frac{1}{M_a} \right)^{-1}$$

where  $M_\pi$  is the molar mass of constituent  $\pi$  and  $R$  is the universal gas constant. The second equation in (2.150) expresses Dalton's law [23]. For the averaging process, remember that dry air, vapour and moist air occupy the same volume fraction  $nS_g$ .

### 2.4.3 Sorption equilibrium

If an oven-dry porous medium is exposed to moist air, the weight of the solid increases because the moisture is adsorbed on the inner surfaces of the pores, starting with the

finest ones. In the cases of interest here, the water is usually present as a condensed liquid that, because of the surface tension, is separated from its vapour by a concave meniscus (capillary water). There is then a relationship between the relative humidity, the water content (saturation) and the capillary pressure in the pores.

The capillary pressure is defined as the pressure difference between the gas phase and the liquid phase, by the capillary pressure equation

$$p^c = p^g - p^w \quad (2.151)$$

where  $p^w$  is the pressure of the liquid phase (water). Gray and Hassanzadeh [21] show that (2.151) is not just a definition, but a derived relationship between two independent quantities,  $p^c$  and  $p^g - p^w$ , at equilibrium.

For the relationship between the relative humidity (RH) and the capillary pressure in the pores, the Kelvin–Laplace law is assumed to be valid:

$$RH = \frac{p^{gw}}{p^{gws}} = \exp\left(\frac{p^c M_w}{\rho^w R \theta}\right) \quad (2.152)$$

The water vapour saturation pressure  $p^{gws}$ , which is a function of the temperature only, can be obtained from the Clausius–Clapeyron equation indicated below, or from empirical formulas such as the one proposed by Hyland and Wexler [24].

Assuming zero contact angle between the liquid phase and the solid phase, as is usually accepted for pore water, the capillary pressure can be obtained through the Laplace equation from the pore radius  $r$ :

$$p^c = \frac{2\sigma}{r} \quad (2.153)$$

where  $\sigma$  is the surface tension.

These considerations apply if the water is present in the pores as a condensed liquid (capillary region). But if the water is present as one or more molecular layers adsorbed on the surface of a solid because of the van der Waals and/or other interactions, the capillary pressure no longer has an obvious meaning, even if it can be retained, referring to the broader concept of water potential or moisture stress (Appendix 2B). In this case a direct relationship between the water content and the relative humidity is assumed to hold, such as the BET equation [24].

#### 2.4.4 Clausius–Clapeyron equation

The Clausius–Clapeyron equation links the water vapour saturation pressure with temperature:

$$p^{gws}(\theta) = p^{gws0} \exp\left[-\frac{M_w \Delta H_{gw}}{R} \left(\frac{1}{\theta} - \frac{1}{\theta_0}\right)\right] \quad (2.154)$$

where  $\theta_0$  is a reference temperature,  $p^{gws}$  is the water vapour saturation pressure at  $\theta$ ,

$p_{\text{gws}0}$  is the water vapour saturation pressure at  $\theta_0$  and  $\Delta H_{\text{gw}}$  is the specific enthalpy of evaporation. The equation is obtained from the second law of thermodynamics and is valid in the vicinity of  $\theta_0$ . In the following, we denote  $T$  as the temperature difference above a reference value  $\theta_0$  such that  $T = \theta - \theta_0$ .

#### 2.4.5 Pore size distribution

As it turns out from equations (2.152) and (2.153), the problem is to know the pore size distribution of the considered porous medium, to relate the size of the largest pore filled (on which the capillary pressure depends) with the actual water content. This relationship is obtained through experimental tests, usually centrifuge tests, sorption isotherm measurements or mercury porosimetry.

The question is somewhat complex, because both the Laplace equation (2.153) and the Kelvin equation (2.152) are obtained from a force equilibrium, evaluated in a cylindrical capillary tube, and the porous medium is considered as a bundle of capillary tubes that do not intersect. The real porous media are more correctly represented as three-dimensional networks of sites (or bodies) interconnected by narrower bonds (or throats) then intrusion, extrusion, adsorption and desorption are subject to hysteresis. In this case the actual value of capillary pressure can only be interpreted as a measurable quantity describing complex interactions between the solid matrix and the adsorbed water. The determination of pore size distribution from sorption isotherms, results of centrifuge tests and mercury porosimetry should follow the percolation theory approach [25].

The Rayleigh distribution or the log-normal distributions are often a good estimation of the pore size and often such distribution can be bimodal or multimodal. As an example, the experimental data for cellular concrete, obtained from Van der Kooi [26], using extracting cells with a porous plate, and the bimodal log-normal distribution used to approximate the data are shown in Figure 2.3.

For soils, we need the Kelvin–Laplace equation (2.152), the Clausius–Clapeyron equation (2.154) and the capillary pressure relationship

$$S_\pi = S_\pi(p^c, T) \quad (2.155)$$

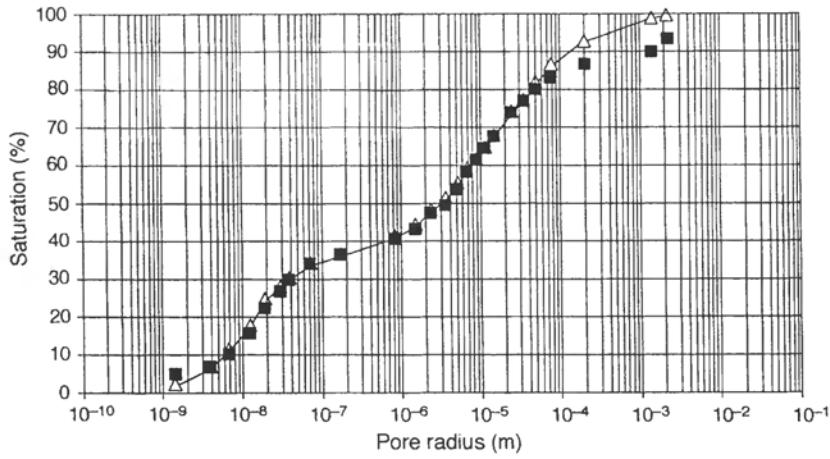
which is directly obtained in the laboratory.

A more complex functional dependence is suggested in the literature [21], including also the interfacial area  $a^{\text{wg}}$ , i.e. the area of the water–air interface per unit volume. Such a relationship would be free of the hysteresis loop, which appears when the dependence on  $a^{\text{wg}}$  is not considered.

#### 2.4.6 Equation of state for water

The equation of state for water has been given by Fernandez [27] as

$$\rho^w = \rho^{w0} \exp [-\beta_w T + C_w (p^w - p^{w0})] \quad (2.156)$$



**Figure 2.3** Pore size distribution for cellular concrete: (■) Van der Kooi data [26], (△) log-normal bimodal distribution (Courtesy F. Van der Kooi)

where the superscript o indicates an initial steady state at standard conditions;  $\beta_w$  is the thermal expansion coefficient and  $C_w$  is the compressibility coefficient. By retaining the first-order terms of the series expansion of (2.156) we obtain

$$\rho^w = \rho^{wo}[1 - \beta_w T + C_w(p^w - p^{wo})] \quad (2.157)$$

and

$$\frac{1}{\rho^{wo}} \frac{D^w \rho^w}{Dt} = \frac{1}{K_w} \frac{D^w p^w}{Dt} - \beta_w \frac{D^w T}{Dt} \quad (2.158)$$

where  $K_w = 1/C_w$  is the bulk modulus of water. Gray and Hassanizadeh [21] have suggested an equation of state that depends on more independent variables but they give no practical expression.

Equation (2.158) can also be obtained from the mass conservation in differential form:

$$\frac{D^w(\rho^w V^w)}{Dt} = 0 \quad (2.159)$$

In fact, by carrying out the differentiation of a product and keeping in mind that  $\rho^w = \rho^w(p^w, T)$ , we can write

$$\frac{1}{\rho^w} \frac{D^w \rho^w}{Dt} = -\frac{1}{V^w} \frac{D^w V^w}{Dt} = \frac{1}{\rho^w} \left( \frac{\partial \rho^w}{\partial p^w} \frac{D^w p^w}{Dt} + \frac{\partial \rho^w}{\partial T} \frac{D^w T}{Dt} \right) \quad (2.160)$$

Since

$$\begin{aligned}\frac{1}{\rho^{wo}} \frac{\partial \rho^w}{\partial p^w} &= \frac{1}{K_w} \\ \frac{1}{\rho^{wo}} \frac{\partial \rho^w}{\partial T} &= -\beta_w\end{aligned}\quad (2.161)$$

it follows that

$$\frac{1}{\rho^{wo}} \frac{D^w \rho^w}{Dt} = \frac{1}{K_w} \frac{D^w p^w}{Dt} - \beta_w \frac{D^w T}{Dt} \quad (2.162)$$

#### 2.4.7 Darcy's law

Generalized to allow for relative permeability, Darcy's law

$$\eta^\pi \bar{\mathbf{v}}^{\pi\alpha} = \frac{k^{\pi\alpha} \mathbf{k}}{\mu^\pi} (-\text{grad } p^\pi + \rho^\pi \mathbf{g}) \quad (2.163)$$

is assumed valid for the transport of both water and gas in slow phenomena;  $\mathbf{k}$  is the permeability tensor of the medium,  $\mu$  is the dynamic viscosity and  $k^{\pi\alpha}$  is the relative permeability, a dimensionless parameter varying from zero to one.

In fact, Darcy's law may be derived for fully saturated porous media by linearising the macroscopic balance equations for fluid momentum under the assumption that temperature gradients have a significant effect only in heat conduction phenomena and that inertial and macroscopic viscous effects are neglected [5] (Section 2.5.2.1). In this case it may be assumed that the momentum exchange term of (2.94) is

$$\rho_\pi \hat{\mathbf{t}}^\pi = -\mathbf{R}^\pi \eta^\pi \bar{\mathbf{v}}^{\pi\alpha} + p^\pi \text{grad } \eta^\pi \quad (2.164)$$

It represents the dissipative part of the fluid–solid exchange of momentum, accounting for the local effects, depending also on the viscosity of the fluid. We assume that  $\mathbf{R}^\pi$  is invertible and define

$$\mathbf{K}^\pi = \eta^\pi (\mathbf{R}^\pi)^{-1} = \frac{\mathbf{k}}{\mu^\pi} (\rho^\pi, \eta^\pi, T) \quad (2.165)$$

When several fluid phases are present, dissipative terms arise at fluid–solid and fluid–fluid interfaces. In this case we assume the permeability of the fully saturated case  $\mathbf{k}$  (also called intrinsic permeability) is modified by the presence of the other fluids. The change depends on the volumetric fraction of the other phases, capillary effects and fluid–fluid momentum exchange terms.

At a macroscopic level we assume the resulting permeability as a product of the intrinsic permeability times the relative permeability  $k^{\pi\alpha}$ ,

$$\mathbf{k}^\pi = k^{\pi\alpha} \mathbf{k} \quad (2.166)$$

as shown in (2.163). The intrinsic permeability varies with the void ratio

$$e = \frac{dv^w + dv^g}{dv^s} \quad (2.167)$$

and degree of saturation. For many situations the change of void ratio may be of secondary importance and  $k^{rw}$  may be satisfactorily defined as a function of the degree of saturation. For each particular porous medium, the relations  $k^{rw}(S_w)$  are either predicted by models based on some more or less realistic capillary assumption, or experimentally determined under laboratory and field conditions. Typical curves of relative permeabilities to water and air may be found in the literature [28].

The relative permeability goes to zero before the saturation reaches the value zero. The water saturation at which the relative permeability goes to zero is called the residual water saturation or the displacement residual water saturation. Among the relationships used in the following are those by Brooks and Corey [29]:

$$\begin{aligned} k^{rw} &= S_e^{(2+3\lambda)/\lambda} \\ k^{ra} &= (1 - S_e)^2 (1 - S_e^{(2+\lambda)/\lambda}) \\ p_c &= \frac{p^b}{S_e^{1/\lambda}} \end{aligned} \quad (2.168)$$

where  $S_e = (S_w - S_{wc})/(1 - S_{wc})$  is the effective saturation,  $S_{wc}$  is the irreducible saturation,  $\lambda$  is the pore size distribution index and  $p^b$  is the bubbling pressure.

Using the second law of thermodynamics, Hassanzadeh and Gray [21] have shown that a new term should be added to Darcy's equation to properly model unsaturated flow. This extra term accounts for changes in energy of the water phase due to changes in saturation and depends on a quantity called wettability potential. It is supposed to be of importance at lower saturation but no quantitative values are yet available.

Darcy's law for the more general case, where inertia forces are not negligible, will be derived in Section 2.5.2.1 from the linear momentum balance equation. Equations (2.164) to (2.166) are also considered valid under these conditions. The generalised version of Darcy's law obtained in this way will be used in soil dynamics.

#### 2.4.8 Fick's law

Diffusive-dispersive mass flux is governed by Fick's law

$$\mathbf{J}_\alpha^\pi = -\rho^\alpha \mathbf{D}_\alpha^\pi \operatorname{grad} \left( \frac{\rho^\pi}{\rho^\alpha} \right) \quad (2.169)$$

where  $\mathbf{D}_\alpha^\pi$  is the effective dispersion tensor,  $\pi$  is the diffusing phase and  $\alpha$  is the phase in which diffusion takes place ( $\alpha = w, g$ ).  $\mathbf{D}_\alpha^\pi$  is a function of the tortuosity factor, which accounts for the tortuous nature of the pathway in soil; because of mechanical dispersion,  $\mathbf{D}_\alpha^\pi$  is also correlated with seepage velocity.

For dry air and water vapour (binary system) we have in particular that

$$\begin{aligned}\mathbf{J}_g^{ga} &= -\rho^g \mathbf{D}_g^{ga} \operatorname{grad} \left( \frac{\rho^{ga}}{\rho^g} \right) \\ \mathbf{J}_g^{gw} &= -\rho^g \mathbf{D}_g^{gw} \operatorname{grad} \left( \frac{\rho^{gw}}{\rho^g} \right)\end{aligned}\quad (2.170)$$

The first equation of (2.150) leads to

$$\operatorname{grad} \left( \frac{\rho^{ga}}{\rho^g} \right) = \operatorname{grad} \left( \frac{\rho^g - \rho^{gw}}{\rho^g} \right) = -\operatorname{grad} \left( \frac{\rho^{gw}}{\rho^g} \right) \quad (2.171)$$

From (2.70) and (2.171) it follows, that

$$\mathbf{D}_g^{ga} = \mathbf{D}_g^{gw} = \mathbf{D}_g. \quad (2.172)$$

Finally, using (2.149) we obtain for binary gas that

$$\begin{aligned}\mathbf{J}_g^{ga} &= -\rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \operatorname{grad} \left( \frac{p^{ga}}{p^g} \right) \\ &= \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \operatorname{grad} \left( \frac{p^{gw}}{p^g} \right) = -\mathbf{J}_g^{gw}\end{aligned}\quad (2.173)$$

Remember that gas diffusion can take place even in the absence of a gas pressure gradient, i.e. when its mass-weighted velocity  $\bar{v}^g$  (2.66) is zero. Again, use of entropy inequality and subsequent linearisation allows more complex expressions [7] to be obtained, including temperature-induced transport.

#### 2.4.9 Stress tensor in the solid phase and total stress

From the entropy inequality written by Hassanzadeh and Gray [22] for unsaturated flow, including interfacial phenomena, it can be shown that the stress vector in the solid phase is

$$\mathbf{t}^s = (1-n)(\mathbf{t}_e^s - \mathbf{I}p^s) \quad (2.174)$$

the pressure in the solid phase is

$$p^s = p^w S_w + p^g S_g \quad (2.175)$$

and

$$\boldsymbol{\sigma}' = (1-n)\mathbf{t}_e^s \quad (2.176)$$

is the effective stress tensor.

Introduction of (2.175) into (2.174) yields

$$\mathbf{t}^s = (1 - n)[\mathbf{t}_e^s - \mathbf{I}(S_w p^w + S_g p^g)] \quad (2.177)$$

The volume fraction  $(1 - n)$  indicates that  $\mathbf{t}^s$  is the stress exerted on the solid phase per unit area of a multiphase medium. The sum of (2.177) and (2.148) written for gas and for water gives the total stress  $\boldsymbol{\sigma}$  acting on a unit area of a multiphase medium:

$$\begin{aligned} \boldsymbol{\sigma} &= \mathbf{t}^s + \mathbf{t}^w + \mathbf{t}^g = (1 - n)[\mathbf{t}_e^s - \mathbf{I}(S_w p^w + S_g p^g)] \\ &\quad - S_w n \mathbf{I} p^w - S_g n \mathbf{I} p^g = (1 - n)\mathbf{t}_e^s - \mathbf{I}(S_w p^w + S_g p^g) \end{aligned} \quad (2.178)$$

This can be put in the usual soil mechanics form, by means of (2.176), as

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' - \mathbf{I}(S_w p^w + S_g p^g) \quad (2.179)$$

or

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + \mathbf{I}(S_w p^w + S_g p^g) \quad (2.180)$$

From (2.180) it follows that the relationship between effective stress and total stress, in partially saturated porous media, is no longer independent of the soil type because of the saturations (2.155); whereas in fully saturated soils, where  $S_g = 0$  and  $S_w = 1$ , the effective stress principle (2.179) is unique for all soil types.

An expression of the form (2.179) was obtained by Bishop and Blight [30] using a phenomenological approach, and the same equation is obtained in Section 2.6.1 and by Schrefler *et al.* [31] by using volume averaging. Volume averaging is not the most appropriate procedure for obtaining macroscopic values of stresses, but because of Delesse's law (Section 2.2.1) it yields acceptable results in this case.

The effective stress, (2.176) or (2.180), is responsible for all major deformations in the skeleton and is linked to the strain rate tensor  $\mathbf{D}^s$  by means of a constitutive relationship

$$\frac{D\boldsymbol{\sigma}'}{Dt} = \mathbf{D}_T[(\mathbf{D}^s - \mathbf{D}_0^s)] \quad (2.181)$$

where

$$\mathbf{D}_T = \mathbf{D}_T(\mathbf{D}^s, \boldsymbol{\sigma}', T) \quad (2.182)$$

is a fourth-order tensor and  $\mathbf{D}_0^s$  represents the increment of all other strains not directly associated with stress changes. Various constitutive models defining the tangent tensor  $\mathbf{D}_T$  will be discussed in Chapter 4.

Assuming a small strain, we can say that

$$\mathbf{D}^s = \boldsymbol{\epsilon} dt \quad (2.183)$$

where  $\boldsymbol{\epsilon}$  is the linear strain tensor.

### 2.4.10 Solid density

When considering the solid phase as compressible, a relationship for the material time derivative of the solid density can be obtained from the mass conservation equation in differential form

$$\frac{D^s(\rho^s V^s)}{Dt} = 0 \quad (2.184)$$

By assuming that the solid density is a function of  $p^s$  (2.175), temperature and the first invariant of the effective stress, we proceed as in Section 2.4.6:

$$\frac{1}{\rho^s} \frac{D^s \rho^s}{Dt} = - \frac{1}{V_s} \frac{D^s V_s}{Dt} = \frac{1}{K_s} \frac{D^s p^s}{Dt} - \beta_s \frac{D^s T}{Dt} - \frac{1}{3(n-1)K_s} \frac{D^s(\text{tr } \sigma')}{Dt} \quad (2.185)$$

where we have kept in mind that

$$\begin{aligned} \frac{1}{\rho^s} \frac{\partial \rho^s}{\partial p^s} &= \frac{1}{K_s} \\ \frac{1}{\rho^s} \frac{\partial \rho^s}{\partial T} &= -\beta_s \\ \frac{1}{\rho^s} \frac{\partial \rho^s}{\partial \mathbf{I}_{1\sigma'}} &= -\frac{1}{3(n-1)K_s} \end{aligned} \quad (2.186)$$

where  $K_s$  is the bulk modulus of the grain material,  $\beta_s$  is the thermal expansion coefficient for the solid and  $\text{tr } \sigma' = \mathbf{I}_{1\sigma'}$  is the first stress invariant.

We introduce now the constitutive relationship for the first stress invariant as

$$\frac{D^s(\mathbf{I}_{1\sigma'})}{Dt} = 3K_T \left( \text{div } \mathbf{v}^s + \frac{1}{K_s} \frac{D^s p^s}{Dt} - \beta_s \frac{D^s T}{Dt} \right) \quad (2.187)$$

where  $K_T$  is the bulk modulus of the skeleton, different from that of the grain material and

$$-\frac{1}{K_s} \frac{D^s p^s}{Dt} \quad (2.188)$$

represents an overall volumetric strain rate caused by uniform compression of particles (as opposed to the skeleton) by the average pressure  $p^s$ . In soils this volumetric strain is relatively insignificant and can be ignored, but it is important in rock mechanics and in concrete, where the compressibility of the solid phase is comparable to that of the skeleton.

Defining further Biot's constant [32] as

$$1 - \alpha = \frac{K_T}{K_s} \quad (2.189)$$

we obtain

$$\frac{1}{\rho^s} \frac{D^s \rho^s}{Dt} = \frac{1}{1-n} \left[ (\alpha - n) \frac{1}{K_s} \frac{D^s p^s}{Dt} - \beta_s (\alpha - n) \frac{D^s T}{Dt} - (1 - \alpha) \operatorname{div} \bar{\mathbf{v}}^s \right] \quad (2.190)$$

For incompressible grain material  $1/K_s = 0$ ,  $\alpha = 1$ . This does not imply that the solid skeleton is rigid, because of rearrangements of the voids. The necessary evolution equation is the mass balance equation for the solid phase (2.58).

### 2.4.11 Fourier's Law

A constitutive assumption for the heat flux is the generalised version of Fourier's law

$$\tilde{\mathbf{q}} = -\chi_{\text{eff}} \operatorname{grad} T \quad (2.191)$$

where  $\chi_{\text{eff}}$  is the effective thermal conductivity tensor and  $\tilde{\mathbf{q}}$  is the heat flux of the multiphase medium, the sum of the partial heat fluxes  $\tilde{\mathbf{q}}^\pi$ .

For isotropic media (2.191) becomes the well-known empirical version of Fourier's law

$$q = -\chi_{\text{eff}} \operatorname{grad} T \quad (2.192)$$

where  $\chi_{\text{eff}}$  is the effective thermal conductivity. The effective thermal conductivity can be predicted theoretically as well as determined experimentally. For a wider description of methods, see Whitaker [33] and Nozad *et al.* [34].

For porous building materials the following linear relationship may be used, which represents with sufficient accuracy the data by Bomberg and Shirtliffe [25]:

$$\chi_{\text{eff}} = \chi_{\text{dry}} \left( 1 + 4 \frac{n S^w \rho^w}{(1-n) \rho^s} \right) \quad (2.193)$$

## 2.5 GENERAL FIELD EQUATIONS

The macroscopic balance laws are now transformed, and the constitutive equations introduced, to obtain the general field equations, which will be used in the subsequent chapters. The averaging symbol, overbar, will be omitted in the remainder of this book, because all quantities belong to the macroscopic situations.

### 2.5.1 Mass balance equation

The macroscopic mass balance equation for the solid phase (2.58), divided by  $\rho^s$  is

$$\frac{(1-n)}{\rho^s} \frac{D^s \rho^s}{Dt} - \frac{D^s n}{Dt} + (1-n) \operatorname{div} \mathbf{v}^s = 0 \quad (2.194)$$

Upon introduction of the relative velocity (2.47) and the material time derivative with respect to the moving solid (2.44), equation (2.60) for liquid water becomes

$$\frac{D^s \rho_w}{Dt} + \mathbf{v}^{ws} \cdot \operatorname{grad} \rho_w + \rho_w \operatorname{div} (\mathbf{v}^s + \mathbf{v}^{ws}) = -\dot{m} \quad (2.195)$$

Introduction of intrinsic phase-averaged densities with the appropriate volume fractions, use of vector identity (2.57) written for water and division by  $S_w \rho^w$  allows us to transform (2.195) into

$$\frac{D^s n}{Dt} + \frac{n}{\rho^w} \frac{D^s \rho^w}{Dt} + \frac{n}{S_w} \frac{D^s S_w}{Dt} + \frac{1}{S_w \rho^w} \operatorname{div}(n S_w \rho^w \mathbf{v}^{ws}) + n \operatorname{div} \mathbf{v}^s = -\frac{\dot{m}}{S_w \rho^w} \quad (2.196)$$

Summation with (2.194), to eliminate  $D^s n / Dt$ , gives

$$\frac{(1-n)}{\rho^s} \frac{D^s \rho^s}{Dt} + \operatorname{div} \mathbf{v}^s + \frac{n}{\rho^w} \frac{D^s \rho^w}{Dt} + \frac{n}{S_w} \frac{D^s S_w}{Dt} + \frac{1}{S_w \rho^w} \operatorname{div}(n S_w \rho^w \mathbf{v}^{ws}) = -\frac{1}{S_w \rho^w} \dot{m} \quad (2.197)$$

Introduction of (2.162) and (2.190) for the material derivatives of the solid and water densities along with (2.175) gives

$$\begin{aligned} & \frac{\alpha - n}{K_s} \frac{D^s}{Dt} (S^w p^w + S^g p^g) - \beta_s (\alpha - n) \frac{D^s T}{Dt} \\ & + \alpha \operatorname{div} \mathbf{v}^s + n \left( \frac{1}{K_w} \frac{D^s p^w}{Dt} - \beta_w \frac{D^s T}{Dt} \right) \\ & + \frac{n}{S_w} \frac{D^s S_w}{Dt} + \frac{1}{S_w \rho^w} \operatorname{div}(n S_w \rho^w \mathbf{v}^{ws}) = -\frac{1}{S_w \rho^w} \dot{m} \end{aligned} \quad (2.198)$$

Because of (2.51) we have

$$\frac{D^s S_g}{Dt} = -\frac{D^s S_w}{Dt} \quad (2.199)$$

Carrying out derivatives of  $p^\pi$ , collecting terms and employing (2.199) yields

$$\begin{aligned} & \left( \frac{\alpha - n}{K_s} S_w^2 + \frac{n S_w}{K_w} \right) \frac{D^s p^w}{Dt} + \frac{\alpha - n}{K_s} S_w S_g \frac{D^s p^g}{Dt} + \alpha S_w \operatorname{div} \mathbf{v}^s \\ & - \beta_{sw} \frac{D^s T}{Dt} + \left( \frac{\alpha - n}{K_s} p^w S_w - \frac{\alpha - n}{K_s} p^g S_w + n \right) \frac{D^s S_w}{Dt} \\ & + \frac{1}{\rho^w} \operatorname{div}(n S_w \rho^w \mathbf{v}^{ws}) = -\frac{\dot{m}}{\rho^w} \end{aligned} \quad (2.200)$$

where

$$\beta_{sw} = S_w [(\alpha - n) \beta_s + n \beta_w] \quad (2.201)$$

For incompressible grains  $\alpha = 1$ ,  $1/K_s = 0$  hence (2.200) may be simplified as follows:

$$\frac{nS_w}{K_w} \frac{D^s p^w}{Dt} + S_w \operatorname{div} \mathbf{v}^s - \beta_{sw} \frac{D^s T}{Dt} + n \frac{D^s S_w}{Dt} + \frac{1}{\rho^w} \operatorname{div} (nS_w \rho^w \mathbf{v}^{ws}) = -\frac{\dot{m}}{\rho^w} \quad (2.202)$$

The mass balance equation for gas as a mixture of dry air and vapour is dealt with next. Material time derivatives with respect to the moving solid (2.44) and the relative velocities (2.47) are introduced in (2.71) to obtain

$$\frac{D^s}{Dt} (nS_g \rho^g) + nS_g \rho^g \operatorname{div} (\mathbf{v}^s + \mathbf{v}^{gs}) + \mathbf{v}^{gs} \cdot \operatorname{grad} (nS_g \rho^g) = \dot{m} \quad (2.203)$$

For the sake of brevity, vector identity (2.57) is introduced into (2.203), then the material time derivative of the first term is carried out and the equation is divided by  $\rho^g S_g$ , yielding

$$\frac{D^s}{Dt} n + \frac{n}{\rho^g} \frac{D^s}{Dt} \rho^g + \frac{n}{S_g} \frac{D^s S_g}{Dt} + \frac{1}{\rho^g S_g} \operatorname{div} (nS_g \rho^g \mathbf{v}^{gs}) + n \operatorname{div} \mathbf{v}^s = \frac{\dot{m}}{\rho^g S_g} \quad (2.204)$$

$D^s n / Dt$  is eliminated by summation of (2.204) with (2.194):

$$\frac{(1-n)}{\rho^s} \frac{D^s \rho^s}{Dt} + \operatorname{div} \mathbf{v}^s + \frac{n}{\rho^g} \frac{D^s \rho^g}{Dt} + \frac{n}{S_g} \frac{D^s S_g}{Dt} + \frac{1}{\rho^g S_g} \operatorname{div} (nS_g \rho^g \mathbf{v}^{gs}) = \frac{\dot{m}}{\rho^g S_g} \quad (2.205)$$

Introduction of (2.149) and (2.150) gives

$$\begin{aligned} & \frac{(1-n)}{\rho^s} \frac{D^s \rho^s}{Dt} + \operatorname{div} \mathbf{v}^s + \frac{n}{\rho^g} \frac{D^s}{Dt} \left[ \frac{1}{\theta R} (p^{ga} M_a + p^{gw} M_w) \right] \\ & + \frac{n}{S_g} \frac{D^s S_g}{Dt} + \frac{1}{\rho^g S_g} \operatorname{div} (nS_g \rho^g \mathbf{v}^{gs}) = \frac{\dot{m}}{\rho^g S_g} \end{aligned} \quad (2.206)$$

The material time derivative of the solid density may be expressed as in the mass balance equation for liquid water by means of (2.190) and (2.175). Furthermore, (2.199) is used and the resulting equation is multiplied by  $S_g$ , yielding

$$\begin{aligned} & \frac{\alpha - n}{K_s} S_w S_g \frac{D^s p^w}{Dt} + \frac{\alpha - n}{K_s} S_g^2 \frac{D^s p^g}{Dt} - \left( n + \frac{\alpha - n}{K_s} p^c S_g \right) \frac{D^s S_w}{Dt} \\ & - \beta_s (\alpha - n) S_g \frac{D^s T}{Dt} + \alpha S_g \operatorname{div} \mathbf{v}^s + \frac{n S_g}{\rho^g} \frac{D^s}{Dt} \left[ \frac{1}{\theta R} (p^{ga} M_a + p^{gw} M_w) \right] \\ & + \frac{1}{\rho^g} \operatorname{div} (n S_g \rho^g \mathbf{v}^{gs}) = \frac{\dot{m}}{\rho^g} \end{aligned} \quad (2.207)$$

For incompressible solid grains this equation is simplified to

$$\begin{aligned} -n \frac{D^s S_w}{Dt} - \beta_s(1-n)S_g \frac{D^s T}{Dt} + S_g \operatorname{div} \mathbf{v}^s + \frac{n S_g}{\rho^g} \frac{D^s}{Dt} \left[ \frac{1}{\theta R} (p^{ga} M_a + p^{gw} M_w) \right] \\ + \frac{1}{\rho^g} \operatorname{div} (n S_g \rho^g \mathbf{v}^{gs}) = \frac{\dot{m}}{\rho^g} \end{aligned} \quad (2.208)$$

Equations (2.207) and (2.200), both with  $\dot{m} = 0$ , were used in a model of heat and mass transfer in deforming porous media with low temperature gradients and negligible phase change phenomena [35] and for pollutant transport analysis [36].

For heat transfer analysis, in partially saturated porous media, it is more convenient to consider the mass balance equation for dry air separately from that of vapour and to sum the mass balance equations for both water species, liquid water and water vapour [37, 38]. In this way the mass rate of water evaporation  $\dot{m}$  disappears from the mass balance equations. But it does require an evolution equation for  $\dot{m}$ , and this will be given by the energy balance equation. Note that, in this way, no constitutive model for the mass rate of water evaporation is needed.

The mass balance equation for dry air (2.62) is transformed in the following equation:

$$\frac{D^g}{Dt} (n S_g \rho^{ga}) + \operatorname{div} \mathbf{J}_g^{ga} + n S_g \rho^{ga} \operatorname{div} \mathbf{v}^g = 0 \quad (2.209)$$

in the same way as in Section 2.3.2.3 we transform (2.63) into (2.76). This equation can also be obtained by subtracting the mass balance equation for vapour (2.76) from that of gas (2.71) and recalling (2.173).

Equation (2.209) is now transformed as the mass balance equation for gas in this section, i.e. material time derivatives for the moving solid (2.44) and relative velocities (2.47) are introduced, then vector identity (2.57) is applied and the time derivatives of the first term are carried out as above. The resulting equation is divided by  $\rho^{ga} S_g$  and summed with (2.194). After introduction of (2.175) and (2.190), we obtain

$$\begin{aligned} \frac{\alpha - n}{K_s} S_g S_w \frac{D^s p^w}{Dt} + \frac{\alpha - n}{K_s} S_g^2 \frac{D^s p^g}{Dt} - \left( n + \frac{\alpha - n}{K_s} S_g p^c \right) \frac{D^s S_w}{Dt} \\ - \beta_s (\alpha - n) S_g \frac{D^s T}{Dt} + S_g \alpha \operatorname{div} \mathbf{v}^s + \frac{S_g n}{\rho^{ga}} \frac{D^s \rho^{ga}}{Dt} + \frac{1}{\rho^{ga}} \operatorname{div} \mathbf{J}_g^{ga} + \frac{1}{\rho^{ga}} \operatorname{div} (n S_g \rho^{ga} \mathbf{v}^{gs}) = 0 \end{aligned} \quad (2.210)$$

And for incompressible solid grains this equation simplifies to

$$\begin{aligned} -n \frac{D^s S_w}{Dt} - \beta_s(1-n)S_g \frac{D^s T}{Dt} + S_g \operatorname{div} \mathbf{v}^s + \frac{S_g n}{\rho^{ga}} \frac{D^s \rho^{ga}}{Dt} \\ + \frac{1}{\rho^{ga}} \operatorname{div} \mathbf{J}_g^{ga} + \frac{1}{\rho^{ga}} \operatorname{div} (n S_g \rho^{ga} \mathbf{v}^{gs}) = 0 \end{aligned} \quad (2.211)$$

We introduce now in (2.211) the constitutive equations (2.149) for  $\rho^{ga}$  and (2.173) for

$\mathbf{J}_g^{ga}$  to obtain

$$\begin{aligned} & -n \frac{D^s S_w}{Dt} - \beta_s(1-n)S_g \frac{D^s T}{Dt} + S_g \operatorname{div} \mathbf{v}^s + \frac{S_g n}{\rho^{ga}} \frac{D^s}{Dt} \left( \frac{M_a}{\theta R} p^{ga} \right) \\ & - \frac{1}{\rho^{ga}} \operatorname{div} \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \operatorname{grad} \left( \frac{p^{ga}}{p^g} \right) \right] + \frac{1}{\rho^{ga}} \operatorname{div} (n S_g \rho^{ga} \mathbf{v}^{gs}) = 0 \end{aligned} \quad (2.212)$$

The way to derive the mass balance equation for vapour is identical to that of dry air, and for the case of incompressible solid grains it results in

$$\begin{aligned} & -n \frac{D^s S_w}{Dt} - \beta_s(1-n)S_g \frac{D^s T}{Dt} + S_g \operatorname{div} \mathbf{v}^s + \frac{S_g n}{\rho^{gw}} \frac{D^s}{Dt} \left( \frac{M_w}{\theta R} p^{gw} \right) \\ & - \frac{1}{\rho^{gw}} \operatorname{div} \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \operatorname{grad} \left( \frac{p^{gw}}{p^g} \right) \right] + \frac{1}{\rho^{gw}} \operatorname{div} (n S_g \rho^{gw} \mathbf{v}^{gs}) = \frac{\dot{m}}{\rho^{gw}} \end{aligned} \quad (2.213)$$

This equation is now multiplied by  $\rho^{gw}$  and added to the mass balance equation of liquid water (2.202), in turn multiplied by  $\rho^w$ . This sum gives the mass balance equation for the water species, liquid and vapour, without mass rate of water evaporation as

$$\begin{aligned} & n(\rho^w - \rho^{gw}) \frac{D^s S_w}{Dt} - \beta_{swg} \frac{D^s T}{Dt} + (\rho^{gw} S_g + \rho^w S_w) \operatorname{div} \mathbf{v}^s \\ & \frac{n \rho^w S_w}{K_w} \frac{D^s p^w}{Dt} + S_g n \frac{D^s}{Dt} \left( \frac{M_w}{\theta R} p^{gw} \right) - \operatorname{div} \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \operatorname{grad} \left( \frac{p^{gw}}{p^g} \right) \right] \\ & + \operatorname{div}(n S_g \rho^{gw} \mathbf{v}^{gs}) + \operatorname{div}(n S_w \rho^w \mathbf{v}^{ws}) = 0 \end{aligned} \quad (2.214)$$

where

$$\beta_{swg} = \beta_s(1-n)(S_g \rho^{gw} + \rho^w S_w) + n \beta_w \rho^w S_w \quad (2.215)$$

In these equations, Darcy's law for the fluid velocities relative to the solid has still to be introduced. This law was introduced in Section 2.4.7 and will be derived again in its generalised form in the next section from the linear momentum balance equations.

## 2.5.2 Linear momentum balance equation

### 2.5.2.1 Fluids

A more suitable form for the linear momentum balance equation for the fluid phases (2.94) is now obtained by introducing kinematic equations and constitutive relationships. Equations (2.48) and (2.49) allow us to write

$$\mathbf{a}^\pi = \mathbf{a}^s + \mathbf{a}^{\pi s} + \operatorname{grad} \mathbf{v}^\pi \cdot \mathbf{v}^{\pi s} \quad (2.216)$$

where  $\mathbf{a}^{\pi s}$  is the relative acceleration. Introduction in (2.94) of (2.216), (2.148) and (2.164) along with the intrinsic phase-averaged density yields

$$\begin{aligned} -\eta^\pi \rho^\pi [\mathbf{a}^s + \mathbf{a}^{\pi s} + \operatorname{grad} \mathbf{v}^\pi \cdot \mathbf{v}^{\pi s}] - \operatorname{div}(\eta^\pi \rho^\pi \mathbf{I}) \\ + \eta^\pi \rho^\pi \mathbf{g} + \eta^\pi \rho^\pi e^\pi (\rho \dot{\mathbf{r}}) - R^\pi \eta^\pi \mathbf{v}^{\pi s} = 0 \end{aligned} \quad (2.217)$$

By neglecting the term that depends on the gradient of the fluid velocity, by neglecting the effects of phase change, and by applying a vector identity analogous to (2.57), for the divergence of the stress tensor in the fluid phase, we obtain the relative velocity of the fluid as

$$\eta^\pi \mathbf{v}^{\pi s} = (\mathbf{R}^\pi)^{-1} \eta^\pi (-\operatorname{grad} p^\pi + \rho^\pi \mathbf{g} - \rho^\pi \mathbf{a}^s - \rho^\pi \mathbf{a}^{\pi s}) \quad (2.218)$$

By introducing (2.165) and (2.166) the linear momentum balance equation for fluids becomes

$$\eta^\pi \mathbf{v}^{\pi s} = \frac{\mathbf{k} k^{r\pi}}{\mu^\pi} [-\operatorname{grad} p^\pi + \rho^\pi (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{\pi s})] \quad (2.219)$$

Finally, neglecting the soil acceleration and the relative acceleration terms, it yields Darcy's law in the form (2.163):

$$\eta^\pi \mathbf{v}^{\pi s} = \frac{\mathbf{k} k^{r\pi}}{\mu^\pi} [-\operatorname{grad} p^\pi + \rho^\pi \mathbf{g}] \quad (2.220)$$

Due to the simplifications introduced, this law is valid as a first approximation for slow flow of a macroscopically inviscid fluid through a porous medium with incompressible grains.

### 2.5.2.2 Solid phase

Taking into account (2.174), (2.175), (2.176) and (2.164), the linear momentum balance equation (2.97) for the solid phase becomes

$$\begin{aligned} \operatorname{div}[\boldsymbol{\sigma}' - \mathbf{I}(1-n)(S_w p^w + S_g p^g)] + (1-n)\rho^s \mathbf{g} \\ - (1-n)\rho^s \mathbf{a}^s + \mathbf{R}^w \eta^w \mathbf{v}^{ws} + \mathbf{R}^g \eta^g \mathbf{v}^{wg} = 0 \end{aligned} \quad (2.221)$$

### 2.5.2.3 Multiphase medium

Through summing the momentum balance equations (2.217), written for water and gas phases respectively, with that of the solid phase (2.221), by taking into account the definition of total stress (2.178), assuming continuity of stress at the fluid–solid interfaces and by introducing the averaged density of the multiphase medium

$$\rho = (1-n)\rho^s + nS_w\rho^w + nS_g\rho^g \quad (2.222)$$

we obtain the linear momentum balance equation for the whole multiphase medium:

$$\begin{aligned} -\rho \mathbf{a}^s - nS_w \rho^w [\mathbf{a}^{ws} + \text{grad } \mathbf{v}^w \cdot \mathbf{v}^{ws}] \\ - nS_g \rho^g [\mathbf{a}^{gs} + \text{grad } \mathbf{v}^g \cdot \mathbf{v}^{gs}] + \text{div } \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \end{aligned} \quad (2.223)$$

Remember that (2.217) is referred to the moving solid phase and that (2.98) has been taken into account.

### 2.5.3 Energy balance equation

For further development we use intrinsic phase-averaged densities in equation (2.126) and divide this equation by the volume fraction  $\eta^\pi$ :

$$\rho^\pi \frac{D^\pi E^\pi}{Dt} = \frac{\mathbf{t}^\pi}{\eta^\pi} : \mathbf{D}^\pi + \rho^\pi h^\pi - \frac{1}{\eta^\pi} \text{div } \tilde{\mathbf{q}}^\pi + \rho^\pi R^\pi \quad (2.224)$$

Remember that the stress tensors (2.148) and (2.174) contain the volume fractions, hence

$$\frac{\mathbf{t}^\pi}{\eta^\pi} \quad (2.225)$$

represents the ‘intrinsic’ part of the stress tensor.

When dealing with phase change it is more convenient to express the energy balance by means of the specific enthalpy  $H$ , defined as follows [23,39]:

$$H^\pi = E^\pi + \frac{\bar{p}^\pi}{\rho^\pi} = E^\pi + \frac{1}{3\rho_\pi} \text{tr } \mathbf{t}^\pi = E^\pi + \frac{1}{\eta^\pi} \text{tr } \mathbf{t}^\pi \frac{1}{3\rho^\pi} \quad (2.226)$$

where  $\bar{p}$  is the intrinsic hydrostatic part (2.225) of the stress tensor (equal to the pressure for fluids, see (2.148) and Chapter 4) and  $1/\rho^\pi$  is the intrinsic specific volume. The differential of the specific energy  $E$  is therefore

$$dE^\pi = dH^\pi - \bar{p}^\pi d\left(\frac{1}{\rho^\pi}\right) - \frac{1}{\rho^\pi} d\bar{p}^\pi \quad (2.227)$$

and the left-hand side of (2.224) can be rewritten as

$$\rho^\pi \frac{D^\pi E^\pi}{Dt} = \rho^\pi \frac{D^\pi H^\pi}{Dt} + \frac{\bar{p}^\pi}{\rho^\pi} \frac{D^\pi \rho^\pi}{Dt} - \frac{D^\pi \bar{p}^\pi}{Dt} \quad (2.228)$$

The specific enthalphy is a function of the absolute temperature  $\theta$  and the pressure  $\bar{p}$  (one-third of the trace of the intrinsic stress tensor or the mean stress) [23]:

$$H^\pi = H^\pi(\bar{p}^\pi, \theta^\pi) \quad (2.229)$$

The differential of this function

$$dH^\pi = \left( \frac{\partial H^\pi}{\partial \theta^\pi} \right)_p d\theta^\pi + \left( \frac{\partial H^\pi}{\partial \bar{p}^\pi} \right)_0 d\bar{p}^\pi \quad (2.230)$$

is further transformed by use of the definition of specific heat at constant pressure

$$C_p^\pi = \left( \frac{\partial H^\pi}{\partial \theta^\pi} \right)_p \quad (2.231)$$

and the following thermodynamic identity [23]:

$$\left( \frac{\partial H^\pi}{\partial \bar{p}^\pi} \right)_0 = \frac{1}{\rho^\pi} - \theta^\pi \left[ \frac{\partial}{\partial \theta^\pi} \left( \frac{1}{\rho^\pi} \right) \right]_p = \frac{1}{\rho^\pi} \left[ 1 - \frac{\theta^\pi}{\rho^\pi} \left( \frac{\partial \rho^\pi}{\partial \theta^\pi} \right)_p \right] \quad (2.232)$$

The first term on the right-hand side of (2.228) is then

$$\rho^\pi \frac{D^\pi H^\pi}{Dt} = \rho^\pi C_p^\pi \frac{D^\pi \theta^\pi}{Dt} + \frac{D^\pi \bar{p}^\pi}{Dt} - \frac{\theta^\pi}{\rho^\pi} \left( \frac{\partial \rho^\pi}{\partial \theta^\pi} \right)_p \frac{D^\pi \bar{p}^\pi}{Dt} \quad (2.233)$$

and taking into account (2.228) and (2.233), equation (2.224) becomes

$$\begin{aligned} \rho^\pi C_p^\pi \frac{D^\pi \theta^\pi}{Dt} + \frac{\bar{p}^\pi}{\rho^\pi} \frac{D^\pi \rho^\pi}{Dt} - \frac{\theta^\pi}{\rho^\pi} \left( \frac{\partial \rho^\pi}{\partial \theta^\pi} \right)_p \frac{D^\pi \bar{p}^\pi}{Dt} &= \frac{1}{\eta^\pi} \mathbf{t}^\pi : \mathbf{D}^\pi \\ &+ \rho^\pi h^\pi - \frac{1}{\eta^\pi} \operatorname{div} \tilde{\mathbf{q}}^\pi + \rho^\pi R^\pi \end{aligned} \quad (2.234)$$

The material time derivative of the intrinsic density  $\rho^\pi$  is now obtained from the mass balance equation (2.60) written for the generic  $\pi$  phase and for intrinsic phase-averaged densities:

$$\frac{D^\pi}{Dt} (\eta^\pi \rho^\pi) + \rho^\pi \eta^\pi \operatorname{div} \mathbf{v}^\pi = \eta^\pi \rho^\pi e^\pi(\rho) \quad (2.235)$$

Carrying out the derivative and dividing the equation by  $\eta^\pi \rho^\pi$  yields

$$\frac{1}{\rho^\pi} \frac{D^\pi \rho^\pi}{Dt} = e^\pi(\rho) - \operatorname{div} \mathbf{v}^\pi - \frac{1}{\eta^\pi} \frac{D^\pi}{Dt} \eta^\pi \quad (2.236)$$

The stress tensor  $\mathbf{t}_m^\pi$  is now expressed as the sum of the deviatoric part and the hydrostatic part:

$$\frac{\mathbf{t}^\pi}{\eta^\pi} = \frac{\mathbf{r}^\pi}{\eta^\pi} - \bar{p}^\pi \mathbf{I} \quad (2.237)$$

and taking into account (2.50), the first term on the right-hand side of (2.234) is transformed as follows:

$$\frac{\mathbf{t}^\pi}{\eta^\pi} : \mathbf{D}^\pi = \frac{\mathbf{t}^\pi}{\eta^\pi} : \operatorname{grad} \mathbf{v}^\pi = \frac{\tau^\pi}{\eta^\pi} : \operatorname{grad} \mathbf{v}^\pi - \bar{p}^\pi \operatorname{div} \mathbf{v}^\pi \quad (2.238)$$

With (2.236) and (2.238) the enthalpy balance (2.234) may be written as

$$\begin{aligned} \rho^\pi C_p^\pi \frac{D^\pi \theta^\pi}{Dt} - \frac{\theta^\pi}{\rho^\pi} \left( \frac{\partial \rho^\pi}{\partial \theta^\pi} \right)_{p^\pi} \frac{D^\pi \bar{p}^\pi}{Dt} &= \frac{\tau^\pi}{\eta^\pi} : \operatorname{grad} \mathbf{v}^\pi + \rho^\pi h^\pi - \frac{1}{\eta^\pi} \operatorname{div} \tilde{\mathbf{q}}^\pi \\ &+ \rho^\pi R^\pi - \bar{p}^\pi e^\pi(\rho) + \frac{\bar{p}^\pi}{\eta^\pi} \frac{D^\pi \eta^\pi}{Dt} \end{aligned} \quad (2.239)$$

Equation (2.127) is now recalled, which after division by  $\eta^\pi$  reads

$$\rho^\pi R^\pi = \rho^\pi [e^\pi(\rho \hat{E}) - e^\pi(\rho) E^\pi + Q^\pi] \quad (2.240)$$

We collect in  $R_H^\pi$  the following exchange terms:

$$\rho^\pi R_H^\pi = \rho^\pi [e^\pi(\rho \hat{E}) + Q^\pi] \quad (2.241)$$

and  $\bar{p}^\pi e^\pi(\rho)$  of (2.239) is combined with  $e^\pi(\rho) E^\pi$  of (2.240), taking into account the definition of specific enthalpy (2.226) to give the following exchange term:

$$\bar{p}^\pi e^\pi(\rho) + \rho^\pi e^\pi(\rho) E^\pi = \rho^\pi e^\pi(\rho) H^\pi \quad (2.242)$$

This represents the latent heat of evaporation for the fluid phases, it is zero for the solid phase. Hence the macroscopic enthalpy balance equation has the form

$$\begin{aligned} \rho^\pi C_p^\pi \frac{D^\pi \theta^\pi}{D\theta^\pi} - \frac{\theta^\pi}{\rho^\pi} \left( \frac{\partial \rho^\pi}{\partial \theta^\pi} \right)_{p^\pi} \frac{D^\pi \bar{p}^\pi}{Dt} &= \frac{\tilde{\tau}^\pi}{\eta^\pi} : \operatorname{grad} \mathbf{v}^\pi + \rho^\pi h^\pi - \frac{1}{\eta^\pi} \operatorname{div} \tilde{\mathbf{q}}^\pi \\ &+ \rho^\pi R_H^\pi - \rho^\pi e^\pi(\rho) H^\pi + \frac{\bar{p}^\pi}{\eta^\pi} \frac{D^\pi \eta^\pi}{Dt} \end{aligned} \quad (2.243)$$

Terms related to viscous dissipation and mechanical work, caused by density variation due to temperature changes and caused by volume fraction changes, can usually be neglected. Here is the final form of the enthalpy balance equation, after multiplication by  $\eta^\pi$ :

$$\rho_\pi C_p^\pi \frac{D^\pi \theta^\pi}{Dt} = \rho_\pi h^\pi - \operatorname{div} \tilde{\mathbf{q}}^\pi + \rho_\pi R_H^\pi - \rho_\pi e^\pi(\rho) H^\pi \quad (2.244)$$

where  $\rho_\pi C_p^\pi$  is the heat capacity of the  $\pi$  phase.

### 2.5.3.1 Enthalpy balance equation for the multiphase medium

We assume that the phases of a partially saturated porous medium are locally in a state of thermodynamic equilibrium. This means that the averaged temperatures of all phases are assumed equal at each point in the multiphase system:

$$\theta^s = \theta^w = \theta^g = \theta \quad (2.245)$$

Because of  $T = \theta - \theta_0$ , with  $\theta_0$  a fixed reference value, all derivatives of  $\theta$  can be substituted with those of  $T$ . After introducing the constitutive equation (2.191) and the definition of the material time derivative (2.43), the macroscopic energy balance equation (2.126) for the solid phase can be rewritten as

$$\rho_s C_p^s \frac{\partial T}{\partial t} + \rho_s C_p^s \mathbf{v}^s \cdot \nabla T = \text{div}(\chi^s \nabla T) + \rho_s h^s + \rho_s R_H^s \quad (2.246)$$

For liquid water we should also take into account the mass source term related to phase change (2.61):

$$\rho_w C_p^w \frac{\partial T}{\partial t} + \rho_w C_p^w \mathbf{v}^w \cdot \nabla T = \text{div}(\chi^w \nabla T) + \rho_w h^w + \rho_w R_H^w + \dot{m} H^w \quad (2.247)$$

Similarly for a gas phase we have

$$\rho_g C_p^g \frac{\partial T}{\partial t} + \rho_g C_p^g \mathbf{v}^g \cdot \nabla T = \text{div}(\chi^g \nabla T) + \rho_g h^g + \rho_g R_H^g - \dot{m} H^{gw} \quad (2.248)$$

where the last term represents a vapour source term only (no phase changes for dry air).

In geomechanics we usually do not consider heat source terms  $\rho_\pi h^\pi$ , which will be further omitted. The convective heat flux in the solid phase is usually insignificant and will be neglected. Adding equation (2.246), (2.247) and (2.248) and also using (2.128), we obtain

$$(\rho C_p)_{\text{eff}} \frac{\partial T}{\partial t} + (\rho_w C_p^w \mathbf{v}^w + \rho_g C_p^g \mathbf{v}^g) \cdot \nabla T - \text{div}(\chi_{\text{eff}} \nabla T) = -\dot{m} \Delta H_{\text{vap}} \quad (2.249)$$

where

$$\begin{aligned} (\rho C_p)_{\text{eff}} &= \rho_s C_p^s + \rho_w C_p^w + \rho_g C_p^g \\ \chi_{\text{eff}} &= \chi^s + \chi^w + \chi^g \\ \Delta H_{\text{vap}} &= H^{gw} - H^w \end{aligned} \quad (2.250)$$

The mass rate of water evaporation is eliminated from equation (2.249) using one of the mass balance equations for water, (2.60), (2.200) or (2.202).

### 2.5.4 Summary of governing equations

Here is a summary of the most important balance equations used in the remaining chapters of the book.

#### 2.5.4.1 Mass balance equations or continuity equations

*Solid*

$$\frac{(1-n)}{\rho^s} \frac{D^s p^s}{Dt} - \frac{D^s n}{Dt} + (1-n) \operatorname{div} \mathbf{v}^s = 0 \quad (2.251)$$

*Liquid water*

With incompressible solid grains:

$$\begin{aligned} \frac{nS_w}{K_w} \frac{D^s p^w}{Dt} + S_w \operatorname{div} \mathbf{v}^s - \beta_{sw} \frac{D^s T}{Dt} + n \frac{D^s S_w}{Dt} \\ + \frac{1}{\rho^w} \operatorname{div} \left\{ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} [-\operatorname{grad} p^w + \rho^w (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{ws})] \right\} = -\frac{\dot{m}}{\rho^w} \end{aligned} \quad (2.252)$$

With compressible solid grains:

$$\begin{aligned} \left( \frac{\alpha - n}{K_s} S_w^2 + \frac{nS_w}{K_w} \right) \frac{D^s p^w}{Dt} + \frac{\alpha - n}{K_s} S_w S_g \frac{D^s p^g}{Dt} + \alpha S_w \operatorname{div} \mathbf{v}^s \\ - \beta_{sw} \frac{D^s T}{Dt} + \left( \frac{\alpha - n}{K_s} p^w S_w - \frac{\alpha - n}{K_s} p^g S_w + n \right) \frac{D^s S_w}{Dt} \\ + \frac{1}{\rho^w} \operatorname{div} \left\{ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} [-\operatorname{grad} p^w + \rho^w (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{ws})] \right\} = -\frac{\dot{m}}{\rho^w} \end{aligned} \quad (2.253)$$

where

$$\beta_{sw} = S_w [(\alpha - n) \beta_s + n \beta_w] \quad (2.254)$$

*Gas*

With incompressible solid grains:

$$\begin{aligned} -n \frac{D^s S_w}{Dt} - \beta_s (1-n) S_g \frac{D^s T}{Dt} + S_g \operatorname{div} \mathbf{v}^s + \frac{nS_g}{\rho^g} \frac{D^s}{Dt} \left[ \frac{1}{\theta R} (p^{ga} M_a + p^{gw} M_w) \right] \\ + \frac{1}{\rho^g} \operatorname{div} \left\{ \rho^g \frac{\mathbf{k} k^{rg}}{\mu^g} [-\operatorname{grad} p^g + \rho^g (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{gs})] \right\} = \frac{\dot{m}}{\rho^g} \end{aligned} \quad (2.255)$$

With compressible solid grains:

$$\begin{aligned} & \frac{\alpha - n}{K_s} S_w S_g \frac{D^s p^w}{Dt} + \frac{\alpha - n}{K_s} S_g^2 \frac{D^s p^g}{Dt} - \left( n + \frac{\alpha - n}{K_s} p^c S_g \right) \frac{D^s S_w}{Dt} \\ & - \beta_s (\alpha - n) S_g \frac{D^s T}{Dt} + \alpha S_g \operatorname{div} \mathbf{v}^s + \frac{n S_g}{\rho^g} \frac{D^s}{Dt} \left[ \frac{1}{\theta R} (p^{ga} M_a + p^{gw} M_w) \right] \\ & + \frac{1}{\rho^g} \operatorname{div} \left\{ \rho^g \frac{\mathbf{k} k^{rg}}{\mu^g} [-\operatorname{grad} p^g + \rho^g (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{gs})] \right\} = \frac{\dot{m}}{\rho^g} \end{aligned} \quad (2.256)$$

When the source term in these two equations is expressed through the vapour phase mass balance equation (2.63), they become the mass balance equations for dry air.

### Water species (liquid water and vapour)

With incompressible solid grains:

$$\begin{aligned} & n(\rho^w - \rho^{gw}) \frac{D^s S_w}{Dt} - \beta_{swg} \frac{D^s T}{Dt} + (\rho^{gw} S_g + \rho^w S_w) \operatorname{div} \mathbf{v}^s \\ & + \frac{n \rho^w S_w}{K_w} \frac{D^s p^w}{Dt} + S_g n \frac{D^s}{Dt} \left( \frac{M_w}{\theta R} p^{gw} \right) - \operatorname{div} \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \operatorname{grad} \left( \frac{p^{gw}}{\rho^g} \right) \right] \\ & + \operatorname{div} \left\{ \rho^{gw} \frac{\mathbf{k} k^{rg}}{\mu^g} [-\operatorname{grad} p^g + \rho^g (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{gs})] \right\} \\ & + \operatorname{div} \left\{ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} [-\operatorname{grad} p^w + \rho^w (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{ws})] \right\} = 0 \end{aligned} \quad (2.257)$$

where

$$\beta_{swg} = \beta_s (1 - n) (S_g \rho^{gw} + \rho^w S_w) + n \beta_w \rho^w S_w \quad (2.258)$$

With compressible solid grains:

$$\begin{aligned} & \left[ \frac{\alpha - n}{K_s} S_w (\rho^{gw} S_g + \rho^w S_w) + \rho^w \frac{n S_w}{K_w} \right] \frac{D^s p^w}{Dt} + \frac{\alpha - n}{K_s} S_g (\rho^{gw} S_g + \rho^w S_w) \frac{D^s p^g}{Dt} \\ & + \left[ \frac{\alpha - n}{K_s} (\rho^{gw} S_g p^c + \rho^w S_w p^w - \rho^w S_w p^c) + n(\rho^w - \rho^{gw}) \right] \frac{D^s S_w}{Dt} \\ & + (\rho^{gw} S_g + \rho^w S_w) \alpha \operatorname{div} \mathbf{v}^s \\ & - \beta_{swg} \frac{D^s T}{Dt} + S_g n \frac{D^s}{Dt} \left( \frac{M_w}{\theta R} p^{gw} \right) - \operatorname{div} \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \operatorname{grad} \left( \frac{p^{gw}}{\rho^g} \right) \right] \\ & + \operatorname{div} \left\{ \rho^{gw} \frac{\mathbf{k} k^{rg}}{\mu^g} [-\operatorname{grad} p^g + \rho^g (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{gs})] \right\} \\ & + \operatorname{div} \left\{ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} [-\operatorname{grad} p^w + \rho^w (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{ws})] \right\} = 0 \end{aligned} \quad (2.259)$$

where (2.258) becomes

$$\beta_{swg} = \beta_s(\alpha - n)(S_g\rho^{gw} + \rho^w S_w) + n\beta_w\rho^w S_w \quad (2.260)$$

#### 2.5.4.2 Linear momentum balance equations

##### Fluids

$$\eta^\pi \mathbf{v}^{\pi s} = \frac{\mathbf{k} k^{\pi s}}{\mu^\pi} [-\operatorname{grad} p^\pi + \rho^\pi (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{\pi s})] \quad (2.261)$$

For dry air and vapour the body force term is usually neglected.

##### Multiphase medium

$$\begin{aligned} & -\rho \mathbf{a}^s - n S_w \rho^w [\mathbf{a}^{ws} + \operatorname{grad} \mathbf{v}^w \cdot \mathbf{v}^{ws}] \\ & - n S_g \rho^g [\mathbf{a}^{gs} + \operatorname{grad} \mathbf{v}^g \cdot \mathbf{v}^{gs}] + \operatorname{div} \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \end{aligned} \quad (2.262)$$

#### 2.5.4.3 Enthalpy balance: multiphase medium

$$(\rho C_p)_{\text{eff}} \frac{\partial T}{\partial t} + (\rho_w C_p^w \mathbf{v}^w + \rho_g C_p^g \mathbf{v}^g) \cdot \operatorname{grad} T - \operatorname{div} (\chi_{\text{eff}} \operatorname{grad} T) = -\dot{m} \Delta H_{\text{vap}} \quad (2.263)$$

where  $\dot{m}$  is obtained either from (2.252) or (2.253).

## 2.6 PHYSICAL APPROACH: EXTENDED BIOT THEORY

The governing equations, using Biot's theory [40–45], are again derived but extended to the case of non-isothermal two-phase flow in deforming porous media. This extension was made [35,46] for the case of slow phenomena. Also, inertia forces are taken into account, as was done by Zienkiewicz *et al.* [47] for the isothermal case with the air phase at atmospheric pressure.

The physical approach works directly with macroscopic variables. However, the distinction between a macroscopic and microscopic domain is not so clear as in the previous sections, because macroscopic variables are sometimes directly used in the microscopic domain.

For the sake of simplicity, small displacements are assumed for the solid phase. The governing equations in the form needed for finite strain analysis are those of the previous section (see also Chapter 14). For all queries regarding the assumed hypotheses the interested reader is referred to the first part of this chapter.

### 2.6.1 The physical model

The voids of the skeleton are filled partly with water and partly with moist air (a mixture of dry air and water vapour), which is referred to as gas. The degree of water saturation  $S_w$  is given as the ratio between the pore space occupied by the water and the total pore volume in a representative elementary volume element (REV); see Figure 2.1. We therefore have

$$S_w = \frac{dv^w}{dv^w + dv^g} \quad (2.264)$$

In the following equations, the super/subscript  $\pi$  refers to the generic phase,  $\pi = s$  to the solid phase,  $\pi = w$  to water and  $\pi = g$  to gas. The volume of the REV is  $dv$ , and the partial volumes are  $dv^s$ ,  $dv^w$ ,  $dv^g$ ; their sum is  $dv$ . The sum of the respective fluid volumes is  $dv^f = dv^w + dv^g$ . The definition of the degree of gas saturation  $S_g$  is

$$S_g = \frac{dv^g}{dv^w + dv^g} \quad (2.265)$$

and both degrees of saturation sum to one, i.e.  $S_w + S_g = 1$ .

The porosity  $n$  is defined as the ratio between the volume of voids and the total volume of the REV:

$$n = \frac{dv^w + dv^g}{dv} \quad (2.266)$$

Stress is defined as tension positive for the solid phase; pore pressure is defined as compressive positive for fluids. The water pressure  $p^w$  and the gas pressure  $p^g$  are related through the capillary pressure  $p^c$ :

$$p^g - p^w = p^c \quad (2.267)$$

The capillary pressure is a function of the water saturation and temperature:

$$p^c = p^c(S_w, T) \quad (2.268)$$

Relation (2.268) is determined experimentally and usually shows hysteresis characteristics, which are ignored in this case. It can be numerically inverted to obtain

$$S_w = S_w(p^c, T) \quad (2.269)$$

The constitutive law of the solid phase is introduced through the concept of effective stress:

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + \mathbf{I}p^s \quad (2.270)$$

which stipulates that the main characteristics of the solid-phase constitutive relation can be written in terms of  $\boldsymbol{\sigma}'$  where  $\boldsymbol{\sigma}$  is the total stress tensor,  $\mathbf{I}$  the second-order unit tensor

and  $p^s$  is the average pressure of both the water and air surrounding the grains. In the case of immiscible two-phase flow, we need a simple averaging technique for the calculation of  $p^s$ , as described in Section 2.2. Thus the modified effective stress principle may be obtained as stated previously. The effective stress resulting from (2.270) is also called the Bishop stress.

$$\begin{aligned}\bar{\sigma} &= \frac{1}{dv} \int_{dv} \sigma dv_m = \frac{1}{dv} \left[ \int_{dv^s} \sigma dv_m + \int_{dv^f} \sigma dv_m \right] \\ &= \frac{dv^s}{dv} \sigma^s + \frac{dv^f}{dv} \left[ \frac{dv^w}{dv^f} \sigma^w + \frac{dv^g}{dv^f} \sigma^g \right] \\ &= (1-n)\sigma^s + n[S_w\sigma^w + S_g\sigma^g]\end{aligned}\quad (2.271)$$

where  $\bar{\sigma}$  is the macroscopic total stress tensor and  $\sigma^\pi$  is the intrinsic phase-averaged stress tensor in the  $\pi$  phase.

For the fluid phases, the stress tensor is given by

$$\sigma^\pi = \tau^\pi - \mathbf{I}p^\pi \quad (2.272)$$

where  $\tau^\pi$  is the shear stress. Under the assumption that the shear stress  $\tau^\pi$  is negligible in fluids, we obtain

$$\bar{\sigma} = (1-n)\sigma^s - n\mathbf{I}[S_g p^g + S_w p^w] \quad (2.273)$$

The term in square brackets represents the intrinsically averaged (or mean) pressure  $p^s$  of the fluid phases, i.e.

$$p^s = (S_g p^g + S_w p^w) \quad (2.274)$$

This weighted pore pressure produces a stress state in the grains, but for the moment we assume the grains will not undergo any deformation due to this stress, i.e. we introduce the hypothesis of incompressible grains. The deformation of the solid skeleton, which depends on the effective stress, will be a function of the grain rearrangement only.

Equation (2.273) can be modified to

$$\begin{aligned}\sigma &= (1-n)\sigma^s - n\mathbf{I}p^s \\ &= (1-n)(\sigma^s + \mathbf{I}p^s) - (1-n)\mathbf{I}p^s - n\mathbf{I}p^s \\ &= \sigma' - \mathbf{I}p^s\end{aligned}\quad (2.275)$$

where the overbar for the total stress has been omitted.

The stress tensor is split into two components: the pore pressure effect and the part which deforms the solid skeleton, i.e. the effective stress. The effective stress is given by

$$\sigma' = (1-n)(\sigma^s + \mathbf{I}p^s) \quad (2.276)$$

Equation (2.275) implies a splitting of the stress tensor similar to that of Terzaghi's

principle and in the presence of several fluid phases, the splitting is

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + \mathbf{I}(S_w p^w + S_g p^g) \quad (2.277)$$

For greater generality, a corrective term known as Biot's constant has to be introduced to account for the deformability of the grains [32]. Therefore a more general expression of the effective stress is assumed as follows:

$$\boldsymbol{\sigma}'' = \boldsymbol{\sigma} + \mathbf{I}\alpha(S_w p^w + S_g p^g) \quad (2.278)$$

where the corrective coefficient  $\alpha$  will be determined in Section 2.6.2 Note how this equation differs substantially from the previous one; for the determination of  $\alpha$  we need the constitutive equations of the solid phase. This effective stress is therefore indicated by  $\boldsymbol{\sigma}''$ , as done by Zienkiewicz *et al.* [47].

A more intuitive way of deriving equation (2.276) follows Bishop [48] and Skempton [49]. Here the microscopic and macroscopic aspects are somewhat mixed. We consider for this purpose the mean stresses, i.e.

$$\hat{\boldsymbol{\sigma}} = \text{tr} \frac{\boldsymbol{\sigma}}{3} \quad \text{and} \quad \hat{\boldsymbol{\sigma}}' = \text{tr} \frac{\boldsymbol{\sigma}'}{3} \quad (2.279)$$

If the pores (Figure 2.4) are filled with water and air, then the surface tension effect means that

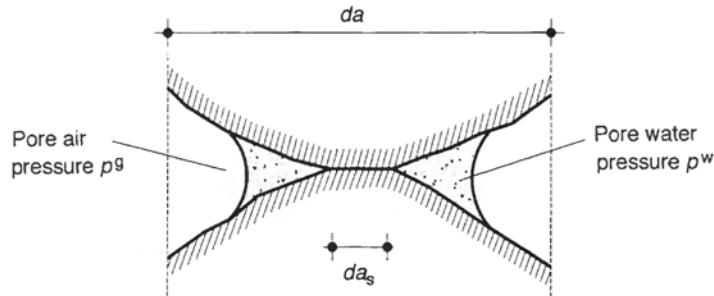
$$p^w < p^g \quad (2.280)$$

If the degree of saturation is relatively low, the water is present as menisci and the corresponding pressure acts over an area  $\chi$  per unit gross area [50]. Consequently, the equivalent pore pressure is given by

$$\chi p^w + (1 - \chi)p^g \quad (2.281)$$

and may also be written as

$$p^g - \chi(p^g - p^w) \quad \text{or} \quad p^w + (1 - \chi)(p^g - p^w) \quad (2.282)$$



**Figure 2.4** Contact area between two grains in the partially saturated case

In analogy with the fully saturated case, Bishop suggested the following expression for the mean effective stress:

$$\hat{\sigma}' = \hat{\sigma} + [p^g - \chi(p^g - p^w)] \quad (2.283)$$

For the full stress tensor, this equation assumes the form

$$\sigma' = \sigma + I[p^g - \chi(p^g - p^w)] \quad (2.284)$$

The coefficient  $\chi$  is not the same as for problems involving shear strength and consolidation [49]. For a given degree of saturation the coefficient  $\chi$  must be determined experimentally for both types of problems.

In general, the pressure is implied as being an absolute value, but many soil mechanics problems use relative pressures, and  $p^g$  is often assumed to be equal to the atmospheric pressure. If this is the case and the atmospheric pressure is assumed to be the reference pressure, then the effective stress principle becomes

$$\sigma' = \sigma + I\chi p^w \quad (2.285)$$

The comparison between (2.277) and (2.284) is noteworthy. Equation (2.277) may be written as a function of the degree of saturation, namely

$$\sigma = \sigma' - I[p^g - S_w(p^g - p^w)] \quad (2.286)$$

and coincides with (2.284) if  $\chi = S_w$ .

The coefficient  $\chi$  is related to the area of contact between solid and fluids, whereas the degrees of saturation  $S_\pi$  depend on the volume occupied by the  $\pi$  phase. We define the volume fraction as

$$\eta^\pi = \frac{dv^\pi}{dv} \quad (2.287)$$

and the cross-sectional area fraction as

$$\alpha^\pi = \frac{da^\pi}{da} \quad (2.288)$$

For the case of microstructurally isotropic constituents, such that  $\alpha^\pi$  is independent of the orientation of the surface, and assuming that

$$\alpha^\pi = \alpha^\pi(\eta^\pi) \quad (2.289)$$

Morland [51] has shown that

$$\alpha^\pi = \eta^\pi \quad (2.290)$$

This conclusion coincides with the basic supposition of Delesse's law (Section 2.2.1). If this assumption regarding the area and volume fraction is valid, then little difference

exists between the definition of Bishop's coefficient  $\chi$  and the degree of water saturation. Bishop's equation (2.283) has been experimentally validated by means of a triaxial tests carried out on soil samples [49].

The assumption

$$\chi = S_w \quad (2.291)$$

is acceptable for many materials, as shown experimentally by Bishop and Blight [30]. However, there are other expressions, e.g. this one used by Gudehus for clay [52]:

$$\chi = S(2 - S) \quad (2.292)$$

If the solid phase is completely surrounded by a single wetting fluid phase, which is the only one in direct contact with the porous medium, then Li and Zienkiewicz [53] use (2.285) with  $\chi = 1$ , as in the case of a fully saturated porous medium. This assumption is widely used when the two fluids are water and bitumen; only water is in contact with the solid grains.

A new expression for the effective stress principle in fully saturated conditions only, involving four parameters, has been proposed by Lade and de Boer [54] and experimentally tested. This formulation, based on principles of mechanics, is valid for all types of material. A distinction is made between the compressibilities of the grain particles and the skeleton due to total stresses and pore pressures.

Two distinct expressions are then obtained for effective stresses in granular material and in solid rock with interconnected pores. For various special conditions these expressions become similar to the expressions found in the literature, particularly the expression used here with Biot's constant  $\alpha$ . Lade and de Boer [54] conclude that Terzaghi's proposed effective stress principle works well for stress magnitudes encountered in most geotechnical applications, but significant deviations occur at very high stresses.

## 2.6.2 Constitutive equations

For a full account of the constitutive relationships, the reader is referred to Section 2.4; we recap only a few equations needed for the remaining part of this section. In particular, an expression for Biot's constant  $\alpha$  is derived, a derivation omitted from Section 2.4.

The averaged pore pressure  $p^s$  of the fluids occupying the void space induces a hydrostatic stress distribution in the solid phase. The ensuing deformation is a purely volumetric strain

$$\boldsymbol{\varepsilon}_v^s = -\frac{p^s}{K_s} \quad (2.293)$$

or in incremental and tensorial form

$$d\boldsymbol{\varepsilon}_v^s = -\mathbf{I} \frac{dp^s}{3K_s} \quad (2.294)$$

where  $K_s$  is the averaged bulk modulus of the solid grains.

The effective stress causes all relevant deformation of the solid skeleton. The constitutive relationship may be written as

$$d\sigma' = D_T[(d\epsilon - d\epsilon^c - d\epsilon_v^s - d\epsilon^0)] \quad (2.295)$$

where  $d\sigma'$  = the stress responsible for all deformations, except for the grain compressibility

$d\epsilon$  = the total strain of the solid

$D_T$  = the tangent constitutive tensor

$$D_T = D_T(\sigma', \epsilon, \dot{\epsilon}) \quad (2.296)$$

$d\epsilon^c = g(\sigma')dt$  accounts for the creep strain

$d\epsilon^0$  = all other strains in the solid skeleton not directly dependent on effective stress

A modified effective stress  $\sigma''$  will now be introduced which will also account for grain compression. This allows the use of Biot's constant  $\alpha$ . By substituting (2.294) into (2.295) and omitting  $d\epsilon^0$  and  $d\epsilon^c$ , for brevity, we obtain

$$\begin{aligned} d\sigma' &= D_T[(d\epsilon - d\epsilon_v^s)] \\ &= D_T[d\epsilon] + D_T[\mathbf{I}] \frac{dp^s}{3K_s} \end{aligned} \quad (2.297)$$

Consequently, the effective stress equation (2.277) and equation (2.274), written in incremental form, become

$$d\sigma = d\sigma' - \mathbf{I}dp^s = d\sigma'' + D_T[\mathbf{I}] \frac{dp^s}{3K_s} - \mathbf{I}dp^s \quad (2.298)$$

where  $\sigma''$  represents the stress responsible for all deformation of the solid.

For the following developments, we use indicial notation and the Kronecker symbol  $\delta_{rs}$  instead of the unit tensor  $\mathbf{I}$ :

$$\begin{aligned} \delta_{rs} &= 1 \quad \text{if } r = s \\ \delta_{rs} &= 0 \quad \text{if } r \neq s \end{aligned} \quad (2.299)$$

It can be immediately verified that (2.298) may be written as given by Zienkiewicz and Shiomi [55]:

$$d\sigma_{ij} = d\sigma''_{ij} - \frac{1}{3} \left( \delta_{mn}\delta_{nm} - \delta_{mn}D_{mnkl}\delta_{kl} \frac{1}{3K_s} \right) \delta_{ij} dp^s \quad (2.300)$$

For an isotropic, elastic material

$$\delta_{mn}D_{mnkl}\delta_{kl} = \frac{3E}{1-2v} = 9K_T \quad (2.301)$$

where  $E$  is Young's modulus,  $v$  is Poisson's ratio and  $K_T$  is the bulk modulus of the

overall skeleton. Hence (2.300) can be written as

$$\begin{aligned} d\sigma_{ij}'' &= d\sigma_{ij}'' - \left(1 - \frac{K_T}{K_s}\right)\delta_{ij}dp^s \\ &= d\sigma_{ij}'' - \alpha\delta_{ij}dp^s \end{aligned} \quad (2.302)$$

where  $\alpha$  is Biot's constant.

The resultant expression of the constitutive relationship is therefore

$$d\boldsymbol{\sigma}'' = \mathbf{D}_T[d\boldsymbol{\varepsilon}] \quad (2.303)$$

The intrinsic pore pressure  $p^\pi$  of the  $\pi$  fluid phase causes a purely volumetric strain of that phase, which is represented by

$$\boldsymbol{\varepsilon}_v^\pi = -\frac{p^\pi}{K_\pi} \quad (2.304)$$

For all other necessary relations see Section 2.4.

### 2.6.3 Governing equations

In the following, a material coordinate system is employed for the solid phase and a spatial coordinate system for the fluid phases; unless specified otherwise, convective terms are neglected. Because of this choice of reference systems, the fluid velocities are conveniently referred to the solid phase by means of the relative velocities for water and gas:

$$\mathbf{v}^{ws} = \mathbf{v}^w - \mathbf{v}^s \quad (2.305)$$

$$\mathbf{v}^{gs} = \mathbf{v}^g - \mathbf{v}^s \quad (2.306)$$

and the accelerations, without convective terms, are

$$\mathbf{a}^w = \mathbf{a}^s + \mathbf{a}^{ws} \quad (2.307)$$

$$\mathbf{a}^g = \mathbf{a}^s + \mathbf{a}^{gs} \quad (2.308)$$

where  $\mathbf{a}^{ws}$  is the acceleration of water relative to the solid phase and  $\mathbf{a}^{gs}$  is the relative acceleration of gas. For the full expression see Section 2.3.1.

#### 2.6.3.1 Linear momentum balance equation: multiphase medium

The linear momentum balance equation for the multiphase system is the sum of the dynamics equations for the individual constituents, but relative to the solid phase via (2.307) and (2.308):

$$-\rho\mathbf{a} - nS_w\rho^w\mathbf{a}^{ws} - nS_g\rho^g\mathbf{a}^{gs} + \operatorname{div}\boldsymbol{\sigma} + \rho\mathbf{g} = 0 \quad (2.309)$$

where

$$\rho = (1 - n)\rho^s + nS_w\rho^w + nS_g\rho^g \quad (2.310)$$

is the averaged density of the multiphase systems;  $\rho^s$  is the intrinsic density of the solid phase, i.e. the density referred to the volume occupied by the solid phase only, as opposed to the volume occupied by the multiphase system;  $\rho^w$  is the intrinsic density of water;  $\rho^g$  is the intrinsic density of gas; and  $\mathbf{g}$  is an acceleration usually related to gravitational effects.

The linear momentum balance equation for each fluid phase yields the generalised form of Darcy's law, where the dissipative terms arising in a multiphase flow system at the interfaces are taken into account through the relative permeabilities:

$$nS_w\mathbf{v}^{ws} = \frac{\mathbf{k}k^{rw}}{\mu_w}[-\text{grad } p^w + \rho^w(\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{ws})] \quad (2.311)$$

$$nS_g\mathbf{v}^{gs} = \frac{\mathbf{k}k^{rg}}{\mu_g}[-\text{grad } p^g + \rho^g(\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{gs})] \quad (2.312)$$

where  $\mathbf{k}$  is the intrinsic permeability,  $k^{rw}$  and  $k^{rg}$  are the relative permeabilities of water and gas. The relative permeabilities vary between 0 and 1 and are a function of the degree of saturation (Section 2.4.7);  $\mu_w$  and  $\mu_g$  are the dynamic viscosities, which are temperature dependent, i.e.  $\mu_\pi = \mu_\pi(T)$ ,  $\pi = w, g$ . In the generalised form of Darcy's law for the gas phase, the body forces are usually neglected. The left-hand side of (2.311) and (2.312) is multiplied by the volume fractions  $nS_\pi$  because Darcy's law is expressed in terms of volume-averaged relative velocities. These are the velocities measured from experimental work.

### 2.6.3.2 Mass balance equations

The mass balance equation for the solid may be written as

$$\frac{\partial(1 - n)\rho^s}{\partial t} + \text{div}[(1 - n)\rho^s\mathbf{v}^s] = 0 \quad (2.313)$$

whereas for the fluid phases we have

$$\frac{\partial(nS_\pi\rho^\pi)}{\partial t} + \text{div}(nS_\pi\rho^\pi\mathbf{v}^\pi) = \pm \dot{m} \quad (2.314)$$

where  $\dot{m}$  is the mass exchange term between the two phases, i.e. the mass rate of evaporation. This is usually negative for water and positive for gas.

By carrying out the time derivatives, neglecting the gradient of  $(1 - n)\rho^s$  as per (2.57), and dividing by  $\rho^s$ , equation (2.313) becomes

$$\frac{1 - n}{\rho^s} \frac{\partial \rho^s}{\partial t} - \frac{\partial n}{\partial t} + (1 - n)\text{div } \mathbf{v}^s = 0 \quad (2.315)$$

We introduce into (2.314) the relative velocities (2.305) and (2.306), carry out the time derivatives and divide the equation by  $S_\pi \rho^\pi$ . After applying the vector identity (2.57) to  $\operatorname{div}(nS_\pi \rho^\pi \mathbf{v}^s)$ , we neglect  $\operatorname{grad}(nS_\pi \rho^\pi)$ . This allows (2.314) to be written as

$$\frac{\partial n}{\partial t} + \frac{n}{\rho^\pi} \frac{\partial \rho^\pi}{\partial t} + \frac{n}{S_\pi} \frac{\partial S_\pi}{\partial t} + \frac{1}{S_\pi \rho^\pi} \operatorname{div}(nS_\pi \rho^\pi \mathbf{v}^{\pi s}) + n \operatorname{div} \mathbf{v}^s = \pm \frac{\dot{m}}{\rho^\pi S_\pi} \quad (2.316)$$

The summation of (2.315) and (2.316) eliminates the time derivative of the porosity and yields the so-called continuity equation of the fluid phases, i.e.

$$\frac{(1-n)}{\rho^s} \frac{\partial \rho^s}{\partial t} + \operatorname{div} \mathbf{v}^s + \frac{n}{\rho^\pi} \frac{\partial \rho^\pi}{\partial t} + \frac{n}{S_\pi} \frac{\partial S_\pi}{\partial t} + \frac{1}{S_\pi \rho^\pi} (nS_\pi \rho^\pi \mathbf{v}^{\pi s}) = \pm \frac{\dot{m}}{\rho^\pi S_\pi} \quad (2.317)$$

This equation is the same as (2.197) with the exception that we have here partial derivatives instead of material time derivatives. For small displacements the two derivatives coincide.

We consider first the continuity equation for water. Introduction of the constitutive relationships for the time derivatives of water density (2.162) and solid density (2.190) along with the pressure in the solid phase (2.274), and multiplication by  $S_w$  allows the continuity equation for water to be written as

$$\begin{aligned} & \left( \frac{\alpha - n}{K_s} S_w^2 + \frac{nS_w}{K_w} \right) \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_w S_g \frac{\partial p^g}{\partial t} + \alpha S_w \operatorname{div} \mathbf{v}^s \\ & - \beta_{sw} \frac{\partial T}{\partial t} + \left( \frac{\alpha - n}{K_s} S_w p^w - \frac{\alpha - n}{K_s} p^g S_w + n \right) \frac{\partial S_w}{\partial t} \\ & + \frac{1}{\rho^w} \operatorname{div}(nS_w \rho^w \mathbf{v}^{ws}) = - \frac{\dot{m}}{\rho^w} \end{aligned} \quad (2.318)$$

where

$$\beta_{sw} = S_w [(\alpha - n)\beta_s + n\beta_w] \quad (2.319)$$

Also  $T$  is the temperature above a reference value and  $\beta_\pi$  is the thermal expansion coefficient. And the relationship

$$\frac{\partial S_g}{\partial t} = - \frac{\partial S_w}{\partial t} \quad (2.320)$$

has been taken into account.

Introduction of Darcy's law (2.311) and the definition of capillary pressure (2.267) yields

$$\begin{aligned} & \left( \frac{\alpha - n}{K_s} S_w^2 + \frac{nS_w}{K_w} \right) \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_w S_g \frac{\partial p^g}{\partial t} + \alpha S_w \operatorname{div} \mathbf{v}^s \\ & - \beta_{sw} \frac{\partial T}{\partial t} + \left( \frac{n - \alpha}{K_s} S_w p^c + n \right) \frac{\partial S_w}{\partial t} \\ & + \frac{1}{\rho^w} \operatorname{div} \left\{ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} [-\operatorname{grad} p^w + \rho^w (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{ws})] \right\} = - \frac{\dot{m}}{\rho^w} \end{aligned} \quad (2.321)$$

Recall that (2.269) implies

$$\frac{\partial S_w}{\partial t} = \frac{\partial S_w}{\partial p^c} \frac{\partial p^c}{\partial t} + \frac{\partial S_w}{\partial T} \frac{\partial T}{\partial t} \quad (2.322)$$

The continuity equation for gas is obtained in the same way. For the time derivative of the gas density, we use (2.149) and (2.150), and we obtain the time derivative of the solid density as above. Furthermore, we multiply (2.317) by  $S_g$  and we take into account the definition of capillary pressure:

$$\begin{aligned} & \frac{\alpha - n}{K_s} S_w S_g \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_g^2 \frac{\partial p^g}{\partial t} - \left( n + \frac{\alpha - n}{K_s} p^c S_g \right) \frac{\partial S_w}{\partial t} \\ & - \beta_s (\alpha - n) S_g \frac{\partial T}{\partial t} + \alpha S_g \operatorname{div} \mathbf{v}^s + \frac{n S_g}{\rho^g} \frac{\partial}{\partial t} \left( \frac{p^g M_g}{\theta R} \right) \\ & + \frac{1}{\rho^g} \operatorname{div}(n S_g \rho^g \mathbf{v}^{gs}) = - \frac{\dot{m}}{\rho^g} \end{aligned} \quad (2.323)$$

where  $M_g$  is the molar mass of gas, given by the third equation in (2.150) and  $\theta$  is the absolute temperature. Again, the introduction of Darcy's law for gas (2.312) yields

$$\begin{aligned} & \frac{\alpha - n}{K_s} S_w S_g \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_g^2 \frac{\partial p^g}{\partial t} - \left( n + \frac{\alpha - n}{K_s} p^c S_g \right) \frac{\partial S_w}{\partial t} \\ & - \beta_s (\alpha - n) S_g \frac{\partial T}{\partial t} + \alpha S_g \operatorname{div} \mathbf{v}^s + \frac{n S_g}{\rho^g} \frac{\partial}{\partial t} \left( \frac{p^g M_g}{\theta R} \right) \\ & + \frac{1}{\rho^g} \operatorname{div} \left\{ \rho^g \frac{\mathbf{k} k^{rg}}{\mu^g} [-\operatorname{grad} p^g + \rho^g (\mathbf{g} - \mathbf{a}^s - \mathbf{a}^{gs})] \right\} = \frac{\dot{m}}{\rho^g} \end{aligned} \quad (2.324)$$

### 2.6.3.3 Energy balance equation

By subtracting the kinetic energy from a global energy balance, the following balance equation of thermal energy may be written for constituent  $\pi$  [56]:

$$\frac{\partial}{\partial t} (\rho^\pi E^\pi) = -\operatorname{div}(\rho^\pi E^\pi \mathbf{v}^\pi) - \operatorname{div} \tilde{\mathbf{q}}^\pi - \bar{p}^\pi \operatorname{div} \mathbf{v}^\pi + \boldsymbol{\tau}^\pi : \operatorname{grad} \mathbf{v}^\pi + \rho^\pi R^\pi \quad (2.325)$$

where  $E^\pi$  is the specific internal energy,  $\boldsymbol{\tau}$  is the deviatoric part of the stress tensor and  $\bar{p}^\pi$  the hydrostatic part (or mean stress; see Chapter 4). The left-hand side represents the rate of accumulation of internal energy in a control volume. The right-hand terms express respectively the rate of internal energy change due to convection, the rate of internal energy change due to conduction, the reversible rate of internal energy increase due to pressure, the irreversible rate of internal energy increase by viscous dissipation, and the contribution of heat sources.

We make use of the following vector identity:

$$\operatorname{div}(\rho^\pi E^\pi \mathbf{v}^\pi) = \rho^\pi E^\pi \operatorname{div} \mathbf{v}^\pi + \operatorname{grad}(\hat{c}^\pi E^\pi) \cdot \mathbf{v}^\pi \quad (2.326)$$

and write (2.325) in the following form:

$$\begin{aligned} \frac{\hat{c}}{\hat{c}_t} (\rho^\pi E^\pi) + \mathbf{v}^\pi \cdot \operatorname{grad}(\rho^\pi E^\pi) + \rho^\pi E^\pi \operatorname{div} \mathbf{v}^\pi &= -\operatorname{div} \hat{\mathbf{q}}^\pi \\ -\bar{p}^\pi \operatorname{div} \mathbf{v} + \tau^\pi : \operatorname{grad} \mathbf{v}^\pi + \rho^\pi R^\pi \end{aligned} \quad (2.327)$$

Taking into account the continuity equation in the following form (see vector identity (2.326) and equation (2.56) under the small-strain assumption):

$$\frac{\partial}{\partial t} \rho^\pi + \mathbf{v}^\pi \cdot \operatorname{grad} \rho^\pi + \rho^\pi \operatorname{div} \mathbf{v}^\pi = \frac{\hat{c}}{\hat{c}_t} \rho^\pi + \operatorname{div}(\rho^\pi \mathbf{v}^\pi) = 0 \quad (2.328)$$

the energy balance equation becomes

$$\rho^\pi \frac{\partial E^\pi}{\partial t} + \rho^\pi \mathbf{v}^\pi \cdot \operatorname{grad} E^\pi = -\operatorname{div} \hat{\mathbf{q}}^\pi - \bar{p}^\pi \operatorname{div} \mathbf{v}^\pi + \tau^\pi : \operatorname{grad} \mathbf{v}^\pi + \rho^\pi R^\pi \quad (2.329)$$

The irreversible part of the internal energy increase by viscous dissipation is neglected in what follows. It is convenient to express (2.327) in terms of temperature and heat capacity instead of internal energy [56]. This is done through the concept of enthalpy (Section 2.5.3). Hence, with sufficient accuracy for geomechanical applications, the energy balance equation can be rewritten as

$$\rho^\pi C_p^\pi \left( \frac{\partial T^\pi}{\partial t} + \mathbf{v}^\pi \cdot \operatorname{grad} T^\pi \right) = -\operatorname{div} \hat{\mathbf{q}}^\pi + \rho^\pi R^\pi \quad (2.330)$$

for  $\pi = s, w$ , which for the purpose of the energy balance only may be considered as incompressible. and

$$\rho^\pi C_p^g \left( \frac{\partial T^g}{\partial t} + \mathbf{v}^g \cdot \operatorname{grad} T^g \right) = -\operatorname{div} \hat{\mathbf{q}}^g + \frac{\partial p^g}{\partial t} + \mathbf{v}^g \cdot \operatorname{grad} p^g + \rho^g R^g \quad (2.331)$$

where  $C_p^\pi$  is the specific heat at constant pressure. We have neglected some insignificant terms which are related to the mechanical work induced by density variations due to temperature changes of the water and solid phases. Also assumed to hold is the following relationship for an ideal gas:

$$\left. \frac{\hat{c} \ln \left( \frac{1}{\rho} \right)}{\hat{c} \ln T} \right|_p = 1 \quad (2.332)$$

Gas pressure changes are usually very slow and their gradients small in geomaterials, hence the time derivatives of gas pressure, as well as the convective terms, are negligible

with respect to the other terms in (2.331). Also, because gas pressure changes are usually small when compared to atmospheric pressure, the specific heat at constant pressure may be utilised instead of the specific heat at constant volume.

A local equilibrium state is assumed to hold, i.e.

$$T^s = T^w = T^g = T \quad (2.333)$$

We also introduce the constitutive equation for heat fluxes (2.191) and add together (2.330) and (2.331), then introduce the appropriate heat sources to obtain the following form of the energy balance equation:

$$\begin{aligned} (\rho C_p)_{\text{eff}} \frac{\partial T}{\partial t} + (\rho^w C_p^w \mathbf{v}^w + \rho^g C_p^g \mathbf{v}^g) \cdot \nabla T \\ - \operatorname{div}(\chi_{\text{eff}} \nabla T) = -\dot{m} \Delta H_{\text{vap}} \end{aligned} \quad (2.334)$$

where the effective thermal capacity  $(\rho C_p)_{\text{eff}}$  and the effective thermal conductivity are defined by (2.250) and the convective heat flux in the solid phase has been neglected.  $\Delta H_{\text{vap}}$  is the latent heat of evaporation; see also (2.250). The mass rate of water evaporation is eliminated from (2.333) by means of the mass balance equation for water (2.317).

## REFERENCES

1. de Boer, R., Ehlers, W., Kowalski S. and Plischka, J. (1991) Porous media: a survey of different approaches. *Forschungsbericht aus dem Fachbereich Bauwesen*, 54, Uninversität-Gesamthochschule Essen.
2. Bedford, A. and Drumheller, D. S. (1983) Theories of immiscible and structured mixtures. *Int. J. Engng Sci.*, **21**, 863–960.
3. Hassanzadeh, M. and Gray, W. G. (1979) General conservation equations for multiphase systems: 1, averaging procedure. *Adv. Water Resources*, **2**, 131–44.
4. Hassanzadeh, M. and Gray, W. G. (1979) General conservation equations for multiphase systems: 2, mass, momenta, energy and entropy equations. *Adv. Water Resources*, **2**, 191–203
5. Hassanzadeh, M. and Gray W. G. (1980) General conservation equations for multiphase systems: 3, constitutive theory for porous media flow. *Adv. Water Resources*, **3**, 25–40.
6. Hassanzadeh, S. M. (1986) Derivation of basic equations of mass transport in porous media, Part 1. Macroscopic balance laws. *Adv. Water Resources*, **9**, 196–206.
7. Hassanzadeh, S. M. (1986) Derivation of basic equations of mass transport in porous media, Part 2. Generalized Darcy's law and Fick's law. *Adv. Water Resources*, **9**, 207–22.
8. Bear, J. and Bachmat, Y. (1984) Transport phenomena in porous media – basic equations, in *Fundamentals of Transport Phenomena in Porous Media*, J. Bear and M. Y. Corapcioglu (eds), Nijhoff, Dordrecht, pp. 5–61.

9. Bear, J. and Corapcioglu, Y. (1981) Mathematical model for regional land subsidence due to pumping. 2. Integrated aquifer subsidence equation for vertical and horizontal displacements, *Water Resources Research*, **17**, 947–58.
10. Simoni, L. and Schrefler, B. A. (1989) FE solution of a vertically averaged model for regional land subsidence. *Int. J. Num. Meth. Engng*, **27**, 215–30.
11. Lemaitre, J. and Chaboche J. L. (1988) *Mécanique des matériaux solides*, Dunod, Paris.
12. Aurialt, J. L. (1991) Dynamic behaviour of porous media, in *Transport Processes in Porous Media*, J. Bear and M. Y. Corapcioglu (eds), Kluwer, Dordrecht, pp. 471–519.
13. Eringen, A. C. and Suhubi, E. S. (1964) Nonlinear theory of simple microelastic solids. *Int. J. Engng Sci.*, **2**, 189–203.
14. Chen, W. F. and Tsui, Y. (1992) Limitations to the large strain theory. *Int. J. Num. Meth. Engng*, **33**, 101–14.
15. Molenkamp, F. (1986) Limits to the Jaumann stress rate. *Int. J. Num. Anal. Meth. Geomech.*, **10**, 151–76.
16. Williams, W. O. (1978) Constitutive equations for flow of an incompressible viscous fluid through a porous medium. *Quart. Appl. Math.*, **Oct**, 255–67.
17. Noll, W. (1993) Theory of interpenetrating continuous bodies, in *Proceedings of Euromech Colloquium 290 on Mechanics of Swelling*, Rhodes, Greece, August 1993.
18. Sampaio, R. and Williams, W. O. (1979) Thermodynamics of diffusing mixtures. *J. de Mécanique*, **18**, 19–45.
19. Coleman, B. D. and Noll, W. (1963) The thermodynamics of elastic materials with heat conduction and viscosity. *Arch. Rational Mech. Anal.*, **13**, 168–78.
20. Gray, W. G. and Hassanizadeh S. M. (1991) Unsaturated flow theory including interfacial phenomena. *Water Resources Research*, **27**, 1855–63.
21. Gray, W. G. and Hassanizadeh, S. M. (1991) Paradoxes and realities in unsaturated flow theory. *Water Resources Research*, **27**, 1847–54.
22. Hassanizadeh, S. M. and Gray, W. G. (1990) Mechanics and thermodynamics of multi-phase flow in porous media including interphase boundaries. *Adv. Water Resources*, **13**, 169–186.
23. Moran, M. J. and Shapiro, H. N. (1993) *Fundamentals of Engineering Thermodynamics*, 2nd edn, Wiley, New York.
24. Organisation (1993) *ASHRAE Handbook: Fundamentals Volume*, ASHRAE, Atlanta.
25. Bomberg, M. and Shirtliffe, C. J. (1978) Influence of moisture gradients on heat transfer through porous building materials in *Thermal Transmission Measurements of Insulation* (ASTM STP 660), R. P. Tye (ed.), ASTM, Materials Park, OH, pp. 211–33.
26. Van der Kooi, F. (1991) Moisture transport in cellular concrete roof. PhD thesis, Delft University.
27. Fernandez, R. T. (1972) Natural convection from cylinders buried in porous media. PhD thesis, University of California.
28. Corey, A. T. (1957) Measurement of water and air permeability in unsaturated soil. *Soil Sci. Soc. Amer. Proc.*, **21**(1), 7–10.
29. Brooks, R. N. and Corey, A. T. (1966) Properties of porous media affecting fluid flow. *J. Irrig. Drain. Div. Amer. Soc. Civ. Engng*, **92**(IR2), 61–68.
30. Bishop, A. W. and Blight, G. E. (1963) Some aspects of effective stress in saturated and partly saturated soils. *Géotechnique*, **13**, 177–97.
31. Schrefler, B. A., Simoni, L., Xikui, L. and Zienkiewicz, O. C. (1990) Mechanics of partially saturated porous media, in *Numerical Methods and Constitutive Modelling in Geomechanics*, C.S. Desai and G. Gioda (eds), Springer-Verlag, Wien, pp. 169–209.
32. Biot, M. A. and Willis, P. G. (1957) The elastic coefficients of the theory of consolidation. *J. Appl. Mech.*, **24**, 594–601.

33. Whitaker, S. (1977) Simultaneous heat mass and momentum transfer in porous media: a theory of drying. *Advances in Heat Transfer*, Vol. 13. Academic Press, New York.
34. Nozad, I., Carbonell, R. G. and Whitaker, S. (1985) Heat conduction in multiphase system 1: theory and experiment for two-phase system. *Chem. Engng Sci.*, **40**, 843–55.
35. Schrefler, B. A., Zhan, X. and Simoni (1995) A coupled model for water flow, airflow and heat flow in deformable porous media. *Int. J. Heat and Fluid Flow*, **5**, 531–47.
36. Schrefler, B. A., D'Alpaos, L., Zhan, X.Y. and Simoni, L. (1994) Pollutant transport in deforming porous media. *Eur. J. Mech. A*, **13**(4), suppl., 175–94.
37. Baggio, P., Bonacina, C. and Strada, M. (1993) Trasporto di calore e di massa nel calcestruzzo cellulare. *La Termotecnica*, **45**, 53–60.
38. Gawin, D., Baggio, P. and Schrefler, B. A. (1995) Coupled heat, water and gas flow in deformable porous media. *Int. J. Num. Meth. Fluids*, **20**, 969–87.
39. Bear, J. and Bachmat, Y. (1991) *Introduction to Modelling of Transport Phenomena in Porous Media*. Kluwer, Dordrecht, Netherlands.
40. Biot, M. A. (1941) General theory of three-dimensional consolidation. *J. Appl. Phys.*, **12**, 155–64.
41. Biot, M. A. (1941) Consolidation settlement under a rectangular load. *J. Appl. Phys.*, **12**, 426–30.
42. Biot, M. A. (1955) Theory of elasticity and consolidation for a porous anisotropic solid. *J. Appl. Phys.*, **26**, 182–85.
43. Biot, M. A. (1956) General solution of the equation of elasticity and consolidation for a porous material. *J. Appl. Mech.*, **23**, 91–96.
44. Biot, M. A. (1956) Theory of deformation of a porous viscoelastic anisotropic soil. *J. Appl. Phys.*, **27**, 459–67.
45. Biot, M. A. (1963) Theory of stability and consolidation of a porous medium under initial stress. *J. Math. Mech.*, **12**, 521–41.
46. Schrefler, B. A. and Zhan, X. Y. (1993) A fully coupled model for water flow and airflow in deformable porous media. *Water Resources Research*, **29**, 155–67.
47. Zienkiewicz, O. C., Xie, Y. M., Schrefler, B. A., Ledesma, A. and Bicanic, N. (1990) Static and dynamic behaviour of soils: a rational approach for quantitative solutions. II. Semi-saturated problems. *Proc. R. Soc. London A*, **429**, 311–21.
48. Bishop, A. V. (1959) The principle of effective stress. *Teknisk Ukeblad*, **39**, 859–63.
49. Skempton, A. W. (1961) Effective stress in soil, concrete and rocks, in *Pore Pressure and Suction in Soils*, Skempton, A. W., ed. Butterworth, London, pp. 4–16.
50. Aitchison, G. D. and Donald, I. B. (1956) Effective stress in unsaturated soils, in *Proceedings of the 2nd Australia–New Zealand Conference on Soil Mechanics*, pp. 192–99.
51. Morland, L.W. (1972) A simple constitutive theory for a fluid-saturated porous solid. *J. Geophys. Res.*, **77**, 890–900.
52. Gudehus, G. (1995) A comprehensive concept for non-saturated granular bodies, in *Proceedings of the 1st International Conference on Unsaturated Soils*, Paris, Balkema, Rotterdam.
53. Li, X. and Zienkiewicz, O. C. (1992) Multiphase flow in deforming porous media and finite element solutions. *Computers and Structures*, **45**, 211–27.
54. Lade, P. V. and de Boer, R. (1997) The concept of effective stress for soil, concrete and rock, *Géotechnique*, **47**, 61–78.
55. Zienkiewicz, O. C. and Shiomi, T. (1985) Dynamic behaviour of saturated porous media: the general Biot formulation and its numerical solution. *Int. J. Num. Anal. Meth. Geomech.*, **8**, 71–96.
56. Bird, R. B.. Stewart, E. W. and Lightfoot, E. N. (1960) *Transport Phenomena*, Wiley, Chichester.

## APPENDIX 2A

Tensorial notation is used in Chapters 2 and 14. In particular, the products are defined as follows.

For second-order tensors  $\mathbf{A}$ ,  $\mathbf{B}$

$$\mathbf{A} \cdot \mathbf{B} \text{ is the tensor whose component } (\mathbf{A} \cdot \mathbf{B})_{rs} \text{ is } (\mathbf{A} \cdot \mathbf{B})_{rs} = \sum_{l=1}^3 A_{rl} B_{ls} \quad (2A.1)$$

$$\mathbf{A} : \mathbf{B} \text{ is the scalar } \sum_{r,l=1}^3 A_{rl} B_{lr} \quad (2A.2)$$

For tensor  $\mathbf{A}$  and vector  $\mathbf{v}$

$$\mathbf{A} \cdot \mathbf{v} \text{ is a vector of components } \sum_{l=1}^3 A_{rl} v_l \quad (2A.3)$$

The product of a fourth-order tensor  $\mathbf{D}$  and a second-order tensor  $\mathbf{A}$  is indicated as

$$\mathbf{B} = \mathbf{D}[\mathbf{A}] \quad \text{with} \quad \mathbf{B}_{rs} = \sum_{l,m=1}^3 D_{rslm} A_{lm} \quad (2A.4)$$

For vectors  $\mathbf{u}$ ,  $\mathbf{v}$  the tensor product or diadic product  $\mathbf{u} \otimes \mathbf{v}$  is defined as

$$\mathbf{u} \otimes \mathbf{v} \cdot \mathbf{a} = \mathbf{u}\mathbf{v} \cdot \mathbf{a} \quad \forall \mathbf{a} \in R^3 \quad (2A.5)$$

Finally the cross-product between vectors  $\mathbf{u}$ ,  $\mathbf{v}$  is indicated as  $\mathbf{u} \times \mathbf{v}$ . The  $i$ th component of  $\mathbf{u} \times \mathbf{v}$  is

$$(\mathbf{u} \times \mathbf{v})_i = \sum_{r,s=1}^3 \epsilon_{irs} u_r v_s \quad (2A.6)$$

where

$$\epsilon_{irs} \begin{cases} 0 & \text{if } i, r, s \text{ are not distinct} \\ 1 & \text{if } (i, r, s) \text{ is an even permutation} \\ -1 & \text{if } (i, r, s) \text{ is an odd permutation} \end{cases}$$

is the Ricci tensor.

The trace or first invariant of tensor  $\mathbf{A}$  is the scalar

$$\text{tr } \mathbf{A} = \sum_{i=1}^3 A_{ii} \quad (2A.7)$$

## APPENDIX 2B

For integration over  $dv$  or  $da$  of averaged quantities, remember that they are constant in  $dv$  and  $da$  and that the following identities for the microscopic functions  $\zeta(\mathbf{r}, t)$  and  $\omega(\mathbf{r}, t)$  hold [1,2]:

$$\bar{\zeta}^\pi = \frac{1}{\langle \rho \rangle_\pi} \int_{dv} \rho (\zeta - \bar{\zeta}^\pi) \gamma^\pi dv_m = \bar{\zeta}^\pi - \bar{\zeta}^\pi = 0 \quad (2B.1)$$

$$\bar{\zeta}^\pi \bar{\omega}^\pi = \bar{\zeta}^\pi \bar{\omega}^\pi = 0 \quad (2B.2)$$

$$\bar{\zeta}^\pi \bar{\omega}^\pi = (\bar{\zeta}^\pi + \bar{\zeta}^\pi)(\bar{\omega}^\pi + \bar{\omega}^\pi)^\pi = \bar{\zeta}^\pi \bar{\omega}^\pi + \bar{\zeta}^\pi \bar{\omega}^\pi \quad (2B.3)$$

$$\begin{aligned} \int_A \bar{\rho \zeta}^\pi dA &= \int_A \left[ \frac{1}{da} \int_{dv} (\rho \bar{\zeta}^\pi) \cdot \mathbf{n}(\mathbf{r}, t) \gamma^\pi da_m \right] dA = \\ &= \int_V \left[ \operatorname{div} \left( \frac{1}{dv} \int_{dv} \rho \bar{\zeta}^\pi \gamma^\pi dv_m \right) \right] dV = \\ &= \int_V \operatorname{div} \left( \langle \rho \rangle_\pi \bar{\zeta}^\pi \right) dV = 0 \end{aligned} \quad (2B.4)$$

$$\begin{aligned} \int_A \bar{(\rho \bar{\zeta}^\pi \bar{\omega}^\pi)}^\pi dA &= \int_A \left[ \frac{1}{da} \int_{dv} (\rho \bar{\zeta}^\pi \bar{\omega}^\pi) \cdot \mathbf{n}(\mathbf{r}, t) \gamma^\pi da_m \right] dA \\ &= \int_V \left[ \operatorname{div} \left( \frac{1}{dv} \int_{dv} \rho \bar{\zeta}^\pi \bar{\omega}^\pi \gamma^\pi dv_m \right) \right] dV \\ &= \int_V \operatorname{div} \left( \langle \rho \rangle_\pi \bar{\zeta}^\pi \bar{\omega}^\pi \right) dV = 0 \end{aligned} \quad (2B.5)$$

## References

1. Hassanizadeh, M. and Gray, W.G. (1979) General conservation equations for multiphase systems: I, averaging procedure. *Adv. Water Resources*, **2**, 131–44.
2. de Boer, R., Ehlers, W., Kowalski S. and Plischka, J. (1991) Porous media: a survey of different approaches. Forschungsbericht aus dem Fachbereich Bauwesen, 54, Universitäts-Gesamthochschule Essen.

## APPENDIX 2C

Equation (2.152) (i.e the Kelvin law) can be rewritten as

$$p^c = -\rho^w \frac{R\theta}{M_w} \ln \left( \frac{p^{gw}}{p_c^{gw}} \right) \quad (2C.1)$$

When a liquid is in equilibrium with its vapour, the specific (mass) Gibbs functions (or thermodynamic potentials) of the two phases are equal and then

$$g^w = g^{gw} \quad \text{and} \quad dg^w = dg^{gw} \quad (2C.2)$$

The differential of the Gibbs function of vapour  $dg^{gw}$  is given by the well-known expression

$$dg^{gw} = -\bar{\lambda}^{gw} d\theta + v^{gw} dp^{gw} \quad (2C.3)$$

where  $v^{gw} = 1/\rho^{gw}$  is the specific (mass) volume of water vapour and  $\bar{\lambda}^{gw}$  is the specific (mass) entropy of water vapour. Assuming that vapour behaves as an ideal gas, i.e.  $p^{gw} = \rho^{gw}\theta R/M_w$  or  $p^{gw} = R\theta/M_w v^{gw}$ , equation (2C.3) becomes

$$dg^{gw} = -\bar{\lambda}^{gw} d\theta + \frac{R\theta}{M_w p^{gw}} dp^{gw} \quad (2C.4)$$

If we integrate (2C.4), considering an isothermal evolution ( $d\theta = 0$ ), from the conditions of equilibrium of vapour with bulk liquid on a flat surface (vapour pressure equal to the vapour tension  $p^{gws}$ ), up to the actual value of vapour pressure inside the porous medium, we obtain

$$g^{gw} - g^{gwo} = \frac{R\theta}{M_w} \ln \left( \frac{p^{gw}}{p^{gws}} \right) \quad (2C.5)$$

Equation (2C.5) gives the difference between the Gibbs function of the water inside the porous medium in equilibrium with vapour at pressure  $p^{gw}$  and the bulk liquid water, at the same temperature. From the thermodynamic point of view, moisture stress or water potential  $\psi$  is then defined as the specific (mass) Gibbs function (or specific thermodynamic potential) of water in a porous medium, referred to pure liquid at the same temperature; it is given by

$$\Psi = g^{gw} - g^{gwo} = \frac{R\theta}{M_w} \ln \left( \frac{p^{gw}}{p^{gws}} \right) \quad (2C.6)$$

Comparing (2C.6) with (2C.1), it is immediately evident that

$$-\frac{p^c}{\rho^w} = \Psi = \frac{R\theta}{M_w} \ln \left( \frac{p^{gw}}{p^{gws}} \right) \quad (2C.7)$$

Assuming the liquid is incompressible ( $\rho^w = \text{const.}$ ), the capillary pressure  $p^c$  is then equal to the water potential  $\psi$  multiplied by a constant ( $-\rho^w$ ). The water potential  $\psi$  is defined for all conditions of water retained inside a porous medium (monolayer surface adsorption, multilayer surface adsorption and capillary suction), whereas the capillary pressure  $p^c$  has a clear physical meaning,  $p^c = p^g - p_w$ , only when liquid water (sometimes called free water) is present.

Redefining the capillary pressure as  $p^c = -\rho^w\Psi$ , equation (2.152) or (2C.1) can be used even for water contents below the capillary region.

# 3

## Numerical Solution for Isothermal Consolidation

### 3.1 INTRODUCTION

Numerical solution procedures are now introduced using the simplest case of fluid–solid interaction, which is isothermal consolidation in small-strain situations. In this context the term *consolidation* denotes slow deformation of the solid phase, accompanied by flow of the pore fluid(s): it is essentially one-phase or multiphase flow in a deforming porous medium, where inertial forces are negligible. In soil mechanics, this phenomenon is known as primary consolidation. Secondary consolidation will be dealt with in Chapter 11. Different solution procedures have been proposed in the literature for the quantitative solution of consolidation problems. They may be subdivided into uncoupled solutions and coupled solutions.

Uncoupled solutions are briefly discussed in Chapter 6, because they have been used, and are still sometimes used, in surface subsidence analysis. In the general case, however, they are irrelevant and we deal here and in the following chapters with coupled finite element techniques only. In particular, the finite element method is used for the discretisation in space and finite difference techniques are used for the time discretisation. This chapter begins with saturated, one-phase flow problems in deforming porous media, then continues with saturated–unsaturated flow problems under the assumption that air can flow without resistance within the pores of the medium and remains always at atmospheric pressure. This last assumption is common for most soil mechanics situations. Finally comes the case where both air and water are flowing within the pores of a deforming porous body.

### 3.2 COUPLED SOLUTION: SATURATED ONE-PHASE FLOW IN A DEFORMING POROUS MEDIUM

#### 3.2.1 Governing equations

Vector notation is used throughout the numerical part of this book. The stress and strain tensors are now written in vector form as

$$\begin{aligned}\boldsymbol{\sigma} &= \{\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yz}, \tau_{zx}\}^T \\ \boldsymbol{\varepsilon} &= \{\varepsilon_x, \varepsilon_y, \varepsilon_z, \gamma_{xy}, \gamma_{yz}, \gamma_{zx}\}^T\end{aligned}\tag{3.1}$$

Under the assumptions of small-strain theory, isothermal equilibrium and negligible inertial forces, we have the following balance equations. First, the linear momentum balance equation for a two-phase medium, i.e. solid plus water, as obtained from (2.262) or (2.309), results in

$$\mathbf{L}^T \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \quad (3.2)$$

where the differential operator  $\mathbf{L}$  is defined as

$$\mathbf{L} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \end{bmatrix} \quad (3.3)$$

and  $\rho = (1 - n)\rho^s + n\rho^w$  (2.310).

Second, the mass balance equation for water, obtained from (2.253) or (2.318) by considering  $S_w = 1$ ,  $k^{rw} = 1$ , the mass rate of evaporation,  $\dot{m} = 0$ , and by neglecting the gradient of the water density, is given by

$$\left( \frac{\alpha - n}{\mathbf{K}_s} + \frac{n}{\mathbf{K}_w} \right) \frac{\partial p^w}{\partial t} + \alpha \operatorname{div} \mathbf{v}^s + \operatorname{div} \left[ \frac{\mathbf{k}}{\mu^w} (-\operatorname{grad} p^w + \rho \mathbf{g}) \right] = 0 \quad (3.4)$$

Together with the balance equations, we take into account the effective stress in the form of (2.278), with  $S_w = 1$ ,  $S_g = 0$ , i.e.

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}'' - \alpha \mathbf{m} p^w \quad (3.5)$$

where instead of the unit tensor  $\mathbf{I}$  we have the vector  $\mathbf{m}$ ,

$$\mathbf{m}^T = [1, 1, 1, 0, 0, 0]^T \quad (3.6)$$

and the constitutive relationship of the solid phase (2.303),

$$d\boldsymbol{\sigma}'' = \mathbf{D}_T d\boldsymbol{\epsilon} \quad (3.7)$$

The matrix  $\mathbf{D}_T$  will be dealt with extensively in the next chapter. Finally, the relationship between the change of total strain  $d\boldsymbol{\epsilon}$  and the change of displacement  $d\mathbf{u}$  is given by

$$d\boldsymbol{\epsilon} = \mathbf{L} d\mathbf{u} \quad (3.8)$$

It can be immediately verified that

$$\operatorname{div} \mathbf{v}^s = \mathbf{m}^T \frac{\partial \boldsymbol{\epsilon}}{\partial t} = \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} \quad (3.9)$$

### 3.2.2 Initial and boundary conditions

Furthermore, it is necessary to define the initial and boundary conditions. The initial conditions specify the full field of displacements and water pressures at time  $t = 0$ :

$$\mathbf{u} = \mathbf{u}_0, p^w = p_0^w \quad \text{in } \Omega \text{ and on } \Gamma \quad (3.10)$$

where  $\Omega$  is the domain of interest and  $\Gamma$  its boundary. The boundary conditions can be either imposed values on  $\Gamma_\pi$ , or fluxes on  $\Gamma_\pi^q$  where the boundary  $\Gamma = \Gamma_\pi \cup \Gamma_\pi^q$ . The imposed boundary values for displacements and water pressures are as follows:

$$\begin{aligned} \mathbf{u} &= \hat{\mathbf{u}} && \text{on } \Gamma_u \\ p^w &= \hat{p}^w && \text{on } \Gamma_w \end{aligned} \quad (3.11)$$

The traction boundary condition for the stresses is

$$\mathbf{I}^T \boldsymbol{\sigma} = \tilde{\mathbf{t}} \quad \text{on } \Gamma_u^q \quad (3.12)$$

where the matrix  $\mathbf{I}$  is related to the unit normal vector  $\mathbf{n} = \{n_x, n_y, n_z\}^T$  by

$$\mathbf{I} = \begin{bmatrix} n_x & 0 & 0 \\ 0 & n_y & 0 \\ 0 & 0 & n_z \\ n_y & n_x & 0 \\ 0 & n_z & n_y \\ n_z & 0 & n_x \end{bmatrix} \quad (3.13)$$

The flux boundary condition for water stems from (2.219) or (2.311), i.e.

$$\rho^w \frac{\mathbf{k}}{\mu^w} (-\operatorname{grad} p^w + \rho^w \mathbf{g})^T \cdot \mathbf{n} = q^w \quad \text{on } \Gamma_w^q \quad (3.14)$$

where  $q^w$  is the imposed mass flux normal to the boundary. Equations (3.12) and (3.14) are the natural boundary conditions for the above balance equations, when their solution is obtained via a ‘weak’ formulation of the problem as is done in the following sections.

## 3.3 SOLUTION OF THE BOUNDARY VALUE PROBLEM

A boundary value problem of the type stated in the previous section can be represented as

$$\mathbf{A}(\mathbf{u}) = \mathbf{C}(\mathbf{u}) + \mathbf{p} = \mathbf{0} \quad \text{in } \Omega \quad (3.15)$$

with

$$\mathbf{B}(\mathbf{u}) = \mathbf{M}(\mathbf{u}) + \mathbf{q} = \mathbf{0} \quad \text{on } \Gamma \quad (3.16)$$

where  $\mathbf{C}$  is an appropriate differential operator and  $\mathbf{p}$  is a known function independent of the field variable  $\mathbf{u}$ .  $\mathbf{M}$  is also an appropriate operator and  $\mathbf{q}$  a known function independent of  $\mathbf{u}$  which is the exact solution of the boundary value problem.

Consider the following integral statement:

$$\int_{\Omega} \mathbf{v}^T \mathbf{A}(\mathbf{u}) d\Omega + \int_{\Gamma} \bar{\mathbf{v}}^T \mathbf{B}(\mathbf{u}) d\Gamma = 0 \quad (3.17)$$

is satisfied for a set of arbitrary functions  $\mathbf{v}$  and  $\bar{\mathbf{v}}$ . This is equivalent to satisfying differential equation (3.15) and boundary condition (3.16). In fact, if (3.15) and (3.16) are satisfied, then (3.17) is true. Conversely, if (3.17) is valid for any arbitrary values of  $\mathbf{v}$  and  $\bar{\mathbf{v}}$ , then (3.15) and (3.16) must be satisfied at all points within and on the boundary of the domain.

We have implicitly assumed that the integrals appearing in (3.17) can be evaluated. This somewhat restricts the class of functions to which  $\mathbf{v}$  and  $\bar{\mathbf{v}}$  must belong. In particular, functions have to be avoided for which any term in the integrals becomes infinite. Further details regarding the restrictions, which do not limit the validity of the above statement, can be found in the literature [1].

Only for the very simplest of problems can an exact solution for (3.17) be obtained. Most practical problems require some form of approximation, achieved here by the finite element method. Hence an approximate solution is sought in the class of functions  $\hat{\mathbf{u}}$ :

$$\mathbf{u} \approx \hat{\mathbf{u}} = \sum_i^r \mathbf{N}_i \mathbf{a}_i = \mathbf{Na} \quad (3.18)$$

which is obtained by introducing a set of trial or shape functions  $\mathbf{N}_i$  in terms of the independent variables (e.g. the coordinates). In (3.18)  $\mathbf{a}_i$  are the parameters which can be identified with the values of the unknowns at defined points (nodes) in the domain  $\Omega$  and are determined so as to obtain a good fit for the solution.

The chosen shape functions are sets of polynomials which must reproduce any well-behaved function. The reader is referred to standard finite element textbooks [1] and to Section 3.5. The domain mapped by the shape functions comprises non-overlapping subdomains  $\Omega_e$  (elements  $e$ ) in which the domain has been divided (piecewise approximation).

A general method is now required for determining the constants which appear in (3.18). When the approximation  $\hat{\mathbf{u}}$  is substituted into (3.15) and (3.16) there remains an error, or residual, i.e.

$$\mathbf{R} = \mathbf{R}_{\Omega} + \mathbf{R}_{\Gamma} = \mathbf{A}(\hat{\mathbf{u}}) + \mathbf{B}(\hat{\mathbf{u}}) \quad (3.19)$$

Therefore, a numerical solution is sought that will minimise the residual over the whole domain  $\Omega$  and the boundary. This will require a zero value for an appropriate number of integrals of the error over  $\Omega$  and  $\Gamma$ , weighted in different ways by weighting functions  $\mathbf{w}$  and  $\bar{\mathbf{w}}$ . And this means

$$\int_{\Omega} \mathbf{w}^T \mathbf{A}(\hat{\mathbf{u}}) d\Omega + \int_{\Gamma} \bar{\mathbf{w}}^T \mathbf{B}(\hat{\mathbf{u}}) d\Gamma = 0 \quad (3.20)$$

where, in general, the functions  $\mathbf{w}$  and  $\bar{\mathbf{w}}$  can be independently chosen. If (3.20) is satisfied for a very large number of arbitrary weighting functions, the approximate solution must approach the exact solution, under the condition that it belongs to the class of functions of expansion (3.18). This process is well known as the weighted residual method.

Equation (3.20) is an approximation to the integral statement (3.17) and results in a set of simultaneous algebraic equations for the unknown parameters  $\mathbf{a}_i$ . The system can be written as

$$\mathbf{K}\mathbf{a} = \mathbf{f} \quad (3.21)$$

with

$$\mathbf{K}_{ij} = \sum_{e=1}^m \mathbf{K}_{ij}^e \quad \mathbf{f}_i = \sum_{e=1}^m \mathbf{f}_i^e \quad (3.22)$$

Possible symmetries can be destroyed by an inappropriate choice of the weighting functions. Furthermore, the fact that integrals must be calculated over the domain results in restrictions on the shape functions, i.e. they must possess at least the same order of derivative of the differential operators  $\mathbf{A}$  and  $\mathbf{B}$ , and these derivatives must be continuous.

### 3.4 APPLICATION OF THE FINITE ELEMENT METHOD

In the following, we choose as basic variables the displacements  $\mathbf{u}$  and pore water pressures  $p^w$ . By applying the integral equation (3.20) to (3.2) and (3.12), we obtain

$$\int_{\Omega} \mathbf{w}^T (\mathbf{L}^T \boldsymbol{\sigma} + \rho \mathbf{g}) d\Omega + \int_{\Gamma} \bar{\mathbf{w}}^T (\mathbf{I}^T \boldsymbol{\sigma} - \bar{\mathbf{t}}) d\Gamma = 0 \quad (3.23)$$

It may also be assumed that the boundary condition (3.11) for  $\mathbf{u}$  on  $\Gamma_u$  is satisfied by the choice of the approximation for  $\mathbf{u}$ . The choice of the weighting functions is limited in such a way that

$$\mathbf{w} = \mathbf{0} \quad \text{on } \Gamma_u \quad (3.24)$$

$$\bar{\mathbf{w}} = -\mathbf{w} \quad \text{on } \Gamma_u^q \quad (3.25)$$

We further apply Green's theorem to the first portion of (3.23), which proves the following result [1]:

$$\int_{\Omega} \phi \frac{\partial \Psi}{\partial x} d\Omega = - \int_{\Omega} \frac{\partial \phi}{\partial x} \Psi d\Omega + \int_{\Gamma} \phi \Psi n_x d\Gamma \quad (3.26)$$

Hence (3.23) becomes

$$-\int_{\Omega} (\mathbf{L}\mathbf{w})^T \boldsymbol{\sigma} d\Omega + \int_{\Gamma} \mathbf{w}^T \mathbf{I}^T \boldsymbol{\sigma} d\Gamma + \int_{\Omega} \mathbf{w}^T \rho \mathbf{g} d\Omega + \int_{\Gamma} \bar{\mathbf{w}}^T (\mathbf{I}^T \boldsymbol{\sigma} - \bar{\mathbf{t}}) d\Gamma = 0 \quad (3.27)$$

and, because of the choice of weighting functions (3.24) and (3.25), is reduced to

$$\int_{\Omega} (\mathbf{L}\mathbf{w})^T \boldsymbol{\sigma} d\Omega = \int_{\Omega} \mathbf{w}^T \rho \mathbf{g} d\Omega + \int_{\Gamma_w^q} \mathbf{w}^T \bar{\mathbf{t}} d\Gamma \quad (3.28)$$

The weighted residual method is now applied to the continuity equation for water flow (3.4) and to its natural boundary condition, (3.14). This yields,

$$\begin{aligned} & \int_{\Omega} \mathbf{w}^{*T} \left\{ \nabla^T \left[ \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] + \alpha \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + \left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} \right\} d\Omega \\ & + \int_{\Gamma_w^q} \bar{\mathbf{w}}^{*T} \left[ \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g})^T \cdot \mathbf{n} - \frac{q^w}{\rho^w} \right] d\Gamma = 0 \end{aligned} \quad (3.29)$$

where the gradient operator used is

$$\nabla = \left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\}^T \quad (3.30)$$

$\nabla^T$  indicates the divergence operator, and where  $\mathbf{w}^*$  and  $\bar{\mathbf{w}}^*$  are arbitrary weighting functions. It is assumed that the boundary condition (3.11) for  $p^w$  on  $\Gamma_w$  is satisfied by the choice of approximation of  $p^w$ . In (3.29) the appearance of second-order derivatives for  $p^w$  necessitates a higher-order continuity for the shape functions. To overcome this limitation, we apply Green's theorem to the terms of the divergence operator. Furthermore, we limit the choice of the weighting functions such that

$$\mathbf{w}^* = 0 \quad \text{on } \Gamma_w \quad (3.31)$$

and

$$\bar{\mathbf{w}}^* = -\mathbf{w}^* \quad \text{on } \Gamma_w^q \quad (3.32)$$

Equation (3.29) may now be written as

$$\begin{aligned} & \int_{\Omega} \left[ -(\nabla \mathbf{w}^*)^T \left( -\frac{\mathbf{k}}{\mu^w} \nabla p^w + \frac{\mathbf{k}}{\mu^w} \rho^w \mathbf{g} \right) + \mathbf{w}^{*T} \alpha \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + \mathbf{w}^{*T} \left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} \right] d\Omega \\ & + \int_{\Gamma_w^q} \mathbf{w}^{*T} \frac{q^w}{\rho^w} d\Gamma = 0 \end{aligned} \quad (3.33)$$

The finite element approximation is now applied to (3.28) and (3.33). As indicated in (3.18), the displacements and pore pressures are expressed in terms of their values  $\mathbf{u}$  and  $p^w$  at a finite number of points in space. This procedure involves the division of the continuum into elements, and the expression of  $\mathbf{u}$  and  $p^w$  within an element in terms of their values at a finite number of points within or on the boundary of that element. In order to ensure continuity of displacements and pore pressures between the elements, it is necessary to place a sufficient number of nodes on the element boundary to satisfy the

shape function  $\mathbf{N}^e$  being used for the elements. The expressions for  $\mathbf{u}$  and  $p^w$  are

$$\mathbf{u} = \mathbf{N}_u \bar{\mathbf{u}} \quad (3.34)$$

and

$$p^w = \mathbf{N}_p \bar{\mathbf{p}}^w \quad (3.35)$$

where  $\bar{\mathbf{u}}$  and  $\bar{\mathbf{p}}^w$  are the vectors of the nodal values of the unknowns. In these equations, global shape functions have been used which already take into account the summation of all element contributions implied in (3.22). The numbers and locations of the nodes are not necessarily the same for  $\mathbf{u}$  and  $p^w$  (section 3.5).

Approximations (3.34) and (3.35) are now introduced into (3.28) and (3.33). Furthermore, by applying the Galerkin method (Bubnov-Galerkin)[2], the weighting functions  $\mathbf{w}$  and  $\mathbf{w}^*$  are replaced by the shape function  $\mathbf{N}_u$  and  $\mathbf{N}_p$  respectively. This gives the following set of equations:

$$\int_{\Omega} (\mathbf{L}\mathbf{N}_u)^T \boldsymbol{\sigma} d\Omega = \int_{\Omega} \mathbf{N}_u^T \rho \mathbf{g} d\Omega + \int_{\Gamma_u^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \quad (3.36)$$

$$\begin{aligned} & \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k}}{\mu^w} \nabla \mathbf{N}_p \bar{\mathbf{p}}^w d\Omega - \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k}}{\mu^w} \rho^w \mathbf{g} d\Omega \\ & + \int_{\Omega} \mathbf{N}_p^T \alpha \mathbf{m}^T \mathbf{L} \mathbf{N}_u \frac{\partial \bar{\mathbf{u}}}{\partial t} + \int_{\Omega} \mathbf{N}_p^T \left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} d\Omega + \int_{\Gamma_w^q} \mathbf{N}_p^T \frac{q^w}{\rho^w} d\Gamma = 0 \end{aligned} \quad (3.37)$$

Taking into account (3.5), these equations are rewritten as

$$\int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega - \mathbf{Q} \bar{\mathbf{p}}^w = \mathbf{f}^u \quad (3.38)$$

$$\mathbf{H} \bar{\mathbf{p}}^w + \mathbf{Q}^T \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{S} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} = \mathbf{f}^p \quad (3.39)$$

where

$\mathbf{B} = \mathbf{L}\mathbf{N}_u$  is the strain operator

$\mathbf{Q} = \int_{\Omega} \mathbf{B}^T \alpha \mathbf{m} \mathbf{N}_p d\Omega$  is the coupling matrix

$\mathbf{H} = \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k}}{\mu^w} \nabla \mathbf{N}_p d\Omega$  is the permeability matrix

$\mathbf{S} = \int_{\Omega} \mathbf{N}_p^T \left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \mathbf{N}_p d\Omega$  is the compressibility matrix [3]

The right-hand terms in (3.38) and (3.39) are given by

$$\begin{aligned} \mathbf{f}^u &= \int_{\Omega} \mathbf{N}_u^T [\rho^s(n-1) + \rho^w n] \mathbf{g} d\Omega + \int_{\Gamma_u^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \\ \mathbf{f}^p &= \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k}}{\mu^w} \rho^w \mathbf{g} d\Omega - \int_{\Gamma_w^q} \mathbf{N}_p^T \frac{q^w}{\rho^w} d\Gamma \end{aligned} \quad (3.40)$$

The first term in (3.38) represents the internal force

$$\mathbf{P}(\bar{\mathbf{u}}) = \int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega \quad (3.41)$$

For the isotropic linear elastic case (Chapter 4) the constitutive relationship (3.7) can be written as

$$\boldsymbol{\sigma}'' = \mathbf{D}_e \boldsymbol{\epsilon} = \mathbf{D}_e L \mathbf{N}_u \bar{\mathbf{u}} = \mathbf{D}_e \mathbf{B} \bar{\mathbf{u}} \quad (3.42)$$

The internal force can therefore be written as

$$\mathbf{P}(\bar{\mathbf{u}}) = \int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega = \int_{\Omega} \mathbf{B}^T \mathbf{D}_e \mathbf{B} d\Omega \bar{\mathbf{u}} = \mathbf{K}_e \bar{\mathbf{u}} \quad (3.43)$$

where  $\mathbf{K}_e = \int_{\Omega} \mathbf{B}^T \mathbf{D}_e \mathbf{B} d\Omega$  is the linear elastic stiffness matrix, which is symmetric in form [1]. However, in problems where the solid-phase behaviour is non-linear, only the tangential stiffness matrix  $\mathbf{K}_T$  can be defined, as will be explained in the next chapter.

$$\begin{aligned} \mathbf{K}_T &= \frac{\partial \mathbf{P}(\bar{\mathbf{u}})}{\partial \bar{\mathbf{u}}} = \int_{\Omega} \mathbf{B}^T \mathbf{D}_T \mathbf{B} d\Omega \\ \frac{\partial \mathbf{P}(\bar{\mathbf{u}})}{\partial t} &= \frac{\partial \mathbf{P}(\bar{\mathbf{u}})}{\partial \bar{\mathbf{u}}} \frac{\partial \bar{\mathbf{u}}}{\partial t} = \mathbf{K}_T \frac{\partial \bar{\mathbf{u}}}{\partial t} \end{aligned} \quad (3.44)$$

Integration of the above matrices usually requires numerical techniques. A standard method is that of Gaussian quadrature [1], where the integrands are evaluated at specific points of the element then weighted and summed. The procedure is carried out in terms of a set of local coordinates [1].

Since the discretisation in space has been carried out, (3.38) and (3.39) now represent a set of ordinary differential equations in time. For convenience, the equations are written in the following form, with the assumption of linear elastic behaviour of the solid skeleton:

$$\begin{bmatrix} \mathbf{K}_e & -\mathbf{Q} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{Q}^T & \mathbf{S} \end{bmatrix} \frac{d}{dt} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix} = \begin{Bmatrix} \mathbf{f}'' \\ \mathbf{f}^p \end{Bmatrix} \quad (3.45)$$

It can be seen immediately that this system is not symmetric. Symmetry can be restored, if  $\mathbf{P}(\mathbf{u})$  is symmetric, by time differentiating (3.38) and by multiplying one set of equations by  $-1$ . This procedure shifts the component matrices horizontally in the global matrix, and is common in the analysis of consolidation problems [4,5]. It will yield the following set of equations:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix} + \begin{bmatrix} -\mathbf{K}_T & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} \end{bmatrix} \frac{d}{dt} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix} = \begin{Bmatrix} -\frac{d}{dt} \mathbf{f}^u \\ \mathbf{f}^p \end{Bmatrix} \quad (3.46)$$

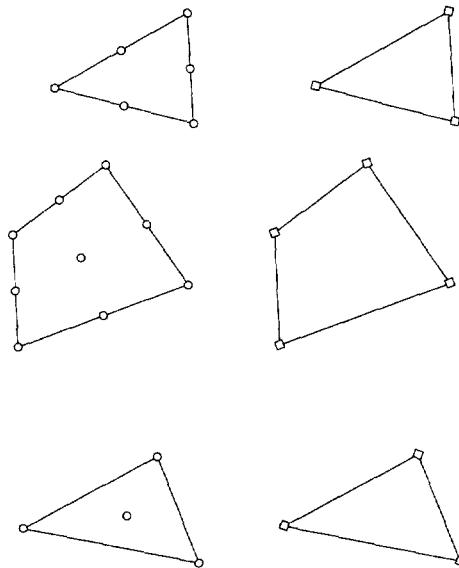
Time differentiation and successive integration is a possible way of introducing computationally non-linear behaviour [6]. This is the reason why  $\mathbf{K}_T$  appears in (3.46); see also equation (3.44). However, care must be exercised in the choice of the initial

conditions because the set of possible solution trajectories is modified through time differentiation. The integration in time will be dealt with in Section 3.6, but the next section considers the choice of possible elements.

### 3.5 CHOICE OF ELEMENTS

Isoparametric elements [1] are used throughout the book, where the coordinates within an element are interpolated using the same shape functions as for the variables. This mapping allows the use of elements of more arbitrary shape than simple forms such as rectangles and triangles.

In (3.34) and (3.35) different shape functions have been used for representing fluid pressure and displacements. This is necessary, especially when approaching the undrained limit state, where the permeability matrix, and often the compressibility matrix, are set to zero [6]. Therefore, in (3.45) and (3.46), zero diagonal terms appear. The matrices to be solved for are then identical to those used in the solution of problems in incompressible elasticity or fluid mechanics [1,6]. In such cases there are limitations for the approximating functions  $\mathbf{N}_u$  and  $\mathbf{N}_p$ , if the Babuska–Brezzi convergence condition is to be satisfied [7–9]. If the undrained limit state is never approached, the choice of elements is wide. Equal-order interpolations for both fields are then common and are used in many examples in this book. However, because of the presence of first-order spatial derivatives in the differential operators,  $C_0$  continuous shape functions should be used [1]. A choice of elements satisfying the convergence criteria for the undrained limit are shown in Figure 3.1 [10].



**Figure 3.1** Elements used for interpolation: (○) displacement  $\mathbf{u}$  and (■) pressure  $p$

### 3.6 DISCRETISATION IN TIME

Finite differences in time are used for the solution of the initial value problem. The method used in this text produces an identical system of equations achieved with the weighted residual method in time with point collocation [1].

We write (3.45) or (3.46) in a more concise form, i.e.

$$\mathbf{B} \frac{d\mathbf{X}}{dt} + \mathbf{C}\mathbf{X} = \mathbf{F} \quad (3.47)$$

where the matrices  $\mathbf{B}$ ,  $\mathbf{C}$  and the vector  $\mathbf{F}$  can be immediately obtained by inspection, and  $\mathbf{X} = \{\mathbf{u}, \mathbf{p}^w\}^T$ . The matrices  $\mathbf{B}$  and  $\mathbf{C}$  are, in general, non-symmetric and depend on  $\mathbf{X}$ . The discretisation in the time domain is carried out by the generalised trapezoidal method, also known as the generalised midpoint rule. This is an implicit method, and for the approximation of

$$\left( \frac{d\mathbf{X}}{dt} \right)_{n+0} = (\mathbf{X}_{n+1} - \mathbf{X}_n)/\Delta t \quad \mathbf{X}_{n+0} = (1 - \theta)\mathbf{X}_n + \theta\mathbf{X}_{n+1} \quad (3.48)$$

then (3.47) at time  $t_{n+0}$  becomes

$$[\mathbf{B} + \theta\Delta t\mathbf{C}]\mathbf{X}_{n+1} = [\mathbf{B} - (1 - \theta)\Delta t\mathbf{C}]\mathbf{X}_n + \Delta t\mathbf{F}_{n+0} \quad (3.49)$$

where  $\Delta t$  is the time step length,  $\mathbf{X}_n$  and  $\mathbf{X}_{n+1}$  are the state vectors at times  $t_n$  and  $t_{n+1}$ , and  $\theta$  is a parameter which has limits  $0 \leq \theta \leq 1$ . The value of  $\theta$  may be obtained from the numerical properties of the scheme, as shown in the next section. The matrices in (3.49) are evaluated at time  $t_{n+0}$ .

Application of this procedure to (3.45) yields

$$\begin{bmatrix} \theta\mathbf{K}_e & -\theta\mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} + \Delta t\theta\mathbf{H} \end{bmatrix}_{n+0} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix}_{n+1} = \begin{bmatrix} (\theta - 1)\mathbf{K}_e & (1 - \theta)\mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} - (1 - \theta)\Delta t\mathbf{H} \end{bmatrix}_{n+0} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix}_n + \begin{Bmatrix} \mathbf{f}^u \\ \Delta t\mathbf{f}^p \end{Bmatrix}_{n+0} \quad (3.50)$$

And (3.46) gives

$$\begin{bmatrix} -\mathbf{K}_T & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} + \Delta t\theta\mathbf{H} \end{bmatrix}_{n+0} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix}_{n+1} = \begin{bmatrix} -\mathbf{K}_T & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} - (1 - \theta)\Delta t\mathbf{H} \end{bmatrix}_{n+0} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix}_n + \Delta t \begin{Bmatrix} -\frac{d\mathbf{f}^u}{dt} \\ \mathbf{f}^p \end{Bmatrix}_{n+0} \quad (3.51)$$

The complete set of equations (3.50), or alternatively (3.51), may be used to determine the values of  $\bar{\mathbf{u}}$  and  $\bar{\mathbf{p}}^w$  at any time relative to their initial values. It can easily be verified

that, at steady-state conditions, the coupling disappears between the equilibrium (or linear momentum balance) equation and the flow (or mass balance) equation.

### 3.7 NUMERICAL PROPERTIES OF THE TIME DISCRETISATION

The basic requirements for a numerical representation of a differential equation are that the approximate solution converges to the exact solution (consistency), and the round-off errors of the initial values do not grow during the solution process (stability). Strict definitions of these criteria can be found in the literature [11–13]. For partial differential equations of an initial value type, there is a phenomenon that has no counterpart in the solutions of ordinary differential equations, i.e. successive refinement of the discretisation may give an approximate solution that is stable, but it may converge to the solution of a different differential equation (inconsistency). One of the best-known examples is the Du Fort–Frankel scheme for a linear parabolic equation [14]. Although this is a finite difference scheme in both time and space, such a phenomenon also occurs for schemes with finite elements in space and finite differences in time [15].

We assume consistency and convergence of the finite element discretisation in space, obtained by a proper choice of elements (Section 3.5); thus we have to establish the convergence of the temporal discretisation. We deal in this section with the linear case; by following the Lax theorem [11], we verify the consistency and stability of the time approximation. If these properties are assured, the numerical algorithm fulfils the basic requirements. The more general non-linear case will be dealt with in Section 3.10.

In the linear case, the matrices  $\mathbf{B}$  and  $\mathbf{C}$  of (3.47) are independent of  $\mathbf{X}$ . But these matrices may be non-symmetric. The total error of the numerical scheme – defined by the difference between the exact solution  $\mathbf{X}(t_{n+1})$  of (3.47) and the numerical solution  $\tilde{\mathbf{X}}_{n+1}$  for time instant  $t_{n+1}$  – can be expressed as follows [16]:

$$\text{Total error} = \mathbf{e}_c + \mathbf{e}_s, \quad (3.52)$$

where  $\mathbf{e}_c = \mathbf{X}(t_{n+1}) - \tilde{\mathbf{X}}_{n+1}$  is the discretisation error of (3.49), and  $\mathbf{e}_s = \tilde{\mathbf{X}}_{n+1} - \bar{\mathbf{X}}_{n+1}$  is the stability error. If the norm of each of these errors is negligible, the total error is also negligible.

Error  $\mathbf{e}_c$  is related to the consistency of (3.49) with the differential equation in time (3.47). The proof of this consistency, also for the non-symmetric matrices  $\mathbf{B}$  and  $\mathbf{C}$  is given for the more general non-linear case in Section 3.10, where the following formula for  $\mathbf{e}_c$  is also derived:

$$\mathbf{e}_c = \frac{1}{2}(1 - 2\theta)\mathbf{O}(\Delta t^2) + \frac{1}{2}\theta^2\mathbf{O}(\Delta t^3) \quad (3.53)$$

where symbol  $\mathbf{A} = \mathbf{O}(x)$  denotes that  $\mathbf{A}$  is of order  $x$ , i.e. there exists a constant  $K$  such that  $\|\mathbf{A}\| < K|x|$ .

In order to examine error  $\mathbf{e}_s$ , we must perform a stability analysis [16]. The amplification factor  $\lambda$  must be identified such that

$$\mathbf{X}_{n+1} = \lambda \mathbf{X}_n \quad (3.54)$$

is a nontrivial solution of (3.49) for  $\mathbf{F} = \mathbf{0}$ . If the following condition is satisfied:

$$|\lambda| < 1 \quad (3.55)$$

then the system (3.49) is stable and the error  $\|\mathbf{e}\| \rightarrow 0$  as  $n \rightarrow \infty$ . We then perform a modal decomposition of (3.49), where  $\mathbf{F}$  has been set to zero. Let  $\mu_1, \mu_2, \dots, \mu_m$  be the distinct complex eigenvalues of  $\mathbf{B}^{-1}\mathbf{C}$ , where  $m$  is the rank of  $\mathbf{B}^{-1}\mathbf{C}$ . Then the homogeneous form of (3.49) decomposes into  $m$  equations:

$$(1 + \theta\Delta t\mu_j)X_{n+1(j)} = (1 - (1 - \theta)\Delta t\mu_j)X_{n(j)} \quad (3.56)$$

where  $j = 1, 2, \dots, m$ , and the  $X_{n(j)}$  denote the scalar values, obtained by premultiplying the vector  $\mathbf{X}_n$  by the eigenvectors  $\Psi_j^T$  of  $\mathbf{B}^{-1}\mathbf{C}$ :

$$X_{n(j)} = \Psi_j^T \mathbf{X}_n \quad (3.57)$$

The condition given by (3.55) can be written as

$$\left| \frac{1 - (1 - \theta)\Delta t\mu_j}{1 + \theta\Delta t\mu_j} \right| < 1 \quad (3.58)$$

for  $j = 1, \dots, m$ . By writing  $\text{Re}[\mu_j] = \mu_R$  and  $\text{Im}[\mu_j] = \mu_I$ , inequality (3.58) can be recast as

$$-2\mu_R < (2\theta - 1)(\mu_R^2 + \mu_I^2)\Delta t \quad (3.59)$$

1. Assume that  $\theta > 1/2$ . If  $\mu_R > 0$  the condition given by (3.59) is satisfied for all  $\Delta t$  and  $\mu_I$ , thus ensuring unconditional stability of equation (3.49). If  $\mu_R \leq 0$  the condition given by (3.59) leads to a conditional stability with a requirement of the form

$$\Delta t > \frac{-2}{2\theta - 1} \frac{\mu_R}{(\mu_R^2 + \mu_I^2)} \quad (3.60)$$

which results in lower bounds for the time step length  $\Delta t$ .

2. Assume that  $\theta < 1/2$ . Conditional stability is achieved for

$$\Delta t < \frac{2}{1 - 2\theta} \frac{\mu_R}{(\mu_R^2 + \mu_I^2)} \quad (3.60)$$

with an upper bound for the time step length  $\Delta t$ . Only when  $\mu_R \leq 0$  is (3.49) unstable.

### 3.8 SATURATED-UNSATURATED FLOW IN A DEFORMING POROUS MEDIUM: ONE-PHASE FLOW

In the present model it is assumed that in the partially saturated zone the air phase flows without resistance and always remains at atmospheric pressure, which is taken as the

reference pressure. In this approach, first used by Neuman for a rigid porous medium [17], it is straightforward to distinguish between the saturated and unsaturated zones: in the saturated zones we have positive pore pressures, whereas this pressure is negative in the unsaturated zone. The free surface which separates the two zones is simply obtained as an isobar, where  $p^w = 0$ . This technique is different from the classical approach, where this surface is treated as a moving boundary.

The assumption of the air phase being at atmospheric pressure in the unsaturated zone is a common one in soil mechanics. Therefore it is feasible to neglect the mass balance equation for the gaseous phase. Hence, the numerical model can be obtained by a simple modification of the matrices involved in the previous model [18]. However, due to the constitutive relations involving the degree of water saturation, relative permeability and pore water pressure, the model is always non-linear, even in the case of linear elastic solid-phase behaviour. The discrete solution of the governing equations is explained in this section and examples are given in Chapter 5.

### 3.8.1 Governing equations

The linear momentum balance equation in terms of the total stresses is

$$\mathbf{L}^T \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \quad (3.62)$$

where the density of the multiphase medium is calculated from the following expression:

$$\rho = (1 - n)\rho^s + nS_w\rho^w \quad (3.63)$$

The mass balance equation for water is obtained from (2.253) or (2.318) by considering  $\rho^g = 0$ ,  $\dot{m} = 0$  and by neglecting the gradient of water density. Furthermore, (3.9) and (3.30) are taken into account, so

$$\begin{aligned} & \left( \frac{\alpha - n}{K_s} S_w^2 + \frac{nS_w}{K_w} \right) \frac{\partial p^w}{\partial t} + \alpha S_w \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + \left( \frac{\alpha - n}{K_s} p^w S_w + n \right) \frac{\partial S_w}{\partial t} \\ & + \nabla^T \left[ \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] = 0 \end{aligned} \quad (3.64)$$

The effective stress principle in this case is given by

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}'' - \alpha \mathbf{m} S_w \rho^w \quad (3.65)$$

and as before, the constitutive relationship for the solid phase is

$$d\boldsymbol{\sigma}'' = \mathbf{D}_T d\boldsymbol{\epsilon} \quad (3.66)$$

Furthermore, we need the capillary pressure relationship, (2.155) or (2.269), which for isothermal conditions and  $\rho^g = 0$  reads

$$S_w = S_w(p^w) \quad (3.67)$$

and the relative permeability

$$k^{rw} = k^{rw}(S_w) \quad \text{or} \quad k^{rw} = k^{rw}(p^w) \quad (3.68)$$

### 3.8.2 Initial and boundary conditions

The initial conditions are almost identical to those described in Section 3.2.2. The only boundary condition which changes is the flux boundary condition

$$\rho^w \frac{\mathbf{k} k^{rw}}{\mu_w} (-\nabla p^w + \rho^w \mathbf{g})^T \cdot \mathbf{n} = q^w \quad \text{on } \Gamma_w^q \quad (3.69)$$

## 3.9 DISCRETISATION OF THE GOVERNING EQUATIONS FOR THE CONSOLIDATION OF PARTIALLY SATURATED SOILS

Standard finite element discretisation techniques, as outlined in Sections 3.3 and 3.4, when applied to the equilibrium equation (3.62) along with boundary condition (3.12), they yield

$$\int_{\Omega} (\mathbf{L} \mathbf{N}_u)^T \boldsymbol{\sigma} d\Omega = \int_{\Omega} \mathbf{N}_u^T \rho \mathbf{g} d\Omega + \int_{\Gamma_w^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \quad (3.70)$$

Introduction of the effective stress principle (3.65) results in

$$\int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega - \mathbf{Q} \bar{\mathbf{p}}^w = \mathbf{f}^u \quad (3.71)$$

where

$$\mathbf{Q} = \int_{\Omega} \mathbf{B}^T S_w \alpha \mathbf{m} \mathbf{N}_p d\Omega$$

is the coupling matrix and the load vector is

$$\mathbf{f}^u = \int_{\Omega} \mathbf{N}_u^T [\rho^s(n-1) + S_w n \rho^w] \mathbf{g} d\Omega + \int_{\Gamma_w^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \quad (3.72)$$

Before discretising (3.64), we take into account that

$$n \frac{\partial S_w}{\partial t} = n \frac{\partial S_w}{\partial p^w} \frac{\partial p^w}{\partial t} = C_s \frac{\partial p^w}{\partial t} \quad (3.73)$$

where  $C_s = \partial S_w / \partial p^w$  is the specific moisture content. Hence the mass balance equation becomes

$$\begin{aligned} & \left[ \frac{\alpha - n}{K_s} S_w \left( S_w + \frac{C_s}{n} p^w \right) + \frac{nS_w}{K_w} + C_s \right] \frac{\partial p^w}{\partial t} + \alpha S_w \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} \\ & + \nabla^T \left[ \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g})^w \right] = 0 \end{aligned} \quad (3.74)$$

The discretisation in space is carried out exactly as for the fully saturated case and yields

$$\mathbf{H} \bar{\mathbf{p}}^w + \mathbf{Q}^T \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{S} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} = \mathbf{f}^p \quad (3.75)$$

where the permeability matrix is

$$\mathbf{H} = \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rw}}{\mu^w} \nabla \mathbf{N}_p d\Omega$$

the compressibility matrix is

$$\mathbf{S} = \int_{\Omega} \mathbf{N}_p^T \left[ \frac{\alpha - n}{K_s} S_w \left( S_w + \frac{C_s}{n} p^w \right) + \frac{nS_w}{K_w} + C_s \right] N_p d\Omega$$

and the right-hand term (flow vector) is

$$\mathbf{f}^p = \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rw}}{\mu^w} \rho^w \mathbf{g} d\Omega - \int_{\Gamma_w^q} \mathbf{N}_p^T \frac{q^w}{\rho^w} d\Gamma \quad (3.76)$$

The structure of the two discretised equations, (3.71) and (3.75), is very similar to that of the fully saturated case, and the time discretisation is carried out as in Section 3.6. The resulting system of equations is formally identical to (3.51) except that the component matrices are now as defined in this section. Now time differentiation of the equilibrium equation does not restore symmetry: the two coupling matrices are no longer the same. However, symmetry can be restored by using staggered solution procedures, which will be explained in Section 6.1.1.

Furthermore, the resulting system of equations is non-linear, even for the case of linear elastic solid behaviour. In general this requires iterations within each time step. The appropriate methods which belong to the fixed-point procedures will be discussed in the next chapter. However, the numerical properties of the time discretisation require further examination.

### 3.10 STABILITY, CONVERGENCE AND CONSISTENCY IN THE NON-LINEAR CASE

We write the system of equations (3.71) and (3.75) in the form of (3.47). And we assume, as before, consistency and convergence of the finite element discretisation in

space. Using a one-step finite difference operator for the time derivative, as given in Section 3.6, we obtain the discrete equation (3.49), which can be written more concisely as

$$\mathbf{A}X_{n+1} - \mathbf{B}X_n - \mathbf{F} = 0 \quad (3.77)$$

where, for the generalised midpoint method, we have

$$\begin{aligned}\mathbf{A} &= \mathbf{B} + \theta\Delta t \mathbf{C} \\ \mathbf{B} &= \mathbf{B} - (1-\theta)\Delta t \mathbf{C} \\ \mathbf{F} &= \mathbf{F}\Delta t\end{aligned} \quad (3.78)$$

and  $\theta$  is a parameter usually in the range [0,1]. Since the problem is non-linear,  $\mathbf{A}$  and  $\mathbf{B}$  depend on  $X$ .

The value of  $X_{n+1}$  is found directly from (3.77) using a fixed-point procedure, i.e.

$$X_{n+1,k+1} = f(X_{n+1,k}) \quad (3.79)$$

where  $k$  is the iteration count. The number of iterations performed per time step is  $K(n)$ .

The total error is defined as

$$\mathbf{e}_{n+1} = X(t_{n+1}) - \tilde{X}_{n+1} \quad (3.80)$$

where  $X(t_{n+1})$  is the exact solution of (3.47) and  $\tilde{X}_{n+1}$  its numerical solution. To simplify calculations, we assume there are no round-off errors generated by the algorithm, except those introduced by the initial values at time  $t_0$ .

If we replace  $X_{n+1}$  and  $X_n$  in (3.77) with the corresponding exact values, we obtain

$$\mathbf{A}X(t_{n+1}) = \mathbf{B}X(t_n) + \mathbf{F} + \mathbf{r}_{n+1} \quad (3.81)$$

where  $\mathbf{r}_{n+1}$  is the local truncation error of (3.77). The calculated numerical solution  $\tilde{X}_{n+1}$  satisfies (3.77). Therefore, we can subtract (3.77) written for  $\tilde{X}_{n+1}$  and  $\tilde{X}_n$  from (3.81). This results in

$$\mathbf{A}\mathbf{e}_{n+1} = \mathbf{B}\mathbf{e}_n + \mathbf{r}_{n+1} \quad (3.82)$$

Neither symmetry nor positive definiteness of the matrices is assumed here. By a recursive application for the index  $n$ , we have from (3.82) that

$$\begin{aligned}\mathbf{e}_{n+1} &= (\mathbf{A}^{-1}\mathbf{B})_{n+1} \dots (\mathbf{A}^{-1}\mathbf{B})_1 \mathbf{e}_{0,0} + \sum_{l=0}^{n-1} (\mathbf{A}^{-1}\mathbf{B})_{n+1-l} \dots \\ &\quad \dots (\mathbf{A}^{-1}\mathbf{B})_{n+1-l} (\mathbf{A}^{-1})_{n-l} \mathbf{r}_{n-l} + (\mathbf{A}^{-1})_{n+1} \mathbf{r}_{n+1}\end{aligned} \quad (3.83)$$

The total error  $\mathbf{e}_{n+1}$  is influenced by the local truncation error  $\mathbf{r}_l$ , where  $l \leq n+1$ , and also by the initial round-off error  $\mathbf{e}_{0,0}$  made at the initial step, because for  $n=0$  and

$K = 0$ , we have

$$\mathbf{e}_{0,0} = \mathbf{X}(t_0) - \tilde{\mathbf{X}}_{0,0} \quad (3.84)$$

with  $\mathbf{X}(t_0) = \mathbf{X}_0$  being the exact solution of (3.77). Thus

$$\|\mathbf{A}^{-1}\mathbf{B}\| < 1 \quad \forall n \quad (3.85)$$

is a sufficient stability condition for the procedure, where  $\|\cdot\|$  is the spectral norm. Consistency is linked to the residual  $\mathbf{r}_l$  and will be dealt with later.

This relationship for the global error requires that within each time step the iterative procedure converges to the exact solution ( $K(n) \rightarrow \infty$ ). If we include the iteration convergence for each time step, the expression for the global error becomes more complicated. Such a relationship, which uses a staggered procedure (Section 6.1.1) as a solution method for non-linear equations, is given in Turska *et al.* [19]. Banach's contraction mapping principle [20,21] may be used for the convergence of (3.79) to a unique solution within one time step. This states that a sufficient condition for the existence of a unique solution is that

$$\|\mathbf{f}'(\mathbf{X})\| < 1 \quad (3.86)$$

where  $\mathbf{f}'(\mathbf{X})$  is the Jacobian matrix of  $\mathbf{f}(\mathbf{X})$ . It is assumed that such a Jacobian matrix can be established [16].

We deal now with the questions of consistency for the non-linear case. For the sake of simplicity, dots will be used to indicate the time derivatives. The temporally discrete model (3.49) may be written in the form

$$\mathbf{X}_{n+1} - [\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}[\mathbf{B} - (1-\theta)\Delta t\mathbf{C}]\mathbf{X}_n - [\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}\mathbf{F}_{n+0}\Delta t = \mathbf{0} \quad (3.87)$$

When we replace  $\mathbf{X}_{n+1}$  and  $\mathbf{X}_n$  by the corresponding exact values, an expression of the following form is obtained:

$$\mathbf{X}(t_{n+1}) - [\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}[\mathbf{B} - (1-\theta)\Delta t\mathbf{C}]\mathbf{X}(t_n) - [\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}\mathbf{F}_{n+0}\Delta t = \mathbf{r}_{n+1} \quad (3.88)$$

where  $\mathbf{r}_{n+1}$  is the local truncation error. If  $\|\mathbf{r}_{n+1}\| \rightarrow 0$  for  $\Delta t \rightarrow 0$ , then the algorithm defined by (3.49) is called consistent. The calculations are further simplified by using the relationships

$$[\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}[\mathbf{B} - (1-\theta)\Delta t\mathbf{C}] = \mathbf{I} - \Delta t[\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}\mathbf{C} \quad (3.89)$$

$$\mathbf{I} - \theta\Delta t[\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}\mathbf{C} = [\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}\mathbf{B} \quad (3.90)$$

We now expand  $\mathbf{X}(t_{n+1})$  and  $\mathbf{X}(t_n)$  in a Taylor series about  $t_{n+0}$ :

$$\begin{aligned} \mathbf{X}(t_{n+1}) &= \mathbf{X}(t_{n+0}) + (1-\theta)\Delta t\dot{\mathbf{X}}(t_{n+0}) + \frac{1}{2}(1-\theta)^2\Delta t^2\ddot{\mathbf{X}}(t_{n+0}) + \mathcal{O}(\Delta t^3) \\ \mathbf{X}(t_n) &= \mathbf{X}(t_{n+0}) - \theta\Delta t\dot{\mathbf{X}}(t_{n+0}) + \frac{\theta^2\Delta t^2}{2}\ddot{\mathbf{X}}(t_{n+0}) + \mathcal{O}(\Delta t^3) \end{aligned} \quad (3.91)$$

And we substitute these expansions into (3.89):

$$\begin{aligned} & \mathbf{X}(t_{n+0}) + (1-\theta)\Delta t \dot{\mathbf{X}}(t_{n+0}) + \frac{1}{2}(1-\theta)^2 \Delta t^2 \ddot{\mathbf{X}}(t_{n+0}) + \mathbf{O}(\Delta t^3) \\ & - [\mathbf{B} + \theta \Delta t \mathbf{C}]^{-1} [\mathbf{B} - (1-\theta)\Delta t \mathbf{C}] \{ \mathbf{X}(t_{n+0}) - \theta \Delta t \dot{\mathbf{X}}(t_{n+0}) \\ & + \frac{1}{2}\theta^2 \Delta t^2 \ddot{\mathbf{X}}(t_{n+0}) + \mathbf{O}(\Delta t^3) \} - [\mathbf{B} + \theta \Delta t \mathbf{C}]^{-1} \mathbf{F}_{n+0} \Delta t = \mathbf{r}_{n+1} \end{aligned} \quad (3.92)$$

Employing (3.90), (3.91) and (3.47), written for  $t_{n+0}$ ,

$$\mathbf{B} \dot{\mathbf{X}}(t_{n+0}) + \mathbf{C} \mathbf{X}(t_{n+0}) - \mathbf{F}_{n+0} = 0 \quad (3.93)$$

we obtain

$$\frac{1}{2}(1-2\theta)\ddot{\mathbf{X}}(t_{n+0})\Delta t^2 + \frac{1}{2}\theta^2[\mathbf{B} + \theta \Delta t \mathbf{C}]^{-1} \mathbf{C} \ddot{\mathbf{X}}(t_{n+0})\Delta t^3 + \mathbf{O}(\Delta t^4) = \mathbf{r}_{n+1} \quad (3.94)$$

As may be seen,  $\|\mathbf{r}_{n+1}\| \rightarrow 0$  for  $\Delta t \rightarrow 0$ . And for

$$\mathbf{X}(t_0) = \mathbf{X}_0 \quad (3.95)$$

i.e. no round-off errors, the discretisation error is defined as follows, using (3.52):

$$\mathbf{e}_{n+1} = \mathbf{X}(t_{n+1}) - \mathbf{X}_{n+1} \quad (3.96)$$

Subtracting (3.88) from (3.87), we obtain

$$\mathbf{e}_{n+1} = [\mathbf{B} + \theta \Delta t \mathbf{C}]^{-1} [\mathbf{B} - (1-\theta)\Delta t \mathbf{C}] \mathbf{e}_n + \mathbf{r}_{n+1} \quad (3.97)$$

or, more concisely,

$$\mathbf{e}_{n+1} = \mathbf{A}_n \mathbf{e}_n + \mathbf{r}_{n+1} \quad (3.98)$$

where

$$\mathbf{A}_n = (\mathbf{B}_{n+0} + \theta \Delta t \mathbf{C}_{n+0})^{-1} [\mathbf{B}_{n+0} - (1-\theta)\Delta t \mathbf{C}_{n+0}] \quad (3.99)$$

A recursive evaluation of the errors, as given by (3.98), yields

$$\begin{aligned} \mathbf{e}_{n+1} &= (\mathbf{A}_n \cdot \mathbf{A}_{n-1} \cdot \dots \cdot \mathbf{A}_0) \mathbf{e}_0 + \sum_{l=1}^n (\mathbf{A}_n \cdot \mathbf{A}_{n-1} \cdot \dots \cdot \mathbf{A}_{n+1-l}) \mathbf{r}_{n+1-l} + \mathbf{r}_{n+1} \\ & \quad (3.100) \end{aligned}$$

For initial values without round-off error  $\mathbf{e}_0$ , we have

$$\mathbf{e}_{n+1} = \sum_{l=1}^n (\mathbf{A}_n \cdot \mathbf{A}_{n-1} \cdot \dots \cdot \mathbf{A}_{n+1-l}) \mathbf{r}_{n+1-l} + \mathbf{r}_{n+1} \quad (3.101)$$

Equation (3.94) is now written as

$$\mathbf{r}_{n+1} = \frac{1}{2}(1 - 2\theta)\mathbf{O}_{n+1}(\Delta t^2) + \frac{1}{2}\theta^2[\mathbf{B} + \theta\Delta t\mathbf{C}]^{-1}\mathbf{CO}_{n+1}(\Delta t^3) + \mathbf{O}_{n+1}(\Delta t^4) \quad (3.102)$$

and is introduced in (3.101), together with (3.99). Since we are interested only in the terms with the lowest power of  $\Delta t$ , we obtain

$$\|\mathbf{e}_{n+1}\| = \frac{1}{2}(1 - 2\theta)\mathbf{O}(\Delta t^2) + \frac{1}{2}\theta^2\mathbf{O}(\Delta t^3) \quad (3.103)$$

Hence  $\|\mathbf{e}_{n+1}\| \rightarrow 0$  if  $\Delta t \rightarrow 0$ , where  $\mathbf{e}_{n+1}$  is easily identified as the consistency error  $\mathbf{e}_c$ .

The stability condition given by (3.85) and the consistency property

$$\mathbf{r}_I = \mathbf{O}(\Delta t^2) \quad (3.104)$$

are sufficient for the discrete equation (3.77) to be globally convergent if the iteration convergence within each time step is satisfied [19]. Condition (3.86) is only sufficient for the existence of a unique solution of (3.79). If this condition is not satisfied, starting with a very specific value of,  $X_0$  one may accidentally hit upon the fixed point at some iteration step. Such cases are rare and should not be taken into account in numerical procedures constructed for general boundary conditions. Thus for each time step, the condition given by (3.86) should be verified. In (3.77) the most recent time value of  $X_n$  appears as a factor, and one wrongly calculated value can propagate the error on all subsequent values of  $X_{n+k}$ ,  $k = 1, 2, \dots$ .

### 3.11 AIRFLOW AND WATER FLOW IN A DEFORMING POROUS MEDIUM

We deal now with a problem which has three degrees of freedom. This type of problem is important when one fluid phase is used to displace another fluid phase from the pores of a deforming porous body. Such a situation may happen when an aquifer is used as a storage for air or gas, or for a tunnel or shaft construction below the water table, where seepage towards the opening is prevented by the use of compressed air. Two-phase flow involving oil and water was solved by Li *et al.* [22], Morgan *et al.* [23] and others have solved air and water flow [24]. Again, the governing equations and their discrete solution are given here, and examples will be presented in Chapter 5.

#### 3.11.1 Governing equations

The linear momentum balance equation in terms of total stresses is

$$\mathbf{L}^T \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \quad (3.105)$$

with the density given by

$$\rho = (1 - n)\rho^s + nS_w\rho^w + nS_g\rho^g \quad (3.106)$$

The mass balance equation for water is obtained from (2.253) or (2.318) by neglecting any phase change ( $\dot{m} = 0$ ) and the gradient of the water density:

$$\left( \frac{\alpha - n}{K_s} S_w^2 + \frac{nS_w}{K_w} \right) \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_w S_g \frac{\partial p^g}{\partial t} + \alpha S_w \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + \left( \frac{\alpha - n}{K_s} p^w S_w - \frac{\alpha - n}{K_s} p^g S_w + n \right) \frac{\partial S_w}{\partial t} + \nabla^T \left[ \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] = 0 \quad (3.107)$$

For gas, considered here as a mixture of dry air and vapour, equation (2.256) or (2.323) is used with a zero phase change term and negligible gradient of gas density:

$$\frac{\alpha - n}{K_s} S_w S_g \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_g^2 \frac{\partial p^g}{\partial t} - \left[ n + \frac{\alpha - n}{K_s} S_g (p^g - p^w) \right] \frac{\partial S_w}{\partial t} + \alpha S_g \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + \frac{n S_g}{\rho^g} \frac{\partial}{\partial t} \left( \frac{p^g M_g}{\hat{\theta} R} \right) + \nabla^T \left[ \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] = 0 \quad (3.108)$$

where  $\hat{\theta}$  indicates the absolute temperature.

The effective stress principle used is given by (2.278):

$$\boldsymbol{\sigma}'' = \boldsymbol{\sigma} + \mathbf{m}^T \alpha (S_w p^w + S_g p^g) \quad (3.109)$$

and the constitutive relationship for the solid phase is

$$d\boldsymbol{\sigma}'' = \mathbf{D}_T d\boldsymbol{\epsilon} \quad (3.110)$$

The capillary pressure relationship now assumes the form of (2.155) or (2.269), written for isothermal conditions:

$$S_w = S_w(p^c) \quad \text{with} \quad p^c = p^g - p^w \quad (3.111)$$

and the relative permeabilities are functions of the respective degrees of saturation.

### 3.11.2 Initial and boundary conditions

The initial conditions specify the full field of displacements, water pressure and air (gas) pressure at time  $t = 0$ , i.e.

$$\mathbf{u} = \mathbf{u}_0, p^w = p_0^w, p^g = p_0^g \quad \text{in } \Omega \text{ and on } \Gamma \quad (3.112)$$

The boundary conditions are

$$\begin{aligned} \mathbf{u} &= \hat{\mathbf{u}} && \text{on } \Gamma_u \\ p^w &= \hat{p}^w && \text{on } \Gamma_w \\ p^g &= \hat{p}^g && \text{on } \Gamma_g \end{aligned} \quad (3.113)$$

The traction boundary conditions for the stresses are unchanged:

$$\mathbf{I}^T \boldsymbol{\sigma} = \bar{\mathbf{t}} \quad \text{on } \Gamma_u^q \quad (3.114)$$

and the flux boundary conditions are

$$\rho_w \frac{\mathbf{k} k^{rw}}{\mu_w} (-\nabla p^w + \rho_w \mathbf{g})^T \cdot \mathbf{n} = q^w \quad \text{on } \Gamma_w^q \quad (3.115)$$

for the water phase, and

$$\rho_g \frac{\mathbf{k} k^{rg}}{\mu_g} (-\nabla p^g + \rho_g \mathbf{g})^T \cdot \mathbf{n} = q^g \quad \text{on } \Gamma_g^q \quad (3.116)$$

for the gaseous phase.

### 3.12 DISCRETISATION OF THE GOVERNING EQUATIONS FOR AIR AND WATER FLOW IN DEFORMING POROUS MEDIA

The basic variables are, as before, displacements, the water pressure  $p^w$  and the gas pressure  $p^g$ . A different choice for the pressure variables will be shown in Chapter 10, where non-isothermal conditions and phase change will be taken into account. Displacement and pressure values are expressed in terms of their nodal values as

$$\begin{aligned} \mathbf{u} &= \mathbf{N}_u \bar{\mathbf{u}} \\ p^w &= \mathbf{N}_p \bar{\mathbf{p}}^w \\ p^g &= \mathbf{N}_p \bar{\mathbf{p}}^g \end{aligned} \quad (3.117)$$

The application of standard finite element discretisation procedures, after introduction of the effective stress principle (3.109), results in the following expression for the equilibrium equation:

$$\int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega - \mathbf{C}_{sw} \bar{\mathbf{p}}^w - \mathbf{C}_{sg} \bar{\mathbf{p}}^g = \mathbf{f}'' \quad (3.118)$$

where

$$\mathbf{C}_{sw} = \mathbf{Q} = \int_{\Omega} \mathbf{B}^T S_w \alpha \mathbf{m} \mathbf{N}^p d\Omega \quad \text{and} \quad \mathbf{C}_{sg} = \int_{\Omega} \mathbf{B}^T S_g \alpha \mathbf{m} \mathbf{N}^p d\Omega \quad (3.119)$$

are the coupling matrices, and the load vector is

$$\mathbf{f}'' = \int_{\Omega} \mathbf{N}_u^T [\rho^s(n-1) + S_w n \rho^w + S_g n \rho^g] \mathbf{g} d\Omega + \int_{\Gamma_u^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \quad (3.120)$$

For the mass balance equations, we take into account that

$$n \frac{\partial S_w}{\partial t} = n \frac{\partial S_w}{\partial p^c} \frac{\partial p^c}{\partial t} = C_s \left( \frac{\partial p^g}{\partial t} - \frac{\partial p^w}{\partial t} \right) \quad (3.121)$$

where  $C_s = \partial S_w / \partial p^c$  is the specific moisture content.

Hence (3.107) for water becomes

$$\begin{aligned} & \left[ \frac{\alpha - n}{K_s} S_w \left( S_w - p^w \frac{C_s}{n} + p^g \frac{C_s}{n} \right) + \frac{nS_w}{K_w} - C_s \right] \frac{\partial p^w}{\partial t} \\ & + \left[ \frac{\alpha - n}{K_s} S_w \left( S_g + p^w \frac{C_s}{n} - p^g \frac{C_s}{n} \right) + C_s \right] \frac{\partial p^g}{\partial t} + \alpha S_w \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} \\ & + \nabla^T \left[ \frac{\mathbf{k} k^{rg}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] = 0 \end{aligned} \quad (3.122)$$

whereas (3.108) reads

$$\begin{aligned} & \left[ \frac{\alpha - n}{K_s} S_g \left( S_g - p^g \frac{C_s}{n} + p^w \frac{C_s}{n} \right) + \frac{nS_g}{\rho^g} \frac{M_g}{\hat{\theta} R} - C_s \right] \frac{\partial p^g}{\partial t} \\ & + \left[ \frac{\alpha - n}{K_s} S_g \left( S_w + p^g \frac{C_s}{n} - p^w \frac{C_s}{n} \right) + C_s \right] \frac{\partial p^w}{\partial t} + \alpha S_g \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} \\ & + \nabla^T \left[ \frac{\mathbf{k} k^{rw}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] = 0 \end{aligned} \quad (3.123)$$

The discretisation in space of (3.122) and (3.123) and their boundary conditions, (3.115) and (3.116), is carried out exactly as for the mass balance equation in the fully saturated state, i.e. the weighted residual statement is again applied to these equations and then Green's theorem is applied to the terms of the divergence operator. Furthermore, the weighting functions are chosen according to (3.31) and (3.32), and the interpolations (3.117) are then applied with  $\mathbf{N}_p$  being the final weighting function. For the mass balance equation of water, this yields

$$\mathbf{C}_{sw}^T \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{P}_{ww} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \mathbf{C}_{wg} \frac{\partial \bar{\mathbf{p}}^g}{\partial t} + \mathbf{H}_{ww} \bar{\mathbf{p}}^w = \mathbf{f}^w \quad (3.124)$$

and for the mass balance equation of gas

$$\mathbf{C}_{sg}^T \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{C}_{gw} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \mathbf{P}_{gg} \frac{\partial \bar{\mathbf{p}}^g}{\partial t} + \mathbf{H}_{gg} \bar{\mathbf{p}}^g = \mathbf{f}^g \quad (3.125)$$

where

$$\begin{aligned} \mathbf{P}_{ww} &= \int_{\Omega} \mathbf{N}_p^T \left[ \frac{\alpha - n}{K_s} S_w \left( S_w - p^w \frac{C_s}{n} + p^g \frac{C_s}{n} \right) + \frac{nS_w}{K_w} - C_s \right] \mathbf{N}_p d\Omega \\ \mathbf{C}_{wg} &= \int_{\Omega} \mathbf{N}_p^T \left[ \frac{\alpha - n}{K_s} S_w \left( S_g + p^w \frac{C_s}{n} - p^g \frac{C_s}{n} \right) + C_s \right] \mathbf{N}_p d\Omega \end{aligned}$$

$$\begin{aligned}
\mathbf{H}_{ww} &= \mathbf{H} = \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rw}}{\mu^w} \nabla \mathbf{N}_p d\Omega \\
\mathbf{C}_{gw} &= \int_{\Omega} \mathbf{N}_p^T \left[ \frac{\alpha - n}{K_s} S_g \left( S_w + p^g \frac{C_s}{n} - p^w \frac{C_s}{n} \right) + C_s \right] \mathbf{N}_p d\Omega \\
\mathbf{P}_{gg} &= \int_{\Omega} \mathbf{N}_p^T \left[ \frac{\alpha - n}{K_s} S_g \left( S_g - p^g \frac{C_s}{n} + p^w \frac{C_s}{n} \right) + \frac{n S_g}{\rho^g} \frac{M_g}{\theta R} - C_s \right] \mathbf{N}_p d\Omega \\
\mathbf{H}_{gg} &= \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rg}}{\mu^g} \nabla \mathbf{N}_p d\Omega \\
\mathbf{f}^w &= \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rw}}{\mu^w} \rho^w \mathbf{g} - \int_{\Gamma_w^q} \mathbf{N}_p^T \frac{q^w}{\rho^w} d\Gamma \\
\mathbf{f}^g &= \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rg}}{\mu^g} \rho^g \mathbf{g} - \int_{\Gamma_g^q} \mathbf{N}_p^T \frac{q^g}{\rho^g} d\Gamma
\end{aligned} \tag{3.126}$$

For the discretisation in time of the three coupled equations (3.118), (3.124) and (3.125), the single-step finite difference operator (3.49) is applied, where matrices  $\mathbf{B}$ ,  $\mathbf{C}$ , and vectors  $\mathbf{X}$ ,  $\mathbf{F}$  are defined as follows:

$$\begin{aligned}
\mathbf{B} &= \begin{bmatrix} \mathbf{K}_T & -\mathbf{C}_{sw} & -\mathbf{C}_{sg} \\ \mathbf{C}_{sw}^T & \mathbf{P}_{ww} & \mathbf{C}_{wg} \\ \mathbf{C}_{sg}^T & \mathbf{C}_{gw} & \mathbf{P}_{gg} \end{bmatrix} \\
\mathbf{C} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{ww} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_{gg} \end{bmatrix} \\
\mathbf{X} &= \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \\ \bar{\mathbf{p}}^g \end{Bmatrix} \quad \mathbf{F} = \begin{Bmatrix} \frac{\partial \mathbf{f}^u}{\partial t} \\ \mathbf{f}^w \\ \mathbf{f}^g \end{Bmatrix}
\end{aligned} \tag{3.127}$$

where  $\mathbf{K}_T$  is the tangential stiffness matrix as defined by equation (3.44).

## REFERENCES

1. Zienkiewicz, O. C. and Taylor, R. L. (1989) *The Finite Element Method*, Vol. 1, McGraw-Hill, London.
2. Galerkin, B. G. (1915) Series solution of some problems of elastic equilibrium of rods and plates (in Russian). *Vestn. Inzh. Tech.*, **19**, 897–908.
3. Zienkiewicz, O. C. and Taylor, R. L. (1991) *The Finite Element Method*, Vol. 2, McGraw-Hill, London.
4. Lewis, R. W. and Schrefler, B. A. (1978) A fully coupled consolidation model of the subsidence of Venice. *Water Resources Research*, **14**(2), 223–30.
5. Norris, V. A. (1980) The elastoplastic analysis of soil consolidation with special reference to kinematic hardening. PhD thesis, University College of Swansea.

6. Zienkiewicz, O. C., Chan, A. H. C., Pastor, M., Paul, D. K. and Shiomi, T. (1990) Static and dynamic behaviour of soils: a rational approach to quantitative solutions. I. Fully saturated problems. *Proc. R. Soc. Lond. A* **429**, 285–309.
7. Babuska, I. (1971) Error bounds for finite element methods. *Num. Math.*, **16**, 322–33.
8. Babuska, I. (1973) The finite element method with Lagrange multipliers. *Num. Math.*, **20**, 179–92.
9. Brezzi, F. (1974) On the existence, uniqueness and approximations of saddle point problems arising from Lagrange multipliers. *RAIRO Mathematical Modelling and Numerical Analysis*, **8**, 129–51.
10. Zienkiewicz, O. C. (1984) Coupled problems and their numerical solution, in *Numerical Methods in Coupled Systems*, R. W. Lewis, P. Bettess and E. Hinton (eds), Wiley, Chichester, pp. 35–38.
11. Richtmyer, R. D. and Morton, K. W. (1967) *Difference Methods for Initial-Value Problems*, Interscience, New York.
12. Ortega, J. M. and Rheinboldt, W. C. (1970) *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, San Diego CA.
13. Morris, J. L. (1983) *Computational Methods in Elementary Numerical Analysis*, Wiley, Chichester.
14. Allen, M. B. III, Herrera, I. and Pinder, G. F. (1988) *Numerical Modelling in Science and Engineering*, Wiley, New York.
15. Farhat, C. and Sobh, N. (1990) A consistency analysis of a class of concurrent transient implicit/explicit algorithms. *Comput. Meth. Appl. Mech. Engng*, **84**, 147–62.
16. Turska, E. and Schrefler, B. A. (1993) On convergence conditions of partitioned solution procedures for consolidation problems. *Comput. Meth. Appl. Mech. Engng*, **106**, 51–63.
17. Neuman, S. P. (1975) Galerkin approach to saturated–unsaturated flow in porous media, in *FE in Fluids*, Vol. 1, R. H. Gallagher, J. T. Oden, C. Taylor and O. C. Zienkiewicz (eds), Wiley, London, pp. 201–17.
18. Schrefler, B. A. and Simoni, L. (1988) A unified approach to the analysis of saturated–unsaturated elastoplastic porous media, in *Numerical Methods in Geomechanics*, G. Swoboda (ed.), Balkema, Rotterdam, pp. 201–12.
19. Turska, E., Wisniewski, K. and Schrefler, B. A. (1994) Error propagation of staggered solution procedures for transient problems. *Comput. Meth. Appl. Mech. Engng*, **114**, 177–88.
20. Kurpel, N. S. (1976) *Projection-Iterative Methods for Solution of Operator Equations*, AMS, Providence, RI.
21. Rall, L. B. (1969) *Computational Solution of Nonlinear Operator Equations*, Wiley, New York.
22. Li, X., Zienkiewicz, O. C. and Xie, Y. M. (1990) A numerical model for immiscible two-phase fluid flow in a porous medium and its time domain solution. *Int. J. Num. Meth. Engng*, **30**, 1195–1212.
23. Morgan, K., Lewis, R. W. and White, I. R. (1980) The mechanisms of ground surface subsidence above compacting multiphase reservoirs and their analysis by the finite element method. *Appl. Math. Model.*, **4**, 215–24.
24. Schrefler, B. A. and Zhan, X. Y. (1993) A fully coupled model for water flow and airflow in deformable porous media. *Water Resources Research*, **29**, 155–67.

---

# 4

## Solid-Phase Constitutive Relationships, Variable Permeabilities and Solution Procedures

### 4.1 INTRODUCTION

A major problem in applying finite element analysis to geotechnical, or reservoir engineering, problems is to provide a realistic representation of the stress–strain characteristics for the porous medium. The choice of an appropriate constitutive relationship may have a significant influence on the numerical results obtained. An important consideration is the time factor involved. If all the loads are applied instantaneously to a soil mass, the resulting deformations will be associated with consolidation and with creep. In this context creep is understood to refer to the deformation which occurs when the effective stresses remain constant.

In this chapter, and in later examples, attention is focused on materials which do not creep. In this case the term *time dependence* relates purely to the consolidation process which has already been dealt with in Chapter 3.

When necessary, creep can be taken into account either by using an elastoviscoplastic constitutive relationship [1], or by means of a creep term. This is dealt with more fully in Chapter 11.

The constitutive relationship may be expressed as a linear relation between small changes in strain,  $\delta\epsilon$  and small changes in effective stress  $\delta\sigma'$ :

$$\delta\epsilon = \mathbf{D}\delta\sigma' \quad (4.1)$$

The matrix  $\mathbf{D}$  is usually symmetric, having a maximum of 21 independent components. These components will, in general, depend on the accumulated strain  $\epsilon$  and effective stress  $\sigma'$ . If the material is elastic and isotropic, the number of independent constants reduces to 2 to match 21.

The possible combinations of soil models are divided into two groups: those with a constitutive law based on plasticity and those without. The latter group contains linear elasticity and variable elasticity. A key distinction between the plasticity and elasticity

groups is that in elasticity groups the strains are recoverable upon decrease in stress, whereas in elastoplastic models the strains are only partly recoverable. Elastoplastic models can therefore be used to simulate all those situations where, on unloading a large increase in stiffness occurs. For instance in the soil mechanics world this can be observed in loading–unloading cycles, or in surface rebound during subsidence.

A crucial point in determining the choice of a suitable soil constitutive model is the ease with which values can be assigned to the constants defining it. It may be possible to reproduce measured behaviour accurately using a model defined by many constants. However, this is of little practical use if the determination of these constants is in itself a research project [2]. This is of particular importance in subsidence problems, where long time histories are involved, requiring expensive trial runs for the assessment of each parameter [3]. An important objective is therefore to minimise the number of constants involved in the choice of a constitutive law.

The state of stress in soil can be considered to consist of two components: deviatoric and hydrostatic stress. This means it is more convenient to work in terms of stress invariants when dealing with isotropic models. These are defined in the next section.

## 4.2 STRESS INVARIANTS

Stress invariants are stress combinations whose values are independent of the orientation of spatial reference axes. The principal stresses  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$ , acting at right angles to each other, are themselves invariants.

The stress invariants used in the following are the mean stress  $p$ , the deviator stress  $q$  and the angular stress invariant  $\theta$ . The definitions apply equally to effective stresses and total stresses:

$$p = -\frac{1}{3}(\sigma_x + \sigma_y + \sigma_z) \quad (4.2)$$

$$\begin{aligned} q^2 &= \sigma_x(\sigma_x - \sigma_y) + \sigma_y(\sigma_y - \sigma_z) + \sigma_z(\sigma_z - \sigma_x) \\ &\quad + 3(\tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2) \end{aligned} \quad (4.3)$$

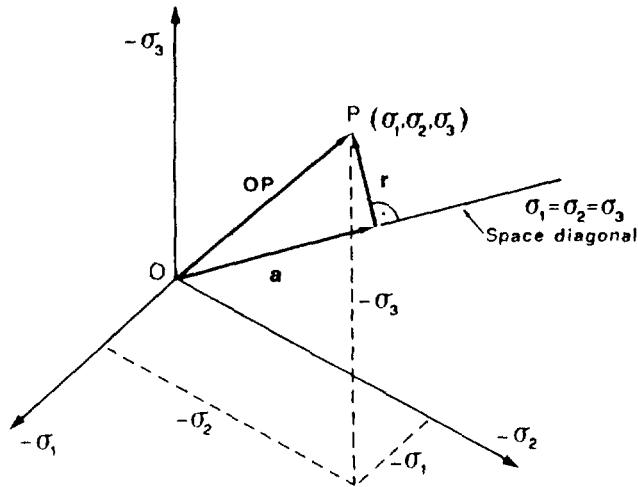
$$\theta = \frac{1}{3} \sin^{-1} \left[ -\frac{27J_3}{2q^3} \right] \quad (4.4)$$

where

$$J_3 = \begin{vmatrix} \sigma_x + p & \tau_{xy} & \tau_{zx} \\ \tau_{xy} & \sigma_y + p & \tau_{yz} \\ \tau_{zx} & \tau_{yz} & \sigma_z + p \end{vmatrix} \quad (4.5)$$

The third stress invariant  $\theta$  was introduced by Nayak and Zienkiewicz [4]. Using this angular representation of the third stress invariant leads to the following expression of the three principal stresses:

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} p \\ p \\ p \end{pmatrix} - \frac{2}{3}q \begin{pmatrix} \sin(\theta_0 - \frac{2}{3}\pi) \\ \sin\theta_0 \\ \sin(\theta_0 + \frac{2}{3}\pi) \end{pmatrix} \quad (4.6)$$



**Figure 4.1** Stress point in the principal stress space

where  $\theta_0$  is a part of the general solution to equation (4.4):

$$\theta = \theta_0 + \frac{2}{3}\pi n \quad (4.7)$$

$n$  = any integer, which lies within the limits  $-\pi/6 \leq \theta_0 \leq \pi/6$ .

Inversely,

$$p = -\frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3) \quad (4.8)$$

$$q^2 = \sigma_1(\sigma_1 - \sigma_2) + \sigma_2(\sigma_2 - \sigma_3) + \sigma_3(\sigma_3 - \sigma_1) \quad (4.9)$$

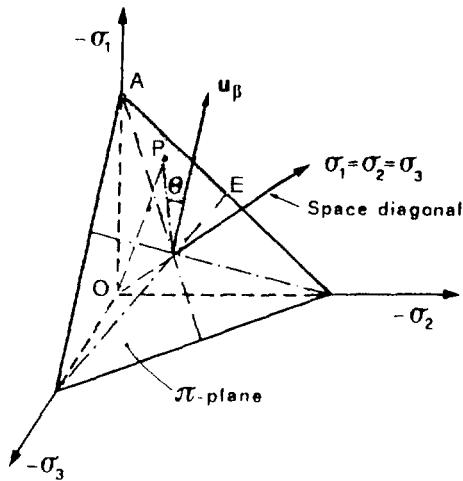
$$\theta_0 = \tan^{-1} \left[ \frac{(\sigma_1 - \sigma_2) - (\sigma_2 - \sigma_3)}{(\sigma_1 - \sigma_3)\sqrt{3}} \right] \quad -\frac{\pi}{6} \leq \theta_0 \leq \frac{\pi}{6} \quad (4.10)$$

In a three-dimensional stress space with  $\sigma_1$ ,  $\sigma_2$ ,  $\sigma_3$ , as orthogonal reference axes, a state of hydrostatic stress, given by  $\sigma_1 = \sigma_2 = \sigma_3$ , will be represented by a line passing through the origin equally inclined to each of the axes. This line is known as the space diagonal and its direction cosines are  $1/\sqrt{3}$ ,  $1/\sqrt{3}$ ,  $1/\sqrt{3}$ .

Let point P (Figure 4.1) represent an arbitrary stress state. With reference to the base axes, the vector  $\mathbf{OP}$  has components  $(-\sigma_1, -\sigma_2, -\sigma_3)$ . This stress vector can be decomposed into two components: component  $\mathbf{a}$  along the space diagonal, and component  $\mathbf{r}$  which lies in the  $\pi$ -plane and is perpendicular to the space diagonal (Figure 4.2).

The projection of  $\mathbf{OP}$  on the space diagonal gives the component  $\mathbf{a}$  as

$$a = \frac{1}{\sqrt{3}}(-\sigma_1 - \sigma_2 - \sigma_3) = p\sqrt{3} \quad (4.11)$$



**Figure 4.2** Hydrostatic and deviatoric components of the stress vector in the  $\pi$ -plane;  $\mathbf{u}_\beta$  lies in the  $\pi$ -plane

The component  $\mathbf{r}$  is given by

$$\begin{aligned}
 r^2 &= (OP)^2 - a^2 \\
 &= (\sigma_1^2 + \sigma_2^2 + \sigma_3^2) - 3p^2 \\
 &= (\sigma_1 + p)^2 + (\sigma_2 + p)^2 + (\sigma_3 + p)^2 \\
 &= \frac{3}{2}q^2
 \end{aligned} \tag{4.12}$$

Therefore, in the principal stress space, a state of stress can be decomposed into a hydrostatic component  $\mathbf{a}$  ( $\equiv$  multiple of the first stress invariant) lying along the space diagonal and a deviatoric component  $\mathbf{r}$  ( $\equiv$  multiple of the second stress invariant) lying in the  $\pi$ -plane. The deviatoric component  $q\sqrt{3/2}$  is the distance of the stress point  $P$  from the space diagonal. Because of the assumption of isotropy, stress functions, such as yield surfaces, are completely defined if taken over the sector  $60^\circ$  of  $\pi$  where  $\sigma_1 > \sigma_2 > \sigma_3$  [5].

It can easily be shown that  $\theta_0$  is the orientation of  $\mathbf{r}$  measured in the  $\pi$ -plane anti-clockwise from the bisector of  $APE$ , indicated as  $\mathbf{u}_\beta$  in Figure 4.2. The geometric representation may be generalised to include the complete  $\pi$ -plane by replacing  $\theta_0$  with  $0$ .

### 4.3 LINEAR ELASTIC ANALYSIS

The early development of numerical methods in geotechnical and reservoir engineering situations was based on the assumption of linear elasticity. In the case of consolidation analysis a linear elastic behaviour of the soil skeleton is usually combined with linear fluid flow behaviour.

Isotropic linear elasticity requires the knowledge of two constants, which are usually Young's modulus  $E$  and Poisson's ratio  $v$ . In the case of a three-dimensional continuum,

the  $\mathbf{D}$  matrix of equation (4.1) becomes

$$\mathbf{D} = \frac{E}{(1+v)(1-2v)} \begin{bmatrix} 1-v & v & v & 0 & 0 & 0 \\ v & 1-v & v & 0 & 0 & 0 \\ v & v & 1-v & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\frac{2v}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\frac{2v}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-\frac{2v}{2} \end{bmatrix} \quad (4.13)$$

The following literature review shows some applications of linear elastic theory to the solution of soil mechanics problems.

In 1965 Clough and Rashid [6] investigated the stress distribution in a semi-infinite elastic half-space subjected to a concentrated point load. Brown and King [7] developed a program for the approximate study of the stability of an embankment. A model for investigation of subsidence, based on the poroelasticity theory, was presented by Geertsma [8]. Sandhu and Wilson [9] were the first to study linear elastic consolidation in a finite element context.

A procedure for determining values of  $E$  and  $v$  from one-dimensional compression tests was developed by Penman *et al.* [10]. The values of the material properties obtained were used for the prediction of the displacement in a dam.

Cole and Burland [11] used a trial-and-error technique by matching numerically the observed displacements of an excavation and a retaining wall. The Young's modulus so obtained was then used to predict stresses and displacements resulting from excavations to be made in the same soil. A similar technique was used by Gambolati *et al.* [12] for the simulation of the subsidence of Venice.

These examples, which are far from exhaustive, demonstrate that linear elastic analyses can give reasonable solutions for a single load path if sufficient care is taken in determining the material properties. This is also true for the investigation of surface subsidence, providing only compaction takes place [3].

Situations involving failure, non-recoverable strains upon unloading, etc., require either a variable elasticity approach or preferably an elastoplastic or viscoplastic constitutive relationship.

#### 4.4 VARIABLE ELASTIC ANALYSIS

In most geotechnical situations the constitutive behaviour of soil is non-linear. The most important contribution factors in consolidation and reservoir analysis are as follows [13]:

1. The bulk stiffness of the porous medium increases as particles are pressed closer together.
2. Clay experiences an abrupt reduction in stiffness when the preconsolidation stress is exceeded.
3. An abrupt increase in stiffness occurs on unloading.

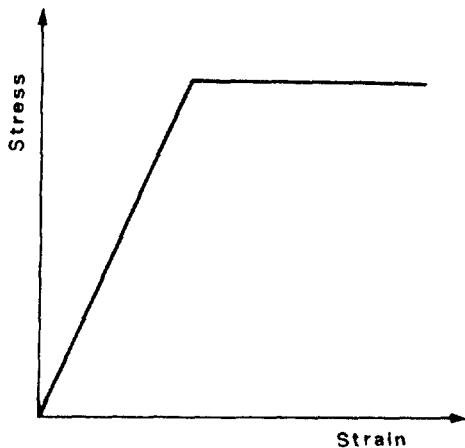
Here are some other important sources of non-linearity:

1. The shear stiffness of the soil reduces with distortion.
2. The tangential shear modulus becomes zero at failure.
3. Stiff clays and dense sands are dilatant; soft clays and loose sands are negatively dilatant.

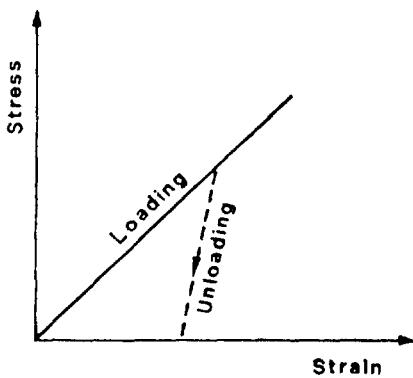
The most widely used stress-strain laws incorporate some of these properties, the precise number depending on the eventual aims of the analysis. Usual classifications distinguish between bilinear models and variable elastic models. From a strictly numerical viewpoint they are all piecewise-linear elastic models.

#### 4.4.1 Bilinear Models

In these models the stress-strain curve is represented by two straight lines, so there are only two possible modulus values for the soil. Bilinear stress-strain curves may have



**Figure 4.3** Bilinear stress-strain curve for the simulation of failure



**Figure 4.4** Bilinear stress-strain curve for loading-unloading

two different shapes:

1. If failure is to be simulated, the shear modulus  $G$  has a relatively large value before failure and a small nominal value after failure (Figure 4.3).
2. If the increase of stiffness upon unloading is to be simulated, the stress-strain curve is of the type shown in Figure 4.4. This is not an elastic behaviour but it is mentioned in this section since computationally it can be simulated by two elastic models.

The linear elastic law may be defined either by Young's modulus  $E$  and Poisson's ratio  $\nu$  or by the bulk modulus  $K$  and the shear modulus  $G$ .

The relations between the moduli are

$$G = \frac{E}{2(1+\nu)} \quad (4.14a)$$

$$K = \frac{E}{3(1-2\nu)} \quad (4.14b)$$

A common approach, if failure is to be simulated, has been to fix  $\nu = \nu_{\text{yield}}$  and vary the modulus  $E$  to match triaxial stress-strain curves. From equation (4.14b) it may be seen that, on yielding,  $K$  becomes very small, resulting in the 'black hole' phenomenon [2]. An alternative approach is to keep  $K$  constant and let  $\nu$  vary along with  $E$ .

A further criterion is needed to define the transition stress for the modulus values. Such a stress could be the Mohr-Coulomb yield stress or the undrained shear strength of the soil (equivalent to the Tresca failure criterion) [14,15]. The combination of a Mohr-Coulomb criterion and constant  $\nu$  model requires the four parameters  $E$ ,  $\nu$ ,  $c$ ,  $\phi$  where  $c$  is the cohesion and  $\phi$  is the friction angle.

A bilinear model of type 2 is useful for the simulation of rebound in subsidence problems. For this purpose it is important to identify the unloading region and its occurrence in time. In the case history of Venice [16,17], where such a bilinear model was applied, it was found that the whole cross-section under investigation expanded once the pore pressure increased again. This behaviour is typical for an open model where the aquifers are recharged.

#### 4.4.2 Variable Elastic Model

##### 4.4.2.1 Hyperbolic Model

A widely used constitutive model is one where the stress-strain curves of sand and clay are approximated by hyperbolae. The hyperbolic equation, first proposed by Kondner [18], is

$$\sigma_1 - \sigma_3 = \frac{\varepsilon_1}{a + b\varepsilon_1} \quad (4.15)$$

which relates the principal stress difference to the major principal strain. The parameters

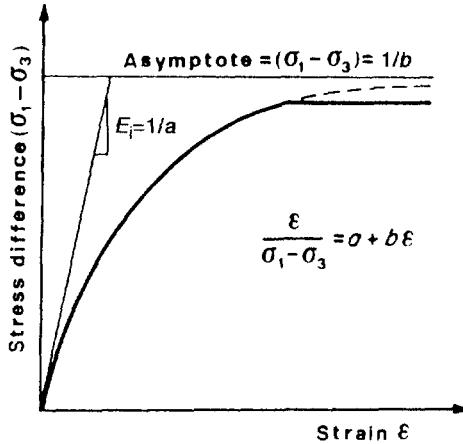


Figure 4.5 Hyperbolic stress-strain curve

$a$  and  $b$  are constants whose values may be determined from conventional (constant  $\sigma_3$ ) triaxial tests. Both  $a$  and  $b$  have physical meanings:  $1/a$  is the initial tangential Young's modulus and  $1/b$  is the ultimate principal stress difference (Figure 4.5).

A refinement to the model introduces the failure stress difference  $R_f/b$ , where  $R_f < 1$  and a cut-off occurs when  $\sigma_1 - \sigma_2$  reaches this value.

The model was further improved by Duncan and Chang [19] by obtaining the relationship for the initial tangent modulus, as given by Janbu [20], from the following expression:

$$E_i = Mp_a \left( \frac{\sigma_3}{p_a} \right)^n \quad (4.16)$$

where  $p_a$  is the atmospheric pressure and  $M$  and  $n$  are constants which are determined from a set of triaxial tests. It was also improved by utilising the Mohr–Coulomb failure criterion for determining the value of  $b$ .

The resulting tangent modulus can be written as

$$E_t = \left[ 1 - \frac{R_f(1 - \sin \phi)(\sigma_1 - \sigma_3)}{2c \cos \phi + 2\sigma_3 \sin \phi} \right]^2 Mp_a \left( \frac{\sigma_3}{p_a} \right)^n \quad (4.17)$$

With  $v = \text{constant}$ , the model requires a knowledge of six parameters, namely  $M$ ,  $n$ ,  $c$ ,  $\phi$ ,  $R_f$  and  $v$ . In this form the model was used to analyse the behaviour of footings on sand and clay. The results showed that the finite element solutions [19] were in good agreement with empirical solutions and applicable theories. This model was also successfully applied for the case of surface subsidence [21].

In their analysis of the Oroville Dam, Kulhawy and Duncan [22] extended the model to incorporate a tangent Poisson's ratio. In this instance eight constants were required to define the model.

#### 4.4.2.2 $E-v$ and $K-G$ Variable Elastic Models

These models, proposed by Naylor [2], define the tangential moduli directly in terms of the effective stress invariants  $p$  and  $q$ . The  $E - v$  model is defined by

$$v = \text{constant} \quad (4.18a)$$

$$E = E_0 + \alpha_E p + \beta_E q \quad (4.18b)$$

It requires four empirical constants  $E_0$ ,  $\alpha_E$ ,  $\beta_E$ ,  $v$  to define it.

The  $K-G$  model is defined by

$$K = K_0 + \alpha_K p \quad (4.19)$$

and

$$G = G_0 + \alpha_G p + \beta_G q \quad (4.20)$$

In this case five constants are needed, which may be obtained from conventional soil tests. A consolidation test should be used to determine the bulk stiffness parameters and a series of consolidated-drained or consolidated-undrained triaxial tests with pore pressure measurements are needed to determine the shear stress parameters. The models were successfully used to reproduce triaxial test results on a compacted boulder clay.

#### 4.4.2.3 Spline Functions

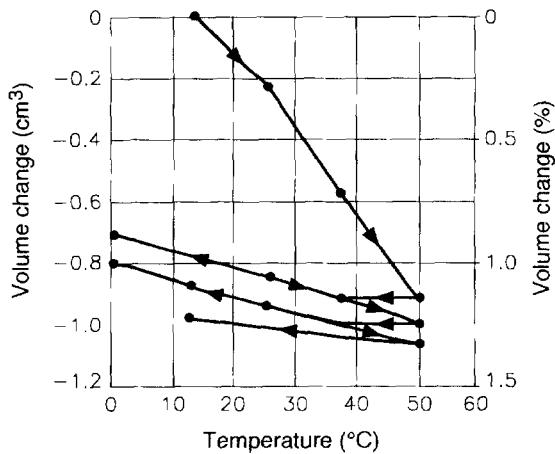
If sufficient experimental data is available, the full stress-strain curve can be defined by interpolation with spline functions [23]. Intermediate values and values of the slope of the curve can be determined by simple numerical calculations. The advantage of this procedure is that the actual experimental data may be represented to any required degree of accuracy. This representation of the stress-strain curve was used by Desai [24] to investigate the load deformation behaviour of a circular footing resting on a cohesive subsoil. The numerical results compared favourably with those obtained using a hyperbolic model and with experimental results.

### 4.4.3 Thermo Elastic Behaviour

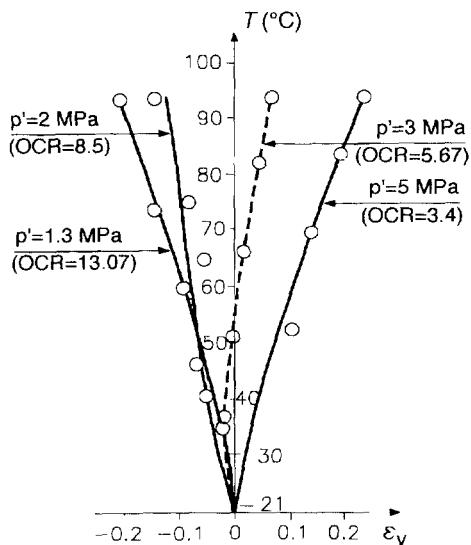
Temperature changes in saturated soils can cause significant volumetric deformations and variations of effective stress [25]. Variations of effective stress are very sensible in undrained conditions, since quick temperature increase can produce quick increase of pore water pressure up to nullify the effective mean stress, or effective pressure  $p$ , and to exhaust the shear strength.

Cyclic thermal loading on drained soils at constant effective stress may result in unrecoverable volume changes at first heating, a phenomenon often termed thermal consolidation, followed by small amounts of reversible deformations (Figure 4.6).

Overall strains are the result of the expansion or contraction of mineral constituents and of pore water, as well as physico chemical interactions, which are usually negligible,



**Figure 4.6** Volume change of clay structure under cyclic thermal loading (Reprinted, by permission, from Campanella, R. G. and Mitchell, J. K., 1968, Influence of temperature variations on soil behaviour, *J. Soil Mech. Found. Div. ASCE*, **94**, 709–34)



**Figure 4.7** Drained heating tests at constant effective stress for various OCR values (Reprinted, by permission, from Hueckel, T. and Baldi, G., 1991, Thermoplasticity of saturated clays: experimental constitutive study, *J. Geotech. Engng.*, **116**, 1778–96)

(e.g. with respect to weight) in sand-like materials but which are prevalent in cohesive materials, especially in the presence of water. The combination of these factors can then result in a thermal behaviour which is expansive at cooling and contractile at heating also in the elastic (reversible) range (Figure 4.7).

The amount of reversible deformation depends on the applied pressure and on the overconsolidation ratio (OCR); the conventional limit value  $\text{OCR}=1.75$ , between

normally consolidated and overconsolidated clays, can be assumed as the limit over which temperature changes have no permanent effect.

The reversible behaviour of Figure 4.7 can be described [26] by the following relationships:

$$\varepsilon_v^{re}(p, p_r, T) = \beta(p, p_r, T)(T_0 - T) \quad (4.21)$$

$$\beta(p, p_r, T) = (\beta_0 + \beta_1 \Delta T) \ln \frac{p}{p_r} + \beta_2 \Delta T \quad (4.22)$$

where the coefficient of thermal expansion  $\beta$  depends on the consolidation pressure  $p_c$  and on the OCR through the reference pressure  $p_r$ ; a reverse of sign for  $\beta$  (4.22) can result, depending on the value of the applied pressure.

Once the pressure is fixed, the coefficient of thermal expansion  $\beta$  depends linearly on the temperature change  $\Delta T = T_0 - T$ . However, for some applications in elastoplasticity, due to the significantly small amount of elastic deformations when compared with the irreversible contributions, the coefficient  $\beta$  can be considered like a material constant; an average value independent of the material characteristics and equal to  $5 \times 10^{-5} \text{ }^{\circ}\text{C}^{-1}$  is often assumed.

#### 4.4.4 Solution Procedures

This section describes commonly used solution algorithms, which incorporate the effect of nonlinear stress-strain laws, as well as a proposed method for simulating the unloading of the soil.

In the secant stiffness method (Figure 4.8) for the one-dimensional case, a first estimation  $\mathbf{u}_0$  of the displacement field due to the external load  $\bar{\mathbf{F}}$  is obtained from the initial trial matrix  $\mathbf{K}_{S0}$ . The resulting strains and stresses follow from this displacement field  $\mathbf{u}_0$  through consistency and the constitutive law; the internal force  $\mathbf{f}_1$  and the new stiffness  $\mathbf{K}_{S1}$  can then be evaluated, to give the new estimation  $\mathbf{u}_1$  of the displacement

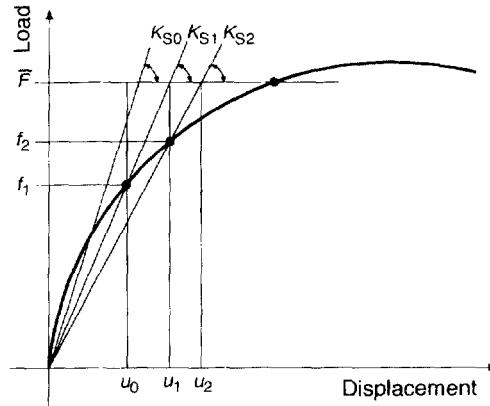
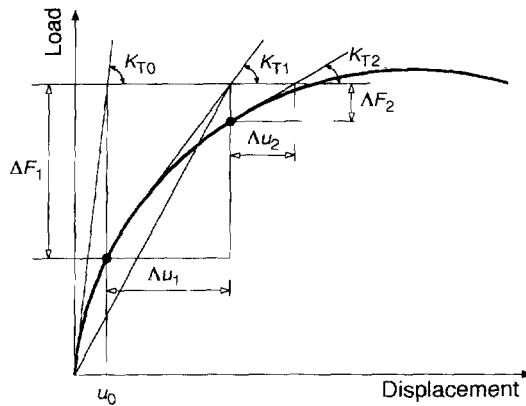


Figure 4.8 Secant stiffness method



**Figure 4.9** Tangent stiffness method

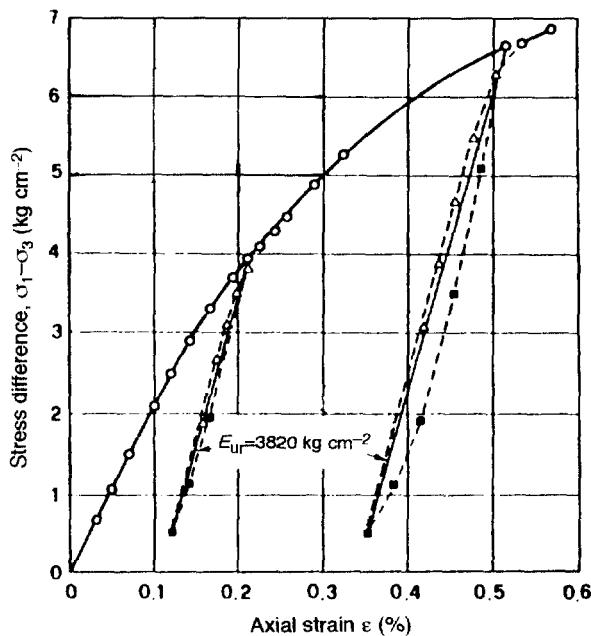
field. The iterative process continues until all convergence criteria are fulfilled. This procedure has been applied in a soil mechanics context [27,28].

The tangential stiffness method employs an incremental procedure for the one-dimensional case (Figure 4.9). The first estimation of the displacement field is obtained from the tangent stiffness matrix  $K_{T0}$ , resulting from the initial elastic moduli of the material. Unbalanced forces  $\Delta F_1 = \mathbf{F} - \mathbf{f}_1$  are then evaluated and the tangent stiffness matrix is updated on the basis of the current strains and stresses to solve the problem for the incremental displacement  $\Delta u_1$ . This (Newton–Raphson) procedure is repeated until the unbalanced forces  $\Delta F_i$  vanish within the given tolerance. At the end of the process, the sequence of incremental displacements  $\Delta u_i$  ( $i = 1, k$ ) is accumulated, and the total displacement is obtained as the sum, i.e.  $\mathbf{u}_0 + \Delta \mathbf{u}_1 + \dots + \Delta \mathbf{u}_k$ . As a variant of this method, the tangent stiffness matrix is evaluated once at the beginning of the process and kept constant in the iterative process, since its modification can prove expensive for large problems (modified Newton–Raphson). However, the computational savings which may be expected are often negated by the increasing number of necessary iterations, since the quadratic rate of convergence, typical of a Newton–Raphson procedure, is lost.

If a consolidation problem is analysed, the tangent modulus is varied as consolidation proceeds and non-linearity progresses. Again, the procedure is based on an incremental load technique.

An efficient and robust procedure of the fixed-point type can be obtained for non-linear problems with convex potential (e.g. partially saturated consolidation with elastic solid behaviour) by combining the above procedures. At each load or time increment a first estimation of the displacement field is obtained from the tangent stiffness matrix  $K_T$ , utilising the material properties at the beginning of the  $n$ th step. These are based on the displacement solution  $\mathbf{u}_n$  of the previous step.

Successive iterations are then carried out by solving  $\mathbf{K}_0 \mathbf{u}_{n+1,i+1} = \mathbf{F}_0$ , with the stiffness matrix  $\mathbf{K}_0$  evaluated at  $\mathbf{u}_0 = (1 - \theta)\mathbf{u}_n + \theta\mathbf{u}_{n+1,i}$ , with  $0 \leq \theta \leq 1$ . The value  $\mathbf{u}_{n+1,i}$  being the current estimation of the displacements (step  $n + 1$ , iteration  $i$ ). Iterations are then continued until two successive solutions for  $\mathbf{u}_{n+1}$  differ by less than a fixed tolerance. The rate of convergence for this procedure is usually linear.



**Figure 4.10** Unloading and reloading of silica under drained triaxial test conditions; for dense sand  $\sigma'_3 = 3 \text{ kg cm}^{-2}$ : (---○) primary loading, (■—■) unloading, (---) reloading (Reprinted, by permission, from Kondner, R.L., 1963, Hyperbolic stress-strain response: cohesive soils. *J. Soil Mech. Found. Div. ASCE*, **89**, 115–43)

Each method has advantages and shortcomings. The secant modulus method has the advantage of being able to model strain-softening behaviour where the stress decreases with increasing strain after reaching a peak value. The disadvantage of this method is that non-zero initial stresses, which can play an important role in many soil mechanics problems, are very difficult to take into account.

These points are reversed for the tangential stiffness method, where initial stresses can easily be accounted for. However, a strain-softening behaviour causes difficulties, as it requires a negative modulus value. This method fits in very well with the sequential construction technique used for embankments and excavation.

The tangential stiffness method, with or without iteration, is used for consolidation and subsidence analyses in several parts of this text. Even a procedure for unloading can be incorporated into the tangential stiffness method by using non-linear stress-strain laws. If unloading occurs, the current tangent stiffness modulus  $E_t$  is replaced by the modulus for unloading-reloading  $E_{ur}$  (Figure 4.10). The only problem is to define a viable criterion to check for unloading [3]. If the whole soil volume either compacts or expands at the same time, which happens in subsidence models with boundaries open to flow, a trial run with the modulus for compaction indicates at what instance in the repeat solution the modulus for unloading-reloading should be applied. An elastoplastic analysis easily overcomes this difficulty.

Most of the soil properties shown in Section 4.4 can be simulated with isotropic variable elastic models. But a dilatancy, which accompanies shearing of soil, cannot be reproduced with such models. This disadvantage is not shared by elastic–plastic models. Another difference is that the strain pattern for a small load increment is largely dictated by the accumulated stress values for elastoplastic laws, whereas for a variable elastic law this is determined by the incremental stresses.

## 4.5 ELASTOPLASTIC MODELS

### 4.5.1 Constitutive Law

A brief summary is given here of the theory required for the analysis of elastoplastic soil behaviour. Both classical elastoplasticity and generalised plasticity will be considered. In the case of classical elastoplasticity both Mohr–Coulomb and critical state yield surfaces will be considered. The critical state component requires an associated flow rule, whereas the Mohr–Coulomb component can also consider a non-associated flow rule. A more extensive development of the theory presented in this section can be found elsewhere, e.g. Zienkiewicz [29], Nayak and Zienkiewicz [4], Naylor [2] and Humpheson [30].

A yield criterion defines the limits of elasticity under any possible combination of stresses. When written in terms of stress components, this is called a yield function and represents a surface in the  $n$ -dimensional stress space, which separates the elastic state from an outer zone of impermissible stress states. The position and size of the surface depends upon the initial yield surface and the hardening law which specifies the manner in which the surface changes during plastic flow. Hardening is considered here to depend only on plastic strain  $\boldsymbol{\varepsilon}^p$ .

The yield surface can therefore be defined as

$$F(\boldsymbol{\sigma}', \boldsymbol{\varepsilon}^p) = 0 \quad (4.23)$$

Irreversible (plastic) straining occurs when the stress state reaches the yield surface, while for  $F < 0$  the behaviour is elastic.

The direction of the plastic strain is defined by the plastic potential surface  $Q(\boldsymbol{\sigma}')$ . The normal to this surface, passing through the current stress point,

$$d\boldsymbol{\varepsilon}^p = d\lambda \frac{\partial Q}{\partial \boldsymbol{\sigma}'} \quad (4.24)$$

where  $d\lambda$  is a yet unknown scalar. If  $F \equiv Q$ , the flow rule is said to be associated.

Differentiating equation (4.23) gives the relationship

$$\left( \frac{\partial F}{\partial \boldsymbol{\sigma}'} \right)^T d\boldsymbol{\sigma}' + \left( \frac{\partial F}{\partial \boldsymbol{\varepsilon}^p} \right)^T d\boldsymbol{\varepsilon}^p = 0 \quad (4.25)$$

If the flow rule (4.24) is taken into account, then equation (4.25) becomes

$$\left(\frac{\partial F}{\partial \sigma'}\right)^T d\sigma' + \left(\frac{\partial F}{\partial \varepsilon^p}\right)^T d\lambda \frac{\partial Q}{\partial \sigma'} = 0 \quad (4.26)$$

Hence

$$d\lambda = -\frac{\{\partial F/\partial \sigma'\}^T d\sigma'}{\{\partial F/\partial \varepsilon^p\}^T \{\partial Q/\partial \sigma'\}} \quad (4.27)$$

Once the elastic limit has been exceeded for an elastoplastic model, the total strain  $d\varepsilon$  can be expressed as the sum of the elastic and the plastic components  $d\varepsilon^e$  and  $d\varepsilon^p$ . This is valid for small-strain plasticity theory, which is assumed in this book and can also be used in finite strain situations when an updated Lagrangian procedure is used (Chapter 14):

$$d\varepsilon = d\varepsilon^e + d\varepsilon^p = \mathbf{D}^e^{-1} d\sigma' + d\lambda \frac{\partial Q}{\partial \sigma'} \quad (4.28)$$

where  $\mathbf{D}^e$  is the tangential elastic stiffness matrix.

In the present context only linear elasticity is considered, so that  $\mathbf{D}^e$  is defined by equation (4.13).

Premultiplying equation (4.28) by  $\{\partial F/\partial \sigma'\}^T \mathbf{D}^e$ , and substituting for  $\{\partial F/\partial \sigma'\}^T d\sigma'$  by means of equation (4.26), leads to the following:

$$\left(\frac{\partial F}{\partial \sigma'}\right)^T \mathbf{D}^e d\varepsilon = -\left(\frac{\partial F}{\partial \varepsilon^p}\right)^T \frac{\partial Q}{\partial \sigma'} d\lambda + \left(\frac{\partial F}{\partial \sigma'}\right)^T \mathbf{D}^e \frac{\partial Q}{\partial \sigma'} d\lambda \quad (4.29)$$

The term  $d\lambda$  is now obtained from equation (4.29) and may be substituted into equation (4.28). Premultiplying by  $\mathbf{D}^e$  and rearranging gives the equation defining the tangent elastoplastic modulus matrix  $\mathbf{D}^{ep}$ :

$$d\sigma' = \left[ \mathbf{D}^e - \frac{\mathbf{D}^e (\partial Q/\partial \sigma') (\partial F/\partial \sigma')^T \mathbf{D}^e}{-(\partial F/\partial \varepsilon^p)^T (\partial Q/\partial \sigma') + (\partial F/\partial \sigma')^T \mathbf{D}^e (\partial Q/\partial \sigma')} \right] d\varepsilon \quad (4.30)$$

Construction of the elastoplastic constitutive matrix requires the vectors  $\partial Q/\partial \sigma$ ,  $\partial F/\partial \sigma'$  and  $\partial F/\partial \varepsilon^p$  to be evaluated. These vectors will be formulated in Sections 4.5.2 and 4.5.3 for the Mohr–Coulomb and critical state yield surfaces.

The surfaces will be defined using the stress invariants  $p$ ,  $q$  and  $\theta$  which were introduced in Section 4.2. It is therefore necessary to evaluate the derivatives of these invariants with respects to  $\sigma'$ .

From the definition of the invariants it follows that

$$\frac{\partial p}{\partial \sigma'} = - \begin{Bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ 0 \\ 0 \end{Bmatrix} \quad (4.31a)$$

$$\frac{\partial q}{\partial \sigma'} = \frac{1}{2q} \begin{Bmatrix} 2\sigma_x - \sigma_y - \sigma_z \\ 2\sigma_y - \sigma_x - \sigma_z \\ 2\sigma_z - \sigma_y - \sigma_x \\ 6\tau_{xy} \\ 6\tau_{yz} \\ 6\tau_{xz} \end{Bmatrix} \quad (4.31b)$$

Differentiating equation (4.4) yields

$$3 \cos \theta_0 d\theta_0 = -\frac{27}{2q^3} \left( dJ_3 - \frac{3J_3}{q} dq \right) \quad (4.31c)$$

so that

$$\frac{\partial \theta}{\partial \sigma'} = \frac{\partial \theta_0}{\partial \sigma'} = \frac{9}{2q^3 \cos 3\theta_0} \left[ \frac{3J_3}{q} - \frac{\partial q}{\partial \sigma'} - \frac{\partial J_3}{\partial \sigma'} \right] \quad (4.31d)$$

For plane strain, plane stress and axial symmetry  $\sigma_{yz} = 0$ , so

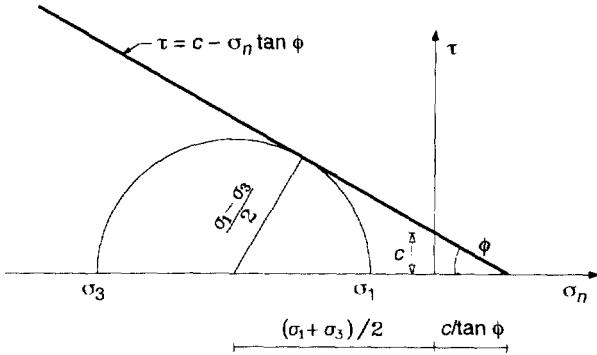
$$|J_3| = (\sigma_z + p)\{(\sigma_x + p)(\sigma_y + p) - \tau_{xy}^2\} \quad (4.32a)$$

and

$$\frac{\partial |J_3|}{\partial \sigma'} = \begin{Bmatrix} (\sigma_y + p)(\sigma_z + p) \\ (\sigma_x + p)(\sigma_z + p) \\ (\sigma_x + p)(\sigma_y + p) - \tau_{xy}^2 \\ -2(\sigma_z + p)\tau_{xy} \\ 0 \\ 0 \end{Bmatrix} + \frac{1}{9} \begin{Bmatrix} q^2 \\ q^2 \\ q^2 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \quad (4.32b)$$

#### 4.5.2 Mohr–Coulomb Yield Surface

Mohr's theory of failure involves the construction of an envelope for all possible circles of stress that can be drawn for a particular problem. These envelopes are generally



**Figure 4.11** Mohr–Coulomb failure envelope

curved but are usually replaced by a straight line. This is equivalent to assuming that the soil conforms to the Coulomb failure criterion [31], which states there is a linear relationship between the shear stress  $\tau$  at failure and the normal stress,  $\sigma_n$  (Figure 4.11):

$$\tau = c + \sigma_n \tan \phi \quad (4.33)$$

where  $c$  = apparent cohesion

$\phi$  = angle of internal friction

From Figure 4.11 it may be deduced that

$$\sin \phi = \frac{(\sigma_1 - \sigma_3)/2}{(c/\tan \phi) + (\sigma_1 + \sigma_3)/2} \quad (4.34)$$

or, by rearranging,

$$(\sigma_1 - \sigma_3) = 2c \cos \phi + (\sigma_1 + \sigma_3) \sin \phi \quad (4.35)$$

where  $\sigma_1$  and  $\sigma_3$  are the major and minor principal stresses at failure. This yield criterion is independent of the intermediate principal stress  $\sigma_2$  and is therefore not a complete generality of the true behaviour. This criticism applies to any yield criterion that can be represented by a single line in a two-dimensional stress space.

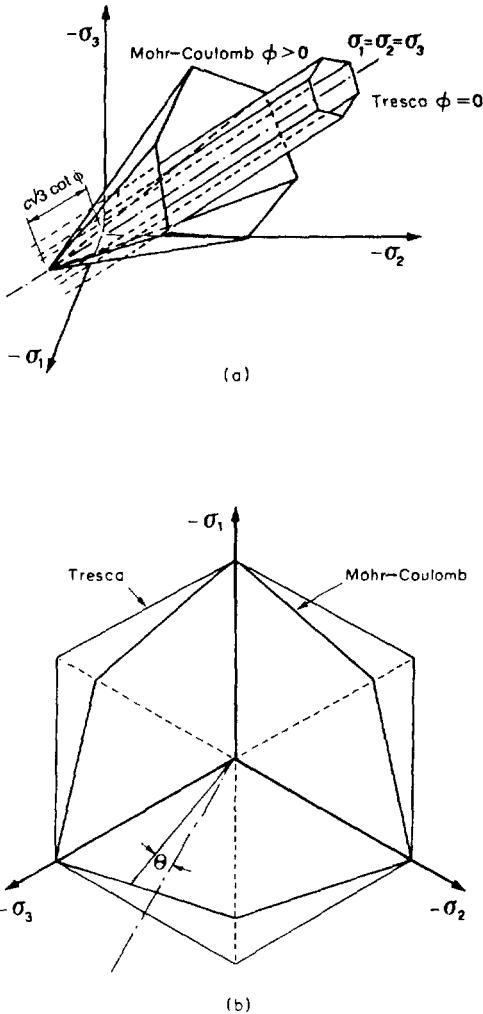
From equation (4.6) the sum and the difference of the maximum and minimum principal stresses are

$$\sigma_1 - \sigma_3 = \frac{2}{\sqrt{3}} q \cos \theta_0 \quad (4.36a)$$

$$\sigma_1 + \sigma_3 = 2p + \frac{2}{3} q \sin \theta_0 \quad (4.36b)$$

which, after substitution into equation (4.35), results in the following expression for the Mohr–Coulomb failure surface:

$$F = (\sqrt{3} \cos \theta_0 - \sin \theta_0 \sin \phi)q - 3p \sin \phi - 3c \cos \phi = 0 \quad (4.37)$$



**Figure 4.12** (a) Mohr-Coulomb and Tresca yield surfaces in the principal stress space; (b) intersection with the  $\pi$ -plane

In a principal stress space, equation (4.37) represents an irregular right hexagonal pyramid, the axis of which lies along the space diagonal (Figure 4.12(a)). The  $\pi$ -plane section of the yield surface is shown in Figure 4.12 (b).

The plastic potential surface,  $Q(p, q, \theta_0) = 0$ , passing through the current stress point can be obtained in a similar manner. The direction of plastic straining in the  $\tau, \sigma_n$  plane is usually expressed by the normal to a line defined as follows,

$$\tau = \bar{c} + \sigma_n \tan \psi \quad (4.38)$$

where the value of  $\bar{c}$  ensures that the line passes through the current stress point, and  $\psi$  is the dilatancy angle.

The potential surface  $Q$ , in terms of  $p$ ,  $q$ ,  $\theta_0$ , can be obtained in the same way as the yield surface and results in the following expression:

$$Q = (\sqrt{3} \cos \theta_0 - \sin \theta_0 \sin \psi)q - 3p \sin \psi - 3\bar{c} \cos \psi = 0 \quad (4.39)$$

Equations (4.37) and (4.39) may be applied to either total or effective stresses by taking into account the appropriate material properties.

When  $\phi = \psi = 0$ , which is the case in the total stress analysis of undrained, saturated soil, both surfaces assume the form of a Tresca prism. In the principal stress space this surface is represented by a regular hexagonal cylinder with its axis also lying along the space diagonal (Figure 4.12).

Strain hardening or softening may be dealt with by making  $c$  a function of volumetric plastic strain by using the following equation:

$$\frac{dc}{d\varepsilon_v^p} = \frac{c}{\chi} \quad (4.40)$$

where  $\chi$  is a constant.

The vectors  $\partial F/\partial \sigma'$ ,  $\partial Q/\partial \sigma'$  and  $\partial F/\partial \varepsilon^p$ , required for the constitutive matrix of equation (4.30) are now derived. The chain rule of differentiation gives

$$\frac{\partial F}{\partial \sigma'} = \frac{\partial F}{\partial p} \frac{\partial p}{\partial \sigma'} + \frac{\partial F}{\partial q} \frac{\partial q}{\partial \sigma'} + \frac{\partial F}{\partial \theta} \frac{\partial \theta}{\partial \sigma'} \quad (4.41)$$

The three derivatives  $\partial F/\partial p$ ,  $\partial F/\partial q$ ,  $\partial F/\partial \theta$  are evaluated as follows:

$$\frac{\partial F}{\partial p} = -3 \sin \phi \quad (4.42a)$$

$$\frac{\partial F}{\partial q} = \sqrt{3} \cos \theta_0 - \sin \theta_0 \sin \phi \quad (4.42b)$$

$$\frac{\partial F}{\partial \theta} = -(\sqrt{3} \sin \theta_0 + \cos \theta_0 \sin \phi)q \quad (4.42c)$$

whereas  $\partial p/\partial \sigma'$ ,  $\partial q/\partial \sigma'$  and  $\partial \theta/\partial \sigma'$  are given by equations (4.31).

Similarly,

$$\frac{\partial Q}{\partial \sigma'} = \frac{\partial Q}{\partial p} \frac{\partial p}{\partial \sigma'} + \frac{\partial Q}{\partial q} \frac{\partial q}{\partial \sigma'} + \frac{\partial Q}{\partial \theta} \frac{\partial \theta}{\partial \sigma'} \quad (4.43)$$

where  $\partial Q/\partial p$ ,  $\partial Q/\partial q$  and  $\partial Q/\partial \theta$  are given by substituting  $\psi$  for  $\phi$  in the above expressions for  $\partial F/\partial p$ ,  $\partial F/\partial q$  and  $\partial F/\partial \theta$ .

Equations (4.37) and (4.40) also lead to

$$\begin{aligned} \frac{\partial F}{\partial \varepsilon^p} &= \frac{\partial F}{\partial \varepsilon_v^p} \frac{\partial \varepsilon_v^p}{\partial \varepsilon^p} + \left[ \frac{\partial F}{\partial \varepsilon_q^p} \frac{\partial \varepsilon_q^p}{\partial \varepsilon^p} = 0 \right] \\ &= \frac{\partial F}{\partial c} \frac{\partial c}{\partial \varepsilon_v^p} \frac{\partial \varepsilon_v^p}{\partial \varepsilon^p} = 3 \cos \phi \frac{\partial c}{\partial \varepsilon_v^p} [-1, -1, -1, 0, 0, 0]^T \\ &= \frac{3c}{\chi} \cos \phi [-1, -1, -1, 0, 0, 0]^T \end{aligned} \quad (4.44)$$

Since  $[-1, -1, -1, 0, 0, 0] \partial Q / \partial \sigma' = \partial Q / \partial p$ , it follows that

$$\left\{ \frac{\partial F}{\partial \varepsilon^p} \right\}^T \hat{c}Q = \frac{3c}{\chi} \cos \phi \frac{\partial Q}{\partial p} = -\frac{qc}{\chi} \cos \phi \sin \psi \quad (4.45)$$

which is needed for the constitutive matrix.

### 4.5.3 Critical State Model

The critical state model is a form of elastoplastic isotropic strain-hardening law. It introduces a distinction between yielding and ultimate collapse by using the concept of a critical state line in conjunction with a strain-dependent yield surface. A soil is at the critical state [32] if, during continuous deformation, there is no change in the void ratio and the effective stress components. In this model a soil undergoing shear deformation can pass through a yield point without collapse and continue to deform until eventually the critical state line is reached, where ideal plasticity conditions exist. The soil continues to deform without further change of void ratio or stress. Starting from two alternative assumptions regarding the dissipation of energy during plastic yielding, Roscoe *et al.* [33] and Schofield and Wroth [34] proposed the ‘Cam clay’ model and Roscoe and Burland [35] the ‘modified Cam clay’ model.

#### 4.5.3.1 Modified Cam Clay Model

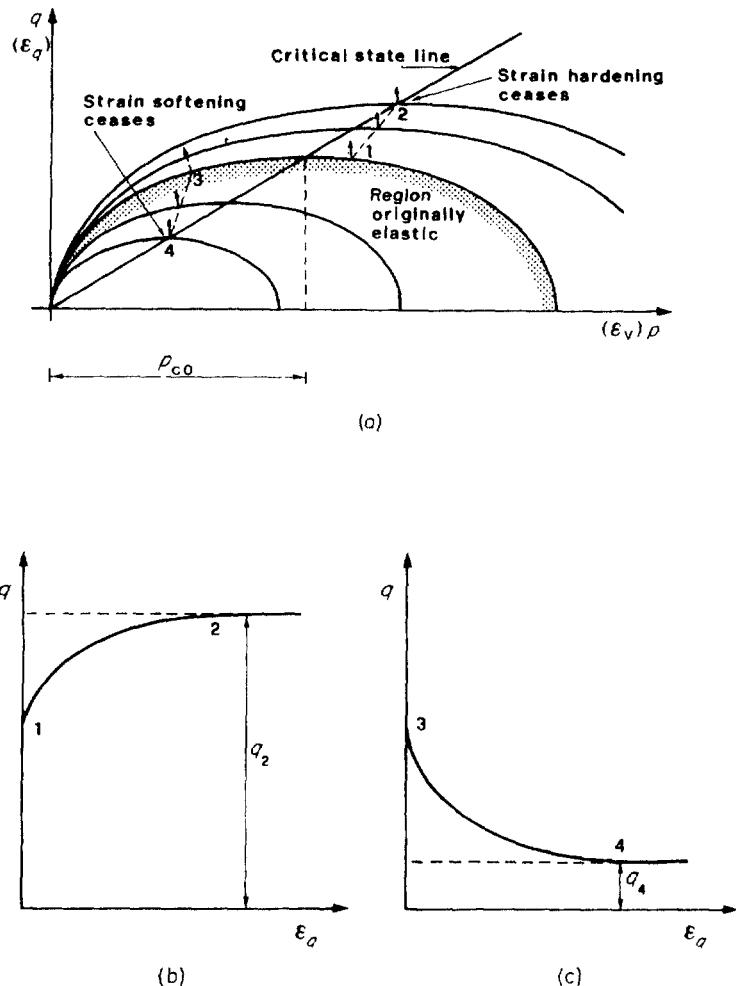
The modified Cam clay model fits experimental data quite satisfactorily and is selected here. Its yield surface is an ellipse in the  $p, q$  plane (Figure 4.13) and is defined by the equation

$$F = \frac{q^2}{M_{cs}^2} - 2pp_c(\varepsilon_v^p) + p^2 = 0 \quad (4.46)$$

where  $M_{cs}$  is the slope of the failure line in the  $p, q$  plane, and  $p_c(\varepsilon_v^p)$  is the current semidiameter of the ellipse in the  $p$ -direction. The full surface is a surface of revolution about the  $q$ -axis and is therefore defined by  $p$  and  $q$  only.

As the ellipse changes in magnitude, the locus of the critical state points is a pyramid with its apex at the origin shown by the critical state line in the plane (Figure 4.13(a)). The flow rule is associative, i.e.  $F \equiv Q$  and the principle of normality therefore applies to the yield surface. Since the surface is smooth, the direction of plastic straining is uniquely defined for every point of the surface. At the intersection of the critical state line and the ellipse, the normal to the yield surface is vertical. Hence at this point no component of plastic volumetric strain exists and all the plastic strain is distortional: the soil can deform at a constant volume.

The yield surface is therefore strain dependent and expands or contracts as the soil hardens or softens. Strain hardening is associated with compaction and strain softening with a volume increase. The initial size of the ellipse is governed by the maximum preconsolidation pressure  $2p_{c0}$  to which the soil has previously been subjected during



**Figure 4.13** Modified Cam clay model in the space of the two stress invariants  $p$  and  $q$ ; (b) strain-hardening behaviour; (c) strain-softening behaviour

its past history. If the soil has been overconsolidated at some time in its history, then  $p_{c0}$  may be quite large and the soil could sustain substantial loads before any yielding occurred. For a stress path of type 1–2 (Figure 4.13(a)) the plastic strain vector normal to the ellipse produces a plastic volumetric decrease which causes the soil to harden. The ellipse expands until eventually position 2 is reached, at which point no further volumetric strain occurs. The soil flows as a frictional fluid with constant volume.

The stress 3–4 shows a strain-softening behaviour due to the expansion of the material. Consequently, the ellipse decreases in size and eventually at point 4 the no-volume-change limit is reached and collapse occurs at constant volume. Many investigators have serious misgivings about the validity of the strain-softening behaviour

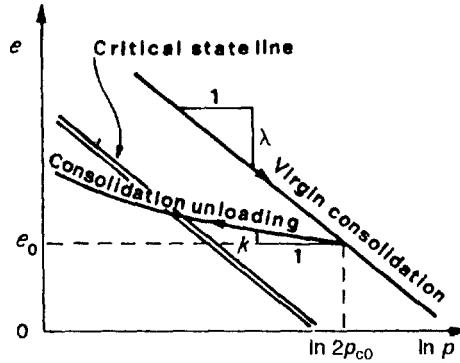


Figure 4.14 Consolidation parameters  $\lambda$  and  $k$

thus described [36]. Combined yield surfaces [30], such as the model used in Section 5.4.2.3, may represent a solution to this problem.

The strain-hardening law uses the consolidation parameter  $\lambda$  and  $k$  obtained by isotropic loading and unloading of normally consolidated soil. The parameter  $\lambda$  is the slope of the void ratio versus  $\ln p$  plot during loading whereas  $-k$  is the initial value of the slope during rebound (Figure 4.14). From the geometry of Figure 4.14 it may be seen that the relationship between the plastic component of the void ratio change and the mean stress  $p$  is given by

$$(e - e_0)^p = -(\lambda - k)(\ln 2p_c - \ln 2p_{c0})$$

or

$$\ln(p_c/p_{c0}) = -\frac{(e - e_0)^p}{(\lambda - k)} \quad (4.47a)$$

The changes in volumetric strain are related to changes in void ratio as follows:

$$\varepsilon_v - \varepsilon_{v0} = -\frac{e - e_0}{1 + e_0} \quad (4.47b)$$

Hence equation (4.47a) may be rewritten in the form

$$p_c = p_{c0} \exp(\varepsilon_v - \varepsilon_{v0})^p / \chi \quad (4.47c)$$

where  $\chi = \lambda - k/1 - e_0$  is an empirical constant. Denoting the accumulated plastic volumetric strain by  $h$ , i.e.

$$h = \int d\varepsilon_v^p = (\varepsilon_v - \varepsilon_{v0})^p \quad (4.47d)$$

the hardening law assumes the form

$$p_c = p_{c0} \exp \frac{h}{\chi} \quad (4.47e)$$

The derivatives  $\partial F / \partial \sigma'$  and  $\partial F / \partial \epsilon^p$  required to set up the constitutive matrix are next obtained. From equation it (4.46) it follows that

$$\frac{\partial F}{\partial \sigma'} = \frac{\partial F}{\partial p} \frac{\partial p}{\partial \sigma'} + \frac{\partial F}{\partial q} \frac{\partial q}{\partial \sigma'} \quad (4.48a)$$

where

$$\frac{\partial F}{\partial p} = 2(p - p_c) \quad (4.48b)$$

$$\frac{\partial F}{\partial q} = \frac{2q}{M_{cs}^2} \quad (4.48c)$$

$\frac{\partial p}{\partial \sigma'}$  and  $\frac{\partial q}{\partial \sigma'}$  are given by equations (4.31).

From equation (4.46)

$$\begin{aligned} \frac{\partial F}{\partial \epsilon^p} &= \frac{\partial F}{\partial \epsilon_v^p} \frac{\partial \epsilon_v^p}{\partial \epsilon^p} + \left[ \frac{\partial F}{\partial \epsilon_q^p} \frac{\partial \epsilon_q^p}{\partial \epsilon^p} = 0 \right] \\ &= \left( -2p \frac{\partial p_c}{\partial \epsilon_v^p} \right) \begin{Bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \end{aligned} \quad (4.49)$$

Equation (4.47c) leads to the following result:

$$\frac{dp_c}{p_c} = \frac{d\epsilon_v^p}{\chi} \quad (4.50)$$

Hence, equation (4.49) becomes

$$\frac{\partial F}{\partial \epsilon^p} = \frac{-2pp_c}{\chi} \begin{Bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \quad (4.51)$$

From equations (4.30), (4.46) and (4.47c) it can be seen that before defining the parameters of the modified Cam clay model, a knowledge of  $E$ ,  $v$ ,  $p_{c0}$ ,  $M_{cs}$ ,  $e_0$ ,  $\lambda$  and  $k$  is required. A more general critical state model involving the third stress invariant is discussed in the next section. This requires one more parameter.

#### 4.5.3.2 $p$ - $q$ - $\theta$ Critical State Model

The generalisation of the model to include the effect of the variation of the third stress invariant  $\theta$  is achieved by letting the slope of the critical state line vary with  $\theta$ . Zienkiewicz *et al.* [36] developed an elliptical model whose  $\pi$ -plane section was the same as that of the Mohr–Coulomb surface. An alternative to that model uses the critical state ellipse only in the subcritical region (to the right of the critical region, to the right of the critical state line) whereas the supercritical region may be cut off by a Mohr–Coulomb surface [30].

The model dealt with here consists of the full ellipse with a Mohr–Coulomb,  $\pi$ -plane section and with the critical state line passing through the origin [37]:

$$F = [p - \alpha_p(\varepsilon_v^p)]^2 + [q/n(\theta)]^2 - a^2(\varepsilon_v^p) = 0 \quad (4.52)$$

where  $\alpha_p(\varepsilon_v^p)$  = the  $p$ -coordinate of the centre of the elliptical surface

$n(\theta)$  = the ratio of the diameters in the  $q$  and  $p$  directions, defined by the Mohr–Coulomb equation for  $c = 0$

$a(\varepsilon_v^p)$  = the semi-diameter of the ellipse in the  $p$ -direction (Figure 4.15)

Plastic yielding is again of the associated form and strain hardening depends on the plastic changes of void ratio or volumetric strain. From inspection of the critical state line passing through the origin, it follows that

$$\frac{d\alpha_p}{\alpha_p} = \frac{da}{a} \quad (4.53)$$

The hardening rule can be obtained in a similar manner [37] as in the previous section,

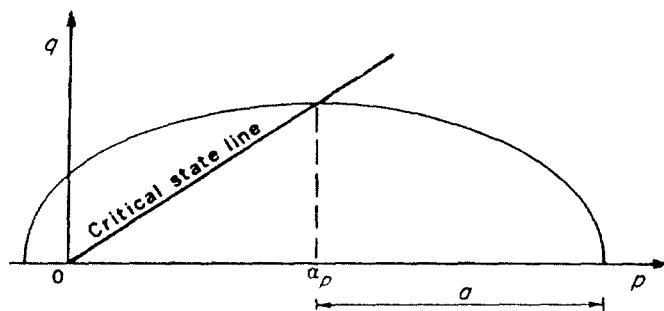


Figure 4.15 The  $p$ - $q$ - $\theta$  critical state model

resulting in

$$\alpha_p = \alpha_{p0} \exp \frac{h}{\chi} \quad (4.54a)$$

and

$$a = a_0 \exp \frac{h}{\chi} \quad (4.54b)$$

Hence

$$\frac{d\alpha_p}{\alpha_p} = \frac{da}{a} = \frac{d\varepsilon_y^p}{\chi} \quad (4.54c)$$

A formula is required for one more term;  $n$  is needed to define the yield surface of equation (4.52).

It is assumed that the ‘critical state’ lies on a Mohr–Coulomb surface for which  $c$  equals zero. From equation (4.37) with  $c = 0$  it follows that

$$M_{cs} = \frac{\partial q}{\partial p} = \frac{3 \sin \phi}{\sqrt{3} \cos \theta_0 - \sin \theta_0 \sin \phi} \quad (4.55)$$

The geometry of Figure 4.15 reveals that

$$n(\theta) = \frac{\alpha_p M_{cs}}{a} = \frac{\alpha_p}{a} \frac{3 \sin \phi}{\sqrt{3} \cos \theta_0 - \sin \theta_0 \sin \phi} \quad (4.56)$$

The derivatives  $\partial F/\partial \sigma'$  and  $\partial F/\partial \varepsilon^p$ , required for the constitutive matrix, are now derived. For equations (4.52) and (4.56) it can be seen that

$$\frac{\partial F}{\partial \sigma'} = \frac{\partial F}{\partial p} \frac{\partial p}{\partial \sigma'} + \frac{\partial F}{\partial q} \frac{\partial q}{\partial \sigma'} + \frac{\partial F}{\partial \theta} \frac{\partial \theta}{\partial \sigma'} \quad (4.57a)$$

where

$$\frac{\partial F}{\partial p} = 2(p - \alpha_p) \quad (4.57b)$$

$$\frac{\partial F}{\partial q} = \frac{2q}{n^2} \quad (4.57c)$$

$$\frac{\partial F}{\partial \theta} = -\frac{2q^2}{n^3} \frac{dn}{d\theta} = -\frac{2q^2}{n^2} \frac{\sqrt{3} \sin \theta_0 + \cos \theta_0 \sin \phi}{\sqrt{3} \cos \theta_0 - \sin \theta_0 \sin \phi} \quad (4.57d)$$

and  $\partial p/\partial \sigma'$ ,  $q/\partial \sigma'$  and  $\partial \theta/\partial \sigma'$  are given by equations (4.31).

From equations (4.52) and (4.54c), and also (4.49a), obtain

$$\begin{aligned}\frac{\partial F}{\partial \boldsymbol{\varepsilon}^P} &= \left[ -2(p - \alpha_p) \frac{d\alpha_p}{d\varepsilon_v^P} - 2a \frac{da}{d\varepsilon_v^P} \right] \begin{Bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{Bmatrix} \\ &= -\frac{2}{\chi} [(p - \alpha_p)\alpha_p + a^2] \begin{Bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{Bmatrix}\end{aligned}\quad (4.58)$$

Finally, the product  $\{\partial F/\partial \boldsymbol{\varepsilon}^P\}^T \partial Q/\partial \boldsymbol{\sigma}'$  appearing in the constitutive matrix is given by

$$\begin{aligned}\left\{ \frac{\partial F}{\partial \boldsymbol{\varepsilon}^P} \right\}^T \frac{\partial Q}{\partial \boldsymbol{\sigma}'} &= -\frac{2}{\chi} [(p - \alpha_p)\alpha_p + a^2] \left( -\frac{\partial Q}{\partial \sigma'_{11}} - \frac{\partial Q}{\partial \sigma'_{22}} - \frac{\partial Q}{\partial \sigma'_{33}} \right) \\ &= -\frac{2}{\chi} [(p - \alpha_p)\alpha_p + a^2] \frac{\partial Q}{\partial p}\end{aligned}\quad (4.59)$$

For associated flow,  $\partial Q/\partial p = \partial F/\partial p = 2(p - \alpha_p)$ , using (4.57b), so

$$\left\{ \frac{\partial F}{\partial \boldsymbol{\varepsilon}^P} \right\}^T \frac{\partial Q}{\partial \boldsymbol{\sigma}'} = -\frac{4}{\chi} [(p - \alpha_p)\alpha_p + a^2](p - \alpha_p) \quad (4.60)$$

#### 4.5.4 Corners of Yield and Potential Surfaces

The Mohr–Coulomb potential surface and the elliptic  $p$ – $q$ – $\theta$  surface show corners for  $\theta_0 = \pm \pi/6$ , where the plastic strain direction is no longer uniquely defined. In the Mohr–Coulomb case there is a further complexity when  $q = 0$ , at the apex of the potential surface. In the case of associative plastic flow, Drucker [37] concluded that the plastic strain direction is confined by the normals to the conjoint potential surfaces. For simplicity, in the models described here, the plastic strain direction is assumed to lie in the planes for which  $\theta_0 = \pm \pi/6$  and at the apex of the Mohr–Coulomb potential surface, in the negative direction of the  $p$ -axis [38].

#### 4.5.5 Generalised Plasticity

Equations governing the behaviour of elastoplastic materials are usually derived from the introduction of a yield surface  $F$  and a potential surface  $Q$  (Section 4.5.1). Then the

direction of the plastic flow (4.24) is defined by the outer normal to the surface  $Q$ , or by its direction in the stress space characterised by the unit vector

$$\mathbf{n}_q = \frac{\partial Q / \partial \boldsymbol{\sigma}'}{|\partial Q / \partial \boldsymbol{\sigma}'|} \quad (4.61)$$

so that

$$d\boldsymbol{\varepsilon}^p = d\lambda' \mathbf{n}_q \quad (4.62)$$

The amount of plastic strain is then obtained from the consistency rule (4.25), which can be simply rewritten as

$$\mathbf{n}_f^T d\boldsymbol{\sigma}' - H d\lambda' = 0 \quad (4.63)$$

after introducing as a counterpart of equation (4.61) the unit vector which represents in a stress space the direction of the outer normal to the yield surface

$$\mathbf{n}_f = \frac{\partial F / \partial \boldsymbol{\sigma}'}{|\partial F / \partial \boldsymbol{\sigma}'|} \quad (4.64)$$

and the plastic modulus

$$H = - \frac{\{\partial F / \partial \boldsymbol{\varepsilon}^p\}^T \{\partial Q / \partial \boldsymbol{\sigma}'\}}{|\partial F / \partial \boldsymbol{\sigma}'| |\partial Q / \partial \boldsymbol{\sigma}'|} \quad (4.65)$$

The magnitude of plastic strains and the definition of the elastoplastic matrix, first introduced in equation (4.30), follow in a straightforward way:

$$d\boldsymbol{\varepsilon}^p = \frac{1}{H} \mathbf{n}_q \mathbf{n}_f^T d\boldsymbol{\sigma}' \quad (4.66)$$

and

$$\mathbf{D}^{ep} = \mathbf{D}^e - \frac{\mathbf{D}^e \mathbf{n}_q \mathbf{n}_f^T \mathbf{D}^e}{H + \mathbf{n}_f^T \mathbf{D}^e \mathbf{n}_q} \quad (4.67)$$

The unit vector  $\mathbf{n}_f$  also allows a discrimination between the plastic loading, which occurs when  $\mathbf{n}_f^T d\boldsymbol{\sigma}' > 0$ , and the elastic unloading, which occurs when  $\mathbf{n}_f^T d\boldsymbol{\sigma}' < 0$ ; the situation  $\mathbf{n}_f^T d\boldsymbol{\sigma}' = 0$  is defined as a neutral loading.

Soil behaviour can then be described by the direct definitions of the vectors  $\mathbf{n}_f$  and  $\mathbf{n}_q$  and of the plastic modulus  $H$  without explicit reference to any yield or potential surface, as first done by Zienkiewicz and Mroz [39] and then by Pastor *et al.* [40] in the framework of the so-called generalised plasticity models.

According to the approach developed by Pastor *et al.* [40], the vectors  $\mathbf{n}_f$  and  $\mathbf{n}_q$  can be defined for saturated soils in the two-dimensional space of the effective stress

invariants  $p$  and  $q$  as follows [41]:

$$\mathbf{n}_f \equiv \frac{1}{\sqrt{1+d_f^2}} \{d_f; 1\}^T \quad d_f = (1+c)(M_f - \eta) \quad (4.68a, b)$$

$$\mathbf{n}_q \equiv \frac{1}{\sqrt{1+d_q^2}} \{d_q; 1\}^T \quad d_q = (1+c)(M_q - \eta) \quad (4.68a, b)$$

where  $c$ ,  $M_f$  and  $M_q$  are material parameters, and  $\eta$  is given by

$$\eta = \frac{q}{p} \quad (4.69)$$

For an isotropic virgin compression path ( $\eta = 0$ ), equation (4.66) becomes

$$d\varepsilon_v^p = \frac{1}{H_0} \frac{dp}{p} \quad (4.70)$$

where  $H_0$  depends on the material characteristics via the parameters  $\lambda$  and  $k$ , which represent the slopes of the virgin loading and unloading line respectively (Figure 4.14), and the initial void ratio  $e_0$ , i.e.

$$H_0 = \frac{1+e_0}{\lambda-k} \quad (4.71)$$

The occurrence of other stress paths can be explicitly accounted for by introducing a function  $H_v \equiv f(\eta)$  such that

$$H = H_v H_0 p \quad (4.72a)$$

where

$$f(\eta) = 1 \quad \text{at} \quad \eta = 0 \quad (4.72b)$$

and

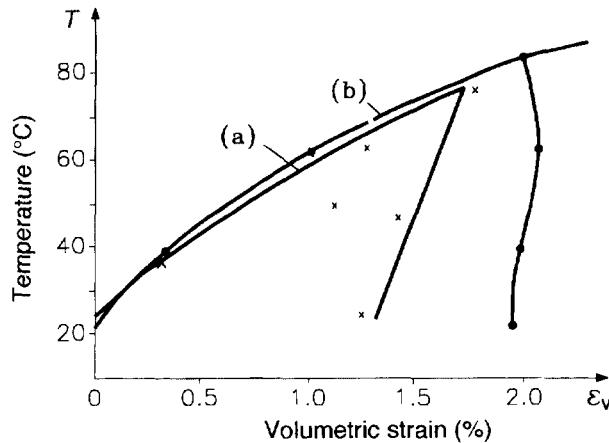
$$f(\eta) = 0 \quad \text{at} \quad \eta = M_f \quad (4.72c)$$

Other factors which affect the plastic response of clays [42] and sands [43] can also be introduced by the additive or multiplicative functions  $H_s$ ,  $H_{dm}$  and  $H_f$ , which respectively account for deviatoric strain hardening, memory of past stress history and possible limit states. The general form of the hardening moduli is

$$H = H_0 p H_f \{H_v + H_s\} H_{dm} \quad (4.73)$$

#### 4.5.6 Thermoplastic Behaviour

The experimental results depicted in Figure 4.16 show that irreversible volume changes occur when a soil is subjected to a temperature increase. These changes are much more



**Figure 4.16** Volumetric strain versus temperature at constant isotropic effective stress, drained test: (a) Pontida clay at 2 MPa, (b) Boom clay at 6 MPa (Reprinted, by permission, from Hueckel, T. and Baldi, G., 1991. Thermoplasticity of saturated clays: experimental constitutive study, *J. Geotech. Engng.*, **116**, 1778–96)

pronounced for normally consolidated clays in the presence of water. This irreversible contraction process, which is similar to the consolidation phenomenon, is essentially due to the breakdown of some intergranular links between the material particles, leading to the rearrangement of voids into a denser structure.

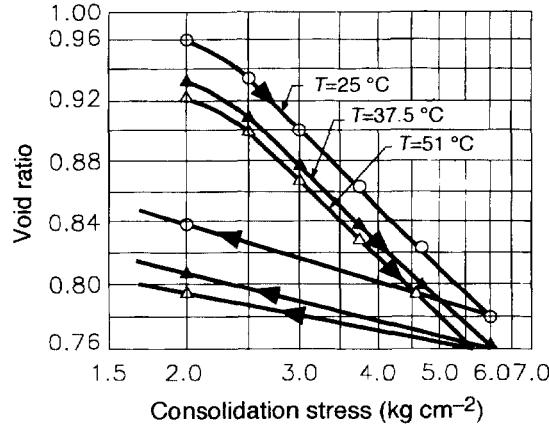
Temperature has a major effect on surface interaction forces, i.e. thermal effects are more pronounced for organic soils (e.g. montmorillonite) than for cohesive inorganic soils (e.g. kaolin or illite); and for the same soil, the amount of consolidation increases with increasing initial void ratio, i.e. looser soils are more compressible.

An increase in temperature also causes a decrease in the shearing strength of the soil structure, since thermal energy acts in conjunction with the shear force at interparticle contacts to increase the probability of bond slippage or failure. However, soil compressibility is quite unaffected by temperature changes [25], as shown by the essentially parallel consolidation curves in Figure 4.17.

On a phenomenological level, the main distinctive properties of the irreversible behaviour of soils during heating are an almost entirely volumetric plastic strain rate, with negligible deviatoric components; the shrinking of the elastic domain and thus the reduction of the yield limit as a function of temperature, called thermal softening; a minor effect on the hardening moduli [45].

For drained heating at constant effective stress, the yield limit must remain fixed in order to sustain the stress level, and thermal softening must then be recovered by strain hardening. In these conditions, a quadratic variation of the irreversible strains with temperature can be assumed (Figure 4.16).

This behaviour may be considered within a thermoplastic theory [44] by introducing the yield function  $F(\sigma', \varepsilon^p, \Delta T)$  which depends also on temperature increments and by considering that the plastic volumetric strains  $\varepsilon_v^p$  are due to both mechanical loading (i.e. effective stress changes) and thermal effects (temperature changes). Then, thermoplastic



**Figure 4.17** Effect of temperature on isotropic consolidation of saturated clay (illite) (Reprinted with permission from Camponella, R. G. and Mitchell, J. K. (1968) [25])

strains are defined via an incremental relationship as a function of stress and temperature increments as follows.

The consistency rule at yielding is as follows,

$$\left\{ \frac{\partial F}{\partial \sigma'} \right\}^T d\sigma' + \left\{ \frac{\partial F}{\partial \epsilon^p} \right\}^T d\epsilon^p + \left( \frac{\partial F}{\partial \epsilon^p} \frac{\partial \epsilon^p}{\partial T} + \frac{\partial F}{\partial T} \right) dT = 0 \quad (4.74)$$

which is similar to equation (4.25).

By taking into account the flow rule, as given by equation (4.24), it is possible to solve equation (4.74) with respect to the plastic multiplier

$$d\lambda = - \frac{\left\{ \frac{\partial F}{\partial \sigma'} \right\}^T d\sigma' + \left( \frac{\partial F}{\partial \epsilon^p} \frac{\partial \epsilon^p}{\partial T} + \frac{\partial F}{\partial T} \right) dT}{\left\{ \frac{\partial F}{\partial \epsilon^p} \right\}^T \frac{\partial Q}{\partial \sigma'}} \quad (4.75)$$

to obtain the plastic strain increment as a function of both  $d\sigma'$  and  $dT$ .

The elastoplastic matrix  $\mathbf{D}^{ep}$  can now be derived. For this purpose, the total strains resulting from the non-isothermal elastoplastic process may be expressed as

$$d\epsilon = d\epsilon^e + d\epsilon^p = [\mathbf{D}^e]^{-1} d\sigma' + \mathbf{m} \frac{\beta}{3} dT + d\lambda \frac{\partial Q}{\partial \sigma'} \quad (4.76)$$

where  $\beta$  is the coefficient of thermal expansion first introduced in Section 4.4.3.

Substituting for  $d\lambda$  (4.75) and rearranging gives

$$d\sigma' = \mathbf{D}^{ep} \left\{ d\epsilon - \left\{ \mathbf{m} \frac{\beta}{3} - \frac{\frac{\partial F}{\partial \epsilon^p} \frac{\partial \epsilon^p}{\partial T} + \frac{\partial F}{\partial T}}{\left\{ \frac{\partial F}{\partial \epsilon^p} \right\}^T \frac{\partial Q}{\partial \sigma'}} \right\} dT \right\} \quad (4.77)$$

where  $\mathbf{D}^{ep}$  results, which is similar in term to (4.30). Note that terms in the brackets on the right-hand side of (4.77), which multiply  $dT$ , just represent the elastic and plastic parts of the thermal strain. A thermoplastic model for saturated clays, based on Cam clay theory [35], has been developed by Hueckel and Borsetto [46].

### 4.5.7 Solution Procedures

In elastoplastic consolidation analyses the Newton–Raphson method, introduced in Section 4.4.4, is often employed, either by the repeated formation and solution of the stiffness matrix as this varies with both the consolidation process and the development of the plastic regions, or by updating the matrix at selected times only.

An alternative to this method is the initial stress method described by Zienkiewicz *et al.* [47]. In this procedure the initial stiffness is assumed to be constant and nodal forces, equivalent to the stress by which the yield surface is exceeded, are elastically redistributed. Since the initial stiffness matrix is used throughout, elastic unloading is automatically achieved. In the context of consolidation analyses this method was first used by Siriwardane and Desai [48].

A fixed-point procedure can also be used for elastoplastic analyses, if corrections are introduced for the nodal forces equivalent to the stress by which the yield surface is exceeded. This was done by Norris [37], and was also successfully used in the elastoplastic code listed in the first edition of this book.

Whichever technique is employed, the accuracy and stability of the global solution is strongly affected by the accuracy and stability of the integration of the stress–strain relation at the Gauss points and by the convergence of the iterative scheme.

The evolution of the stresses, which depends on the evolution of the internal variables accounting for the irreversibility of the plastic flow, is governed by equation (4.28), which can be rewritten as

$$d\sigma' = \mathbf{D}^e \{d\epsilon - d\epsilon^p\} \quad (4.78)$$

where

$$d\epsilon^p = d\lambda' \mathbf{n}_q \quad (4.79)$$

The flow rule (4.79) may also follow from a potential surface  $Q$  (Section 4.5.1).

In any numerical computation, the continuum procedure is approximated by a sequence of small but finite (as opposed to infinitesimal) time steps. Depending on the type of integration scheme, the stress increment within each step can then be evaluated as follows. Care must be taken to evaluate stress changes which are based on displacements after convergence has been achieved, as non-physical stress reversal and unloading [49] can occur during the iterative procedure.

#### 4.5.7.1 Explicit Algorithms

The elastoplastic matrix  $\mathbf{D}^{ep}(\sigma'_n, \epsilon_n^p)$ , as defined by equation (4.30) or by equation (4.65) and evaluated from the known values of stresses  $\sigma'_n$  and plastic strains  $\epsilon_n^p$  at the

beginning of each step, can be used to update the stresses after integration over a load increment or time step  $[t_n, t_{n+1}]$

$$\boldsymbol{\sigma}'_{n+1} = \boldsymbol{\sigma}'_n + \Delta\boldsymbol{\sigma}' \quad (4.80)$$

$$\Delta\boldsymbol{\sigma}' = \mathbf{D}^e(\boldsymbol{\sigma}'_n, \boldsymbol{\varepsilon}_n^p)\Delta\boldsymbol{\varepsilon} \quad (4.81)$$

The explicit integration scheme described previously may be improved by using more accurate integration rules, e.g. a modified Euler scheme, a Runge–Kutta scheme or by sub-incrementation strategies [50–52].

#### 4.5.7.2 Implicit Algorithms

Implicit (return-mapping) integration rules, introduced by Taylor and Simo [53], are widely employed in soil plasticity [54–56].

According to this approach, equation (4.78) is written as

$$\Delta\boldsymbol{\sigma}' = \mathbf{D}^e\{\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^p\} \quad (4.82)$$

$$\text{or } \boldsymbol{\sigma}'_{n+1} = \boldsymbol{\sigma}'_{tr,n+1} - \mathbf{D}^e\Delta\boldsymbol{\varepsilon}^p \quad \boldsymbol{\sigma}'_{tr,n+1} = \boldsymbol{\sigma}'_{n+1} + \mathbf{D}^e\Delta\boldsymbol{\varepsilon} \quad (4.83a, b)$$

where the elastic trial stress  $\boldsymbol{\sigma}_{tr,n+1}$  represents the actual stress at the time  $t_{n+1}$  when no plastic flow develops.

After integration over a load increment, or time step  $[t_n, t_{n+1}]$ , equation (4.79) becomes

$$\Delta\boldsymbol{\varepsilon}^p = \Delta\lambda'[(1-\theta)\mathbf{n}_{q,n} + \theta\mathbf{n}_{q,n+1}] \quad (4.84)$$

which is a simple approximation and stable for  $\theta = 1$  (backward difference). With this latter choice, equations (4.84) and (4.83a) result in

$$[\mathbf{D}^e]^{-1}\{\boldsymbol{\sigma}'_{n+1} - \boldsymbol{\sigma}'_{tr,n+1}\} + \Delta\lambda'\mathbf{n}_{q,n+1} = 0 \quad (4.85)$$

This equation can be solved iteratively along with the consistency rule enforced by equation (4.62) to obtain the actual values of  $\boldsymbol{\sigma}'_{n+1}$  and  $\Delta\lambda'$ .

The tangent matrix, resulting from the linearization of equations (4.85) and (4.62), is now given as

$$\mathbf{K}_{n+1,i} = \begin{bmatrix} [\mathbf{D}^e]^{-1} & \left[ \frac{\partial \mathbf{n}_q}{\partial \boldsymbol{\sigma}'} \right]_{n+1,i} \Delta\lambda'_i & \mathbf{n}_{q,n+1,i} \\ \mathbf{n}_{n+1,i}^T & -H_{n+1,i} \end{bmatrix} \quad (4.86)$$

which has to be updated with successive iterations  $i$ .

#### 4.5.7.3 Consistent Stiffness Matrix

The ‘continuum’ elastoplastic tangent matrix, given by (4.30) or (4.67), is known to destroy the asymptotic quadratic convergence of the Newton–Raphson method. Hence, to improve the computational performance of the integration schemes, particularly when the size of the time step is large, Simo and Taylor [57] proposed to substitute the ‘continuum’ elastoplastic tangent matrix as follows.

At step  $n + 1$  equations (4.78) and (4.79) can be written as

$$d\sigma' = \mathbf{D}^e \{d\epsilon - d\epsilon^p\} \quad (4.87)$$

and

$$d\epsilon^p = d\lambda' \mathbf{n}_{q,n+1} + \lambda' \left[ \frac{\partial \mathbf{n}_q}{\partial \sigma'} \right]_{n+1} d\sigma' \quad (4.88)$$

which correspond to a complete linearization of equation (4.79).

Combining these relations gives

$$d\sigma' = \mathbf{D}_{n+1}^* \{d\epsilon - d\lambda' \mathbf{n}_{q,n+1}\} \quad (4.89)$$

where

$$\mathbf{D}_{n+1}^* = \left[ [\mathbf{D}^e]^{-1} + \left[ \lambda' \frac{\partial \mathbf{n}_q}{\partial \sigma'} \right]_{n+1} \right]^{-1} \quad \text{or} \quad \mathbf{D}_{n+1}^* = \left[ \mathbf{I} + \mathbf{D}^e \left[ \lambda' \frac{\partial \mathbf{n}_q}{\partial \sigma'} \right]_{n+1} \right]^{-1} \mathbf{D}^e \quad (4.90)$$

The consistency rule given by equation (4.63) yields

$$\{\mathbf{n}_{f,n+1}\}^T d\sigma' - H_{n+1} d\lambda' = 0 \quad (4.91)$$

equations (4.91) and (4.89) give

$$d\lambda' = \left( \frac{\mathbf{n}_f^T \mathbf{D}^* d\epsilon}{H + \mathbf{n}_f^T \mathbf{D}^* \mathbf{n}_q} \right)_{n+1} \quad (4.92)$$

and

$$d\sigma' = \left[ \mathbf{D}^* - \frac{\mathbf{D}^* \mathbf{n}_q \mathbf{n}_f^T \mathbf{D}^*}{H + \mathbf{n}_f^T \mathbf{D}^* \mathbf{n}_q} \right]_{n+1} d\epsilon \quad (4.93)$$

where the matrix in square brackets on the right-hand side of equation (4.93) is the expression of the algorithmic (consistent) elastoplastic matrix, which retains the property of asymptotic quadratic convergence of the Newton–Raphson iterative procedure.

## 4.6 PARTIALLY SATURATED MODELS

Partially saturated soils are characterised by the relative content of pore fluids, usually water and gases such as dry air and/or water vapour. These quantities can be directly

measured or can be related to the relative saturations  $S_w$  and  $S_g$  ( $S_w + S_g = 1$ ); e.g. for water content

$$w = S_w e \frac{\rho_w}{\rho_s} \quad (4.94)$$

where  $\rho_w$  and  $\rho_s$  represent the densities of the water and the soil skeleton and  $e$  is the void ratio.

The relative saturation of water depends on the capillary pressure introduced in equation (2.151), which also coincides with the matrix suction  $s$ , through relationship (2.155).

For isothermal conditions, relative saturation and suction can be linked by this relationship [58]:

$$S_w = 1 - m \tanh(ls) \quad (4.95)$$

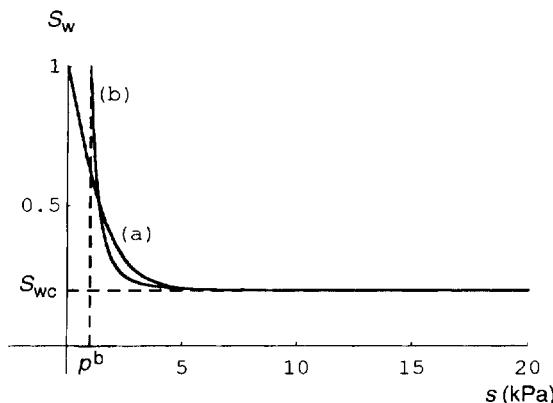
or by this relationship [59]:

$$S_w = S_{wc} + (1 - S_{wc}) \left[ \frac{p^b}{s} \right]^n \quad (4.96)$$

where  $l$ ,  $m$ ,  $n$ ,  $p^b$  and  $S_{wc}$  are material constants.

The material parameter  $S_{wc}$ , introduced in equation (2.168), represents the irreducible saturation, i.e. the limiting value of  $S_w$  as the suction approaches infinity (Figure 4.18, curve b). The corresponding value according to equation (4.95) is given by  $1 - m$  (Figure 4.18, curve a).

The bubbling pressure  $p^b$ , also introduced by equation (2.168), represents the lowest value of suction which corresponds to full saturation ( $S_w = 1$ ). According to equation (4.95), it equals zero independently of the material characteristics.



**Figure 4.18** Relative saturation of water versus suction in partially saturated soils: (a) from Alonso *et al.* [58], (b) from Brooks and Corey [59]

Suction variations directly affect the stresses as measured by equation (2.180). These are sometimes referred to as Bishop stresses, which are responsible for the major deformations of the soil skeleton.

Volume changes can be related to the first invariant of the above stresses, which can be obtained as

$$p = \bar{p} + S_w s \quad (4.97)$$

where  $\bar{p} = p - p^g$ , represents the mean total stress in excess of air pressure. This value is sometimes called the mean net stress and is employed as an alternative measure of stress in partially saturated soils [60]. In fact, a number of constitutive models used in engineering practice [61–63] makes almost exclusive use of total stress in excess of the air pressure and suction.

It can be easily verified that  $p$  and  $\bar{p}$  as defined by equation (4.97) are identical in the case of full saturation since in this case  $s = 0$ .

#### 4.6.1 Elastic Behaviour

The first constitutive models for the description of volume or saturation changes were based on state surfaces in a  $(\bar{p}, q, s)$  space [64], where  $\bar{p}$ ,  $s$  and  $q$  respectively represent net mean stress, suction and deviatoric stresses. Constitutive equations in the same space were given by Lloret and Alonso [65].

However, the essential features of the elastic behaviour of partially saturated soils can be described equally well by a model based on the modified effective stress concept (2.180), together with a non-linear elastic constitutive relationship and capillary pressure relationship for each particular soil [66]. This model makes use of the fact that effective stresses and suction are independent variables, and is defined in the  $(p, q, s)$  space.

A direct link between the reversible changes of volumetric strain and effective pressure (4.97) can be introduced, as in consolidation analyses, by assuming that

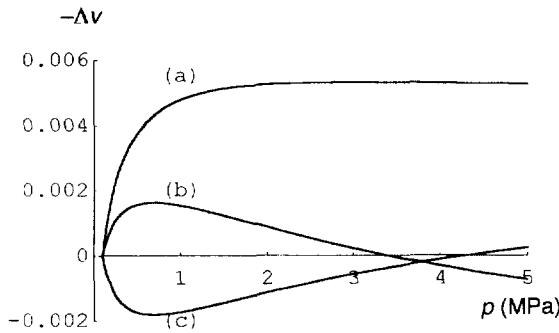
$$d\varepsilon_v = \kappa \frac{dp}{p} \quad (4.98)$$

where the parameter  $\kappa$  is a material characteristic.

The substitution of  $p$  by  $\bar{p}$  and  $s$  via equations (4.97) and (4.95) or (4.96) enables a state surface to be defined, after integration, one that closely reproduces the state surface introduced by Matyas and Radhakrishna [64], as shown by Bolzon and Schrefler [66].

A dependence of the parameter  $\kappa$  on the suction  $s$  and/or pressure  $p$  can also be considered to account for the observed soil stiffness changes with suction. Also, the behaviour of loose soils, which may show both swelling and collapse with suction variations, can be accounted for by this fact.

These phenomena are usually related to the irreversible behaviour of soil, and are better described within the context of an elastoplastic constitutive theory. However, if conditions of monotonic loading from the virgin state exist, then the constitutive equations can be integrated as for non-linear elasticity.



**Figure 4.19** Volume changes due to suction decrease at constant effective pressure: (a)  $a_1 = 0.3 \text{ MPa}^{-1}$ ,  $a_0 = -0.015 \text{ MPa}^{-1}$ ; (b)  $a_1 = 0.3 \text{ kPa}^{-1}$ ,  $a_0 = -0.15 \text{ MPa}^{-1}$ ; (c)  $a_1 = -0.3 \text{ MPa}^{-1}$ ,  $a_0 = 0.15 \text{ MPa}^{-1}$

A variety of behavioural patterns can then be easily simulated by assuming

$$\kappa = \frac{k}{1 + [a_0 + a_1 \exp(-p)]s} \quad (4.99)$$

where  $\kappa$  reduces to  $k$  for fully saturated conditions ( $s = 0$ ).

The curves a, b and c of Figure 4.19 show the volume changes resulting from suction decrease at constant effective pressure as a function of the effective pressure itself, by keeping constant all the data but varying the parameters  $a_0$  and  $a_1$ .

#### 4.6.2 Plastic Behaviour

Some typical features of the irreversible behaviour of partially saturated soils are shown in Figure 4.20 and 4.21.

Figure 4.20 plots the specific volume changes, under isotropic compression and at constant suction, of a compacted kaolin [62] as a function of mean total stress in excess of air pressure. If the suction is kept constant, the relationship between the pressure and volume changes of the partially saturated soil is similar to that for saturated soils [68], but the soil compressibility does depend on suction.

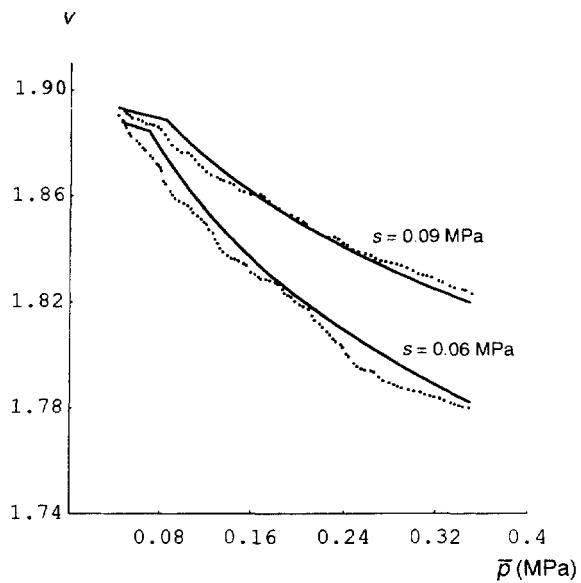
Figure 4.21 represents the specific volume changes of clay [69] under suction changes at constant external pressure. On suction decrease (wetting) the material exhibits first swelling and then collapse.

This behaviour can be described in a generalised plasticity framework [70] by using Bishop's mean stress and suction as pressure values.

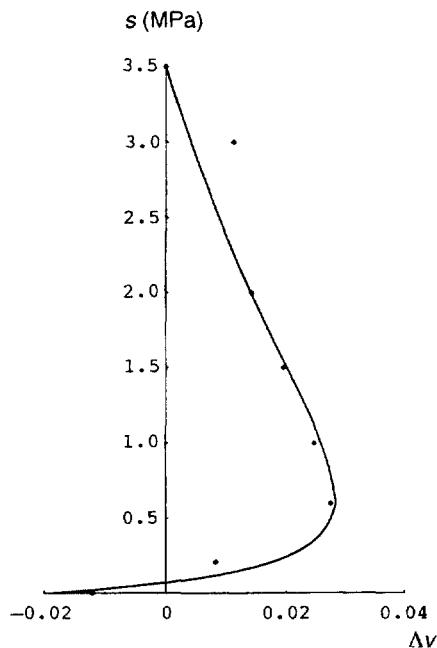
The dependence of soil stiffness on suction can be modelled by modifying the hardening modulus under isotropic conditions (4.65 and 4.70) through the introduction of a multiplicative function  $H_w$ , which in its simplest form linearly relates changes in stiffness with changes in suction:

$$H = H_0 H_w p \quad H_w = 1 + as \quad (4.100)$$

where  $a$  is a material parameter.



**Figure 4.20** Specific volume versus mean total stress over air pressure for isotropic compression at constant suction: experimental data from Alonso *et al.* [58]



**Figure 4.21** Specific volume changes versus suction for a suction change at constant pressure: experimental data from Escario and Sáez [69]

The slope of the virgin loading line changes accordingly (4.71 and 4.98):

$$\lambda(s) = \frac{1 + e_0}{H_0 H_w} + \kappa = \frac{1 + e_0}{H_0(1 + as)} + \kappa \quad (4.101)$$

The function  $H_w$  can be refined, if required, to fit different experimental data. As an example, a two-parameter exponential function can be used [62]:

$$H_w = [1 + b_1(\exp(-b_2s) - 1)]^{-1} \quad (4.102)$$

In some cases, a dependence of  $H_w$  on the pressure  $p$  must also be assumed in order to describe the behaviour of loose soils which exhibit a maximum of collapse on wetting at some critical pressure [67]. This feature can be dealt with by the above formulation by assuming that the dependence of the parameter  $a$  on the effective mean stress  $p$  through the function

$$a = a_1 \exp(-p) - a_2 \quad (4.103)$$

Hence, as in equation (4.99)

$$H = H_0[1 + (a_1 \exp(-p) - a_2)s]p \quad (4.104)$$

The ratio between the material constants  $a_1$  and  $a_2$  gives the critical value of pressure at which the maximum collapse is detected, and the total amount of collapse depends on the absolute value of these parameters.

The isotropic yield limit of partially saturated soils also increases with suction [71]. This effect can be accounted for once the initial yield stress  $p_{y0}$  for the saturated condition and some values of the yield limit at different suction values  $p_y(s)$  are given by experiments. A linear regression of data can often be considered satisfactory [70], and it is possible to define that parameter  $i$  such that

$$p_y(s) = p_{y0} + is \quad (4.105)$$

Since the evolution of the yield surface during the plastic process is governed by plastic strains, a reference value  $p_c$  of Bishop stress can also be introduced to characterise the situation when no plastic strains develop as a consequence of suction changes. In this case the isotropic yield limit changes according to the following relationship:

$$\frac{p_y(s)}{p_c} = \left[ \frac{p_{y0}}{p_c} \right]^{H_w(s)} = \left[ \frac{p_{y0}}{p_c} \right]^{\frac{\lambda(s) + \kappa}{\lambda(s) + \kappa}} \quad (4.106)$$

which is analogous to the evolution law introduced by Alonso *et al.* [62] in terms of the net mean stress instead of the Bishop stress.

## 4.7 VARIATION OF PERMEABILITY

Until now only the non-linear relationship between relative permeability and saturation has been taken into account. However, there is experimental evidence that, even in saturated soil with only one fluid phase flowing, the permeability is not constant.

In groundwater flow literature, the product

$$\mathbf{K} = \frac{\mathbf{k}\gamma^w}{\mu^w} \quad (4.107)$$

is usually called permeability, where  $\mu^w$  is the dynamic viscosity of water and  $\gamma^w$  is the specific weight of water.

De Wiest [72] has shown that  $\gamma^w$  is a function of the pressure

$$\gamma^w = \gamma_0^w \exp\left(\frac{p_0 + p}{K_w}\right) \quad (4.108)$$

where  $K_w$  is the bulk modulus of water.

The effect of the dependence of the permeability  $\mathbf{K}$  on the pressure, via the specific weight of water, was investigated by Gambolati [73] in the case of one-dimensional vertical flow of groundwater.

It was found that the effects on the pressure field are negligible unless very thick formation ( $>10^4$  m) and very high boundary pressures ( $>5 \times 10^5$  MPa) are considered. Such values are extremely unrealistic and therefore this source of non-linearity is disregarded.

The following research workers investigated other forms of non-linear permeability in their numerical models.

Finol and Farouq Ali [74] considered the case of a black oil system, where the water was immobile. The variation of porosity and permeability with pressure satisfied the following relationship:

$$\phi^{m+1} = \phi^m[1 + c_p(p_o^{m+1} - p_o^m)] \quad (4.109a)$$

$$k^{m+1} = k^m[1 + c_k(p_o^{m+1} - p_o^m)] \quad (4.109b)$$

with

$$c_p = \frac{c_m + (1 - \phi)c_r}{\phi} \quad (4.109c)$$

where  $c_m$  = the uniaxial compaction coefficient

$c_r$  = the rock matrix compressibility

$c_k$  = a coefficient of permeability reduction

$p_o$  = the pressure in the oil phase

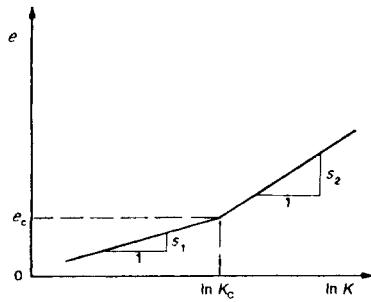
The uniaxial compaction coefficient  $c_m$  is defined as the formation compaction per unit change in pore pressure reduction and can be obtained from laboratory compressibility data by

$$c_m = \frac{1}{3} \frac{1+v}{1-v} (1-\beta)c_b \quad (4.110)$$

where  $v$  = Poisson's ratio

$\beta$  = the ratio of rock matrix and rock bulk compressibility

$c_b$  = the bulk compressibility



**Figure 4.22** Relationship between void ratio and the logarithm of the permeability

Because of the lack of appropriate experimental data,  $c_k$  was arbitrarily taken to be equal to the uniaxial compaction coefficient.

In the case of one-dimensional consolidation, Monte and Kritzen [75] found experimentally that a bilinear relationship between void ratio  $e$  and logarithm of permeability  $K$  represents soil behaviour fairly well. Such a relationship is shown in Figure 4.22.

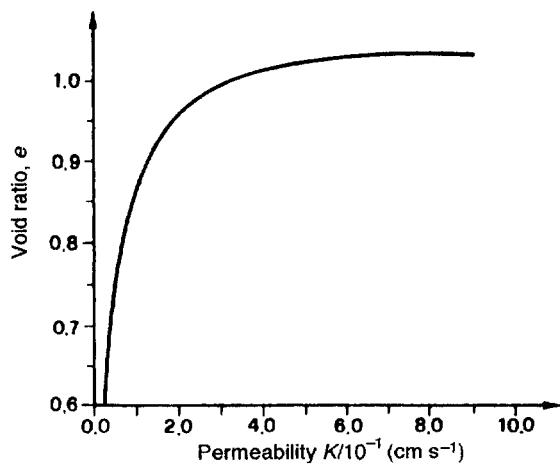
If  $K_c$  and  $e_c$  are the critical values of permeability and void ratio at which the slope changes (Figure 4.22) and if the initial void ratio is  $e_0$ , then for a void ratio  $e$  the permeability  $K$  is given by

$$\ln \left( \frac{K}{K_c} \right) = \frac{e - e_c}{S} \quad (4.111)$$

where

$$S = S_1 \quad (e < e_c)$$

$$S = S_2 \quad (e > e_c)$$



**Figure 4.23** Relationship between void ratio and permeability, obtained experimentally

Since for more than one dimension the variation of  $K$  is less certain, this procedure has been used numerically by Norris [37] for the vertical direction only, which is assumed to be the direction of maximum strain.

The procedure utilised in this text was proposed by Lewis *et al.* [76], where an experimental permeability–void ratio relationship was determined in the laboratory. An example of such a relationship is shown in Figure 4.23.

This relationship is represented by as many pairs of data points as required and intermediate values are obtained by interpolation. The model is fully defined by the following settlement–void ratio relationship:

$$\delta = d \frac{e_1 - e_2}{1 + e_1} \quad (4.112)$$

where  $\delta$  = the average settlement of each element

$d$  = the initial depth

$e_1$  = the initial void ratio

$e_2$  = the final void ratio

Both relationships apply to one-dimensional consolidation. Their validity to model the variation of the permeability  $K$  in more than one dimension is less certain and requires further research. The variable permeability scheme is of practical use only if the expected consolidation is important; examples are given in Chapter 14.

## 4.8 CONCLUSIONS

This chapter has outlined linear and non-linear elastic and elastoplastic constitutive relationships suitable for application in consolidation analysis and modelling of surface subsidence. Variable permeability schemes have also been discussed. The state is now set for the application of the procedures outlined in Chapters 2 and 3, together with the constitutive relationships. This is done in the following three chapters. Chapter 5 deals with physical loading of soil layers. Existing solutions are extensively used for testing the programs based on these procedures. In Chapter 6 and 7 the theory is applied to case histories of surface subsidence and to parametric investigations. The difficulties encountered in such simulations of observed field histories will be emphasised.

## REFERENCES

1. Aubry, D. and Kodaissi, E. (1984) Two-phase implicit scheme for viscoplastic soil consolidation, in *Numerical Methods for Transient and Coupled Problems*, R. W. Lewis, E. Hinton, B. Bettess and B. A. Schrefler (eds). Pineridge Press, Swansea, pp. 498–508.
2. Naylor, D. J. (1975) Non-linear finite element models for soils., PhD thesis, University College of Swansea.
3. Schrefler, B. A., Lewis, R. W. and Majorana, C. E. (1982). The problem of rebound in subsidence models, in *Finite Element Flow Analysis*, T. Kawai (ed.), University of Tokyo Press, Tokyo pp. 689–96.

4. Nayak, G. C. and Zienkiewicz, O. C. (1972) Convenient form of stress invariants for plasticity. *Proc. ASCE*, **98** (ST4), 949–53.
5. Hill, R. (1950) *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford.
6. Clough, R. W. and Rashid, Y. (1965) Finite element analysis of axi-symmetric solids. *J. Engng Mech. Div. ASCE*, **91** (EM1), 71–85.
7. Brown, C. B. and King, I. P. (1966) Automatic embankment analysis: equilibrium and stability conditions. *Géotechnique*, **16**(3), 209–19.
8. Geertsma, J. (1966) Problems of rock mechanics in petroleum production engineering, in *Proceedings of the 1st Conference of the International Society of Rock Mechanics*, Lisbon, Vol. 1, pp. 585–94.
9. Sandhu, R. S. and Wilson, E. L. (1969) Finite element analysis of seepage in elastic media. *Proc. J. Engng Mech. Div. ASCE*, **95** (EM3), 641–52.
10. Penman, A. D. M., Burland, J. B. and Charles, J. A. (1971) Observed and predicted deformations in a large embankment dam during construction. *Proc. Inst. Civ. Eng.*, **49**, 1–21.
11. Cole, N. H. and Burland, J. B. (1972) Observations of retaining wall movements associated with large excavations, in *Proceedings of the 5th European Conference on Soil Mechanics and Foundation Engineering*, Madrid.
12. Gambolati, G., Gatto, P. and Freeze, R. A. (1974) Mathematical simulation of the subsidence of Venice. 2 Results. *Water Resources Research*, **10**(3), 563–77.
13. Naylor, D. J., Pande, G. N., Simpson, B. and Tabb, R. (1981) Critical state model, in *Finite Elements in Geotechnical Engineering*, Pineridge Press, Swansea, Ch. 7.
14. Dunlop, P. and Duncan, J. M. (1970) Development of failure around excavated slopes. *J. Soil Mech. Found. Engng Div. ASCE*, **96** (SM2), 471–93.
15. Hoyaux, B. and Landanyi, B. (1970) Cavitational stress field around a tunnel in soft clay. *Can. Geotech. J.*, **7**, 54–61.
16. Lewis, R. W. and Schrefler, B. A. (1978) A fully coupled consolidation model of the subsidence of Venice. *Water Resources Research*, **14**, 223–30.
17. Schrefler, B. A. (1978) Investigation of land subsidence in the lower Po Valley using a finite element consolidation model, in *Proceedings of the xx<sup>th</sup> ICES Users Group Worldwide Conference*, Padova; *ICES J.*, **10**(3), 39–57.
18. Kondner, R. L. (1963) Hyperbolic stress-strain response: cohesive solids. *J. Soil. Mech. Found. Div. ASCE*, **89** (SM1), 115–43.
19. Duncan, J. M. and Chang, C. Y. (1970) Nonlinear analysis of stress and strain in soils. *J. Soil Mech. Found. Div. ASCE*, **96** (SM5), 1629–53.
20. Janbu, N. (1963) Soil Compressibility as determined by oedometer and triaxial tests, in *Proceedings of the European Conference on Soil Mechanics and Foundation Engineering*, Wiesbaden, Vol. 1, pp. 19–25.
21. Mazzalai, P., Ricceri, G. and Schrefler, B. A. (1978) Studio della subsidenze nel Delta Padano, in *Proceedings of the conference I problemi della subsidenze nella politica del territorio e della difesa del suolo*, Pisa, Vol. 4, pp. 50–64.
22. Kulhawy, F. H. and Duncan, J. M. (1972) Stresses and movements in Oroville dam. *Proc. ASCE*, **98** (SM7), 653–65.
23. Desai, C. S. and Holloway, D. M. (1971) Discussion on ‘Nonlinear analysis of stress and strain in soils’ by Duncan and Chang. *J. Soil Mech. Found. Engng Div. ASCE*, **97** (SM5), 817–20.
24. Desai, C. S. (1971) Non-linear analysis using spline functions. *J. Soil. Mech. Found. Engng Div. ASCE*, **97** (SM10), 1461–80.
25. Campanella, R. G. and Mitchell, J. K. (1968) Influence of temperature variations on soil behaviour. *J. Soil Mech. Found. Div. ASCE*, **94** (SM3), 709–34.
26. Baldi, G., Hueckel, T. and Pellegrini, R. (1988) Thermal volume changes of the mineral–water system in low-porosity clay soils. *Can. Geotech. J.*, **25**, 807–25.

27. Girijavallabhan, C. V. and Resse, L. C. (1968) Finite element method for problems in soil mechanics. *J. Soil Mech. Found. Engng Div ASCE*, **94** (SM2), 473–96.
28. Duncan, J. M., Monismith, C. L. and Wilson, E. L. (1968) Finite element analyses of pavements, in *Proceedings of the 1968 Annual Meeting of the Highway Research Board*.
29. Zienkiewicz, O. C. (1977) *The Finite Element Method*, McGraw-Hill, London.
30. Humphezon, C. (1976) Finite element analysis of elasto-visco-plastic soils. PhD thesis, University College of Swansea.
31. Coulomb, C. A. (1776) Essai sur une application des règles des maximis et minimis à quelques problèmes de statique relatifs à l'architecture. *Mem. Acad. Roy. Pres. Divers Sav.* **5**, 7.
32. Roscoe, K. H. and Poorooshah, H. B. (1963) A theoretical and experimental study of strains in triaxial compression tests on normally consolidated clays. *Géotechnique*, **13**(1), 12–34.
33. Roscoe, K. H. Schofield, A. N. and Thurairajah, A. (1963) Yielding of clays in states wetter than critical. *Géotechnique*, **13**(3), 211–40.
34. Schofield, A. N. and Wroth, C. P. (1968) *Critical State Soil Mechanics*, McGraw-Hill, New York.
35. Roscoe, K. H. and Burland, J. B. (1968) On the generalised stress-strain behaviour of 'wet' clay, in *Engineering Plasticity*, Heyman, J., Lechie, F. A. Cambridge University Press, Cambridge, pp. 535–609.
36. Zienkiewicz, O. C. Humphezon, C. and Lewis, R. W. (1975) Associated and non-associated viscoplasticity and plasticity in soil mechanics. *Géotechnique*, **25**(4), 671–89.
37. Drucker, D. C. (1959) A definition of stable inelastic material. *Trans. ASME*, **26**, 00–00.
38. Norris, V. A. (1980) The elasto-plastic analysis of soil consolidation with special reference to kinematic hardening. PhD thesis, University College of Swansea.
39. Zienkiewicz, O. C. and Mroz, Z. (1984) Generalized plasticity formulation and applications to geomechanics, in *Mechanics of Engineering Materials*, C. S. Desai and R. H. Gallagher (eds), Wiley, Chichester, pp. 655–79.
40. Pastor, M., Zienkiewicz, O. C. and Chan, A. H. C. (1990) Generalized plasticity and the modelling of soil behaviour. *Int. J. Num. Anal. Meth. Geomech.*, **14**, 151–90.
41. Nova, R. (1982) A constitutive model for soil under monotonic and cycling loading, in *Soil Mechanics: Transient and Cycling Loads*, G. N. Pande and O. C. Zienkiewicz (eds), Wiley, Chichester, pp. 343–75.
42. Zienkiewicz, O. C., Leung, K. H. and Pastor, M. (1985) Simple model for transient soil loading in earthquake analysis. I. Basic model and its application. *Int. J. Num. Anal. Meth. Geomech.*, **9**, 453–476.
43. Pastor, M., Zienkiewicz, O. C. and Leung, K. H. (1985) Simple model for transient soil loading in earthquake analysis II. Non-associative models for sands. *Int. J. Num. Anal. Meth. Geomech.*, **9**, 477–98.
44. Prager, W. (1958) Non-isothermal plastic deformation. *Bol. Konink. Nederl. Acad. Wet.*, **8**(61/3), 176–82.
45. Hueckel, T. and Baldi, G. (1991) Thermoplasticity of saturated clays: experimental constitutive study. *J. Geotech. Engng.*, **116**(12), 1778–96.
46. Hueckel, T. and Borsetto, M. (1991) Thermoplasticity of saturated soils and shales: constitutive equations. *J. Geotech. Engng.*, **116**(12), 1765–77.
47. Zienkiewicz, O. C., Valliappan, S. and King, I. P. (1968) Elasto-plastic solutions of engineering problems: initial stress finite element approach. *Int. J. Num. Meth. Engng*, **1**, 75–100.
48. Siriwardane, H. J. and Desai, C. S. (1981) Two numerical schemes for nonlinear consolidation. *Int. J. Num. Meth. Engng.*, **17**, 405–26.
49. Zienkiewicz, O. C. and Taylor, L. R. (1992) *The Finite Element Method*, Vol. 2, McGraw-Hill, New York.
50. Nayak, G. C. and Zienkiewicz, O. C. (1972) Elasto-plastic stress analysis: generalization for various constitutive relations including strain softening. *Int. J. Num. Meth. Engng.*, **5**, 113–35.

51. Sloan, S. W. (1987) Substepping schemes for the numerical integration of elasto-plastic stress-strain relations. *Int. J. Num. Meth. Engng.*, **24**, 893–911.
52. Chan, A. H. C. (1988) A unified finite element solution to static and dynamic problems in geomechanics. PhD thesis, University College of Swansea.
53. Taylor R. L. and Simo, J. C. (1986) A return mapping algorithm for plane stress elasto-plasticity. *Int. J. Num. Meth. Engng.*, **22**, 649–76.
54. Borja, R. I. and Lee, S. R. (1990) Cam-clay plasticity, part II: implicit integration of elasto-plastic constitutive relations. *Comp. Meth. Appl. Mech. Engng.*, **78**, 49–72.
55. Borja, R. I. (1991) Cam-clay plasticity, part I: implicit integration of constitutive relations based on a nonlinear elastic stress predictor. *Comp. Meth. Appl. Mech. Engng.*, **88**, 225–40.
56. Alawaji, H., Runesson, K., Sture, S. and Axelsson, K. (1992) Implicit integration in soil plasticity under mixed control for drained and undrained response. *Int. J. Anal. Num. Meth. Geomech.*, **16**, 737–56.
57. Simo, J. C. and Taylor, R. L. (1985) Consistent tangent operators for rate dependent elasto-plasticity. *Comp. Meth. Appl. Mech. Engng.*, **48**, 101–18.
58. Alonso, E. E., Gens, A. and Josa, A. (1990). A constitutive model for partially saturated soils *Géotechnique*, **40**(3), 405–30.
59. Brooks, R. N. and Corey, A. T. (1966) Properties of porous media affecting fluid flow. *J. Irrig. Drain. Div. Amer. Soc. Civ. Engng.*, **92** (IR2), 61–68.
60. Fredlund, D. G. and Morgenstern, N. R. (1977) Stress state variables for unsaturated soils. *J. Geotech. Engng Div. ASCE*, **103** (GT5), 447–66.
61. Josa, A. (1988) An elastoplastic model for partially saturated soils. PhD thesis (in Spanish). ETSICCP, Barcelona.
62. Alonso, E. E., Gens, A. and Josa, A. (1990) A constitutive model for partially saturated soils. *Géotechnique*, **40** (3), 405–30.
63. Josa, A., Balmaceda, A., Alonso, E. E. and Gens, A. (1992) An elastoplastic model for partially saturated soils exhibiting a maximum of collapse, in *Computational Plasticity: Fundamentals and Applications*, Owen D. R. J. Oñate E. and Hinton E. (eds), Pineridge Press, Swasea in association with Centre International de Methods Numerical en Ingenieria, Barcelona pp. 815–26.
64. Matyas, E. L. and Radhakrishna, H. S. (1968) Volume change characteristics of partially saturated soils. *Géotechnique*, **18**(4), 432–48.
65. Lloret, A. and Alonso, E. E. (1985) State surfaces for partially saturated soils, in *Proceedings of the 11th International Conference on Soil Mechanics and Foundation Engineering*, San Francisco, Vol. 2, pp. 557–62.
66. Bolzon, G. and Schrefler, B. A. (1993) State surfaces of partially saturated soils: an effective pressure approach, in *Proceedings of Euromech Colloquium 290: Mechanics of Swelling*.
67. Lloret, A. and Alonso, E. E. (1980) Consolidation of unsaturated soil including swelling and collapse behaviour. *Géotechnique*, **30**(4), 449–77.
68. Karube, D. (1986) New concept of effective stress in unsaturated soil and its proving tests. *ASTM Symposium on Advanced Triaxial Testing of Soil and Rock*, ST **977**, 539–52.
69. Escario, V. and Sáez, J. (1973) Measurement of the properties of swelling and collapsing soils under controlled suction, in *Proceedings of the 3rd ICES*, Haifa, pp. 195–200.
70. Bolzon, G., Schrefler, B. A. and Zienkiewicz, O. C. (1996) Elastoplastic soil constitutive laws generalized to partially saturated states. *Géotechnique*, **46**, 279–289.
71. Fredlund, D. G., Morgenstern, N. R. and Widger, R. A. (1978) The shear strength of unsaturated soils. *Can. Geotech. J.*, **15**, 313–21.
72. De Wiest, R. J. M (1966) On the storage coefficient of ground water flow. *J. Geophys. Res.*, **71**(4), 1117–22.
73. Gambolati, G. (1973) Equation for one-dimensional vertical flow of groundwater. 2. Validity range of the diffusion equation, *Water Resources Research*, **9**(5), 1385–95.

74. Finol, A. and Farouq Ali, S. M. (1975) Numerical simulation of oil production with simultaneous ground subsidence. *SPEJ*, **15**, 411–24.
75. Monte, J. L. and Kritzen, R. J. (1976) One-dimensional mathematical model for large-strain consolidation. *Géotechnique*, **26**(3), 495–510.
76. Lewis, R. W., Roberts, G. K. and Zienkiewicz, O. C. (1976) A non-linear flow and deformation analysis of consolidation problems, in *Proceedings of the 2nd International Conference on Numerical methods in Geomechanics*, Blacksburg WV, pp. 1106–18.

# 5

## Verification of Elastic and Elastoplastic Consolidation Programs

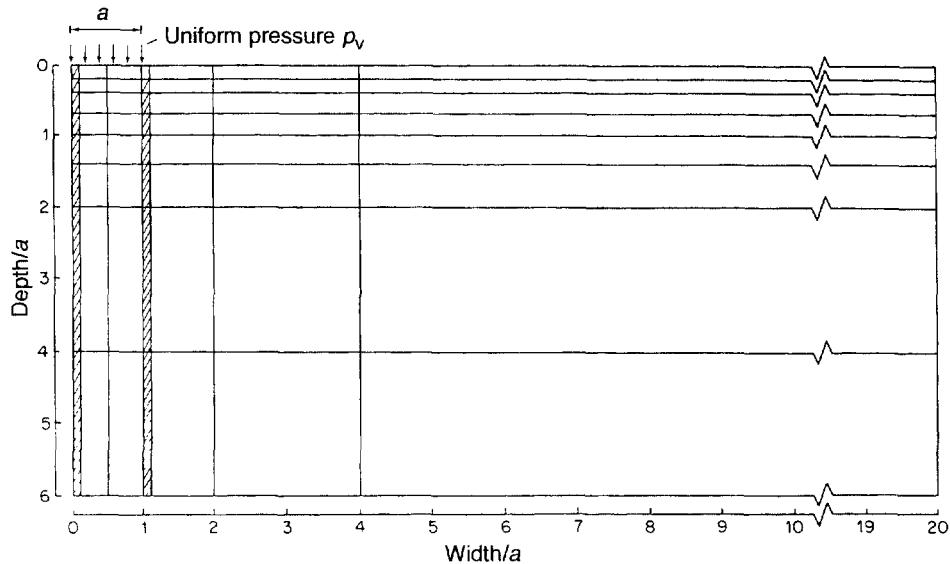
### 5.1 INTRODUCTION

This chapter presents a set of tests to check the essential features of elastic and elastoplastic consolidation programs for a wide spectrum of applications.

The treatment covers fully and partially saturated conditions; along with the consolidation aspect, it considers limiting cases such as drained and undrained situations. The tests contained in Sections 5.2 to 5.4 were devised by Norris [1]. First, a comparison is made between the program predictions for elastic soil behaviour and exact solutions cited by Poulos and Davis [2]. Both plane strain and radial symmetry solutions are considered for drained and undrained conditions. Since an exact solution is not expected from finite element programs, further assurance of their correctness has been obtained by performing the same tests with an established elastoviscoplastic finite element program developed by Cormeau [3] and Humpheson [4]. Next the consolidation aspect of the programs is validated using an elastic analytical solution for a footing [5].

Elastoplastic analyses cannot in general be so rigorously checked. Exact and limit solutions for failure loads are available only for the case of simple yield surfaces (e.g. von Mises, Tresca and Mohr–Coulomb) and displacements can be predicted exactly only for simple geometry. It was decided that the most reliable way of testing the consolidation programs was by comparison with the elastoviscoplastic program. Elastoviscoplastic predictions of failure compare well with analytical solutions [6], showing that for Mohr–Coulomb analyses the program is very reliable. The same has not been shown for analyses involving the critical state ellipse, for which a simple triaxial computation serves for comparison with a hand calculation. The consolidation programs and the viscoplastic program have been developed along completely different lines yet they give similar results for a footing test; this is strong evidence of the validity of both models.

The consolidation aspect of the elastoplastic program is then validated by comparison with a one-dimensional and a two-dimensional plane strain solution published by Siriwardane and Desai [7]. The effects on pore pressure and settlements of considering a variable permeability are investigated in Section 5.6 by means of two examples [8]. One of them allows for comparison with existing experimental data.



**Figure 5.1** Finite element mesh used for computer analyses: Gauss points to the left of shaded elements are used to represent states below the centre and edge of the loaded area

Finally, a benchmark problem for partially saturated situation is shown in Section 5.7. The solutions obtained with the model of Section 3.8, where only one-phase flow is modelled, are compared with those of the full model of Section 3.11. Attention is paid to the transition from fully saturated to partially saturated states. This transition causes some difficulties when using the full two-phase flow formulation of Section 3.11.

All the analyses other than those of triaxial tests and up to Section 5.5 employ the finite element mesh shown in Figure 5.1 (modified for the consolidation tests). The meshes of the subsequent examples are shown in the respective sections.

## 5.2 ELASTIC SOLUTIONS FOR DRAINED AND UNDRAINED CONDITIONS

Fully drained and undrained conditions are contained as limit situations [6] in the general formulation given in Chapters 2 and 3.

### *Fully Drained, Steady-state Conditions*

Provided the pressures and displacements reach constant values, the second equation of (3.46) reduces to

$$\mathbf{H}\mathbf{p}^w = \mathbf{f}^p \quad (5.1)$$

and the pressures can be independently determined, assuming of course that the permeability is known and is independent of strain.

The first equation of (3.46) is then given by

$$\mathbf{K}d\bar{\mathbf{u}} = d\mathbf{f} + \mathbf{Q}d\bar{\mathbf{p}}^w \quad (5.2)$$

Equation (5.2) essentially represents an incremental form of a non-linear mechanics problem with pore pressure forces  $\mathbf{Q}d\bar{\mathbf{p}}^w$  added in. Standard programs and methods used in such solutions can be directly applied once the constitutive relations are known. For a simple elastic medium, integration gives

$$\mathbf{K}\bar{\mathbf{u}} = \mathbf{f} + \mathbf{Q}\bar{\mathbf{p}}^w \quad (5.3)$$

showing the ‘forces’ due to pore pressure.

#### *Instantaneous Load (undrained conditions)*

If equations (3.46) are multiplied by  $dt$  and  $dt \rightarrow 0$ , we note that all finite terms disappear:

$$\begin{bmatrix} -\mathbf{K} & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} \end{bmatrix} \begin{Bmatrix} d\bar{\mathbf{u}} \\ d\bar{\mathbf{p}}^w \end{Bmatrix} = \begin{Bmatrix} -d\mathbf{f} \\ \mathbf{0} \end{Bmatrix} \quad (5.4)$$

Again this is an incremental form of the well-known nearly incompressible (or, if  $\mathbf{S} = 0$ , incompressible) solid formulation, and once again for linear elastic behaviour, integration will yield

$$\begin{bmatrix} -\mathbf{K} & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}} \\ p^w \end{Bmatrix} = \begin{Bmatrix} -\mathbf{f}^u \\ \mathbf{0} \end{Bmatrix} \quad (5.5)$$

a form developed originally by Herrmann [9] and used subsequently by others [10] for the solution of incompressible elastic solids.

In Sections 5.2.1 and 5.2.2 computer analyses for both plane strain and radial symmetry conditions are compared with exact solutions and with the results given by the elastoviscoplastic program.

#### **5.2.1 Plane Strain, Uniform Loading**

One solution was obtained for drained conditions with Poisson’s ratio ( $v$ ) equal to 0.2 and one for undrained conditions with an effective Poisson’s ratio of 0.4 (taking into account the overall compressibility of the saturated soil). Vertical and horizontal displacements below the edge of the loading strip are compared with both the results of the viscoplastic program and the values given by an exact analysis in Table 5.1 [11]. Vertical, bulk and shear stresses at various depths below the edge of the loading strip are compared in Figure 5.2 for  $v$  equal to 0.2.

The results show that the computer programs are in almost exact agreement with each other. They agree well with the analytical solution regarding stresses computed at Gauss

**Table 5.1** Comparison of computed and analytical results for displacements beneath strip and circular loading<sup>a</sup>

Plane strain	Radial symmetry	Drained	Undrained	E	v	K <sub>f</sub> /η	E <sub>e</sub>	v <sub>e</sub>	Computed results × 10 <sup>-2</sup>			Analytical results × 10 <sup>-2</sup>			
									δ <sup>V</sup> <sub>edge</sub>	δ <sup>H</sup> <sub>edge</sub>	δ <sup>V</sup> <sub>centre</sub>	δ <sup>V</sup> <sub>edge</sub>	δ <sup>H</sup> <sub>edge</sub>	δ <sup>V</sup> <sub>centre</sub>	
*	*	*		200	0.2	0	200	0.2	2.06	0.571	2.88	1.91	0.573	—	
*		*	*	200	0.1	341	254.6	0.4	1.21	0.128	1.78	1.16	0.143	—	
*	*	*		200	0.3	0	200	0.3	0.923	0.242	1.59	0.916	—	1.54	
*	*	*		200	0.45	0	200	0.45	0.772	0.0562	1.35	0.748	—	1.27	
*		*	*	200	0.1	796	263.6	0.45	0.586	0.0433	1.03	0.567	—	0.967	
									0.526	0.0057	0.932				
*		*	*	200	0.1	8778	271.8	0.495		0.524	0.0074	0.924	—	—	—
*				—	—	—	271.8	0.5	—	—	—	—	—	0.935	

<sup>a</sup> Where two sets of computed results are given, the upper set is from the consolidation program, the lower set from the viscoplastic program. In all other cases the two programs gave results which were identical, to three significant figures. The notation is as follows:

E = elastic modulus of soil skeleton

E<sub>e</sub> = effective total elastic modulus

v = Poisson's ratio of the soil skeleton

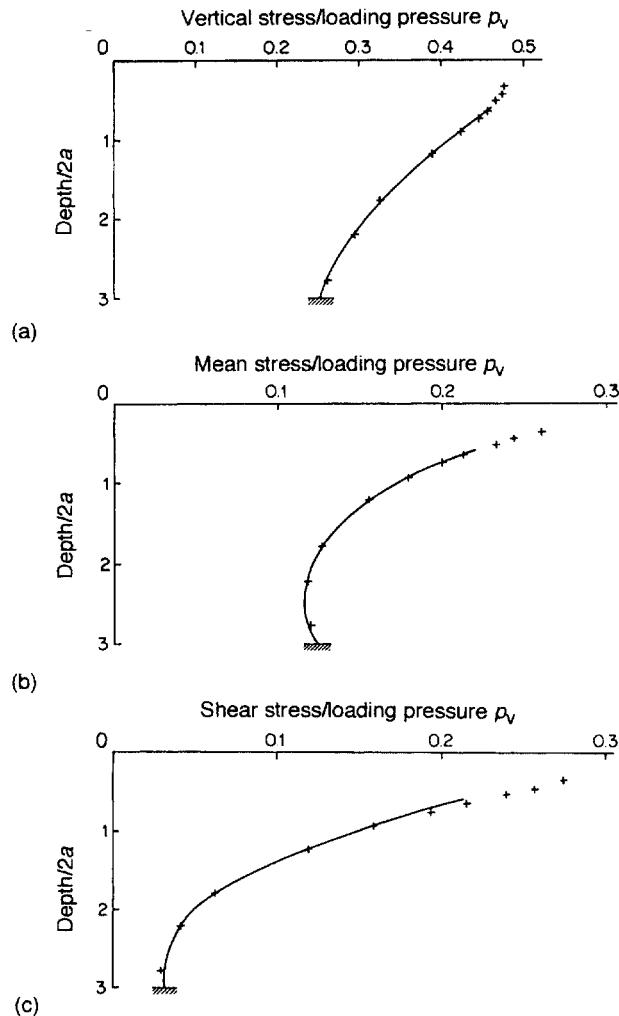
v<sub>e</sub> = effective total Poisson's ratio

K<sub>f</sub>/η = effective fluid bulk modulus

δ = displacement

V = vertical

H = horizontal

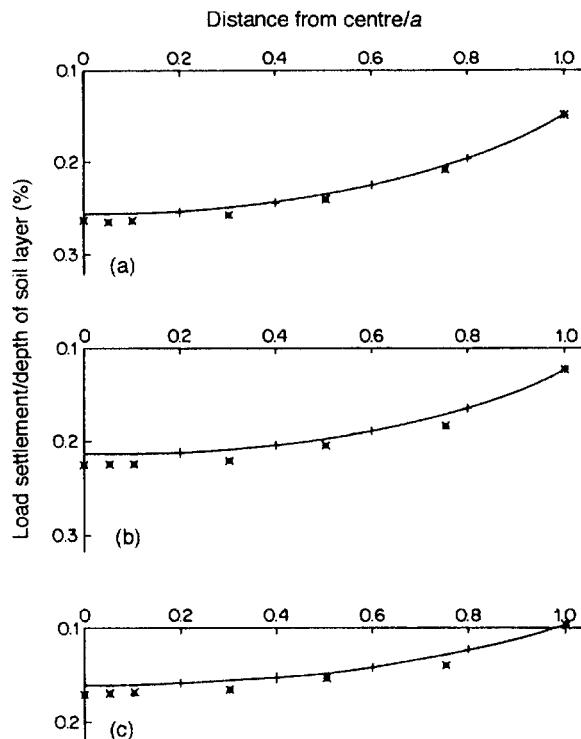


**Figure 5.2** Elastic analysis of stresses beneath the edge of a strip loading,  $\nu = 0.2$ ,  $E/p_v = 100$ : (—) analytical solution [2]; (+ + +) consolidation program and elastoviscoplastic program

points but edge displacements are greater in the vertical direction by up to 7.5% and smaller in the horizontal direction.

### 5.2.2 Radially Symmetric, Uniform Loading

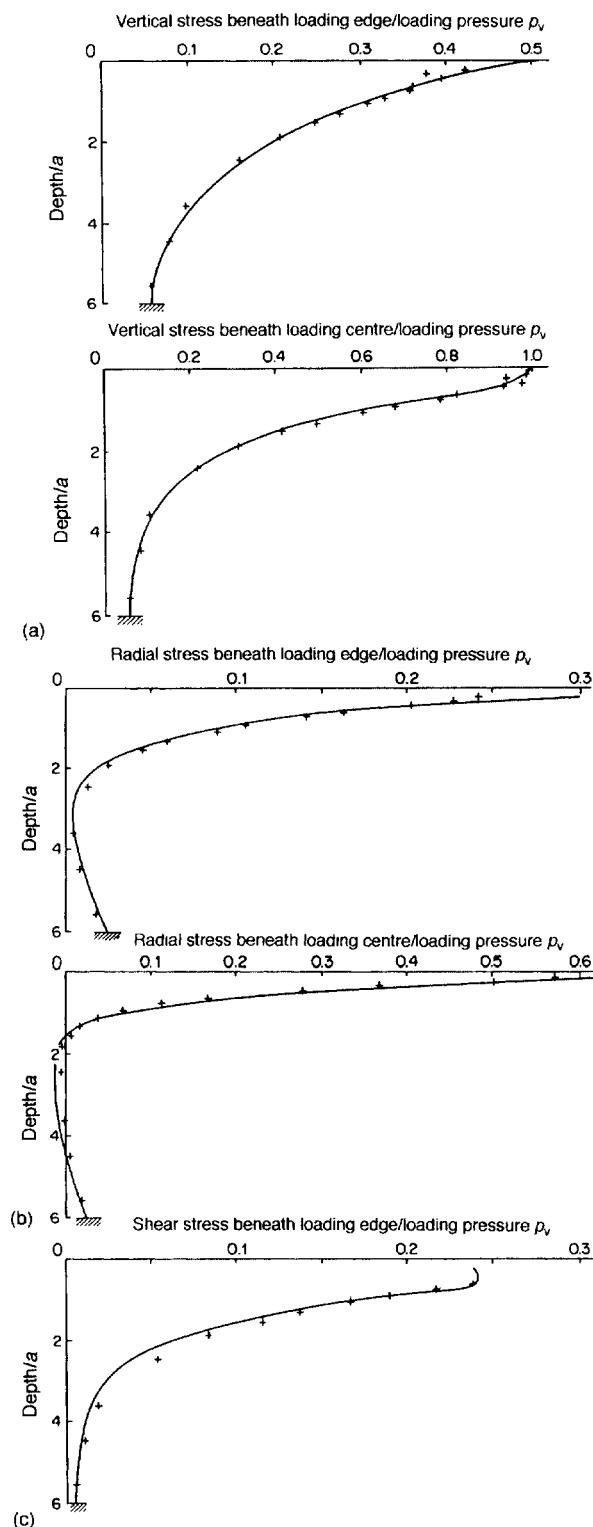
Two solutions were obtained for drained conditions with Poisson's ratio equal to 0.3 and 0.45, and a further two for undrained conditions with effective Poisson's ratios of 0.45 and 0.495. For  $\nu = 0.3$  and 0.45 the vertical displacements beneath the edge and centre of the loading are compared with the results of the viscoplastic program and with



**Figure 5.3** Elastic analysis of settlement for a circular loading area with (a)  $v = 0.3$  (drained),  $E/p_v = 100$ ; (b)  $v = 0.45$  (drained),  $E/p_v = 100$ ; (c)  $v = 0.45$  (undrained),  $E/p_v = 100$ : (+ + +) analytical solution [12]; (---) consolidation program; (x x x) elastoviscoplastic program

Milovic's analytical results (Table 5.1) [12]. Here, for  $v = 0.495$ , the vertical edge displacements given by the computer programs are also compared and their central vertical displacements are positioned alongside the analytical solution of Ueshita and Meyerhof [13] for  $v = 0.5$ . In the case of  $v = 0.3$  (drained) and  $v = 0.45$  (both drained and undrained) the computed vertical displacements beneath the whole loading area are compared with Milovic's analytical results in Figure 5.3. Stresses are compared for  $v = 0.3$  at various depths below the edge and centre of the loading in Figure 5.4.

The results show that the computer programs agree almost exactly, except for the undrained analysis with an effective Poisson's ratio of 0.495, in which the difference is about 1 percent. The analytical displacements again differ markedly, the computed vertical displacements being greater by up to 6 percent. Vertical and radial stresses show good agreement, but the computed shear stresses at middle depths are higher by up to 30 percent, though they are in good agreement at top and bottom.

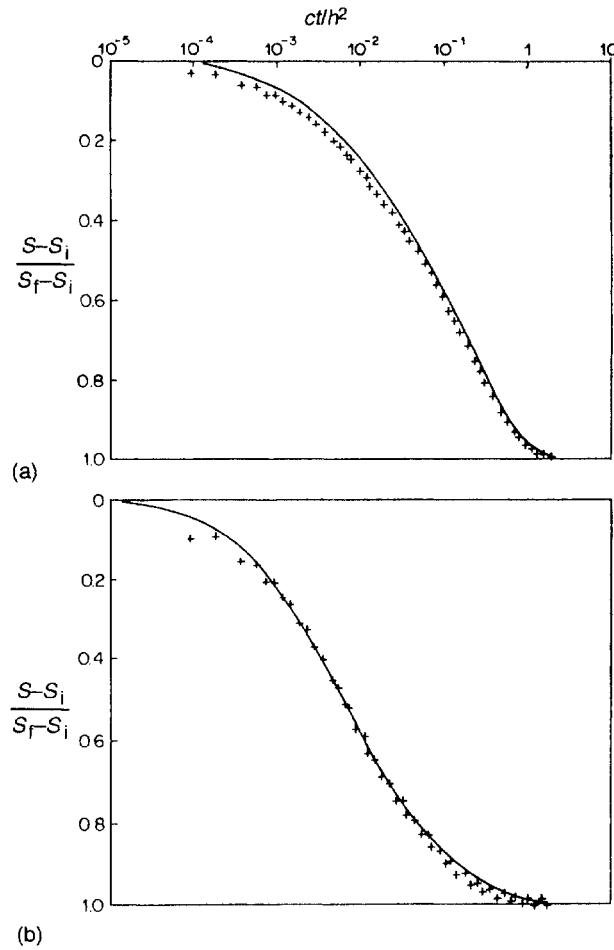


**Figure 5.4** Elastic analysis of stresses beneath a circular loading area,  $\nu = 0.3$ ,  $E/p_v = 100$ : (—) analytical solution [12]; (+ + +) consolidation program and elastoviscoplastic program

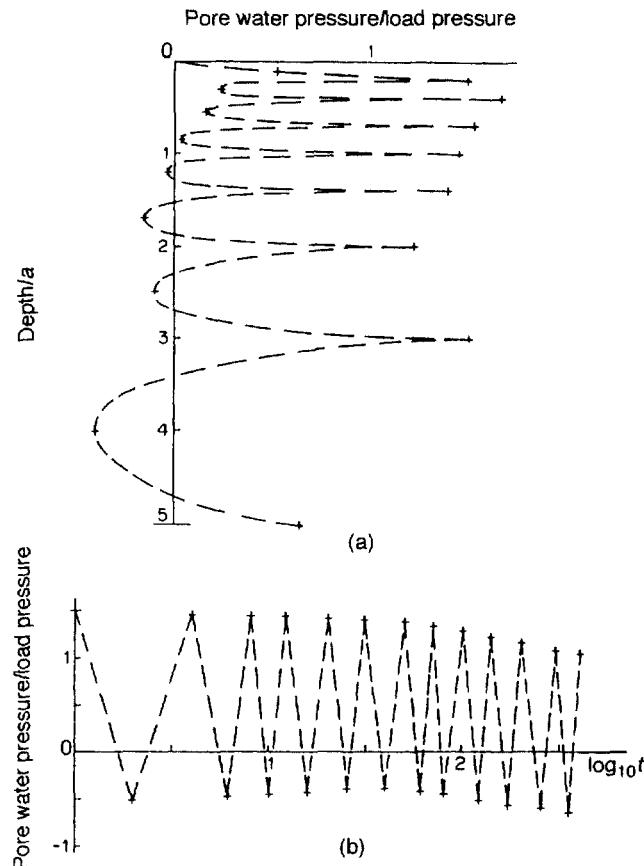
### 5.3 ELASTIC ANALYSIS OF CONSOLIDATION UNDER STRIP AND CIRCULAR UNIFORM LOADING

A comparison has been made between settlement predictions of a consolidation program and analytical solutions published by Booker [5]. In order to do so, the finite element mesh of Figure 5.1 has been slightly modified by reducing the depth of elements 43–49 from 2.0 to 1.0, hence the ratio depth/ $a$  from 6 to 5.

The solutions are compared in Figure 5.5. For both strip and circular loadings the initial settlement is about 4 percent too large, correcting gradually in the case of the strip loading but very quickly for the circular loading. There is a faster rate of settlement in the latter case, which is responsible for the more marked oscillation.



**Figure 5.5** Elastic analysis of settlement for a footing,  $h/a = 5$ ,  $v = 0.25$ : (a) plane strain and (b) radial symmetry. Key: (—) analytical solution [5]; (+ + +) consolidation program. Notation:  $c$  = coefficient of consolidation,  $h$  = total depth of soil,  $S$  = settlement,  $t$  = time,  $v$  = Poisson's ratio,  $i$  = initial,  $f$  = final



**Figure 5.6** Variation of pore water pressure for an elastic, plane strain analysis: (a) with depth, along centreline at  $t = 0$ , (+) nodal value at  $t = 0$ ; (b) with time, at depth =  $a$  on centreline, (+) nodal value at end of time step

For the plane strain case, Figure 5.6(a) shows the initial nodal pore pressures below the centre of loading, whereas Figure 5.6(b) shows the variation with time of the pore pressure at a single node. Notice that oscillation of pore pressures occurs both in space and time. The spatial oscillation can be attributed to the inability of the element shape functions to model the sharp change in pore pressure at the upper boundary, and may be reduced but not eliminated by using smaller elements in this region. A smoothing matrix was proposed by Reed [14] to eliminate this spatial oscillation.

Likewise the oscillation in time is a time boundary effect due to the inability of the time-stepping scheme to model the initial sudden change in pore pressure. Such oscillations may be reduced very significantly by the simple device of using the mean of pore pressures, displacements and effective stresses at times  $\Delta t_1$  and  $\Delta t_1 + \Delta t_2$  to restart the time-stepping scheme at time  $(\Delta t_1 + \Delta t_2)/2$ . However, in elastic analyses these oscillations of pore pressure at the nodes do not usually affect the displacements seriously; the displacements are governed by the much more accurate pore pressures at the Gauss points.

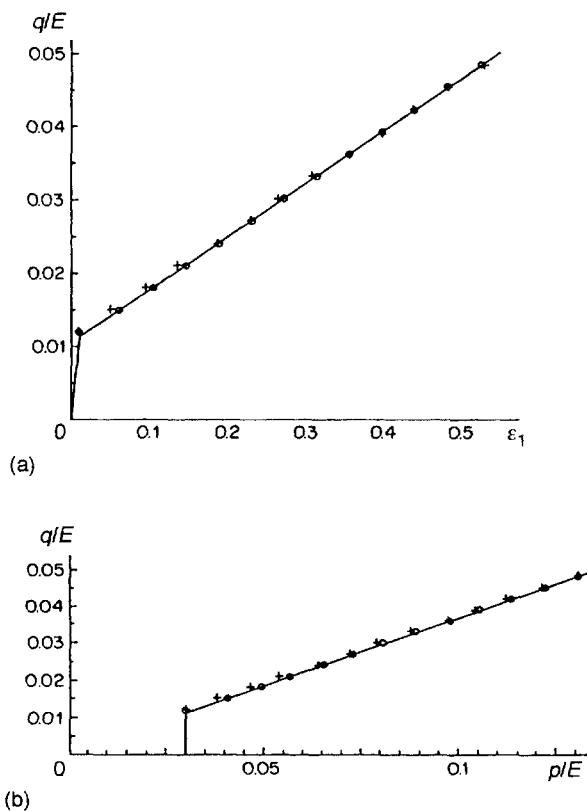
## 5.4 ELASTOPLASTIC SOLUTIONS

The consolidation and viscoplastic program results are first compared with hand calculations for the case of simple triaxial tests. The two programs are then compared with each other for strip loading analyses.

### 5.4.1 Undrained Triaxial Tests on Normally Consolidated Soil

#### 5.4.1.1 Mohr–Coulomb Analysis

This analysis was performed for undrained conditions, because a drained analysis would cause failure to occur at the onset of plasticity, the stress point remaining stationary. Figure 5.7 shows the resulting stress paths in terms of  $p$  and  $q$ , together with the Mohr–Coulomb line. Notice that the stress points computed by the consolidation program lie almost exactly on the Mohr–Coulomb line, whereas those of the viscoplastic program lie slightly above. This result is to be expected in the case of a linear, non-hardening



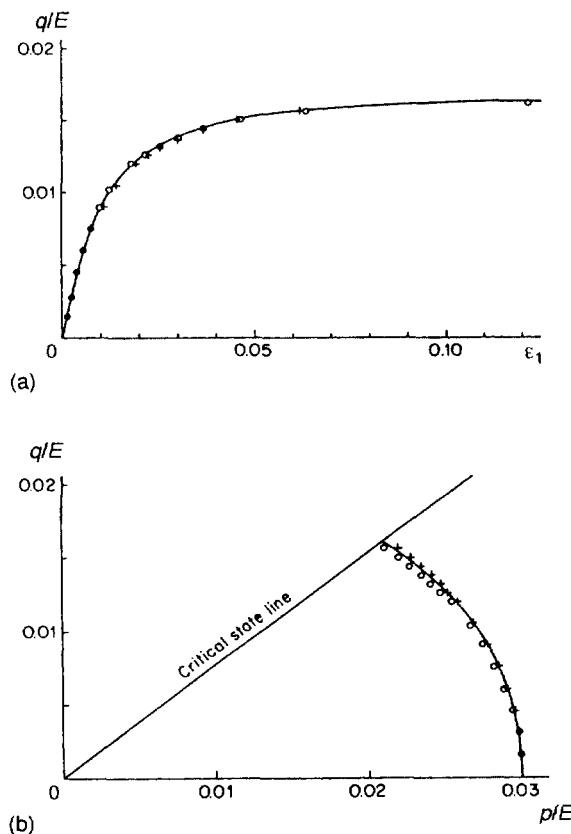
**Figure 5.7** Mohr–Coulomb analysis of undrained triaxial test,  $\nu = 0.2$ ,  $c = 0$ ,  $\phi = 10^\circ$ : (○) consolidation program, (+) viscoplastic program, (—) hand calculation

yield surface for stress conditions involving no corner effects. During a tangential stiffness analysis the stress point remains on the yield surface during a load increment, inducing no errors. For the viscoplastic analysis a stress point may not quite return to the yield surface, so the stress will be slightly high and the strain slightly low, as may be seen in the stress-strain plot. The strains computed by the consolidation program are very exact, in keeping with the stresses.

#### 5.4.1.2 Critical State Ellipse Analysis

Figure 5.8 compares the results obtained by the consolidation program and the viscoplastic program with those of a hand calculation. The stress path shown in Figure 5.8(b) could be hand-calculated exactly; the deviatoric strains, however, were calculated incrementally, using the same increments of deviatoric stress as the computer programs but obtaining mean stress values  $p$  from the hand-calculated stress path.

The consolidation program predicts lower values of mean stress, hence a smaller increase in the size of the yield surface. One would expect this since the plastic modulus



**Figure 5.8** Critical state ellipse analysis of undrained triaxial test,  $v = 0.2$ ,  $c = 0.0$ ,  $\phi = 20^\circ$ ,  $\chi = 20$ : (.) consolidation program, (+) viscoplastic program, (—) hand calculation

is formed at the start of each increment, before the incremental hardening has taken place. The viscoplastic program predicts stresses outside the yield surface, as for the Mohr–Coulomb analysis.

The picture is more complicated in the case of deviatoric strain prediction. The strains computed by the consolidation program are very slightly greater than those calculated by hand, in keeping with the smaller size of the yield surface. Those of the viscoplastic program, however, are greater at relatively low deviatoric stresses and less at high deviatoric stresses than those of the consolidation program. Both effects can be attributed to the fact that the consolidation program forms the stiffness matrix at the start of the increment, whereas the viscoplastic program bases it on the higher deviatoric stress given by an initial elastic analysis; this stress remains constant throughout the following iterations for this particular test. Thus the ratio  $d\epsilon_q^p/d\epsilon_v^p$  is higher for the viscoplastic program, particularly in the middle range of stresses. At higher stresses its effect is less significant than the effect of the lower plastic modulus of the consolidation program. The difference between the predictions of the two programs would obviously decrease with the size of the load increments.

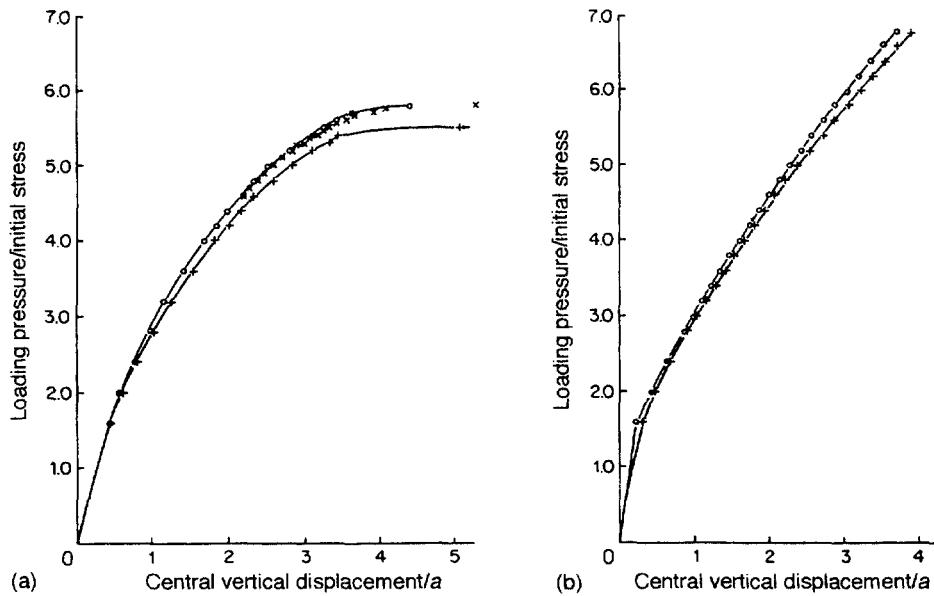
### 5.4.2 Drained and Undrained Analyses of Strip Loading

These tests were carried out using the element mesh and loading shown in Figure 5.1, but with an initial overburden pressure and accompanying isotropic stress.

#### 5.4.2.1 Mohr–Coulomb Analysis

Figure 5.9 shows both drained and undrained analyses performed by the consolidation and viscoplastic programs. The consolidation program predicts smaller displacements, as one would expect. This is because the stiffness moduli are determined by the stresses at the start of each load increment, when fewer Gauss points have experienced the onset of plasticity. The accompanying smaller strains induce a slower spread of plastic straining with the result that the predicted failure load in the drained case is a few percent higher. Failure cannot occur when an associated plastic flow rule is used in undrained analysis; the dilatant plastic strains allow the stresses to increase indefinitely.

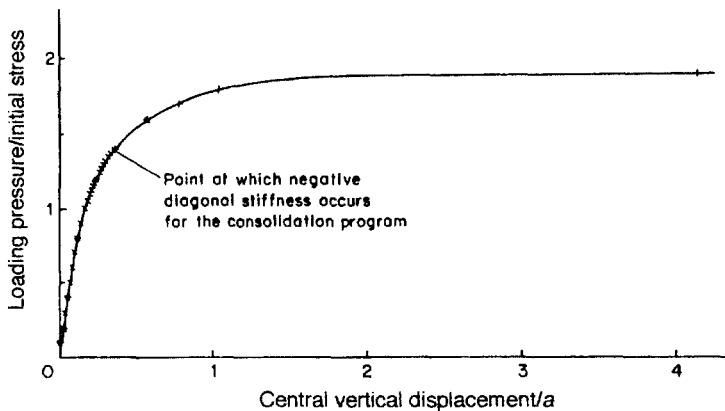
Note how in Figure 5.9(b) the consolidation analysis was performed with only one iteration, so the first load increment was analysed elastically. The drained consolidation analysis of Figure 5.9(a) was also performed with only one iteration, but this made no difference in the first load increment as no stresses exceeded the elastic limit. This analysis was repeated with two iterations per load increment, so the tangential stiffness matrix was reformulated taking account of the onset of plasticity at the Gauss points during each increment. The resulting displacements were almost the same as for one iteration, casting doubt on the usefulness of this technique in the present context, especially as it requires twice as much computer time. A further analysis was performed with the load increments halved and one iteration per load increment, so the computer time was much the same as for the two-iteration analysis. Significantly larger strains were predicted for high loading, indicating that this method of analysis is more useful than having two iterations per load increment. The results for high loading only are shown in Figure 5.9(a).



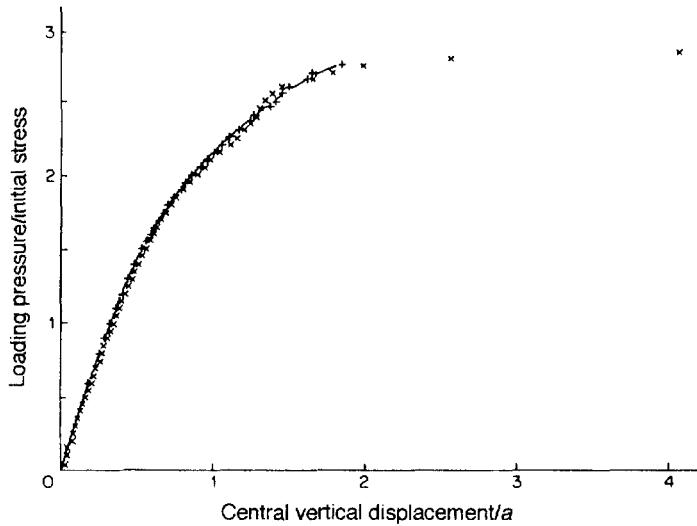
**Figure 5.9** Mohr–Coulomb analysis of strip loading with  $\nu = 0.2$ ,  $c = 0.0$ ,  $\phi = 20^\circ$ , initial isotropic stress  $= E/10$ : (a) drained and (b) undrained. Key: ( $\times$ ,  $\circ$ ) consolidation program, (+) viscoplastic program

#### 5.4.2.2 Critical State Ellipse Analysis

The results of this analysis are shown in Figure 5.10. Using small load increments for the consolidation program, the loading–displacement plot was almost identical with that of the viscoplastic program up to a loading/initial stress ratio of 1.4, i.e. 74 percent of the value at failure predicted by the viscoplastic program.



**Figure 5.10** Critical state analysis of undrained strip loading:  $\nu = 0.2$ ,  $c = 0.0$ ,  $\phi = 20^\circ$ ,  $\chi = 100$ , initial isotropic stress  $= E/10$



**Figure 5.11** Analysis of drained strip loading using a critical state ellipse with Mohr–Coulomb cut-off;  $v = 0.2$ ,  $c = 0.0$ ,  $\phi = 20^\circ$ ,  $\chi = 100$ , initial isotropic stress =  $E/10$ : ( $x$ ) consolidation program, ( $+ - +$ ) viscoplastic program

#### 5.4.2.3 Analysis by Critical State Ellipse with Mohr–Coulomb Cut-off ( $c = 0$ )

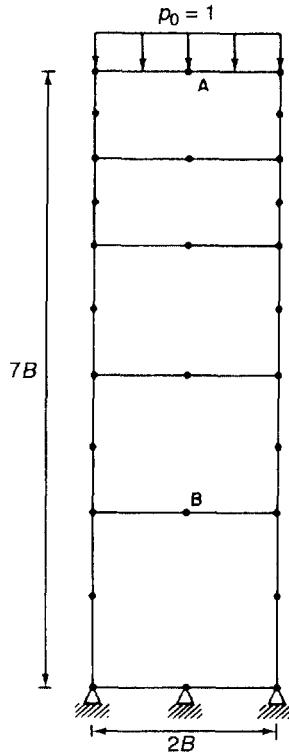
For this analysis a combined critical state model was used, where the supercritical region was cut off by a Mohr–Coulomb surface [4]. In such a combined model the elliptical surface meets the Mohr–Coulomb surface on the critical state surface in a single  $\pi$ -plane where both must have the same cross-section.

As may be seen from Figure 5.11, the consolidation and viscoplastic programs agree well until the loading–displacement curve approaches failure. Strain softening has been deliberately suppressed by using Mohr–Coulomb surface with  $c = 0$ . Apart from confirming the reliability of the consolidation program, this result illustrates an oscillatory tendency which had not been observed in the Mohr–Coulomb analyses. It has been found that such oscillation is very significant when large load increments are used but becomes less obvious as the increment sizes are reduced. It seems likely that this phenomena is related to the strain-hardening property of the critical state ellipse surface, but again further investigation would be useful.

## 5.5 ELASTOPLASTIC ANALYSIS OF CONSOLIDATION UNDER UNIFORM LOADING AND STRIP LOADING

Elastoplastic solutions are shown for one-dimensional and two-dimensional plane strain consolidation.

In this case a comparison has been made between the settlement predictions obtained by Siriwardane and Desai [7] and those obtained with the consolidation program listed



**Figure 5.12** Finite element mesh for the one-dimensional consolidation problem

in the first edition of this book. Siriwardane and Desai also used the initial stress approach in addition to the tangential stiffness matrix approach.

The critical state model described in Section 4.5.3.1 has been used as the constitutive relationship for both examples.

### 5.5.1 One-dimensional Consolidation

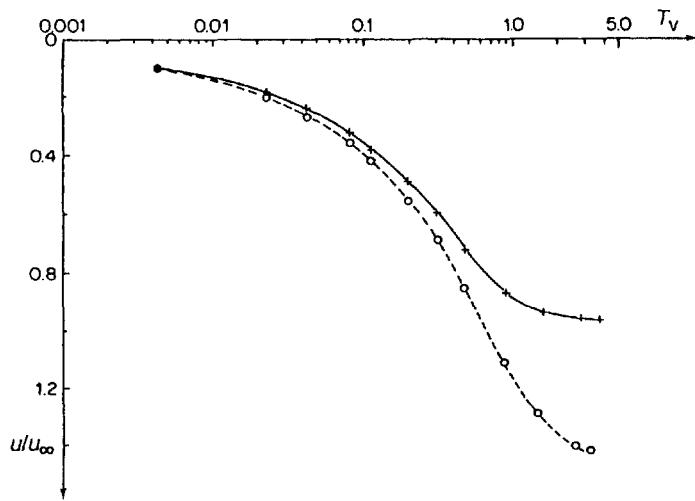
The finite element mesh is shown in Figure 5.12. Isoparametric eight-noded elements have been employed, instead of triangular elements as in the work of Siriwardane and Desai [7]. The linear elastic behaviour for this example has been described previously by other authors numerically and in a closed form [15,16].

The properties of the materials are as follows:

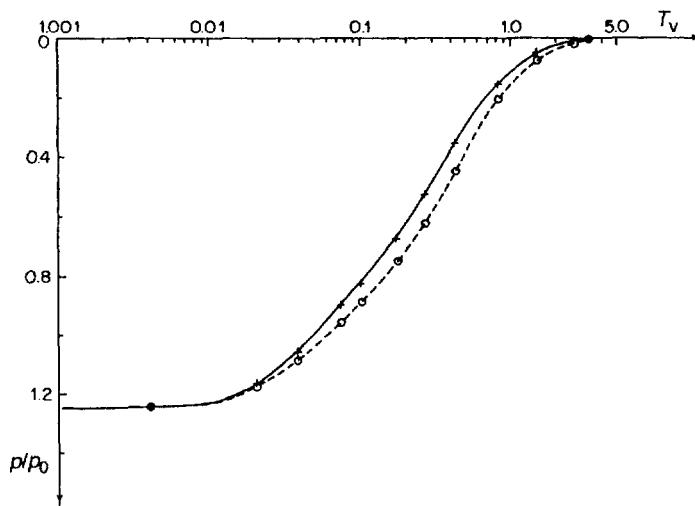
$$\begin{aligned} E_0 &= 6000 \quad v = 0.4 \quad M = 1.05 \quad \lambda = 0.14 \\ \kappa &= 0.05 \quad e_0 = 0.9 \quad K_x = K_y = 4.0 \times 10^{-6} \end{aligned}$$

An external surface load \$p\_0 = 1\$ was applied at the top surface of the model.

The calculated settlements and pore pressures are compared in Figures 5.13 and 5.14. The pore pressures are non-dimensionalised with respect to \$p\_0\$, and the settlements with



**Figure 5.13** One-dimensional consolidation surface settlement versus time ( $T_v = c_v t / H^2 = 0.10496 \times 10^{-4}t$ ). Elastic analysis: (—) Siriwardane–Desai, (+ + +) FE program of first edition. Elastoplastic analysis: (- - -) Siriwardane–Desai, (○ ○ ○ ○) FE program of first edition



**Figure 5.14** One-dimensional consolidation: pore pressure versus time at node 0.2  $H$  from bottom (node B) ( $p_0$  is the initial pressure,  $T_v = 0.10496 \times 10^{-1}t$ ). Elastic analysis: (—) Siriwardane–Desai, (+ + +) FE program of first edition. Elastoplastic analysis: (- - -) Siriwardane–Desai, (○ ○ ○ ○) FE program of first edition

respect to  $u_s$ , the ultimate settlement predicted by Terzaghi's theory. The non-dimensional time factor  $T_v = c_v t / H^2$  was used as the abscissa;  $E_0$  was used to calculate  $c_v$  [7].

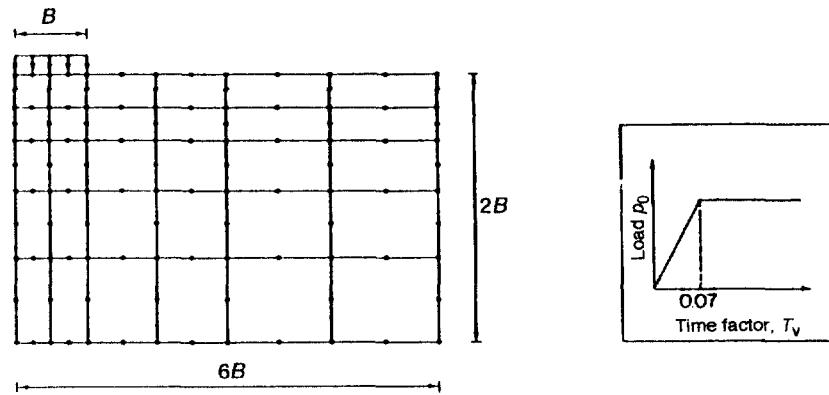
The inclusion of the plasticity effects decreases the magnitude of dissipation of pore water pressure for almost the whole time span. The predicted pore pressures are therefore higher than those from an elastic analysis. The settlements are also larger than those predicted in the elastic case.

### 5.5.2 Two-dimensional Consolidation

The demonstration problem and the finite element mesh are shown in Figure 5.15. The problem was solved using the following parameters:

$$\begin{aligned} E_0 &= 13000 \quad v = 0.4 \quad M = 1.05 \quad \lambda = 0.14 \\ \kappa &= 0.05 \quad e_0 = 0.9 \quad K_x = K_y = 4.0 \times 10^{-5} \end{aligned}$$

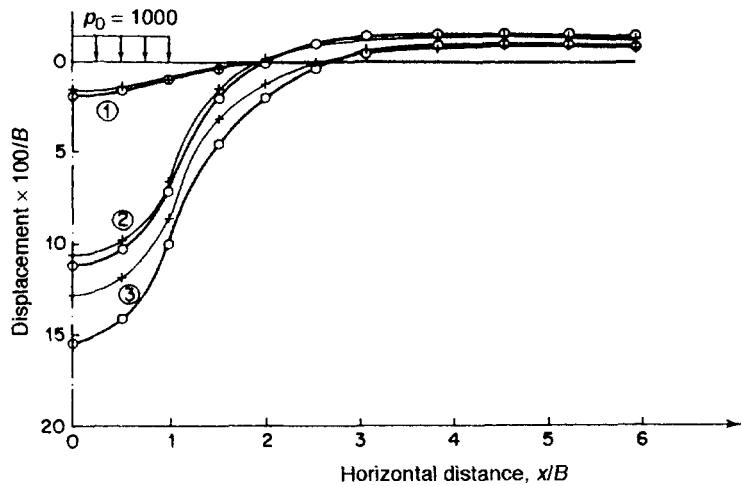
The load applied at the top surface of the model for a length  $B = 10$  varies with time as indicated in Figure 5.15. The time-stepping scheme is identical to that of Siriwardane



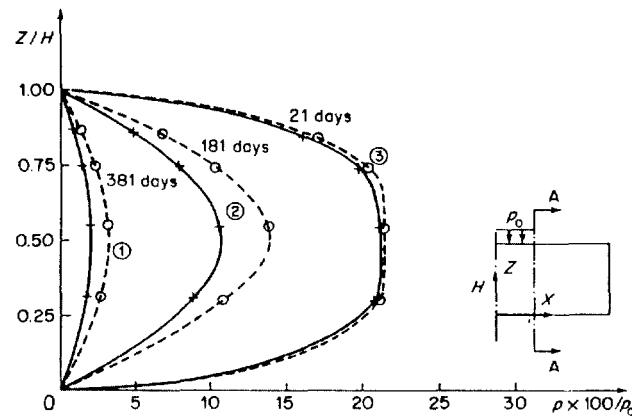
**Figure 5.15** Finite element mesh for plane strain elastoplastic consolidation and load history

**Table 5.2** Time steps for different time intervals

Time interval	Number of time steps
0.1	10
1.0	10
2.0	10
5.0	10
10.0	15
20.0	35



**Figure 5.16** Surface settlements versus horizontal distance at three different time values ( $T_v = 0.27857 \times 10^{-2}t$ ): (1)  $T_v = 0.017$ , (2)  $T_v = 0.142$ , (3)  $T_v = 0.782$ . Elastic analysis: (—) Siriwardane–Desai, (+ + +) FE program of first edition. Elastoplastic analysis: (—) Siriwardane–Desai, (---) FE program of first edition



**Figure 5.17** Pore pressure versus depth along section A–A at three different time values: (1)  $T_v = 0.059$ , (2)  $T_v = 0.504$ , (3)  $T_v = 1.060$ . Elastic analysis: (—) Siriwardane–Desai, (+ + +) FE program of first edition. Elastoplastic analysis: (---) Siriwardane–Desai, (----) FE program of first edition

and Desai [7] and is indicated in Table 5.2. The resulting settlements and pore pressures are compared in Figures 5.16 and 5.17 respectively. The same conclusions regarding settlements and pore pressure dissipation apply as in the previous example.

The results of this section show that the two independently developed computer programs are in almost exact agreement.

## 5.6 LINER AND NON-LINEAR ELASTIC CONSOLIDATION WITH VARIABLE PERMEABILITY

The variable permeability scheme used in the following two examples [8] is that proposed by Lewis *et al.* [6] and was described in Section 4.7.

The constitutive relationships applied are respectively the hyperbolic model of Section 4.4.2.1 and the linear elastic model. The main purpose of the first example is to show the effect of variable permeability on settlement and pore pressure, with no reference to particular experimental data. Whereas the results of the second example are compared with the existing experimental data. No account was made of finite deformation (Chapter 14).

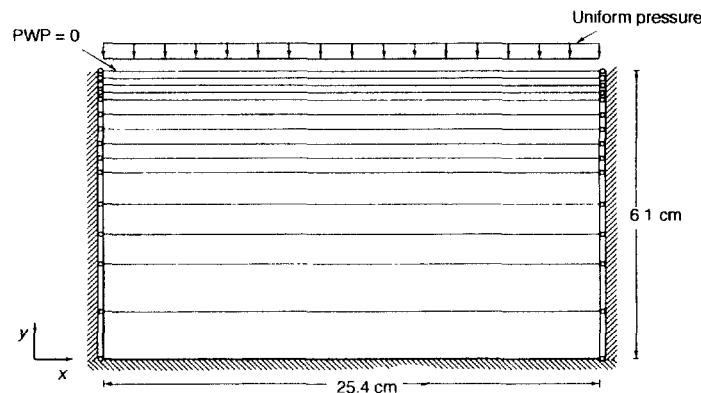
### 5.6.1 Consolidation of Swansea Blue Clay in a Rowe Consolidation Cell

The graph of void ratio versus permeability for this soil is shown in Figure 4.23. The drained parameters used for the hyperbolic model are as follows:

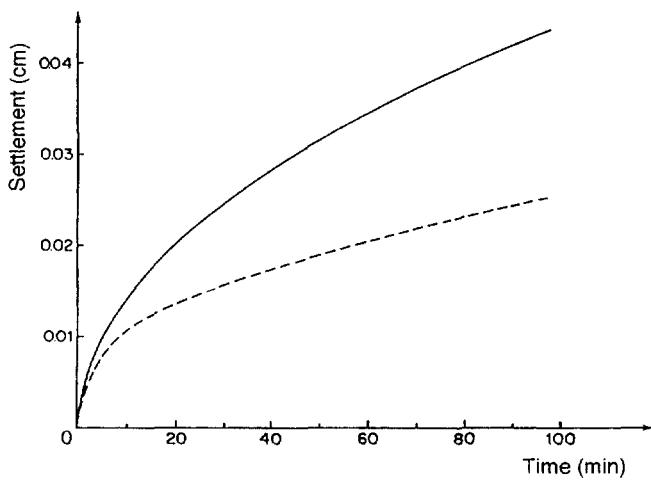
$$\begin{aligned} E_i &= 2.6 \text{ MPa} & \phi &= 8.0^\circ \\ R_f &= 0.75 & v &= 0.2 \\ C &= 22.8 \text{ kPa} & K &= 4.83 \times 10^{-6} \text{ cm min}^{-1} \end{aligned}$$

A surface traction of  $-0.172 \text{ MPa}$  is applied; Figure 5.18 shows the boundary conditions and the finite element mesh.

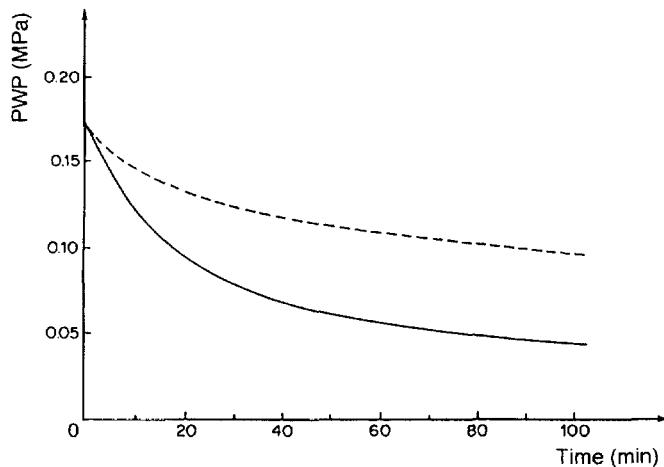
Figure 5.19 compares the settlements versus time obtained with variable permeability and with constant permeability. The pore pressure versus time for both solutions is indicated in Figure 5.20. Both figures show the results up until 100 min. Ultimately there should be no difference between the two solutions, as permeability affects only the rate of consolidation.



**Figure 5.18** Finite element mesh for the Rowe consolidation cell



**Figure 5.19** Settlement versus time for a clay, predicted by the non-linear consolidation program: (—) constant permeability, (---) variable permeability



**Figure 5.20** Pore pressure versus time for a clay, predicted by the non-linear consolidation program: (—) constant permeability, (---) variable permeability

For the time period considered, the variable scheme leads to a considerably higher pore pressure. This is an important factor in assessing the shear strength of a soil in the early stages of construction (although construction loads would not be ‘instantaneous’ upon the soil, so this effect would be reduced but not eliminated). Conversely the constant permeability scheme will achieve an upper bound on settlements using less computer time than the variable scheme.

### 5.6.2 Consolidation of Kaolin in a Rowe Consolidation Cell

Experimental consolidation data (settlement-time and pore water pressure-time) was available in this case, but unfortunately no void ratio permeability data or reliable non-linear parameters. It was therefore decided to use a linear elastic consolidation model with a variable permeability scheme. A comparison with a constant permeability scheme was also carried out. For the variable permeability model, a slightly lower average Young's modulus was used than in the constant permeability analysis, which compensates for the increased stiffness due to a lower permeability. Furthermore, by using the hypothetical graph of void ratio versus permeability (Figure 5.21), virtually identical settlement results were obtained with both numerical schemes. The material properties used in the linear elastic program were as follows:

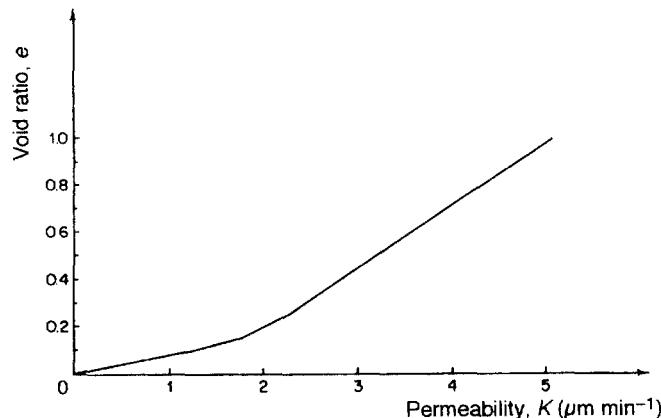
Constant permeability

$$\begin{aligned}E_a &= 2.14 \text{ MPa} \\v &= 0.2 \\K &= 0.71 \times 10^{-4} \text{ cm min}^{-1}\end{aligned}$$

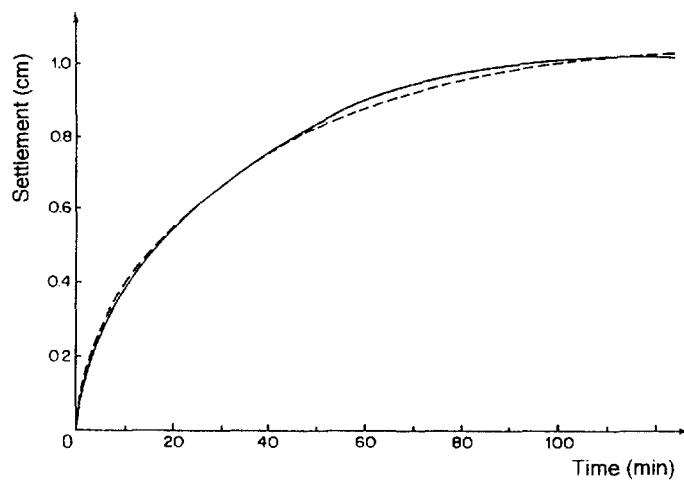
Variable permeability

$$\begin{aligned}E_a &= 2.07 \text{ MPa} \\v &= 0.2 \\K_i &= 0.79 \times 10^{-4} \text{ cm min}^{-1} \\e_i &= 1.0\end{aligned}$$

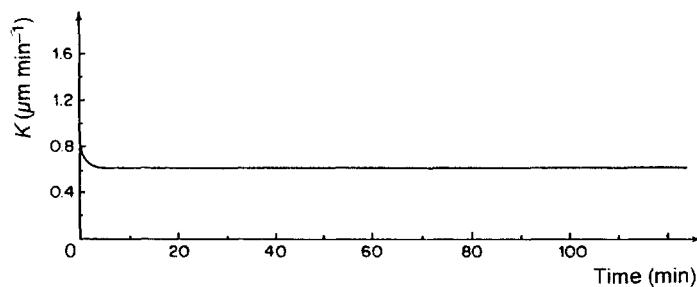
Figure 5.22 compares the experimental data with the numerical results for the settlements versus time. Figure 5.23 shows the permeabilities versus time obtained from the applied scheme. And Figure 5.24 compares the experimental pore pressure history



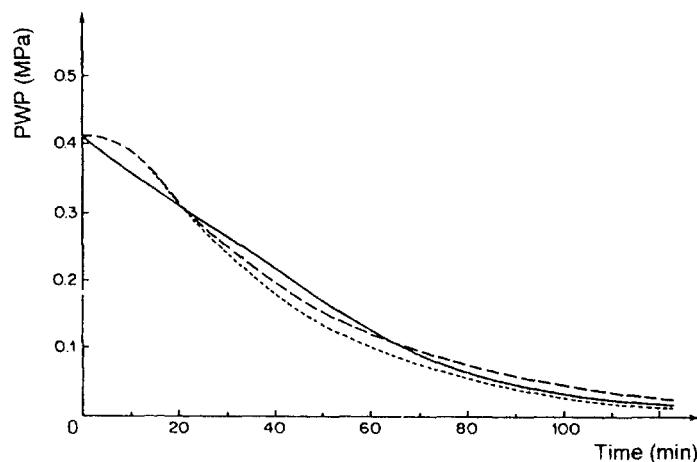
**Figure 5.21** Void ratio versus permeability for kaolin



**Figure 5.22** Settlement versus time: (—) experimental results, (- - -) both computed solutions



**Figure 5.23** Resulting permeability versus time



**Figure 5.24** Pore water pressure versus time: (—) experimental results; (· · · ·) linear  $E$ , constant  $K$ ; (- - -) linear  $E$ , variable  $K$

with the numerical pore pressure history. Notice that the settlement results are good; the two computed solutions are virtually identical.

But the pore pressure results from the two computed solutions do vary; the variable permeability solution gives the better overall fit. However, it must be remembered that the  $E$  versus  $K$  values were chosen in order to give the best possible results.

From the previous two examples it is clear that variable permeability may have considerable effects on the rate of consolidation, hence it may improve the accuracy of the model. The method suffers from the fact that it requires extra data to be obtained from tests. Also, extreme care is needed when obtaining this data, as a small change of void ratio may cause enormous variations in the permeability.

Besides the measurement aspects of the data, another drawback is that the permeability scheme somewhat depends on the size of the elements used. For example, if a very shallow element is used directly under the load in the Rowe cell experiment, assuming drainage takes place at the top of the cell, then this element will suffer a great fall in permeability in the first few minutes, and as it controls the outflow of fluid, it will predict a lower settlement than if a deeper element had been used.

The variable permeability scheme is of practical use only if the expected rate of consolidation is important, and in this case it should be included in the model.

## 5.7 MULTIPHASE FLOW IN POROUS MEDIA: A BENCHMARK PROBLEM FOR NON-SATURATED FLOW

In this example, partially saturated flow in deforming porous media is investigated under the assumption that, either air remains at atmospheric pressure in the unsaturated zone, or there is flow of both water and air. Furthermore, the problem of passing from fully to partially saturated conditions (and vice versa) is addressed in detail. For fully saturated conditions we have one mass balance equation for water, and in partially saturated conditions we have one for water and one for air. The transition from one state to the other is not always straightforward.

The proposed benchmark is based on an experiment performed by Liakopoulos [17] on a column of Del Monte sand and instrumented to measure the moisture tension at several points along the column during its desaturation due to gravitational effects. Before the start of the experiment, water was continuously added from the top and allowed to drain freely at the bottom through a filter, until uniform flow conditions were established. At the start of the experiment, the water supply ceased and the tensiometer readings were recorded, as well as the outflow and outflow rate at the bottom. The porosity and hydraulic properties of Del Monte sand were measured by Liakopoulos [17] in an independent set of experiments; they are given in Table 5.3.

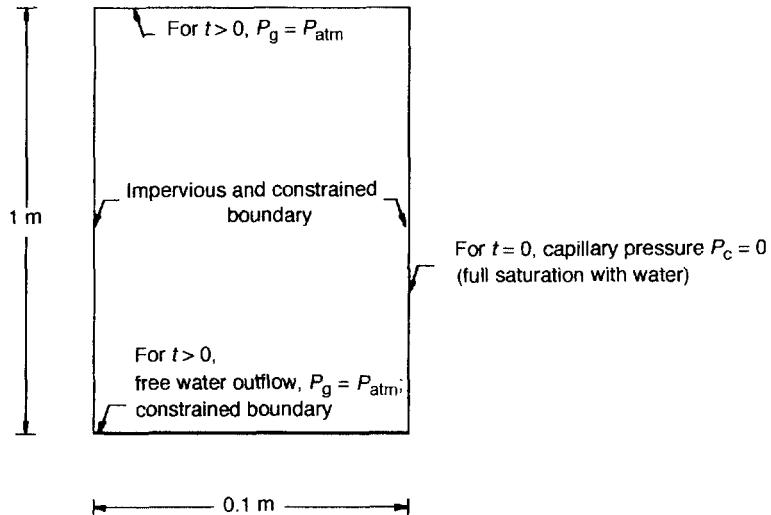
The equations used by Liakopoulos for the saturation–capillary pressure and relative permeability–saturation relationships, valid for saturation  $S \geq 0.91$ , have the following form:

$$S = 1 - 1.9722 \times 10^{11} p_c^{2.4279}$$

$$K_{rl} = 1 - 2207(1 - S)^{1.0121}$$

**Table 5.3** Schematic description of the benchmark

Young's modulus	$E = 1.3 \text{ MPa}$
Poisson's ratio	$\nu = 0.4$
Solid grain density	$\rho^s = 2000 \text{ kg m}^{-3}$
Liquid density	$\rho^w = 1000 \text{ kg m}^{-3}$
Porosity	$p_0 = 0.2975$
Intrinsic permeability	$k = 4.5 \times 10^{-13} \text{ m}^2$
Water viscosity	$\mu^w = 1 \times 10^{-3} \text{ Pa s}$
Air viscosity	$\mu^g = 1.8 \times 10^{-5} \text{ Pa s}$
Gravitational acceleration	$g = 9.806 \text{ m s}^{-2}$
Atmospheric pressure	$P_{\text{atm}} = 101325 \text{ Pa}$

**Figure 5.25** The Liakopoulos test problem

This test problem has been solved previously by Liakopoulos [17], Narasimhan and Whitherspoon [18], Schrefler and Simoni [19], Zienkiewicz *et al.* [20], Schrefler and Zhan [21], Gawin *et al.* [22], as well as by Gawin and Schrefler [23] to check their numerical models. The initial conditions are: for  $t = 0$ :  $p_c = 0$  for all the nodes, which corresponds to a steady flow of water through the sand column (Figure 5.25). Furthermore, a state of mechanical equilibrium is assumed for  $t = 0$ . All the displacements are related to these initial displacements which correspond to the equilibrium state.

The boundary conditions are as follows:

For the lateral surface:  $q_T = 0, u_h = 0$

where  $u_h$  is the horizontal displacement of the soil.

For the top surface:  $p_g = p_{\text{atm}}$

where  $p_{\text{atm}}$  is the atmospheric pressure,  $T = 293.15 \text{ K}$ .

For the bottom surface:  $p_g = p_{\text{atm}}, p_c = 0$  for  $t > 0, T = 293.15 \text{ K}, u_h = u_v = 0$

where  $u_v$  is the vertical displacement soil.

Liakopoulos did not measure the mechanical parameters of the soil, therefore the Young's modulus of the soil was assumed to be  $E = 1.3 \text{ MPa}$ , Poisson's ratio was taken as  $\nu = 0.4$  and Biot's constant was taken as  $\alpha = 1$  [19,21,22].

The relative permeability of the gas phase was assumed to be as given by Brooks and Corey [24]:

$$K_{rg} = (1 - S_e)^2 (1 - S_e^{5/3})$$

$$S_e = (S - 0.2)/(1 - 0.2)$$

with an additional lower limit for  $K_{rg}$ , as explained in the following. Figure 5.25 and Table 5.3 show schematically the initial and boundary conditions defining the problem, as well as the complete set of physical data used in the calculations.

The transition from a fully saturated condition to a partially saturated state during the desaturation process can be solved using one of the following methods:

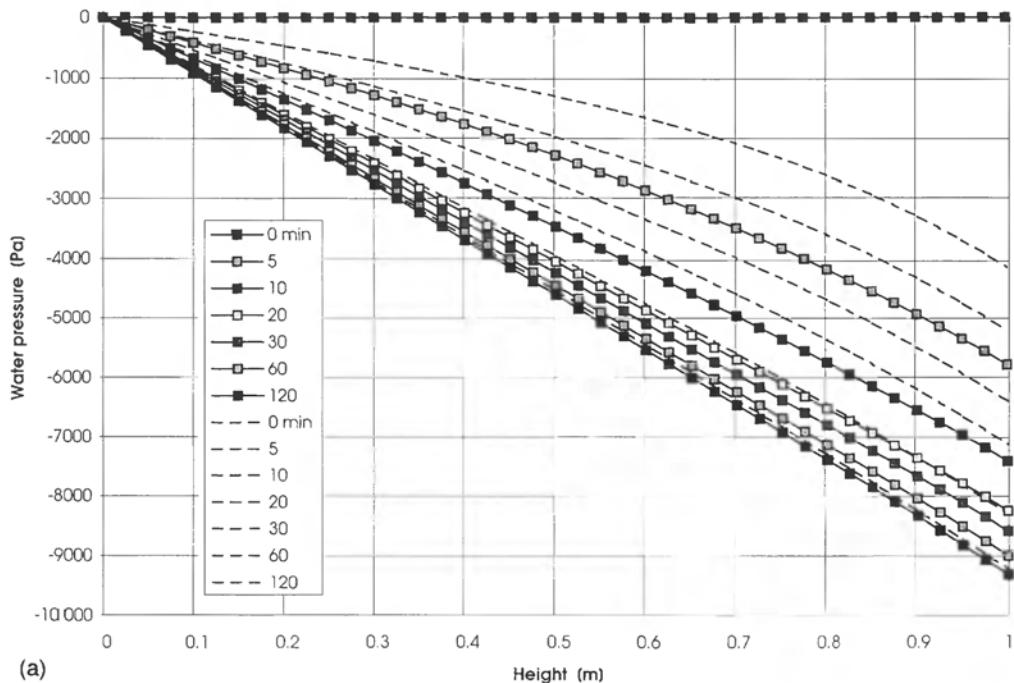
1. The gas pressure value in the partially saturated zone is assumed to be equal to the atmospheric pressure (one-phase flow approach), which implies an infinite value of gas relative permeability. This assumption was used by Schrefler and Simoni [19], Schrefler and Zhan [21] and by Gawin *et al.* [22,23].
2. A very small, but finite, value of gas relative permeability exists even for the fully saturated state. This physically implies that a flow of gas exists for the fully saturated state and can be obtained by appropriate modification to the capillary pressure–saturation relationship [21] and/or the relative permeability curve [23]. This approach was used by Schrefler and Zhan [21] and Gawin and Schrefler [23], and in this case the gas-phase continuity equation is always maintained.
3. A ‘switching’ between the one-phase flow and full two-phase solutions could occur at a certain value of the saturation, e.g. at  $S_w = 1$ , or at the value corresponding to the bubbling pressure, which has a physical justification. Also, the application of an additional lower limit for the gas relative permeability is usually necessary in order to avoid oscillations in the pressure solution. This approach was proposed by Gawin *et al.* [22] and Gawin and Schrefler [23].

The full set of results for methods 1 and 3 can be found in the literature [25], both for deformable and undeformable material (i.e. a Young's modulus value of  $10^6$  times greater), where these assumptions are made:

1. One-phase flow, i.e. the gas pressure was assumed equal to the atmospheric pressure in the partially saturated zone.

2. Two-phase flow with switching at  $p_c = 0$  ( $S = 1$ , no gas flow for the fully saturated state) and an additional lower limit for the gas relative permeability of  $K_{rg\min} = 0.0001$ . Additionally, a case with a higher value for the lower limit of the gas relative permeability of  $K_{rg\min} = 0.0005$  was solved in order to analyse its effect on the results obtained.
3. Two-phase flow, with switching occurring at  $p_c = 2000 \text{ Pa}$  ( $S = 0.998$ , no gas flow below the bubble point pressure) and an additional lower limit for the gas relative permeability of  $K_{rg\min} = 0.0001$ . Also, a case with a lower limit for the gas relative permeability of  $K_{rg\min} = 0.0005$ , as well as with switching at  $p_c = 6875 \text{ Pa}$  ( $S = 0.96$ ) were solved in order to analyse the influence of these parameters on the results.

For the numerical calculations, the column of soil was divided into 10 and 20 four-, eight- and nine-node isoparametric finite elements of equal size; various time-step dimensions were assumed in the time domain. The values used were  $\Delta\tau = 10 \text{ s}$ , 1 s and 0.5 s respectively, which gave practically the same results. Furthermore,  $3 \times 3$  and  $2 \times 2$



**Figure 5.26** (a) Water pressure versus height for deformable material, Initial Conditions:  $p_c = 0$ ,  $K_{rg} > 0.0001$  (solid lines);  $p_g = p_{atm}$  (dashed lines). (b) Gas pressure versus height for deformable material, (IC):  $p_c = 0$ ,  $K_{rg} > 0.0001$ . (c) Capillary pressure versus height for deformable material, (IC):  $p_c = 0$ ,  $K_{rg} > 0.0001$  (solid lines);  $p_g = p_{atm}$  (dashed lines). (d) Vertical displacement versus height for deformable material, (IC):  $p_c = 0$ ,  $K_{rg} > 0.0001$  (solid lines);  $p_g = p_{atm}$  (dashed lines). (e) Saturation versus height for deformable material, (IC):  $p_c = 0$ ,  $K_{rg} > 0.0001$  (solid lines);  $p_g = p_{atm}$  (dashed lines)

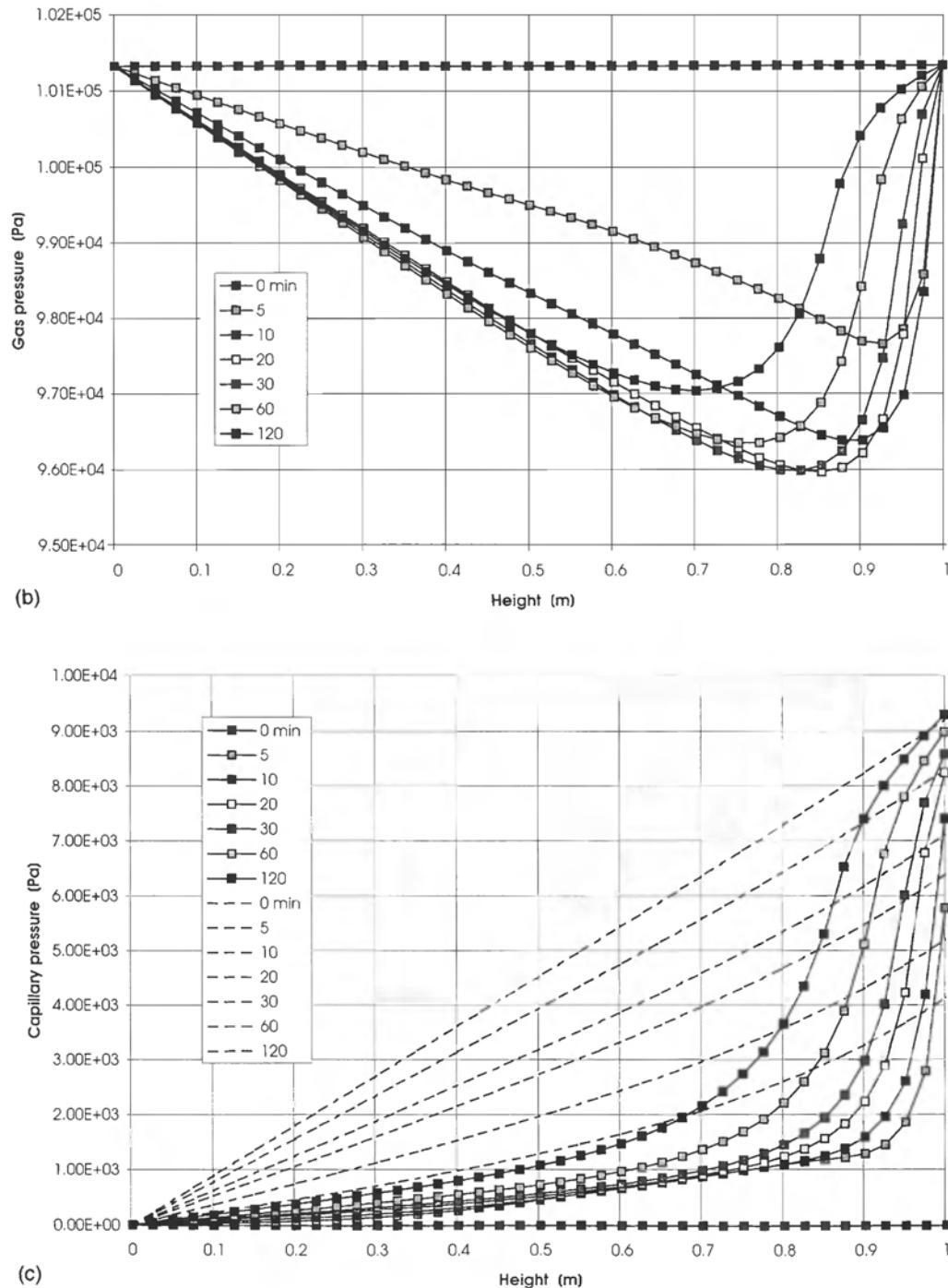
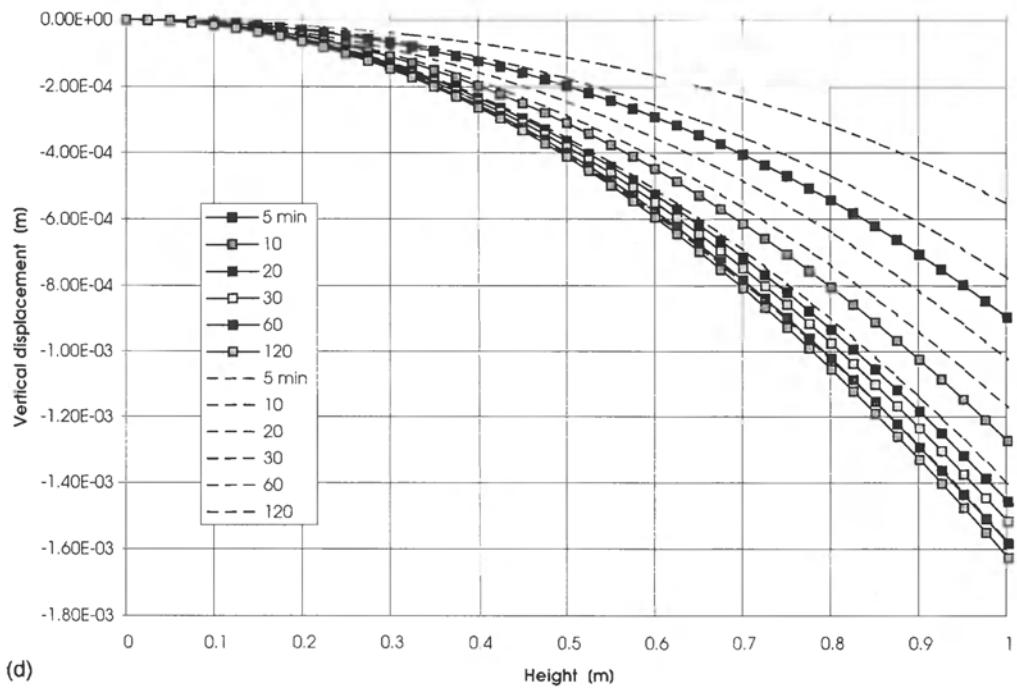
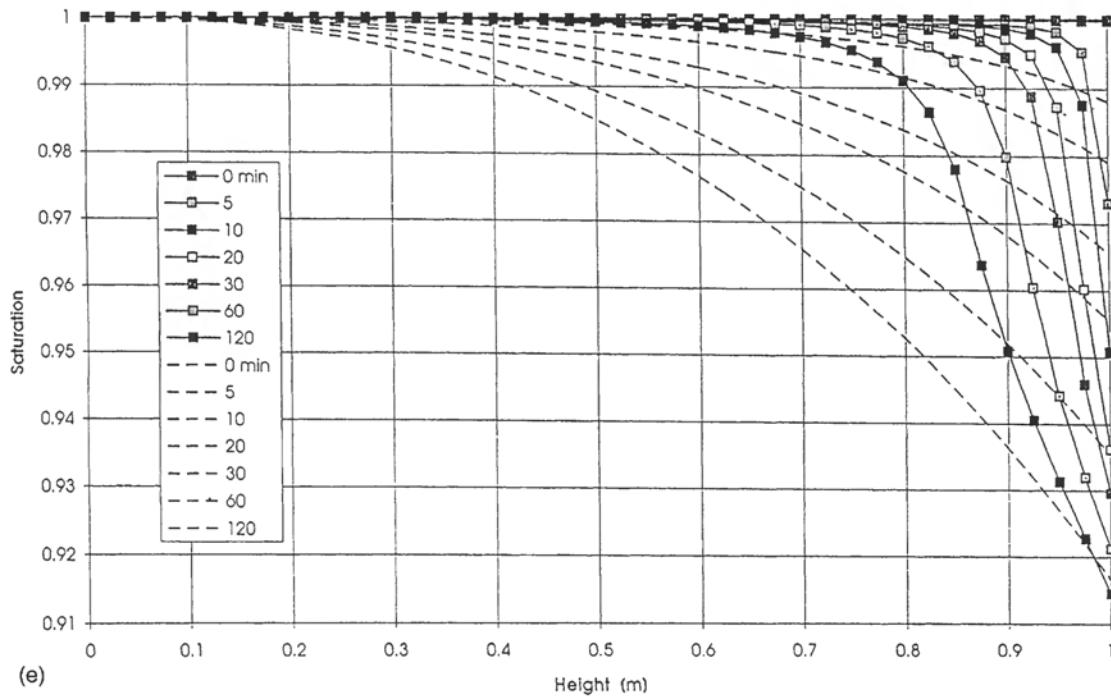


Figure 5.26 (continued)



(d)



(e)

**Figure 5.26** (continued)

Gauss point integration schemes were used without visible differences in the solution obtained. However, the different methods utilised in switching from one-phase flow to two-phase flow strongly influence the results [25].

Here we show the results for assumptions 1 and 2 with  $K_{rg\min} = 0.0001$  using a deformable material; the experimental parameters, are as follows:

Space mesh: 20 eight-node isoparametric finite elements of equal size

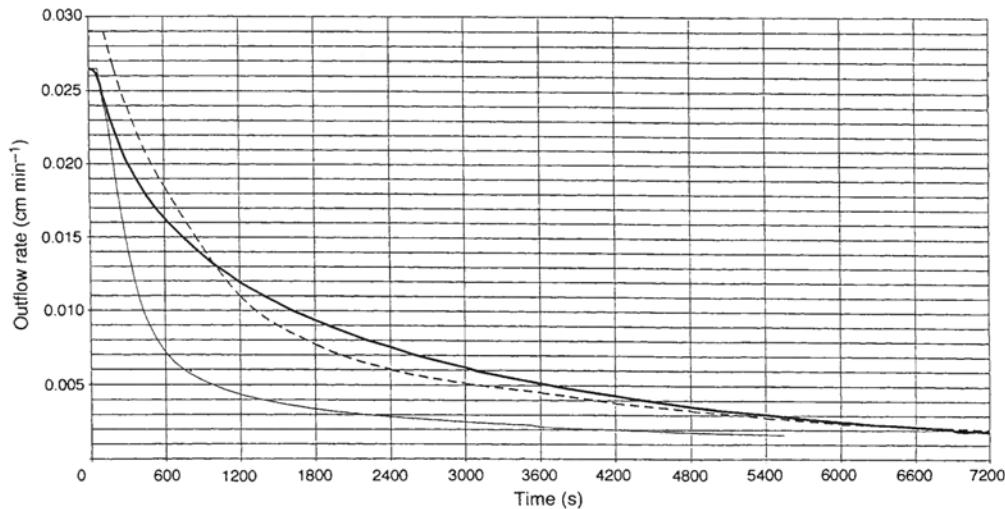
Time mesh: 3600 time steps of  $\Delta\tau = 1$  s, then 360 time steps of  $\Delta\tau = 10$  s

Integration:  $3 \times 3$  Gauss point

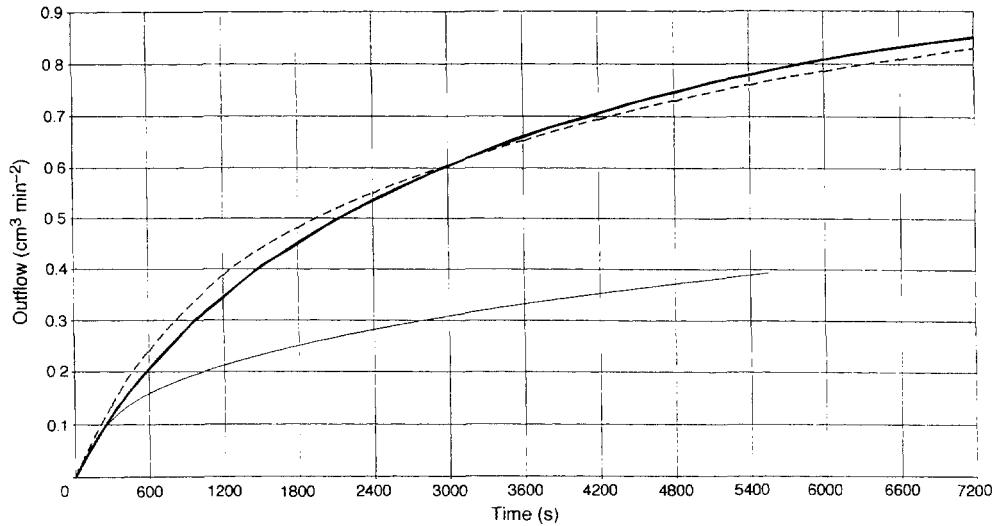
For the different solution methods, Figure 5.26 compares graphs of relative water pressure, absolute gas pressure (only for two-phase flow), capillary pressure, vertical displacements and the saturation of water.

The time histories of the outflow rate and the total water outflow via the base of the column under the two assumptions – one-phase flow where  $p_g = p_{atm}$  (solid line) and two-phase flow with the lower limit for relative gas permeability of  $K_{rg\min} = 0.0001$  (fine line) – are compared with the experimental data given by Liakopoulos (broken line) in Figs. 5.27 and 5.28. After the supply of water ceases, i.e. at  $t = 0$ , there is a gradually decreasing outflow rate, which corresponds to the real physical conditions for the Liakopoulos experiment. The outflow is due not only to the desaturation process, but also to the ‘squeezing’ effect related to soil deformability. This fact was first observed by Narasimhan and Witherspoon [2], who were able to obtain pressure profiles similar to the experimental ones only by accounting for the deformability of the soil.

Finally, during the flow, compaction of the sand takes place and the bulk stiffness increases as particles are pressed closer together. By assuming a constant Poisson’s ratio,



**Figure 5.27** Outflow rate at the bottom, (IC):  $p_c = 0$ , (—),  $p_g = p_{atm}$  (—)  $K_{rg} > 0.0001$ ; experiment (---)



**Figure 5.28** Outflow rate at the bottom, (IC):  $p_c = 0$ . (—)  $p_g = p_{atm}$ ; (—)  $K_{rg} > 0.0001$  experiment (- - -)

this results in a variable Young's modulus, which allows a more accurate representation of the measured pressure profiles at the initial stages of the test.

## 5.8 CONCLUSIONS

A series of test for fully and partially saturated consolidation and for drained and undrained situations have been presented in this chapter. In particular, elastic drained and undrained numerical analyses, for plane strain and radial symmetry conditions, have been compared with analytical solutions. A comparison of the predicted pore pressure dissipation with exact elastic analysis shows good agreement. Elastoplastic analyses have also been carried out and successfully compare with analytical solutions of triaxial tests in the case of Mohr–Coulomb and critical state yield criteria. Other tests of this nature are required, involving a variation of the third stress invariant. For a Mohr–Coulomb analysis there was good agreement with the viscoplastic program of Cormeau [3]. In the case of the critical state ellipse model, an incomplete solution was obtained due to the development of a negative diagonal term in the solution of the global equations. The problem did not occur when the combination model was used, with  $c = 0$ , in order to ensure that the occurrence of strain softening was unlikely. Further investigation is required to determine what factors are important in producing negative stiffness values. Some results can be found in Section 14.7.4, where strain localisation is briefly addressed.

The consolidation aspect for fully saturated situations was tested by comparison with existing numerical solutions. The variable permeability scheme was also validated by comparison with experimental data. Finally, multiphase flow in deforming porous media

was investigated and the results of several approaches were compared. The solutions outlined in these chapters and some new procedures are now applied in the next three chapters for the investigation of surface subsidence due to fluid withdrawal.

## REFERENCES

1. Norris, V. A. (1980) The elastoplastic analysis of soil consolidation with special reference to kinematic hardening. PhD thesis, University College of Swansea.
2. Poulos, H. G. and Davis, E. H. (1974) *Elastic Solutions for Soil and Rock Mechanics*, Wiley, Chichester.
3. Cormeau, I. C. (1976) Viscoplasticity and plasticity in the finite element method. PhD thesis, University of Wales, Swansea.
4. Humpheson, C. (1976) Finite element analysis of elastoviscoplastic soils. PhD thesis, University of Wales, Swansea.
5. Booker, J. R. (1973), The consolidation of a finite layer subject to surface loading. Report R223, University of Sydney.
6. Lewis, R. W., Roberts, G. K. and Zienkiewicz, O. C. (1976) A non-linear flow and deformation analysis of consolidation problems, in *Numerical Methods in Geomechanics*, Vol. 2, C. S. Desai (ed.), ASCE, New York, pp. 1106–18.
7. Siriwardane, H. J. and Desai, C. S. (1981) Two numerical schemes for nonlinear consolidation. *Int. J. Num. Meth. Engng.*, **17**, 405–26.
8. Roberts, G. K. (1976) Some aspects of finite element analysis of consolidation of soils. MSc thesis, University of Wales, Swansea.
9. Herrmann, L. R. (1965) Elasticity equations for incompressible, or nearly incompressible materials by variational theorem. JAIAA, American Institute of Aeronautics and Astronautics **3**, 1896–1905.
10. Christian, J. T. (1968) Undrained stress distribution by numerical methods. *J. Soil Mech. Found. Div. ASCE*, **94**, 1333–45.
11. Poulos, H. G. (1967) Stresses and displacements in an elastic layer underlain by a rough rigid base. *Geotechnique*, **17**, 378–410.
12. Milovic, D. M. Contraintes et déplacements dans une couche élastique d'épaisseur limitée, produits par une fondation circulaire. *Le Génie Civil*, **147**(5), 281–5.
13. Ueshita, K. and Meyerhof, G. G. (1968) Surface displacement of an elastic layer under uniformly distributed loads. *Highway Res. Record*, **228**, 1–10.
14. Reed, M. B. (1984) An investigation of numerical errors in the analysis of consolidation by finite elements. *Int. J. Num. Anal. Meth. Geomech.*, **8**, 243–57.
15. Sandhu, R. S., Liu, H. and Singh, K. J. (1977) Numerical performance of some finite element schemes for analysis of seepage in porous elastic media. *Int. J. Num. Anal. Meth. Geomech.*, **1**, 177–94.
16. Gibson, R. E., Schiffman, R. L. and Pu, S. L. (1970) Plane strain and axially symmetric consolidation of a clay layer on a smooth impervious base. *Q. J. Mech. Appl. Math.*, **23**, 505–20.
17. Liakopoulos, A. C. (1965) Transient flow through unsaturated porous media. PhD thesis, University of California, Berkeley.
18. Narasimhan, T. N. and Witherspoon, P. A. (1978) Numerical model for saturated–unsaturated flow in deformable porous media. 3, applications. *Water Resources Research*, **14**, 1017–34.
19. Schrefler, B. A. and Simoni, L. (1988) A unified approach to the analysis of saturated–unsaturated elastoplastic porous media, in *Numerical Methods in Geomechanics*, G. Swoboda (ed.), Balkema, Rotterdam, pp. 205–12.

20. Zienkiewicz, O. C., Xie, Y. M., Schrefler, B. A., Ledesma, A. and Bicanic N. (1990) Static and dynamic behaviour of soils: a rational approach to quantitative solutions, II. Semi-saturated problems. *Proc. R. Soc. Lond. A*, **429**, 311–21.
21. Schrefler, B. A. and Zhan Xiaoyong (1993) A fully coupled model for water flow and airflow in deformable porous media. *Water Resources Research*, **29**, 155–67.
22. Gawin, D., Baggio, P. and Schrefler B. A. (1995) Coupled heat, water and gas flow in deformable porous media. *Int. J. Num. Meth. Fluids*, **20**, 969–87.
23. Gawin, D. and Schrefler, B. A. (1996) Thermo-hydro-mechanical analysis of partially saturated porous materials. *Engng Comp.*, **13**(7), 113–43.
24. Brooks, R. N. and Corey A. T. (1966) Properties of porous media affecting fluid flow. *J. Irrig. Drain. Div. Amer. Soc. Civ. Engng.*, **92**(IR2), 61–68.
25. Gawin, D. and Schrefler B. A. (1996) Modelling of saturated–unsaturated transition in porous media, in *Proceedings of the GIMC-AMC Conference*, Padua.

# 6

## Modelling Subsidence: Numerical Aspects and Problems of Regional Scale

### 6.1 INTRODUCTION

From a physical viewpoint, we may have surface subsidence connected with the change of hydraulic equilibrium in systems comprising aquifers and aquitards and/or hydrocarbon reservoirs when there is extensive groundwater withdrawal and/or oil and gas pumppage. The intrinsic nature of the phenomenon is three-dimensional and interaction occurs between the solid and fluid (water, gas, oil) phases. Moreover, the problem is usually defined on a regional scale, i.e. the horizontal distances of interest are much greater than the thickness of the aquifer/aquitard or reservoir systems, and involves very long time spans. The timescale is in fact controlled by the filtration velocity and the pumping history.

The mathematical description of subsidence, i.e. a thermo-hydro-mechanical problem defined for a multicomponent medium, is governed by the following governing equations, as shown in Chapter 2:

1. One equilibrium equation for each constituent and for the whole multiphase medium.
2. One continuity equation for each constituent.
3. One energy balance equation for the system.

Each one of these equations for the generic constituent contains, in its general form, terms representing the effects of all the present fields. In real applications there is an obvious necessity to reduce the model complexity by introducing simplifying assumptions, and partial models are generally used in practice. The analysts have to carefully ascertain the admissibility and the correctness of such simplifications, in order to obtain meaningful numerical models and reliable forecasts. The problem may be simplified in several ways: it is possible to operate on the numerical model (two different techniques of this kind are presented in the following) or it is possible to operate on the theoretical model, e.g. by introducing suitable simplifying assumptions. One of the most common ways is obtained by solving separately the governing equations of the single fields, i.e. by *uncoupling* them.

For a numerical model, we want to stress the meaning of *coupling*, which is not commonly understood. We use the definition given by Zienkiewicz [1]:

Coupled formulations are those applicable to multiple domains and dependent variables which usually (but not always) describe different physical phenomena and in which

- neither domain can be solved separately from the other and
- neither set of dependent variables can be explicitly eliminated.

Sometimes, in closed-form solutions, it is possible to eliminate one variable by the inversion of a matrix of order one. In these cases the distinction between coupled and uncoupled solutions is rather delicate, but always possible and clear.

In subsidence problems, uncoupling is obtained by introducing into the equation of one field an estimated value of the terms representing the effect of the other fields. Such an approach is possible when dealing with isothermal consolidation, fully saturated conditions and where the constitutive relationships for each component are independent of the behaviour of the other field variables. Uncoupling is in fact possible when the problem is rather simple. Often the assumption of uncoupled fields may lead to correct results. However, it should not be concluded that all subsidence problems are uncoupled. The correctness of this particular approach should be carefully assessed each time.

For isothermal consolidation the opportunity to use uncoupled models, instead of coupled ones, is still a matter of debate [2–5]. Coupling is generally admitted when the forcing function is applied on the displacement field, e.g. a physical load on a foundation [6], but is questioned if forcing functions are applied to the flow field (e.g. outflow in subsidence problems). For the latter case, some authors state that the coupling effects between the stress and flow fields is not very important and may be disregarded when analysing long-term problems, which occur typically in land subsidence [7].

As a reference frame for the following discussion, we refer to the case where coupling is mainly questionable, e.g. isothermal fully saturated conditions. On the basis of the analysis of standard test problems, the importance of coupling appears in almost all situations. Compared with uncoupled solutions, coupled solutions generally lead to smaller displacements; but this is not always true, as illustrated by the last example in Chapter 7. If the same soil parameters are used, then coupled and uncoupled solutions generate different time histories of the state variables.

In a survey of the various models employed for the study of land subsidence due to the withdrawal of fluids from underlying aquifers and/or gas and oil from reservoirs, Corapcioglu [8] subdivided the models based on Terzaghi's effective stress principle [9] into two basic approaches. One approach stems from Biot's theory [10], where a simultaneous solution is sought for both the pore pressure and strain in the solid matrix. In the second approach, Terzaghi's theory is implemented in a two-step procedure [11,12]. First the pressure distribution is obtained either by solving a fluid flow equation in a two- or three-dimensional domain, or from field measurements. Land subsidence is then calculated assuming one-dimensional consolidation theory. The flow equation is usually derived by assuming vertical displacements only and a constant total stress [8]. These assumptions produce uncoupling, in Biot's formulation, between the flow and equilibrium equations.

For this discussion, we start from Biot's equations. By properly reducing the equations presented in Chapter 2, we have the following results, as in Chapter 3:

1. The equilibrium equation (linear momentum balance) for the two-phase medium composed of water and soil. This equation, on ignoring the convective and inertial terms, can be written as

$$\operatorname{div} \boldsymbol{\sigma} + \rho \mathbf{g} = \mathbf{0} \quad (6.1)$$

Also, the constitutive relationship, which accounts for Terzaghi's effective stress principle is given by

$$d\boldsymbol{\sigma} = \mathbf{D}_T d\boldsymbol{\epsilon} - \alpha \mathbf{m} dp^w \quad (6.2)$$

2. The continuity equation for the mixture of water and solid which takes the form

$$\left[ \frac{n}{K_w} + \frac{\alpha - n}{K_s} \right] \frac{\partial p^w}{\partial t} + \alpha \mathbf{m} \frac{\partial \boldsymbol{\epsilon}}{\partial t} + \operatorname{div} \left[ \frac{\mathbf{k}}{\mu^w} (-\operatorname{grad} p^w + \rho^w \mathbf{g}) \right] = 0 \quad (6.3)$$

The displacement vector  $\mathbf{u}$  of the solid skeleton and the pore pressure  $p^w$  are chosen as the basic variables of the problem. The tangent stiffness matrix  $\mathbf{D}_T$  is dependent on the level of effective stress and on the total strain of the skeleton. Furthermore, the permeability matrix may also be considered as strain dependent. In such a case, however, it is impossible to consider the two fields as being uncoupled.

The standard spatial discretization of equations (6.1) and (6.3), taking into account (6.2) as in Chapter 3, yields the following system of equations, similar to (3.46):

$$\begin{bmatrix} -\mathbf{K} & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} \end{bmatrix} \frac{d}{dt} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{Bmatrix} = \begin{Bmatrix} -\frac{d\mathbf{f}^u}{dt} \\ \mathbf{f}^p \end{Bmatrix} \quad (6.4)$$

where  $\bar{\mathbf{u}}$  and  $\bar{\mathbf{p}}^w$  are the nodal values of the basic variables. A definition of the other terms in (6.4) can be found in Chapter 3. After integration in the time domain, e.g. using the generalised midpoint rule, both displacements and pore pressures can be obtained simultaneously for the whole time history. This is the general procedure applied throughout this book, unless stated otherwise. It can be easily verified that the coupling between the two fields occurs only during the time transient and disappears when steady-state conditions are reached.

A few other approaches commonly found in the literature will now be discussed in some detail. A coupled two-step procedure, or two single-field analyses, can be derived from the time-discretised form of (6.4) by eliminating either  $\bar{\mathbf{u}}$  or  $\bar{\mathbf{p}}^w$ . This was done by Krause [13] and Smith [14] for the case of a linear elastic porous solid and for an incompressible pore fluid. The equations used by these authors were somewhat different from (6.4): the first equation is written in integral form, whereas the second omits the capacity matrix  $\mathbf{S}$  because of the incompressibility of the pore fluid. If subsidence (compaction) history is the main concern and if we know the outflow rate  $Q$  (volume of water per unit area per unit time), we can use the so-called *displacement formulation* [14], which is obtained by eliminating  $\bar{\mathbf{p}}^w$ . After integration in the time domain, the

resulting displacements are obtained from

$$\begin{aligned} (\mathbf{Q}\mathbf{H}^{-1}\mathbf{Q}^T + \theta\Delta t\mathbf{K})\bar{\mathbf{u}}_{n+1} &= (\mathbf{Q}\mathbf{H}^{-1}\mathbf{Q}^T - (1-\theta)\Delta t\mathbf{K})\bar{\mathbf{u}}_n \\ &\quad + \Delta t(\theta\mathbf{f}_{n+1}^u + (1-\theta)\mathbf{f}_n^u + \mathbf{Q}\mathbf{H}^{-1}\mathbf{f}^p) \end{aligned} \quad (6.5)$$

where the subscript  $n$  refers to the generic time station in which the solution is calculated. In this chapter, a comma does not represent differentiation but only distinguishes between two different symbols. Krause's solution is almost identical to (6.5). The pressure is then obtained from the displacement field by numerical differentiation, as can be seen from the second equation of (6.4) with no  $\mathbf{S}$  matrix.

If the main interest is in calculating the drawdown, then the displacement vector is eliminated from the simplified set of equation (6.4) and a *pressure formulation* [13] is obtained after discretisation in the time domain:

$$\begin{aligned} (\mathbf{Q}^T\mathbf{K}^{-1}\mathbf{Q} + \theta\Delta t\mathbf{H})\bar{\mathbf{p}}_{n+1}^w &= (\mathbf{Q}^T\mathbf{K}^{-1}\mathbf{Q} - (1-\theta)\Delta t\mathbf{H})\bar{\mathbf{p}}_n^w \\ &\quad + \Delta t(\mathbf{f}^p - \mathbf{Q}^T\mathbf{K}^{-1}\dot{\mathbf{f}}^u) \end{aligned} \quad (6.6)$$

The equilibrium equation (the first equation in (6.4) in integral form) then yields directly the displacement field. In equation (6.5) or (6.6) the number of unknowns is reduced with respect to equation (6.4), but the resulting matrices are generally full, so the cost of forming and reducing them is high for linear problems and prohibitive for non-linear problems. Furthermore, for non-linear problems, it is necessary to solve both fields at each time step and check the convergence of the solution. Both the *displacement* and *pressure formulations* are fully coupled solutions and can be used as two-step procedures, if we are concerned with both fields.

However, alternative approaches exist in the literature for solving linear subsidence problems in particular. The most common is the uncoupled two-step procedure based on Terzaghi's theory. Another common approach is based on the nucleus-of-strain concept, which is widely used for steady-state thermoelasticity problems [15,16]. In this case the continuity equation for water [17] is

$$\begin{aligned} \operatorname{div}[-\mathbf{k}^w \operatorname{grad} p^w + \rho^w \mathbf{g}] &= \frac{n}{K_w} \frac{\partial p^w}{\partial t} \\ &\quad + \tilde{\alpha} \left[ \frac{\partial p^w}{\partial t} + \frac{1-2v}{\pi} \int_{\Omega} \frac{\partial p^w}{\partial t} \frac{2(z+z_1)^2 - (x-x_1)^2 - (y-y_1)^2}{[(x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2]^{5/2}} dx_1 dy_1 dz_1 \right] \end{aligned} \quad (6.7)$$

where  $\tilde{\alpha}$  represents the vertical compressibility of the soil,  $x_1, y_1, z_1$  are the coordinates of the tension centre and  $x, y, z$  the coordinates of the generic point. The integral term added in the continuity equation (6.7), on the basis of simplified physical assumptions, accounts for the effect of the volumetric strain of the solid skeleton and is of the type

$$\int_{\Omega} \varepsilon^*(P, \xi) \frac{\partial p^w}{\partial t} d\Omega \quad (6.8)$$

where  $\varepsilon^*$  may be regarded as the *Green function* of the homogeneous half-space steady state (hence uncoupled) problem, not the transient coupled problem. In such a formulation  $\varepsilon^*$  does not depend on time, but only on the point where the effects are calculated say  $P$ , and on the point where the strain nucleus is located, say  $\xi$ . In general, Green functions depend on the particular case under analysis and are usually much more complicated as the time variable must appear (sink case [18,19], load case [20]). Moreover, the Green function is strictly related to a particular problem, e.g. homogeneous half-space, hence (6.7) cannot be used in general. Equation (6.7) can be numerically integrated, and the displacement field can then be obtained using a solid mechanics algorithm, which considers pressure changes as a body force. In this way the solution for the flow field is sought independently from that of the displacement field, i.e. the problem is assumed to be uncoupled, owing to the introduction of the hypotheses leading to (6.7). This simplification is not always possible, especially in real applications where often the constitutive relationship of each field variable depends on the variables of the other fields. As an example, we may cite the dependence of the permeability parameter on the deformation of the solid skeleton [21].

### 6.1.1 More About Coupling and Staggered Procedures

From the previous definition of coupling, it follows that, in subsidence problems due to withdrawal of groundwater fluids, it is only meaningful to speak of coupled or uncoupled models if the outflow (pumping rate) is specified. If the pressure drop is given in the whole reservoir, e.g. from field measurements, then a simple displacement analyser (e.g. a solid mechanics code which considers pressure as a body force) gives the correct surface subsidence, provided it takes into account three-dimensional strain effects and non-linear material behaviour, when necessary. In this case the concept of coupling or uncoupling has no real meaning.

We consider the case when the outflow is specified. Models based on the general theory, presented in the previous chapters, are generally coupled. But they may uncouple in particular situations, e.g. when steady-state conditions are considered or in the case of a single axisymmetric aquifer with fluid withdrawal from a fully penetrating well. This last case will be discussed in further detail in a later section. Finally, two-step models have to be considered as coupled if the three-dimensional strain effects are correctly included in the flow equation and in the equation(s) yielding land subsidence. However, if three-dimensional strain effects are taken into account, but volumetric strain is a function of pore pressure only, then the problem is still uncoupled.

From an inspection of the *transient* coupled system of equations (6.1) to (6.3) it follows that a condition for uncoupling is that the underlying stress problem is statically determinate, i.e. that the volumetric strain  $\varepsilon_v = \varepsilon_x + \varepsilon_y + \varepsilon_z$ , which appears in the continuity equation differentiated with respect to time, can be directly expressed as a function of the pore pressure through the equilibrium equations and the constitutive relationships, without invoking geometrical compatibility. The continuity equation can be interpreted as an additional compatibility equation in the sense that the displacement field has to be *compatible*, not only according to the usual physical interpretation of compatibility (i.e. the deformed body has to be continuous), but also with regard to the

conservation of the fluid present in a reference volume element (i.e. the quantity of fluid present has to vary according to the transport law).

In real problems it is not always easy to judge whether the stress problem is statically determinate. This may be achieved by means of simplifying assumptions, for instance when the total stress is assumed as one-dimensional, constant in time and due to overburden only. For this case

$$\varepsilon_v = \frac{1-2\nu}{E} (\sigma_z + 3p^w) \quad (6.9)$$

Hence it is possible to express the capacity term  $\dot{\varepsilon}_v$  as a function of  $\dot{p}^w$  and consequently to eliminate the displacement field in the water continuity equation. A particular case of the previous statement is the Terzaghi–Jacob theory of consolidation [9]. Similar scenarios arise when the horizontal displacements are constant or zero (horizontal strains constant in time or zero) and the total stress  $\sigma_z(t)$  due to an overburden load is known.

A simple way of verifying *a posteriori* the coupling of the problem in question is achieved by using a staggered solution procedure. Since staggered procedures will be used on several occasions in the text, a detailed explanation is given as follows.

As shown in Chapter 3, equation (6.4) can be concisely written as

$$\mathbf{B}\dot{\mathbf{X}} + \mathbf{C}\mathbf{X} = \mathbf{F} \quad (6.10)$$

with all symbols easily identified. After discretisation in time, using the generalised midpoint rule, equation (6.10) becomes (3.77):

$$\mathbf{A}\mathbf{X}_{n+1} - \mathbf{B}\mathbf{X}_n - \mathbf{F} = 0 \quad (6.11)$$

Equation (6.11) is called the *monolithic equation*. The following partitioning for matrix  $\mathbf{A}$  is introduced:

$$\mathbf{A} = \mathbf{A}^L + \mathbf{A}^R \quad (6.12)$$

$$\mathbf{A}^L = \begin{vmatrix} \mathbf{a}_{11} & \mathbf{0} \\ 0 & \mathbf{a}_{22} \end{vmatrix} \quad \mathbf{A}^R = \begin{vmatrix} \mathbf{0} & \mathbf{a}_{12} \\ \mathbf{a}_{21} & \mathbf{0} \end{vmatrix} \quad (6.13)$$

Then equation (6.11) can be written as

$$\mathbf{A}^L\mathbf{X}_{n+1} = \mathbf{B}\mathbf{X}_n + \mathbf{F} - \mathbf{A}^R\mathbf{X}_{n+1} \quad (6.14)$$

After some arrangements, the partitioned form is obtained as

$$\mathbf{x}_{n+1} = -\mathbf{G}_1\mathbf{y}_{n+1} + \mathbf{H}_1\mathbf{X}_n + \mathbf{f}_1 \quad (6.15a)$$

$$\mathbf{y}_{n+1} = -\mathbf{G}_2\mathbf{x}_{n+1} + \mathbf{H}_2\mathbf{X}_n + \mathbf{f}_2 \quad (6.15b)$$

where

$$\begin{aligned} \mathbf{X}_{n+1} &= [\mathbf{x}_{n+1,K}, \mathbf{y}_{n+1,K}]^T \\ \mathbf{f} &= [\mathbf{f}_1, \mathbf{f}_2]^T \quad \mathbf{f}_1 = \mathbf{a}_{11}^{-1} \mathbf{F}_1 \quad \mathbf{f}_2 = \mathbf{a}_{22}^{-1} \mathbf{F}_2, \\ \mathbf{G}_1 &= \mathbf{a}_{11}^{-1} \mathbf{a}_{12} \quad \mathbf{G}_2 = \mathbf{a}_{22}^{-1} \mathbf{a}_{21} \\ \mathbf{H}_1 &= \mathbf{a}_{11}^{-1} [\mathbf{b}_{11}, \mathbf{b}_{12}]^T \quad \mathbf{H}_2 = \mathbf{a}_{22}^{-1} [\mathbf{b}_{21}, \mathbf{b}_{22}]^T \end{aligned} \quad (6.16)$$

The staggered iteration process is generally defined as follows. First, the value of  $\mathbf{x}_{n+1}$ , appearing on the right-hand side of equation (6.1.15b) is predicted by assuming a linear constitutive behaviour. Therefore  $\mathbf{x}_{n+1}^{(p)} = \mathbf{x}_{n+1,0}$ , and from equation (6.1.15b) the value of  $\mathbf{y}_{n+1,0}$  is calculated, then substituted into equation (6.1.15a), which is now used to solve for  $\mathbf{x}_{n+1,1}$ . Now  $\mathbf{x}_{n+1,1}$  is used as a predictor in equation (6.1.15b) and the whole process is repeated until a required tolerance is attained. The final step is calculate  $\mathbf{y}_{n+1,K}$ , then

$$\mathbf{X}_{n+1} = [\mathbf{x}_{n+1,K}, \mathbf{y}_{n+1,K}]^T$$

The predictor  $\mathbf{x}_{n+1,0}$  has the form

$$\mathbf{x}_{n+1,0} = \sum_{i=0}^m \beta_i \mathbf{x}_{n-1,K} \quad (6.17)$$

where  $K$  is the number of performed iterations.

The numerical properties of the staggered procedure will be investigated in Chapter 10.

The strength of coupling is analysed as follows. Equation (6.11) in extended form (the first equation is used in integral form) may be written as follows:

$$\begin{vmatrix} \theta \mathbf{K} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} + \theta \Delta t \mathbf{H} \end{vmatrix} \left\{ \begin{array}{l} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{array} \right\}_{n+1} = \begin{vmatrix} (\theta - 1) \mathbf{K} & -(\gamma \theta - 1) \mathbf{Q} \\ \gamma \mathbf{Q}^T & \mathbf{S} - (1 - \gamma \theta) \Delta t \mathbf{H} \end{vmatrix} \left\{ \begin{array}{l} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{array} \right\}_n - \gamma \begin{vmatrix} \mathbf{0} & -\theta \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{0} \end{vmatrix} \left\{ \begin{array}{l} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \end{array} \right\}_{n+1} + \left\{ \begin{array}{l} \mathbf{f}^u \\ \Delta t \mathbf{f}^p \end{array} \right\} \quad (6.18)$$

where  $\theta$  is the parameter for the temporal discretisation and the following type of predictor for the unknowns has been introduced [22]:

$$\mathbf{X}_{P,n+1} = (1 - \gamma) \mathbf{X}_n + \gamma \mathbf{X}_{n+1} \quad (6.19)$$

with  $0 \leq \gamma \leq 1$  and in general  $\gamma = \theta$ .

The staggered scheme starts in this particular case by solving the first equation of system (6.18) with  $\gamma = 0$  (last solution), then the second equation is solved with the calculated value of  $\mathbf{u}_{n+1}$  and with the actual value of  $\gamma$ . Henceforth, the procedure is iterative until the convergence is achieved within the same time step. The next time step proceeds in a similar manner.

The iteration count within each time step in this staggered procedure can be assumed as a measure of the strength of the coupling between the fields: if the coupling is weak,

the convergence of the iterative process will be fast [1]. But note that, for the same convergence limit, the absolute number of iterations depends on the asymptotic convergence rate of the staggered solution, i.e. on the spectral radius of the iterative matrix; see also equation (6.51). This depends in turn on the spatial and temporal discretisations. Moreover, the number of iterations also depends on the initial error, which is related to the type of predictor used. Hence, to obtain a measure of the coupling strength, we have to eliminate the effect of the spatial and temporal discretisations. This may be achieved by controlling the iteration count for different discretisations of the same problem, or by using the same discretisation for different problems defined in the same domain. Section 6.5 describes an application that contains useful benchmark problems.

Before continuing, let us introduce some useful numerical techniques, which reduce the computational effort in solving numerical subsidence models. All these techniques will be applied here and in later chapters.

## 6.2 PROBLEMS OF REGIONAL SCALE: VERTICALLY AVERAGED MODELS

Real problems of surface subsidence due to the withdrawal of underground fluids are usually defined in domains of regional scale. The traditional approach to such three-dimensional problems is to apply simplified two-dimensional solutions, i.e. axisymmetry, plane strain or plane stress. Axisymmetric models are extensively applied in the next chapter, as three-dimensional numerical models, especially for fully coupled formulations, present a number of unknowns which, together with the necessity to consider long time spans, may preclude an effective and economic numerical analysis. However, these assumptions are sometimes too drastic. For instance, the hypothesis of a two-dimensional model can hardly be justified in the case of withdrawal from an aquifer of varying thickness in the presence of arbitrarily distributed wells. An alternative approach is a mathematical model based on the application of an averaging technique, in which governing equations are integrated (or averaged) over the smallest dimension of the domain (thickness). The numerical solution of the resulting discretised equations requires a partitioning procedure for linear cases, too. The advantage lies in the fact that a two-dimensional problem ensues, in which the three-dimensional nature of the problem is accurately accounted for. This may be done exactly or approximately, depending on a knowledge of the boundary conditions involved in the averaging procedure. This type of mathematical model was first introduced by Bear and Corapcioglu [23] and its numerical solution, using matrix partitioning, is given in the literature [24].

### 6.2.1 Spatially Averaged Quantities: The Megascopic Level

In averaging theories, governing equations for the mechanical behaviour of porous media are obtained by applying some kind of averaging technique to the equations established in the real non-homogeneous domain of the porous medium (microscopic level), as shown in Chapter 2. This is performed by integrating the equations over a

representative elementary volume (REV), in which the non-homogeneous nature of the medium is accounted for. The dimensions of the REV must satisfy both the lower and upper bounds, which respectively depend on the characteristic length of the microscopic and macroscopic inhomogeneities. Starting from the macroscopic level, a similar procedure is employed here to obtain the equations, or quantities, defined at the *megascopic level*, i.e. when the REV is a large volume of the porous medium, so at least one of its dimensions is comparable with the smallest dimension of the entire domain. The analogy to the jump conditions, existing at interfaces between components occurring at the microscopic level, may also occur at the macroscopic level. For instance, we may have a thick layer of a porous medium which has embedded layers of differing permeability values. Bounds for the dimension of the REV still exist, but they are only upper bounds and depend on the characteristic length over which changes in the values of field variables must be known. We may assume one of the dimensions of the REV to be equal to the smallest dimension of the domain (thickness). The other two dimensions of the REV are small enough to prevent significant variations of the field variables in their directions. This requires that the thickness  $B$  of the domain is sufficiently small in comparison with other dimensions of the porous medium. In this situation the field variables are averaged quantities over the thickness. Furthermore, equilibrium and continuity equations are valid for the whole thickness.

We refer to a generic field variable  $\psi(\mathbf{x}, t)$  as a *macroscopic variable* and to its mean value  $\hat{\psi}(\bar{\mathbf{x}}, t)$  over the REV as a *megascopic variable*. A megascopic variable is defined in the usual manner as

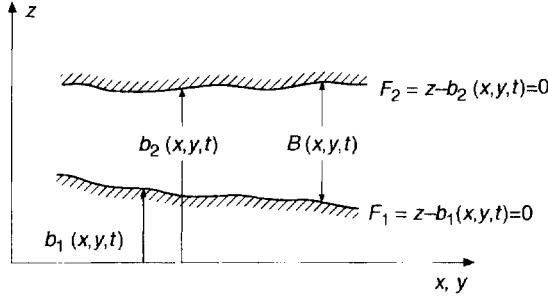
$$\hat{\psi}(\bar{\mathbf{x}}, t) = \frac{1}{U_0} \int_{U_0} \psi(\mathbf{x}, t) dV \quad (6.20)$$

where  $U_0$  is the volume of the REV, and  $\bar{\mathbf{x}}$  and  $\mathbf{x}$  represent the megascopic and the macroscopic scale (domain) respectively. It is worth noting that  $\psi(\mathbf{x}, t)$  is defined in  $R^3$ , say the domain  $\Omega(x, y, z)$ , whereas  $\hat{\psi}(\bar{\mathbf{x}}, t)$  is defined in  $R^2$ , say the domain  $\Omega'(x, y)$ . Due to the assumed dimension of the REV in the plane normal to the thickness, we may also write

$$\hat{\psi}(\bar{\mathbf{x}}, t) = \frac{1}{B} \int_{b_1}^{b_2} \frac{1}{\partial A} \int_{\partial A} \psi(\mathbf{x}, t) dz dA = \frac{1}{B} \int_{b_1}^{b_2} \psi(\mathbf{x}, t) dz \quad (6.21)$$

where  $b_1$  and  $b_2$  are functions of  $x$  and  $y$  only (and  $t$ , if necessary) and define the upper and lower limiting surface of the domain (Figure 6.1 explains the nomenclature). In this procedure, attention must be paid to the definition of the mean value of a variable which is proportional to a derivative of a field variable, e.g. a strain component. In order to define the mean value, according to previous hypotheses on the dimensions of the REV, we use Leibnitz's rule:

$$\frac{\partial \hat{\psi}}{\partial t} = \frac{1}{B} \int_{b_1}^{b_2} \frac{\partial \psi}{\partial t} dz = \frac{1}{B} \left[ \frac{\partial}{\partial t} (B \hat{\psi}) - \psi|_{b_2} \frac{\partial b_2}{\partial t} + \psi|_{b_1} \frac{\partial b_1}{\partial t} \right] \quad (6.22)$$



**Figure 6.1** Nomenclature

where the average value  $\hat{\psi}$  is defined by equation (6.21) and

$$B(x, y, t) = b_2(x, y, t) - b_1(x, y, t) = F_2(x, y, z, t) - F_1(x, y, z, t) \quad (6.23)$$

Notation  $\psi|_{b_i}$  indicates the value of variable  $\psi$  at limiting surface  $b_i$ .

In equation (6.22) boundary conditions appear at the upper and lower surfaces of the domain. This fact is not usual in homogenisation problems, because the dimensions of the REV are appropriately smaller than the domain. But some simplifications are possible in equation (6.22) when the field variable assumes a nearly uniform value over the thickness, as will be shown below. The three-dimensional nature of the problem is accounted for exactly only when these boundary conditions are known.

## 6.2.2 Macrolevel Governing Equations

Governing equations of the proposed mathematical model may be obtained by integrating the balance equations over the thickness. Attention is limited to the fully saturated isothermal case, as the non-saturated case has a strictly local nature and the temperature field requires knowledge outside the stratum over which we integrate the balance laws.

### 6.2.2.1 Equilibrium Equation for the Two-Phase Medium

The linear momentum balance condition for the two-phase medium is expressed by its weak formulation (3.38). This equation is integrated over the thickness, under the assumption of no inertial forces. The work done by boundary tractions  $\hat{\mathbf{t}}$  at the upper and lower limiting surfaces is accounted for in a separate term. Averaged values are used in writing the resulting equilibrium condition, and integration is over the domain in which variations occur for integrated terms, i.e.  $\Omega'(x, y)$ . Hence the equilibrium equation takes the following form:

$$\begin{aligned} & - \int_{\Omega'} B \delta \hat{\boldsymbol{\epsilon}}^T \mathbf{D}_T \hat{\boldsymbol{\epsilon}} d\Omega + \int_{\Omega'} B \rho \delta \hat{\mathbf{u}}^T \mathbf{g} d\Omega + \int_{\Omega'} B \delta \hat{\boldsymbol{\epsilon}}^T \alpha \mathbf{m} \hat{p}^w d\Omega + \int_{\Omega'} B \delta \hat{\boldsymbol{\epsilon}}^T \mathbf{D}_T \hat{\boldsymbol{\epsilon}}^0 d\Omega \\ & + \int_{\Omega'} B \delta \hat{\boldsymbol{\epsilon}}^T \mathbf{D}_T \hat{\boldsymbol{\epsilon}} d\Omega + \int_{\Gamma'} \delta \hat{\mathbf{u}}^T \hat{\mathbf{t}} d\Gamma + \int_{\Gamma''} \delta \hat{\mathbf{u}}^T \hat{\mathbf{t}} d\Gamma = 0 \end{aligned} \quad (6.24)$$

where  $\hat{\epsilon}$  and  $\hat{u}$  are the mean values over the thickness of the strains and displacements. Their volume averages are calculated according to the previous statements. Moreover,  $\Gamma'$  represents the contour of  $\Omega'(x, y)$  and  $\Gamma''$  represents the upper and lower limiting surfaces  $F_1$  and  $F_2$ . Note that the constitutive equation is also defined at the megascopic level.

The averaged values of the strain tensor components are next obtained using equations (6.21) and (6.22). A comma implies differentiation:

$$\begin{aligned}\hat{\epsilon}_x &= \hat{u}_{x,x} + \frac{1}{B} [\hat{u}_x B_{,x} + u_x|_{F_2} F_{2,x} - u_x|_{F_1} F_{1,x}] \\ \hat{\epsilon}_y &= \hat{u}_{y,y} + \frac{1}{B} [\hat{u}_y B_{,y} + u_y|_{F_2} F_{2,y} - u_y|_{F_1} F_{1,y}] \\ \hat{\epsilon}_z &= \hat{u}_{z,z} + \frac{1}{B} [\hat{u}_z B_{,z} + u_z|_{F_2} F_{2,z} - u_z|_{F_1} F_{1,z}] \\ \hat{\gamma}_{xy} &= \hat{u}_{x,y} + \hat{u}_{y,x} + \frac{1}{B} [\hat{u}_x B_{,y} + u_x|_{F_2} F_{2,y} - u_x|_{F_1} F_{1,y} + \hat{u}_y B_{,x} + u_y|_{F_2} F_{2,x} - u_y|_{F_1} F_{1,x}] \\ \hat{\gamma}_{xz} &= \hat{u}_{z,x} + \frac{1}{B} [\hat{u}_z B_{,x} + u_z|_{F_2} F_{2,x} - u_z|_{F_1} F_{1,x} + u_x|_{F_2} F_{2,z} - u_x|_{F_1} F_{1,z}] \\ \hat{\gamma}_{yz} &= \hat{u}_{z,y} + \frac{1}{B} [\hat{u}_z B_{,y} + u_z|_{F_2} F_{2,y} - u_z|_{F_1} F_{1,y} + u_y|_{F_2} F_{2,z} - u_y|_{F_1} F_{1,z}]\end{aligned}\quad (6.25)$$

As can be seen from equation (6.25), in addition to the mean value of the displacements, we must consider derivatives of the thickness  $B(x, y, t)$ , derivatives of both the upper and lower boundary surfaces of the domain  $b_i(x, y, t)$  and values of the displacement field at these boundaries. But these equations may be simplified depending on the boundary conditions.

It is interesting to note that the third equation in (6.25) simplifies to

$$\hat{\epsilon}_z = \frac{1}{B} [u_z|_{F_2} - u_z|_{F_1}] \quad (6.26)$$

due to the independence of  $\hat{u}_z$  and  $B(x, y, t)$  from  $z$ , resulting in a linear distribution of  $\hat{u}_z$  over the thickness. Moreover, if the bottom surface is fixed, it is possible to write

$$\epsilon_z = \frac{\delta}{B} \quad \hat{u}_z = \frac{\delta}{2} \quad (6.27)$$

where  $\delta(x, y, t)$  is the compaction of the domain.

### 6.2.2.2 Fluid-Phase Behaviour

The mass balance equation of the fluid phase is obtained by integrating over the thickness the point mass balance equation presented in Chapter 2. Some simplifications are introduced in the present formulation, in particular:

1. Vertical lines are assumed equipotential, i.e.  $\hat{p}^w \approx p^w|_{b_2} \approx p^w|_{b_1}$ .
2. The first invariant of the effective stress is assumed constant along the thickness, i.e.  $\hat{I}_1 \approx I_1|_{b_2} \approx I_1|_{b_1}$ .

Integration of (6.3) over the thickness of the stratum, also accounting for fluxes  $q_1$  and  $q_2$  at the lower and upper limiting surfaces, yields

$$\left[ \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right] B \frac{\partial \hat{p}^w}{\partial t} + \alpha B \mathbf{m} \frac{\partial \hat{\epsilon}}{\partial t} - B \operatorname{div} [\hat{\mathbf{k}}^w \operatorname{grad} \hat{p}^w - \hat{p}^w \hat{\mathbf{b}}] + q_1 + q_2 = 0 \quad (6.28)$$

Note that all terms in (6.28) vary in the  $x$  and  $y$  direction only, and the validity of Darcy's law at the megalevel is implied for transport of the liquid phase.

The solution of the problem is obtained from (6.24) and (6.28), where the unknowns are the mean value of displacements  $\hat{u}_x$  and  $\hat{u}_y$ , the mean value  $\hat{p}^w$  of pressure at every point of  $\Omega'(x, y)$  and the values of displacements at the upper and lower surfaces of the domain. The last terms may be known in part through the boundary conditions. In conclusion, the problem is defined in a two-dimensional domain but still exhibits an original three-dimensional nature, which results from the assumed unknowns.

### 6.2.3 Implementation of the Numerical Model

For the numerical solution, the following simplifications are made:

1. The lower surface  $F_1 = 0$  is assumed fixed and rigid. Both surfaces  $F_1 = 0$  and  $F_2 = 0$  are assumed to be smooth. These assumptions are acceptable for real problems.
2. Distribution over the thickness of the displacement components  $\hat{u}_x$  and  $\hat{u}_y$  depends strongly on the boundary conditions at the upper and lower limiting surfaces. If these surfaces are smooth, the displacements are nearly uniform over the thickness. If the surfaces are not smooth, or the aquifer is embedded in a half-space with different hydraulic properties, the horizontal displacements are almost negligible (Section 6.5).
3. Limiting surfaces are assumed impermeable. This implies that the fluid flux is mainly horizontal. Possible sources or sinks are modelled as being uniformly distributed over the thickness.

Due to the fact that

$$u_{zz} = 0, B_{1z} = 0 \quad F_{iz} = 1 \quad (i = 1, 2) \quad (6.29)$$

and assuming

$$\hat{u}_\alpha \approx u_\alpha|_{b_1} \approx u_\alpha|_{b_2} \quad (\alpha = x, y) \quad (6.30)$$

for the reasons given earlier, we can express the strain-displacements relationship

(6.25) in the following matrix form:

$$\begin{Bmatrix} \hat{\varepsilon}_x \\ \hat{\varepsilon}_y \\ \hat{\gamma}_{yy} \\ \hat{\varepsilon}_z \\ \hat{\gamma}_{xz} \\ \hat{\gamma}_{yz} \end{Bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \\ 0 & 0 & \frac{1}{B} \\ 0 & 0 & \frac{1}{2} \frac{\partial}{\partial x} \\ 0 & 0 & \frac{1}{2} \frac{\partial}{\partial y} \end{bmatrix} \begin{Bmatrix} \hat{u}_x \\ \hat{u}_y \\ \delta \end{Bmatrix} \quad (6.31)$$

The constitutive equation is the usual one for an elastic, isotropic medium in a three-dimensional case. Using the same interpolation functions for displacements and thickness, and appropriate interpolation functions for pore pressures, the standard Galerkin application of the finite element discretisation in space may be applied to (6.24) and (6.28), yielding the following system of partial differential equations in time:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \begin{Bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{p}}^w \end{Bmatrix} + \begin{bmatrix} -\mathbf{K} & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} \end{bmatrix} \frac{d}{dt} \begin{Bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{p}}^w \end{Bmatrix} = \begin{Bmatrix} -\frac{d\hat{\mathbf{f}}_1}{dt} + \mathbf{C} \\ \hat{\mathbf{f}} \end{Bmatrix} \quad (6.32)$$

where displacement vector  $\hat{\mathbf{u}}$  contains nodal components  $\hat{u}_x$ ,  $\hat{u}_y$  and nodal compaction  $\delta$ . The matrices of equation (6.32) are given by equations (3.40).

Note that, to obtain (6.32), equation (6.24) has been differentiated with respect to time for numerical purposes only. In this way a symmetric coefficient matrix is obtained, which is computationally advantageous both in terms of memory and time.

Equation (6.32) is solved using the following procedure :

1. The displacement vector is ordered as follows:

$$\begin{Bmatrix} \hat{u}_x \\ \hat{u}_y \\ \delta \end{Bmatrix} = \{u_{1x}, u_{1y}, u_{2x}, u_{2y}, \dots, u_{nx}, u_{ny}, \delta_1, \delta_2, \dots, \delta_n\}^T \quad (6.33)$$

2. The element matrices  $\mathbf{K}$ ,  $\mathbf{Q}$ ,  $\mathbf{Q}^T$  are rewritten as follows:

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 \\ \mathbf{K}_2^T & \mathbf{K}_3 \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2^T \end{bmatrix} \quad \mathbf{Q}^T = [\mathbf{Q}_1^T \quad \mathbf{Q}_2^T] \quad (6.34)$$

where the index 2 refers to the variables  $\delta_1, \dots, \delta_n$ . The matrices  $\mathbf{S}$  and  $\mathbf{H}$  do not change with respect to the usual formulation.

The general system of equations (6.32) then assumes the following form, where time derivatives are denoted by a dot:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{p}}^w \\ \boldsymbol{\delta} \end{Bmatrix} + \begin{bmatrix} -\mathbf{K}_1 & \mathbf{Q}_1 & -\mathbf{K}_2 \\ \mathbf{Q}_1^T & \mathbf{S} & \mathbf{Q}_2 \\ -\mathbf{K}_2^T & \mathbf{Q}_2^T & -\mathbf{K}_3 \end{bmatrix} \frac{d}{dt} \begin{Bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{p}}^w \\ \boldsymbol{\delta} \end{Bmatrix} = \begin{Bmatrix} -\frac{d\hat{\mathbf{f}}_1}{dt} + \dot{\mathbf{C}} \\ \hat{\mathbf{f}} \\ \hat{\mathbf{f}}_2 \end{Bmatrix} \quad (6.35)$$

or, more concisely,

$$\mathbf{B}\dot{\mathbf{X}} + \dot{\mathbf{C}}\dot{\mathbf{X}} = \mathbf{F} \quad (6.36)$$

A determination of the compaction  $\boldsymbol{\delta}$  is greatly simplified when plane incremental total stress is assumed [25]. In this case we have

$$\frac{d}{dt}(\hat{\sigma}'_z - \hat{p}^w) = 0 \quad (6.37)$$

which results in

$$\frac{d}{dt} \left[ \lambda(\hat{u}_{x,x} + \hat{u}_{y,y}) + (\lambda + 2\mu) \frac{\delta}{B} - \hat{p}^w \right] = \frac{d}{dt} \left[ \lambda\hat{\varepsilon}_v + 2\mu \frac{\delta}{B} - \hat{p}^w \right] = 0 \quad (6.38)$$

where  $\lambda$  and  $\mu$  are Lamé's constants and  $\hat{\varepsilon}_v$  represents the averaged volumetric strain. Equation (6.38) is used instead of the third equation of (6.35).

Using a single-step formula for integration within the time domain, the final form of the completely discretised system is

$$[\mathbf{B} + \theta\Delta t_k \bar{\mathbf{C}}]\bar{\mathbf{X}}^{t_k + \Delta t_k} = [\mathbf{B} - (1 - \theta)\Delta t_k \bar{\mathbf{C}}]\bar{\mathbf{X}}^{t_k} - \mathbf{F}_{n+0}\Delta t_k \quad (6.39)$$

The matrix  $\mathbf{B}$  is now partitioned according to

$$\mathbf{B} = \mathbf{B}^I + \mathbf{B}^E \quad (6.40)$$

where

$$\begin{aligned} \mathbf{B}^I &= \begin{bmatrix} -\mathbf{K}_1 & \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{Q}_1^T & \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{K}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1^I & \mathbf{0} \\ \mathbf{0} & -\mathbf{K}_3 \end{bmatrix} \mathbf{B}^E = \begin{bmatrix} \mathbf{0} & \mathbf{0} & -\mathbf{K}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_2 \\ -\mathbf{K}_2^T & \mathbf{Q}_2^T & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{B}_1^E \\ \mathbf{B}_1^{E^T} & \mathbf{0} \end{bmatrix} \\ \bar{\mathbf{C}} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{aligned} \quad (6.41)$$

After substituting (6.40) into (6.39),  $\mathbf{B}^E$  is transferred to the right-hand side. This allows one to split the problem into a two-dimensional problem of finding  $\hat{\mathbf{u}}_x$ ,  $\hat{\mathbf{u}}_y$  and  $\hat{\mathbf{p}}^w$  and a

one-dimensional problem of finding  $\hat{\boldsymbol{\delta}}$ . A suitable predictor is introduced for the unknowns at time  $t_k + \Delta t_k$  on the right-hand side. We use, as previously, an extrapolator of the type

$$x_{s+1} = (1 - \gamma)x_s + \gamma x_{s+1} \quad (6.42)$$

where  $0 \leq \gamma \leq 1$  and  $s$  is the iteration within a time step. If  $\gamma = 0$  we have a one-term predictor, given by the last solution.

The solution of (6.39) may now be performed in the following way. In a typical time step  $\hat{\boldsymbol{\delta}}_p^{t_k + \Delta t_k}$  is predicted using (6.42) and introduced into the modified analyser for  $\hat{\mathbf{u}}_w^{t_k + \Delta t_k}$  and  $\hat{\mathbf{p}}_w^{t_k + \Delta t_k}$ :

$$[\mathbf{B}_1^T + 0\Delta t_k \bar{\mathbf{C}}_1] \bar{\mathbf{X}}^{t_k + \Delta t_k} = -\mathbf{B}_1^E(\hat{\boldsymbol{\delta}}_p^{t_k + \Delta t_k} - \hat{\boldsymbol{\delta}}^{t_k}) + [\mathbf{B}_1^T - (1 - \theta)\Delta t_k \bar{\mathbf{C}}_1] \bar{\mathbf{X}}^{t_k} - \mathbf{F}_1 \Delta t_k \quad (6.43)$$

where the new terms are defined as follows:

$$\bar{\mathbf{X}} = \begin{Bmatrix} \hat{\mathbf{u}}_v \\ \hat{\mathbf{u}}_v \\ \hat{\mathbf{p}}_w \end{Bmatrix} = \begin{Bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{p}}_w \end{Bmatrix} \quad \mathbf{F}_1 = \begin{Bmatrix} -\dot{\mathbf{f}}_1 + \hat{\mathbf{C}} \\ \hat{\mathbf{f}} \end{Bmatrix} \quad (6.44)$$

The calculated values  $\bar{\mathbf{X}}^{t_k + \Delta t_k}$  are then used on the right-hand side of the second analyser, which gives  $\hat{\boldsymbol{\delta}}$ :

$$-\mathbf{K}_3 \hat{\boldsymbol{\delta}}^{t_k + \Delta t_k} = -\mathbf{B}_1^{E^T} (\bar{\mathbf{X}}_p^{t_k + \Delta t_k} - \bar{\mathbf{X}}^{t_k}) - \mathbf{K}_3 \hat{\boldsymbol{\delta}}^{t_k} + \mathbf{F}_2 \Delta t \quad (6.45)$$

Due to the chosen predictor for the compaction, at least at the start of the integration procedure and in the presence of changes in time step length, iterations are generally needed within the time steps to achieve convergence. This is now analysed briefly.

At the start of the procedure,  $\hat{\boldsymbol{\delta}}_{s+1}^{t_k + \Delta t_k}$  is obtained from (6.42) by putting  $\gamma = 0$  and  $\hat{\boldsymbol{\delta}}^{t_k}$  is known from the initial conditions. Successive iterations within the time step and subsequent time steps are solved with the desired value of  $\gamma$ . The typical equation to be solved is

$$[\mathbf{B}_1^T + 0\Delta t_k \bar{\mathbf{C}}_1] \bar{\mathbf{X}}_{s+1}^{t_k + \Delta t_k} = -\mathbf{B}_1^E((1 - \gamma)\hat{\boldsymbol{\delta}}_s^{t_k + \Delta t_k} - \gamma \hat{\boldsymbol{\delta}}_{s+1}^{t_k + \Delta t_k}) + \tilde{\mathbf{g}}(\hat{\mathbf{u}}, \hat{\mathbf{p}}_w, \hat{\boldsymbol{\delta}}) \quad (6.46)$$

where the terms not updated in the iteration are collected in  $\tilde{\mathbf{g}}(\hat{\mathbf{u}}, \hat{\mathbf{p}}_w, \hat{\boldsymbol{\delta}})$ . This equation yields

$$\bar{\mathbf{X}}_{s+1}^{t_k + \Delta t_k} = -[\mathbf{B}_1^T + 0\Delta t_k \bar{\mathbf{C}}_1]^{-1} \mathbf{B}_1^E((1 - \gamma)\hat{\boldsymbol{\delta}}_s^{t_k + \Delta t_k} - \gamma \hat{\boldsymbol{\delta}}_{s+1}^{t_k + \Delta t_k}) + \tilde{\mathbf{g}}'(\hat{\mathbf{u}}, \hat{\mathbf{p}}_w, \hat{\boldsymbol{\delta}}) \quad (6.47)$$

By substituting (6.47) into the analyser (6.45), the operation accelerates the rate of convergence of the problem, and by collecting in  $\mathbf{L}(\hat{\mathbf{u}}, \hat{\mathbf{p}}_w, \hat{\boldsymbol{\delta}})$  terms remaining constant within the iteration, equation (6.45) may be rewritten with the help of some algebra:

$$\hat{\boldsymbol{\delta}}_{s+1}^{t_k + \Delta t_k} = (1 - \gamma)[-\mathbf{K}_3 - \gamma \mathbf{G}]^{-1} \mathbf{G} \hat{\boldsymbol{\delta}}_s^{t_k + \Delta t_k} + \mathbf{L}(\hat{\mathbf{u}}, \hat{\mathbf{p}}_w, \hat{\boldsymbol{\delta}}) \quad (6.48)$$

with

$$\mathbf{G} = \mathbf{B}_1^E [ \mathbf{B}_1^I + \theta \Delta t_k \bar{\mathbf{C}}_1 ]^{-1} \mathbf{B}_1^E \quad (6.49)$$

The iteration matrix of (6.48) is

$$\mathbf{E} = (1 - \gamma) [ -\mathbf{K}_3 - \gamma \mathbf{G} ]^{-1} \mathbf{G} \quad (6.50)$$

The iterative method converges if the iteration matrix  $\mathbf{E}$  has real eigenvalues and if the largest eigenvalue  $\rho(\mathbf{E})$  satisfies the condition

$$\rho(\mathbf{E}) < 1 \quad (6.51)$$

and the asymptotic rate of convergence is

$$R = -\log(\rho(\mathbf{E})) \quad (6.52)$$

Equations (6.51) and (6.52) give conditions for the choice of  $\gamma$  or for the number of iterations to be performed.

When using (6.38) to determine the compaction, a further advantage is gained in operating at the Gauss points only. The compaction values are then calculated at the nodes of each element using a smoothing technique [26]. However, oscillations in time may occur near sinks or sources if the mesh is not sufficiently refined. This is due to the strong variation of pore pressure in these areas. But the oscillations decrease very rapidly and the results are acceptable at a limited distance from the applied fluxes. In doing so, it is not necessary to assemble and solve a general system of equations, hence the procedure is very fast. This procedure is applied to the last example of the next chapter.

Here are a few more comments on the procedure:

1. The applied predictor is usually the last solution. Vertical displacements are assumed equal to zero for the first time step.
2. Iterations are necessary during the first time steps or during changes in the value of the time step. The number of iterations observed from experience is usually small.
3. The advantage of the proposed method lies in the possibility of studying a three-dimensional problem using a modified two-dimensional analyser. This means that brick elements with 16 nodes (using linear interpolation in  $z$ ) and 64 degrees of freedom (d.o.f.) are replaced by elements with 8 nodes and 24 d.o.f. (the remaining 8 d.o.f. are separately treated).
4. Compared with a two-dimensional analyser, this approach allows solution of a three-dimensional problem by a limited increment of d.o.f. plus an iteration phase. But remember that field variables are known only through their mean value over the aquifer thickness.

### 6.3 FAR-FIELD BOUNDARY CONDITIONS

This section deals with infinite elements, with which it is possible to obtain a more precise simulation of thermo-hydro-mechanical problems defined in very large domains.

Such interaction problems between two or three fields (mechanical, thermal and hydraulic) often involve large domains that may be assumed as *unbounded*. Typical examples are land subsidence analyses and the analysis of heat sources buried in soil (Chapter 8). Numerical modelling of such problems sometimes produces incorrect answers, due to the difficulty of correctly fixing the boundary conditions at large distances.

The usual technique within the finite element method is the so-called *truncation approach*, where a far-field boundary is introduced at some distance from the central area. Conditions valid at infinity are then usually imposed at this fictitious boundary [27,28]. The accuracy of the simulation depends on the location of the far-field boundary, and this dependence may differ for each interacting field. Generally speaking, more accurate results are obtained by increasing the size of the modelled domain. However, a larger domain means more degrees of freedom (d.o.f.). Also, interaction problems usually involve the sum of the d.o.f. of the single interacting fields, resulting in very large matrices. Furthermore, these problems are often time dependent with long time histories. Hence the ensuing numerical models are very large, both in space and in time, and analysts often attempt to reduce the costs by truncating the mesh as close as possible to the area of interest. This procedure is not easy to judge and a trial- and -error analysis should be performed to assess the influence of the boundary and its assumed conditions on the numerical model. General criteria do not exist, due to the fact that the coupled solution is strongly dependent on the response characteristics of each interacting field. Usually the matrices relative to the different interactive fields present eigenvalues which differ greatly from each other. For this reason the accuracy of the solution depends on the natural timescale of each field and on the time span of interest for the analysis. The following situations can be observed [29]:

1. The errors do not always affect all the state variables and not necessarily even the state variables of the field where the forcing function is applied.
2. The errors may occur after a certain time span, depending on the intrinsic velocity of each field. Sometimes the time span of the trial- and -error analysis is too short to reveal them.
3. The comparison with known normalised solutions may provide the correct shape of the solution, but errors may still influence the amplitudes.

Hence a check for only one variable, or only over an initial time span, or a comparison with known normalised solutions may lead to unjustified confidence in the results. Appropriate modelling techniques are therefore mandatory for acceptable results. Partitioned solution procedures, as shown in the previous sections, reduce the size of the problem to be solved at each time step [30,31]. Boundary integral methods [32], analytical matching of the far-field solution [33] and similar procedures represent valuable means of reducing the number of d.o.f. of the discretised problem in space. However, only few solutions are known for coupled field problems with overlapping domains. Furthermore, the necessary fundamental solutions are not always available.

### 6.3.1 Infinite Elements

Within the framework of the finite element method, infinite elements (IE) seem to offer particular advantages in solving this type of problems; a general presentation is available

in the literature [34]. Here we use mapped infinite elements (MIE), which were first applied in coupled problems by Lewis *et al.* [35] and Schrefler and Simoni [36,37]. MIE are at the same time efficient and simple to use in quasi-static problems, such as isothermal or non-isothermal consolidation of soils. Their use allows for a more realistic assessment of the boundary conditions and reduces the number of d.o.f. required for accurate numerical modelling.

There are essentially two infinite element formulations currently in use for quasi-static problems: decay function infinite elements and mapped infinite elements. In the first case a finite element is extended so that it stretches to infinity in the direction of one local axis (say  $\xi$ ). The shape functions  $M_i(\xi, \eta)$  of the original finite element in the  $\eta$ -direction are retained, whereas in the  $\xi$ -direction they are simply multiplied by a decay function  $f_i(\xi)$ :

$$N_i(\xi, \eta) = M_i(\xi, \eta)f_i(\xi) \quad (6.53)$$

under the conditions that at node  $i$  the shape function  $N_i$  must be equal to unity and tend to the far-field value at infinity. The decay function ensures the finiteness condition of the solution and has to be chosen to obtain a reasonable description of the physics of the problem. One of the most widely used decay functions is the exponential function:

$$f_i(\xi) = \exp \frac{\xi_i - \xi}{L} \quad (\text{for } \xi \geq 0) \quad (6.54)$$

where the presence of  $\xi_i$  ensures the unit value at node  $i$  and  $L$  is a length which determines the severity of the decay. Another popular choice is a *reciprocal function* of the type

$$f_i(\xi) = \left[ \frac{\xi_i - \xi_0}{\xi - \xi_0} \right]^n \quad (6.55)$$

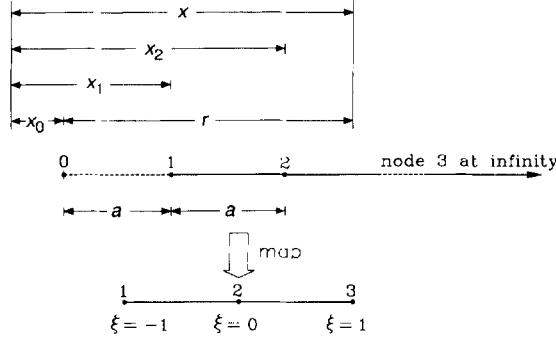
where  $\xi_0$  is the distance from some origin. This must be outside the infinite element, hence opposite the side that extends to infinity. Extension to the case where a finite element is stretched to infinity in more than one direction is straightforward. The use of these types of elements requires particular quadrature formulae, e.g. Gauss–Laguerre or Gauss–Hermite, because the domain of the element is  $[-1, \infty]$ .

As far as the mapped type of element is concerned, we consider the one-dimensional case shown in Figure 6.2, where the infinite element, ranging from node 1 to node 3 (at infinity) through node 2, is mapped onto the parent element defined by the local coordinate  $-1 \leq \xi \leq 1$ .

Point 0 is called the *pole* of transformation and is positioned arbitrarily, with the only restriction that, once positioned, the location of node 2 is defined by

$$x_2 = 2x_1 - x_0 \quad (6.56)$$

The interpolation from local to global coordinates is obtained by using the standard



**Figure 6.2** Global-to-local mapping of a one-dimensional infinite element

finite element method for isoparametric elements:

$$x(\xi) = \sum_{i=1}^2 M_i(\xi) x_i \quad (6.57)$$

where the summation extends to *finite* nodes only and the mapping functions are

$$M_1(\xi) = \frac{-2\xi}{1-\xi} \quad M_2(\xi) = \frac{1+\xi}{1-\xi} \quad (6.58)$$

From these relations it follows that in global coordinates the positions  $x = x_1, x_2, \infty$  correspond to the points  $\xi = -1, 0, 1$ . Moreover, as

$$M_1(\xi) + M_2(\xi) = 1 \quad (6.59)$$

the transformation does not depend on the adopted reference system.

The field variable  $\psi$  (displacement, pressure or temperature) is now interpolated using the standard shape functions  $N_i$ :

$$\psi(\xi) = \sum_{i=1}^{n-1} N_i(\xi) \psi_i = \alpha_0 + \alpha_1 \xi + \alpha_2 \xi^2 + \alpha_3 \xi^3 + \dots + \alpha_n \xi^n \quad (6.60)$$

where  $\alpha_i$  contains the nodal values of  $\psi_i$ .

Solving (6.57) for  $\xi$  yields

$$\xi = 1 - \frac{2a}{x - (x_1 - a)} = 1 - \frac{2a}{r} \quad (6.61)$$

where  $r$  indicates the distance from pole 0 to a general point belonging to the element, and

$$a = x_2 - x_1 = x_1 - x_0 \quad (6.62)$$

By introducing (6.61) into (6.60) the expression of the field variable  $\psi$  as a function of the global coordinate  $r$  is obtained as

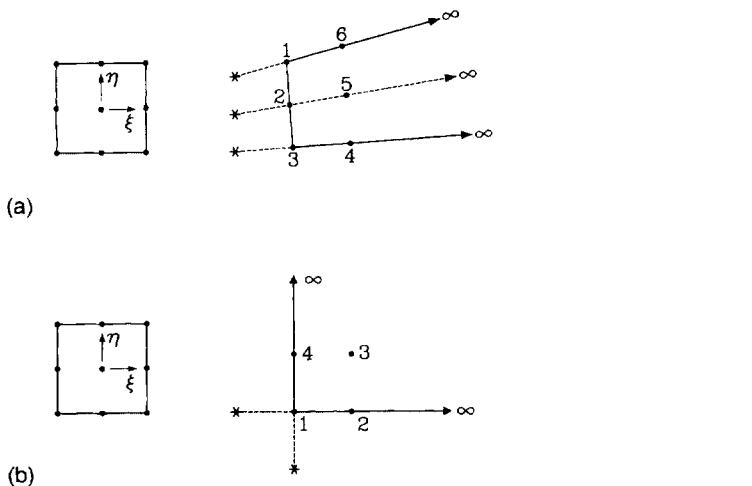
$$\psi(r) = \beta_0 + \frac{\beta_1}{r} + \frac{\beta_2}{r^2} + \frac{\beta_3}{r^3} + \dots + \frac{\beta_n}{r^n} \quad (6.63)$$

which represents the polynomial approximation defined in the unlimited domain  $x$  corresponding to a polynomial of degree  $n$  defined in the limited domain  $\xi$ .

These relations reveal the part played by pole 0 and also its position: the type of decay depends on the choice of the pole and on the shape functions used in (6.60). The variable  $\psi$  vanishes at infinity by setting  $\beta_0 = 0$ , i.e. by restricting the summation in (6.60) to the finite nodes. In such a way we lose the possibility of representing constant values of the field variables and it is impossible to represent a constant strain field for the infinite elements. This situation is physically consistent, but from the mathematical viewpoint it is impossible to prove the existence of *h-convergence*, even if compatibility of elements is preserved. Hence the convergence of the approximate solution can be proved by operating just on the degree of the approximating polynomial (*p-convergence*).

This procedure can be easily extended to the case of two or three dimensions, by simply using shape function products. Elements may be generated which are either singly or doubly infinite. For example, the infinite element in Figure 6.3(a) can be considered as a 9-noded Lagrangian element, in which the 3 nodes corresponding to  $\xi = 1$  are at infinity. The relative mapping functions  $\hat{M}_i(\xi, \eta)$  are given by (6.58) in the infinite direction multiplied by the usual quadratic shape functions  $N_i$  in the  $\eta$  direction:

$$\begin{aligned} \hat{M}_1(\xi, \eta) &= M_1(\xi)N_1(\eta) & \hat{M}_2(\xi, \eta) &= M_1(\xi)N_2(\eta) & \hat{M}_3(\xi, \eta) &= M_1(\xi)N_3(\eta) \\ \hat{M}_4(\xi, \eta) &= M_2(\xi)N_3(\eta) & \hat{M}_5(\xi, \eta) &= M_2(\xi)N_2(\eta) & \hat{M}_6(\xi, \eta) &= M_2(\xi)N_1(\eta) \end{aligned} \quad (6.64)$$



**Figure 6.3** Two-dimensional mapped infinite elements: (a) Lagrangian biquadratic singly infinite element; (b) Lagrangian biquadratic doubly infinite element

Similar expressions can be obtained for the element in Figure 6.3(b), which is an element extending to infinity in two directions. In this element all nodes corresponding to  $\xi = 1$  and  $\eta = 1$  are at infinity. The mapping functions are the product of (6.58) in the  $\xi$  and  $\eta$  directions, and the shape functions are properly chosen among those of the 9-noded finite Lagrangian element. Elements like these (doubly infinite elements) can be used in corner areas between elements extending to infinity in different directions. Further information and details of three-dimensional elements can be found in the literature [36,38].

Both the geometry and the expansion of the field variable of the mapped infinite element depend on the same pole, or group of poles in the case of more than one dimension. When choosing the pole(s), we must take into account both the geometry and the physical characteristics of the problem. In particular, we take three things into account:

1. The pole(s) must be external to the infinite element.
2. The element sides extending to infinity must be parallel or divergent to avoid the overlapping of elements and preserve mapping uniqueness.
3. To ensure continuity across common sides between infinite and finite elements or adjacent infinite elements, the numbers and positions of the connecting nodes must be the same.

Once these restrictions have been taken into account, the shifting of the pole(s) involves a variation in the type of decay for the field variable only. As far as the numerical integration is concerned, the usual formulae for finite element codes can be used (e.g. Gauss-Legendre).

The choice of the decay parameters of the approximations is the main concern of the analyst, who must appreciate the physics of the problem to be solved. Thus infinite elements represent a powerful tool in dealing with unbounded domain problems. These elements are applied in Section 6.5 and also in Chapter 8.

#### **6.4 A COUPLED SOLUTION FOR THE SETTLEMENT ABOVE GAS RESERVOIRS**

The full solution for the coupled equations of three-phase immiscible flow will be dealt with in Chapter 8. A simplified solution can be obtained by considering the mass balance equation in integral form in some part of the domain, e.g. in a hydrocarbon reservoir. This allows us to easily solve an important problem in petroleum engineering, the deformation induced in the overburden, underburden and in the reservoir itself by withdrawal of underground fluids. The procedure employed is valid for hydrocarbon reservoirs which have a limited lateral extension and thickness when compared to the depth of burial. This is the case in several gas reservoirs in subsiding areas [39] which were investigated by the authors.

The fluid production from the reservoir is assumed to be effected via several wells producing at the same rate and evenly spaced throughout the reservoir. This assumption ensures a reasonably uniform pressure drop within the reservoir. The average pore pressure in the reservoir, together with the production data, must satisfy at each instant

the cumulative material balance equation [40,41]:

$$W_e = N_p(B_o - R_{so}B_g) + G_pB_g + W_pB_w - N(B_t - B_{ti}) - G(B_g - B_{gi}) \\ - F - \frac{B_{ti}}{1 - S_{wi}}(C_f - S_w C_w)(\bar{p}_i - \bar{p}) \quad (6.65)$$

where

- $C_f$  = effective rock compressibility
- $C_w$  = water compressibility
- $B_g$  = gas formation volume factor
- $B_o$  = oil formation volume factor
- $B_t$  = total formation volume factor
- $B_w$  = water formation volume factor
- $F$  = cumulative fluid injection
- $G$  = initial free gas in place
- $G_p$  = cumulative gas production
- $N$  = initial oil in place
- $N_p$  = cumulative oil production
- $\bar{p}$  = average reservoir pressure, average pressure at initial oil-water contact
- $R_{so}$  = gas-oil solution ratio
- $S_w$  = average interstitial water saturation
- $W_e$  = water influx from the adjacent aquifer and from leaky aquitards
- $W_p$  = cumulative water production
- $i$  = subscript: denotes initial conditions.

The right-hand side represents the production of oil, gas and water, the expansion of rock and fluids in place, and influx and injection terms. The material balance equation is essentially a volumetric balance at reservoir condition for

1. The fluids contained initially in the reservoir
2. The portion of those fluids still remaining at a later date
3. The encroaching water from the adjacent aquifer
4. The water drained from other bounding strata
5. Injected water and gas

The generalised form of the material balance equation enables an investigation to be made of surface subsidence in the cases of gas-condensate reservoirs, undersaturated oil reservoirs with simultaneous drives such as dissolved gas drive, gas-cap drive and water drive. For the sake of simplicity, the numerical development will be carried out for the case of a waterdrive gas reservoir [39].

The material balance equation in this case can be written as follows [42]:

$$GB_{gi} = (G - G_p)B_g + W_e - B_w W_p \quad (6.66)$$

and in incremental form as

$$(G - G_p)\Delta B_g + \Delta W_e = \Delta G_p B_g + B_w \Delta W_p + \Delta B_w W_p \quad (6.67)$$

**Table 6.1** Compressibility factors for methane

Temp (K)	Pressure (MPa)											
	0.1	0.5	1.0	2.0	4.0	6.0	8.0	10.0	20.0	30.0	40.0	50.0
250	0.9972	0.9841	0.9678	0.9356	0.8694	0.8035	0.7403	0.6889	0.6953	0.8593	1.0383	1.2172
300	0.9982	0.9915	0.9828	0.9663	0.9342	0.9042	0.8773	0.8548	0.8280	0.9140	1.0417	1.1812
350	0.9988	0.9954	0.9905	0.9821	0.9657	0.9513	0.9390	0.9293	0.9226	0.9775	1.0678	1.1751
400	0.9995	0.9976	0.9957	0.9908	0.9833	0.9771	0.9721	0.9691	0.9783	1.0258	1.0968	1.1821
450	0.9999	0.9996	0.9991	0.9965	0.9941	0.9923	0.9917	0.9922	1.0128	1.0577	1.1195	1.1916
500	1.0000	1.0000	1.0000	1.0003	1.0009	1.0021	1.0043	1.0068	1.0335	1.0780	1.1347	1.1990

The formation volume factor for natural gas [40] is obtained from the state equation of ideal gas as

$$B_g = \frac{p_{sc}}{\theta_{sc}} \left( \frac{Z\theta}{p} \right) \quad (6.68)$$

and

$$\Delta B_g = \frac{p_{sc}}{\theta_{sc}} \left[ \Delta Z \left( \frac{\theta}{p} \right) - \left( \frac{Z\theta}{p^2} \right) \Delta p + \left( \frac{p_{sc}}{T_{sc}} \right) Z \Delta \theta \right] \quad (6.69)$$

where  $Z(p, \theta)$  is the compressibility factor,  $\theta$  is the absolute temperature,  $p$  is the gas pressure (subscript omitted) and sc stands for surface conditions. Data for the compressibility factor may be found in Standing [43] and is indicated for methane in Table 6.1.

The material balance equation is applicable only to the reservoir as a whole because of possible internal gas migration from one position to another. The equation is obviously also valid for volumetric reservoirs. This relationship is utilised in what follows to obtain the average reservoir pressure at each time step when the gas production and the water inflow are known.

This average reservoir pressure is then applied as a boundary condition at the reservoir boundary. The whole subsiding volume is then analysed using the fully coupled equations (6.4). These equations yield the flow of water across the reservoir boundary which is required in the material balance equations, (6.66) and (6.67), and also the deformation of the reservoir itself. With a suitable mesh this procedure yields a constant pressure distribution in the reservoir at each time step. If the formation thickness is great, then it is also possible to vary the reservoir pressure linearly with depth. The solution obviously takes into account the horizontal displacements in the whole subsiding volume.

The technique chosen for the transient solution of the simultaneous non-linear equations, (6.4) and (6.66), is a single-step iterative method. The logic of the scheme is as follows:

1. At time  $t$  the water influx into the reservoir, the cumulative gas production, the reservoir conditions and the water and gas production for the time step  $\Delta t$  are all known. Using an iterative procedure, the cumulative material balance equation yields the resulting mean value of the pore pressure in the reservoir:

$$\frac{p}{Z(p, \theta)} = \frac{(G - G_p)\theta}{GZ(p_i, \theta_i) \left( \frac{\theta_i}{p_i} \right) - \left( \frac{\theta_{sc}}{p_{sc}} \right) W_e - B_w W_p} \quad (6.70)$$

2. Check the incremental material balance equation (6.67).
3. Apply the new value of reservoir pore pressure as a boundary condition for the coupled equations (6.4) and calculate the displacement, pore pressure field and the water influx into the reservoir for the next time step.

If excess pore pressures are employed instead of actual pressures, then care must be taken in the proper choice of the reservoir pressures. The procedures adopted for the

solution of the three simultaneous non-linear equations need some comment. A straightforward application of the procedure outlined above implies that the variation of the equivalent bulk modulus of the pore fluid, due to the upward displacement of the gas–water contact surface, is neglected as well as the change of submerged weight due to the same cause.

A better approximation can be obtained by updating these material properties for the appropriate reservoir elements when the water–gas contact surface moves through them. An approximate instantaneous position of the contact surface can be obtained from the total water influx, the porosity and the reservoir volume.

The chosen form of the material balance equation (6.66) neglects some minor factors such as the variation in volume of the interstitial water with pressure, the change of porosity with pressure and the evolution of gas dissolved in the interstitial water with a decrease in pressure. But these factors may be included in the material balance relationship if warranted by the precision of available data.

The procedure described in this section will be applied in Section 7.4.

## 6.5 SINGLE-AQUIFER WITHDRAWAL

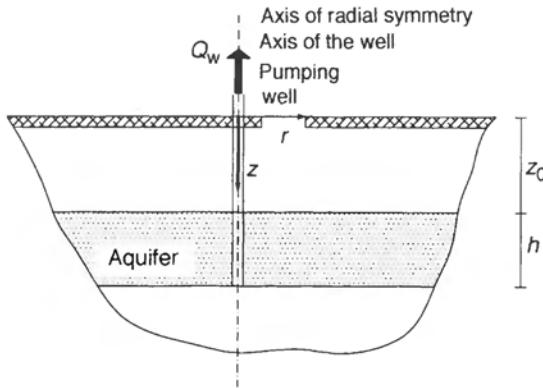
The compaction of a single aquifer pumped from a fully penetrating well is now discussed in some detail. Numerical solutions are compared with analytical ones, which are also given. These are useful benchmark problems for testing consolidation programs. But the principal aim is to analyse the differences between coupled and uncoupled solutions for this problem and to draw useful conclusions for subsidence analysis. The aquifer is overlain and underlain by sediments with lower permeability values, is laterally unbounded and conditions of axial symmetry are assumed to exist. Two types of solution can be found in the literature for the linear elastic case: one solves the problem as an isolated aquifer with radial flow only, and the other considers the aquifer embedded in a mechanically homogeneous half-space, with zero permeability outside the aquifer.

For comparison the full Biot equations (6.4) are solved for an isolated aquifer and an embedded aquifer with the following parameters (Figure 6.4):

Thickness of aquifer	$h = 10 \text{ m}$
Radius of well	$r_0 = 0.3 \text{ m}$
Compressibility of water	$\beta_w = 5 \times 10^{-10} \text{ N m}^{-2}$
Specific weight of water	$\gamma_w = 9800 \text{ N m}^{-3}$
Poisson's ratio for solid	$\nu = 0.1$
Permeability of medium	$k_w = 10^{-4} \text{ m s}^{-1}$
Porosity	$n = 0.2$
Young's modulus	$E = 2.2 \times 10^7 \text{ N m}^{-2}$
Pumping rate (fully penetrating well)	$Q = 6.81 \text{ m}^3 \text{ h}^{-1}$

For the isolated aquifer, the top and bottom surfaces are assumed impervious and vertical displacements at the bottom and horizontal displacements at  $r = r_0$  are restrained.

In order to simulate the usual condition of an impermeable surrounding half-space, the permeability outside the embedded aquifer is assumed to be  $10^{-8}$  times smaller, but



**Figure 6.4** Embedded aquifer in a half-space

the Young's modulus is taken to be the same. This case is called HOMOG in the following analyses.

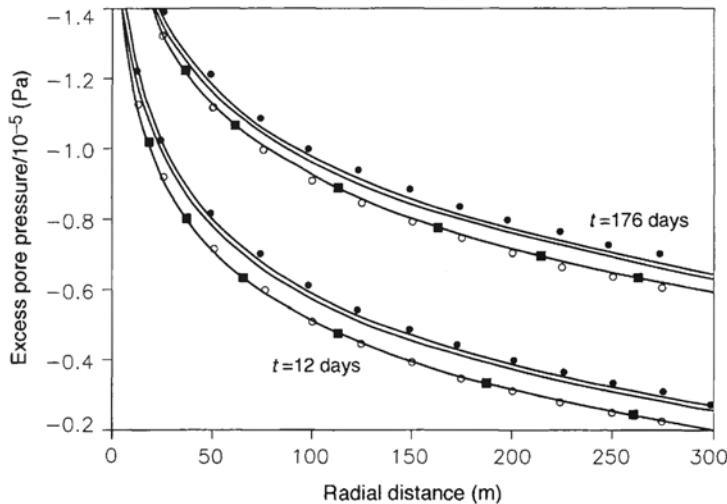
Several cases were solved in which the depth  $z_0$  of burial of the aquifer ranges from  $w = z_0/h = 0$  to  $w = z_0/h = 10$ . In order to assess the effect of horizontal displacement boundary conditions, a further analysis was performed for the embedded aquifer, which was assumed to lie on a smooth bedrock with variable overburden.

The far-field behaviour for all the cases was approximated by means of infinite elements, since predictions are known to be sensitive to the position of the lateral boundary assumed in the numerical model (Section 6.3). In this case the poles are located on the axis of symmetry. The domain below the embedded aquifer is also modelled by finite and infinite elements, with poles located on the lower boundary of the aquifer. The problem assesses the importance of some of the simplifying assumptions introduced in the conventional approaches.

### 6.5.1 Isolated Aquifer

The analytical solution for an isolated aquifer is quite simple under the assumptions of axisymmetric conditions, uniform distribution over the thickness of the aquifer for pore pressure and for horizontal displacements and plane incremental total stress [44]. In this way the governing equations are reduced to a system of two differential equations for the unknown volumetric strain  $\varepsilon_v$  and pore pressure  $p^w$ . The solution is sought in terms of  $p^w$  by obtaining the volumetric strain from the equilibrium equation, i.e. by inverting the stiffness matrix (in this case of order 1). According to the definitions outlined in Section 6.1, it is a coupled problem. This will also be shown below by the iteration count. The analytical solution for the pressure field given by Theis [45] is strictly related to the uncoupled Terzaghi–Jacob one-dimensional theory. By accounting for the deformability of the aquifer through the storativity coefficient  $S$  [25], the drawdown  $v$  in the Theis solution is given by

$$v = \frac{Q}{4\pi T} E_1 \left( \frac{r^2 S}{4Tt} \right) \quad (6.71)$$



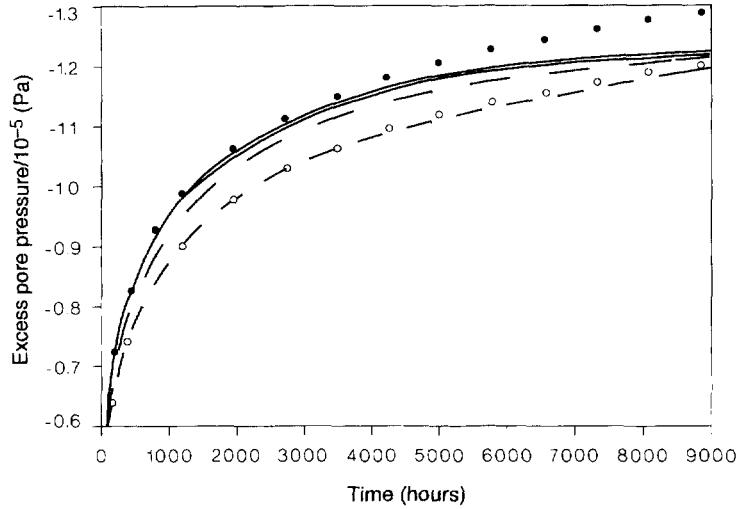
**Figure 6.5** Numerical and analytical results for excess pore pressure versus radial distance: (—) embedded aquifer, numerical; (-■-) isolated aquifer, numerical; (●) equation (6.71), one-dimensional consolidation; (○) equation (6.71), three-dimensional consolidation

where  $T$  is the transmissivity coefficient [25],  $t$  is the time variable and  $r$  is the radial distance. We recall that the storativity and transmissivity coefficients are respectively defined as the product of the aquifer thickness and the averaged values of the capacity term and permeability. Furthermore,  $E_1$  is the exponential integral (*well function*):

$$E_1(x) = \int_x^{\infty} \frac{1}{u} e^{-u} du \quad (6.72)$$

As a consequence of these assumptions, the storativity coefficient in the Theis solution corresponds to one-dimensional consolidation. In a coupled analytical solution by Bear and Corapcioglu [23] the drawdown has the same expression as (6.71), but the coefficient of consolidation corresponds to true three-dimensional conditions. The results, in terms of excess pore pressures, are therefore different (Figure 6.5), which confirms the coupled nature of the problem. The differences are approximately 15 percent for  $r = 5h$  and  $t = 12$  days, but they increase with increasing distance from the well and decrease with time. Figure 6.5 also shows the results obtained numerically via equation (6.4).

The assumption of a one-dimensional coefficient of consolidation corresponds to anticipating in time the correct coupled solution, i.e. it leads to a higher drop in pore water pressure (Figure 6.6). Hence the use of these pore pressures in a two-step uncoupled model results in higher compactions with respect to the exact solution. When the boundary conditions for the isolated aquifer change, e.g. horizontal displacements vanish for the upper or lower surface, or the aquifer is assumed to be embedded in a semi-infinite medium, the corresponding solutions in the considered time span lie between the two limiting cases of coupled and uncoupled solution (Figure 6.5 and 6.6).



**Figure 6.6** Numerical and analytical results for excess pore pressure versus time at  $r = 50\text{ h}$ : (—) embedded aquifers, numerical; (----) isolated aquifer, numerical; (- - -) aquifer on bedrock ( $w = 1.0$ ), numerical; (●) equation (6.71), one-dimensional consolidation; (○) equation (6.71), three-dimensional consolidation

These assumptions correspond to some sort of constraint on the horizontal displacements, and their influence on the drawdown are not easily predicted. Hence it is very difficult to recognise *a priori* the possibility of uncoupling. The results presented in Figures 6.5 and 6.6 confirm this statement.

For the uncoupled solution (one-dimensional consolidation), the isolated aquifer compaction is given as follows [44]:

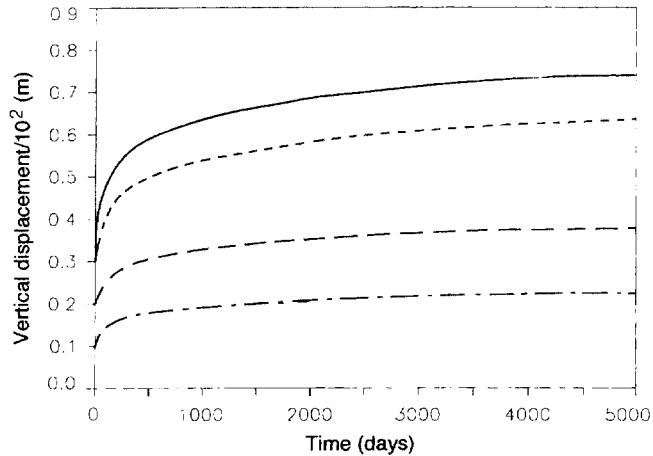
$$\delta = \frac{Q}{4\pi C_v} E_1 \left( \frac{r^2}{4C_v t} \right) \quad (6.73)$$

where  $C_v$  is the coefficient of one-dimensional consolidation of the isotropic medium and is defined as

$$C_v = \frac{k_w}{\alpha' + n/K_w} \quad (6.74)$$

where  $\alpha'$  is the aquifer compressibility. As a function of Lamé's constants,  $\lambda$  and  $\mu$ , this takes the form  $(\lambda + \mu)^{-1}$  in a one-dimensional context and  $(\lambda + 2\mu)^{-1}$  in a three-dimensional situation. The coupled solution [44] gives the compaction as

$$\delta = \frac{Q}{8\pi C_v} E_1 \left( \frac{r^2}{4C_v t} \right) \quad (6.75)$$



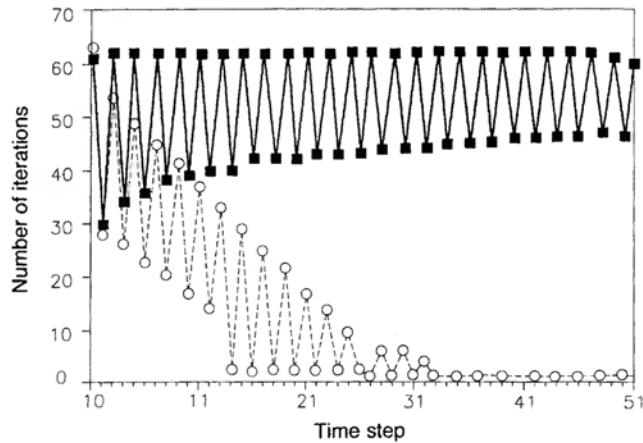
**Figure 6.7** Analytical solutions for vertical displacements of an isolated aquifer versus time at  $r = 5 h$  from the well: (—) equation (6.73),  $v = 0.1$ ; (- - -) equation (6.75),  $v = 0.1$ ; (- -) equation (6.73),  $v = 0.4$ ; (- · -) equation (6.75),  $v = 0.4$

and the horizontal displacement (in the radial direction) as

$$u_r = -\frac{Q}{16\pi C_v h} \left[ E_1(u) - \frac{1 - e^{-u}}{u} \right] \quad (6.76)$$

where  $u = r^2/4\pi C_v t$  and the coefficient of consolidation is calculated assuming three-dimensional consolidation. By comparing the solution of (6.73) and (6.75), it follows that, for the same pumpage, equation (6.75) gives approximately  $(1 - v)$  times the value of  $\delta$  as given by (6.73), where  $v$  is Poisson's ratio. This is shown in Figure 6.7 where the two equations are evaluated for  $v = 0.1$  and  $v = 0.4$ . The differences are relevant; they can be of the order of 70 percent but they decrease with decreasing Poisson's ratio.

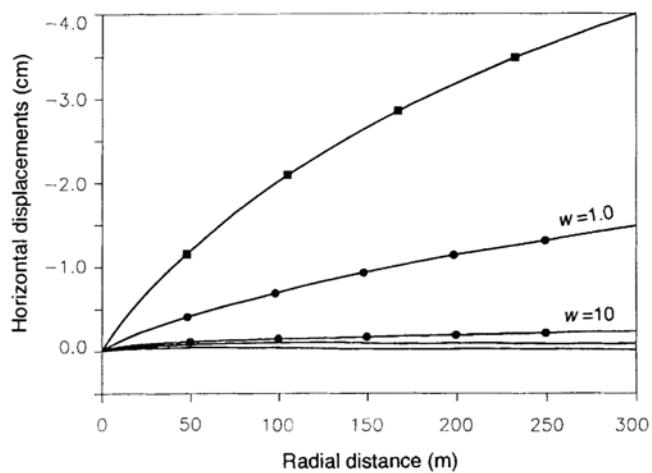
The coupled nature of the pumped isolated aquifer problem is also revealed by the iteration count when applying the partitioned procedure of equation (6.18). A comparison is made between the cases when the forcing function is applied to the displacement field (coupled situation [6]) and when the aquifer is pumped by a fully penetrating well. The same spatial and temporal discretisations are used in both cases, giving a spectral radius of 0.8202 for the iterative matrix. This numerical model used a mesh truncated at  $R = 100h$  from the axis of symmetry; the pumped aquifer then showed a steady-state condition after  $t = 200$  days, which was not present for the aquifer of infinite radial extent [29]. During the whole transient situation, the number of iterations within each time step is far from negligible (Figure 6.8) and becomes small (one or slightly more than one, due to the chosen measure of numerical accuracy) only when steady-state conditions are reached. At the onset of the phenomenon, the number of iterations is the same as for an identical aquifer loaded in a circular area with a radius equal to the thickness of the aquifer. This last problem is coupled and the iteration count also confirms the coupled nature of the transient pumped aquifer with three-dimensional consolidation.



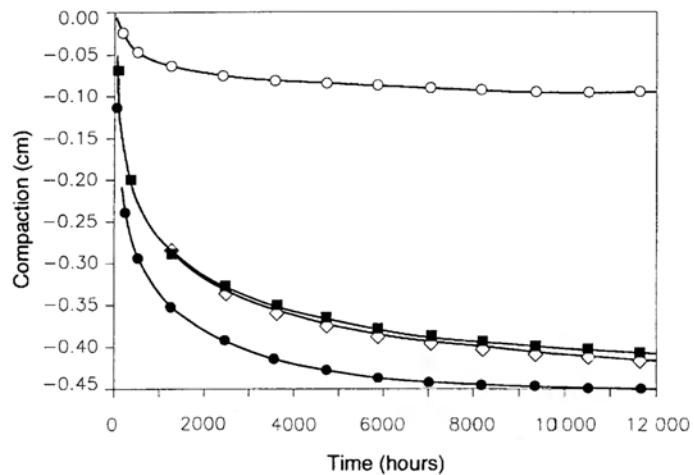
**Figure 6.8** Iteration count for the analysis of an isolated aquifer under (■) applied load and (○) imposed flux

### 6.5.2 Embedded Aquifer

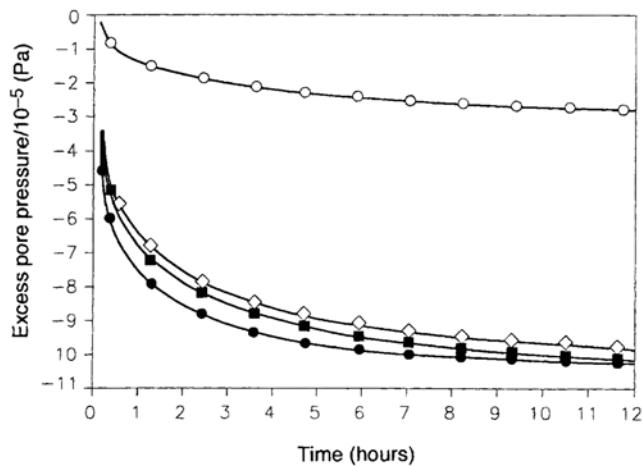
Figure 6.9 shows the horizontal displacements at the top of the aquifer for different situations; the results were obtained numerically using equation (6.4). The aquifer resting on bedrock with a small depth of burial indicates horizontal displacements which are not negligible, but they are negligible for an embedded aquifer in a mechanically homogeneous half-space. According to what was discussed in Section 6.1.1, this last situation is *de facto* uncoupled and it should not be concluded that subsidence problems are in general uncoupled. Real subsidence problems are often much more complicated than this academic problem (mechanically homogeneous half-space). Compared with



**Figure 6.9** Mean horizontal displacements versus radial distance at  $t = 176$  days: (●) aquifers on bedrock, (—) embedded aquifers, (■) isolated aquifer



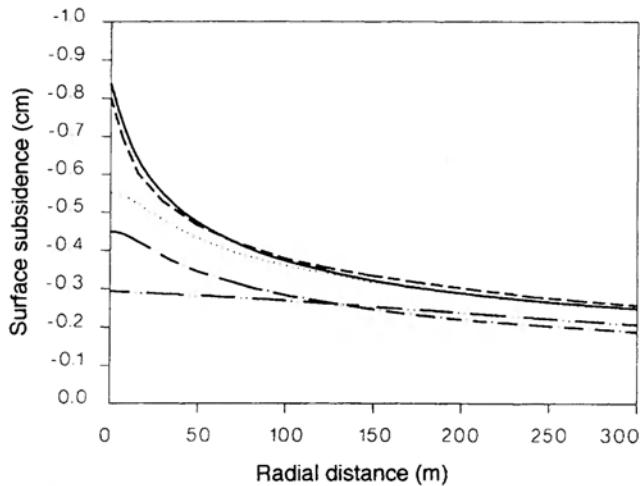
**Figure 6.10** Numerical results for compaction versus time of single aquifers at  $r = 10\text{ h}$ : (■) NHOMOG1, (○) NHOMOG2, (●) HOMOG, (◇) isolated aquifer



**Figure 6.11** Numerical results for excess pore pressure versus time at  $r = 10\text{ h}$  (■) NHOMOG1, (○) NHOMOG2, (●) HOMOG, (◇) isolated aquifer

the aquifer, if it is assumed that the surrounding medium has different hydraulic and mechanical characteristics, then the changes are dramatic. The results for the following two cases are shown:

1. Mechanically non-homogeneous conditions with a Young's modulus in the area surrounding the aquifer that is 80 times smaller than the Young's modulus in the aquifer; the external permeability is  $10^{-8}$  times smaller (NHOMOG1).
2. The Young's moduli are the same as in the previous case and the permeability in the surrounding half-space is assumed to be  $2.5 \times 10^{-3}$  times smaller (NHOMOG2).



**Figure 6.12** Land subsidence versus radius at  $t = 176$  days: (---) isolated aquifer; (—) embedded aquifer,  $w = 0.0$ ; (· · · ·) embedded aquifer,  $w = 1.0$ ; (- · - ·) embedded aquifer,  $w = 10.0$ ; (- - -) aquifer on bedrock,  $w = 1.0$

The chosen values were suggested by the usual assumption of zero external permeability and by a real case illustrated in the next chapter.

Using an isolated aquifer as a reference, Figs. 6.10 and 6.11 compares the time histories for compaction and for pore pressure at a distance of  $10h$  from the well using the two cases and a depth of burial  $z_0 = 10h$ . The differences in behaviour are remarkable, e.g. steady-state conditions are obtained for NHOMOG2 in a much shorter time than for all the other cases. This fact is mainly due to the diffusion of the pumping effect outside the aquifer, which is usually neglected in uncoupled solutions.

However, the surface subsidence of the cases under analysis are shown in Figure 6.12, and they are very different. This is due to the fact that with increasing depth of burial, the upward displacement of the bottom surfaces also increases when the aquifer is embedded in a homogeneous half-space. So, in real subsidence problems, hardly any conclusions can be drawn from these analytical models.

Only single aquifers have been considered up to now. The next chapter analyses real subsidence problems, and where possible it compares them with coupled and uncoupled solutions.

## 6.6 CONCLUSIONS

This chapter has introduced a clear definition of coupling, and on this basis it has discussed the possibility of using uncoupling numerical procedures. The analyses of a single aquifer under differing mechanical and hydraulic conditions, pumped from a fully penetrating well, indicate that coupling occurs for most situations where outflow is specified and should be solved as such. Uncoupling occurs only under very special circumstances, e.g. in the case of a static determinacy. This happens very seldom in

practice; one exception is an embedded pumped aquifer in a mechanically homogeneous half-space. It is always possible to calibrate an uncoupled model to fit existing data, but the price to pay is inconsistency of the model, i.e. different material properties or different soil models, in the hydrological and consolidation analyses. Even if the calibration is satisfactory then long-term predictions are often proven wrong. The unwarranted use of uncoupling may lead to very substantial errors in quite common examples, e.g. Figures 6.10 and 6.11; it cannot therefore be justified.

The following points are worth noting:

1. In subsidence problems it is only meaningful to speak of coupled or uncoupled models if the outflow rate is given as an input.
2. It is possible to obtain coupled solutions, based on a general formulation for the pressure field or the displacement field, through elimination of one of the field variables. These models can advantageously substitute for other two-step models.
3. For embedded aquifers, with or without underlying bedrock, the use of a one-dimensional consolidation coefficient may or may not be acceptable for subsidence analysis. It depends on the mechanical boundary conditions because the horizontal strains are relevant only in certain cases.
4. The depth of burial does not influence aquifer compaction and drawdown, but it is relevant in the transfer of the aquifer or reservoir compaction to surface subsidence.
5. The mechanical model of a homogeneous half-space bears little or no relevance to real subsidence problems.

This chapter also presents three interesting and useful numerical techniques. The first, based on an averaging technique over a particular representative volume element, allows a fully three-dimensional analysis to be economically performed in particular situations where the average behaviour of thick formations is of interest. The second deals with infinite elements, which are useful in introducing more accurate boundary conditions when analysing very large domains. The third shows a solution for two-phase flow where the mass balance equation for the gas phase is already integrated over the reservoir volume. All these techniques will be used in the following chapters.

## REFERENCES

1. Zienkiewicz, O. C. (1984) Coupled problems and their numerical solution, in R. W. Lewis, P. Bettess and E. Hinton (eds), *Numerical Methods in Coupled Systems*, Wiley, Chichester, pp. 35–58.
2. Lewis, R. W. and Schrefler, B. A., (1987) *The Finite Element Method in the Deformation and Consolidation of Porous Media*, Wiley, Chichester.
3. Lewis, R. W., Schrefler, B. A. and Simoni, L. (1991) Coupling versus uncoupling in soil consolidation. *Int. J. Num. Anal. Meth. Geomech.*, **15**, 533–48.
4. Gambolati, G. (1991) Comments on ‘Coupling versus uncoupling in soil consolidation’. *Int. J. Num. Anal. Meth. Geomech.*, **16**, 833–35.
5. Lewis, R. W., Schrefler, B. A. and Simoni, L. (1992) Authors’ reply ‘Coupling versus uncoupling in soil consolidation’. *Int. J. Num. Anal. Meth. Geomech.*, **16**, 835–37.
6. Desai, C. S. and Saxena, S. K. (1977) Consolidation analysis of layered anisotropic foundations. *Int. J. Anal. Meth. Geomech.*, **1**, 5–23.

7. Gambolati, G., Sartoretto, F., Rinaldo, A. and Ricceri, G. (1987) A boundary element solution to land subsidence above 3-D gas/oil reservoirs. *Int. J. Num. Anal. Meth. Geomech.*, **11**, 489–502.
8. Corapcioglu, M. Y. (1984) Land subsidence: a state-of-the-art review, in *Fundamentals of Transport Phenomena in Porous Media*, J. Bear and M. Y. Corapcioglu (eds), Nijhoff, Dordrecht, pp. 371–444.
9. Terzaghi, K. (1925) *Erdbaumechanik auf bodenphysikalischer Grundlage*, Franz Deutike, Wien.
10. Biot, M. A. (1941) General theory of three-dimensional consolidation. *J. Appl. Phys.*, **12**, 155–64.
11. Helm, D. C. (1975) One-dimensional simulation of aquifer system compaction near Pixley, California. I, constant parameters. *Water Resources Research*, **11**, 465–78.
12. Gambolati, G. and Freeze, R. A. (1973) Mathematical simulation of the subsidence of Venice. I, theory. *Water Resources Research*, **9**, 721–33.
13. Krause, G. (1978) Finite element schemes for porous elastic media. *J. Engng Mech. Div. ASCE*, **104** (EM3), 605–20.
14. Smith, I. M. (1978) Transient phenomena of offshore foundations, in *Numerical Methods in Offshore Engineering*, O. C. Zienkiewicz, R. W. Lewis and K. G. Stagg (eds), Wiley, Chichester, pp. 483–513.
15. Nowacki, W. (1962) *Thermoelasticity*, Pergamon, Oxford.
16. Geertsma, J. Land subsidence above compacting oil and gas reservoirs, *J. Pet. Technol.*, **25**, 734–44.
17. Gambolati, G. (1977) Deviation from the Theis solution in aquifers undergoing three-dimensional consolidation. *Water Resources Research*, **13**, 62–68.
18. Cheng, A. H. D. and Liggett, J. A. (1984) Boundary integral equation method for linear porous-elasticity with applications to soil consolidation. *Int. J. Num. Meth. Engng*, **20**, 255–78.
19. Kanok-Nukulchai, W. and Chau, K. T. (1988) Fundamental solutions and boundary element for saturated porous halfspace in *Computational Mechanics'88* S. N. Atluri and G. Yagawa (eds), Springer-Verlag, 63, vi. 1–2.
20. McNamee, J. and Gibson, R. E. (1960) Plane strain and axially symmetric problems of the consolidation of a semi-infinite clay stratum, *Q. J. Mech. Appl. Math.*, **13**, 210–27.
21. Lewis, R. W., Roberts, G. K. and Zienkiewicz, O. C. (1976) A non-linear flow and deformation analysis of consolidation problems, in *Proceedings of the 2nd International Conference on Numerical Methods in Geometry*, Blacksburg WV, pp. 1106–18.
22. Saetta, A., Schrefler, B. A. and Vitaliani, R. (1991) Solution strategies for coupled consolidation analysis in porous media. *Int. J. Metodos Numericos para Calculo y Diseno en Ingenieria*, **7** (1), 55–66.
23. Bear, J. and Corapcioglu, Y. (1981) Mathematical model for regional land subsidence due to pumping. 2. Integrated Aquifer subsidence equations for vertical and horizontal displacements. *Water Resources Research*, **17**, 947–58.
24. Simoni, L. and Schrefler, B. A. (1989) FE solution of a vertically averaged model for regional land subsidence. *Int. J. Num. Meth. Engng*, **27**, 215–30.
25. Verruijt, A. (1969) Elastic storage of aquifers, in *Flow Through Porous Media*, R. J. M. DeWiest (ed.) Academic Press, New York, pp. 331–76.
26. Reed, M. B. (1984) An investigation of numerical errors in the analysis of consolidation by finite elements. *Int. J. Num. Anal. Meth. Geomech.*, **8**, 243–57.
27. Zienkiewicz, O. C. (1982) Finite element method for coupled problems in engineering, in *Proceedings of the International Conference on Finite Element Methods*. Science Press, Beijing, China, pp. 1–15.

28. Lewis, R. W., Majorana, C. E. and Schrefler, B. A. (1986) A coupled finite element model for the consolidation of nonisothermal elastoplastic porous media. *Transport in porous media*, **1**, 155–78.
29. Simoni, L. and Schrefler, B. A. (1989) Numerical modelling of unbounded domains in coupled field problems. *Meccanica*, **24**, 98–106.
30. Park, K. C. and Felippa, C. A. (1983) Partitioned analysis of coupled systems, in *Computational Methods for Transient Analysis*, T. Belytschko and T. J. R. Hughes (eds), Elsevier, Amsterdam.
31. Schrefler, B. A. (1985) A partitioned solution procedure for geothermal reservoir analysis. *Comm. Appl. Num. Meth.*, **1**, 53–56.
32. Beer, G. and Watson, J. O. (1992) *Introduction to Finite and Boundary Element Methods for Engineers*. Wiley, Chichester.
33. Ausaverri, M., Huerta, A. and Kriegsman, G. (1993) Condiciones de contorno artificiales para dominios infinitos: aplicacion a un problema de filtracion, in *Metodos numericos en ingenieria*, F. Navarrina and M. Casteleiro (eds), SEMNI, Barcelona, pp. 511–20.
34. Bettes, P. (1992) *Infinite Elements*, Penshaw Press, Cleaton.
35. Lewis, R. W., Morgan, K. and Roberts, P. M. (1985) Infinite element modelling of heat losses during thermal recovery processes, in *Proceedings of the 8th SPE Symposium on Reservoir Simulation*, Dallas TX, pp. 191–200.
36. Simoni, L. and Schrefler, B. A. (1987) Mapped infinite elements in soil consolidation. *Int. J. Num. Meth. Engng.*, **24**, 513–27.
37. Schrefler, B. A. and Simoni, L. (1987) Nonisothermal consolidation of unbounded porous media using mapped infinite elements. *Comm. Appl. Num. Meth.*, **3**, 445–52.
38. Marques, J.M.M.C. and Owen, D.R.J. (1984) Infinite elements in quasi-static materially nonlinear problems. *Computers and Structures*, **18**, 739–51.
39. Schrefler, B. A., Lewis, R. W. and Majorana, C. (1981) Subsidence above volumetric and waterdrive gas reservoirs. *Int. J. Num. Meth. Fluids*, **1**, 101–15.
40. Craft, B. C. and Hawkins, M. F. (1969) *Applied Petroleum Reservoir Engineering*, Prentice Hall, Englewood Cliffs NJ.
41. Crichlow, H. B. (1977) *Modern Reservoir Engineering: A Simulation Approach*, Prentice Hall, Englewood Cliffs NJ.
42. Slider, H. C. (1976) *Practical Petroleum Reservoir Engineering Methods*, Petroleum Publications, Tulsa OK.
43. Standing, M. B. (1952) *Volumetric and Phase Behaviour of Oil Field Hydrocarbon Systems*, Reinhold, New York.
44. Bear, J. and Corapcioglu, M. Y. (1991) Mathematical model for regional land subsidence due to pumping. 1. Integrated aquifer subsidence equations based on vertical displacement only. *Water Resources Research*, **17**(4), 937–46.
45. Theis, C. V. (1935) The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Eos Trans. AGU*, **16** 519–24.

# 7

## Modelling Subsidence: Case Studies

### 7.1 INTRODUCTION

The solution procedures outlined in previous chapters have been applied to several cases of recorded surface subsidence due to the withdrawal of underground fluids. These case histories cover four regions in the lower Po valley of north-east Italy (Figure 7.1).

The first case involves the historical city of Venice, where the settlement is due to removal of water underlying aquifers in the neighbouring industrial area of Marghera. The second involves the Polesine area and the Po Delta, where considerable surface subsidence was recorded due to the removal of water and gas from the underlying reservoir sediments. In this instance gas was dissolved in the water and the separation between methane and water occurred at the surface. Sufficiently documented subsidence bowls of this area are modelled in the succeeding sections by means of the fully coupled consolidation equation (6.4).

In the Venice model a linear elastic constitutive relationship was used, whereas in the Po–Polesine model a non-linear elastic and an elastoplastic analysis was carried out as the necessary material properties were available. The problem of subsidence rebound has been investigated in detail. Equation (6.4) was used in both models.

The third area of interest is Ravenna, where both water and gas are being produced. The water is withdrawn from the upper region of the sediments, and gas is produced from deep reservoirs onshore and offshore. It is conceivable that the resulting settlements may permanently modify the existing shoreline. For the parametric investigations carried out in this case the procedure outlined in Section 6.4 has been applied.

The fourth model deals with the subsidence in the geothermal area of the Euganean Hills. It used the two-dimensional model with variable thickness of the aquifer system, obtained by vertically averaging the governing equations (Section 6.2).

Each of the cases and the applied mathematical models will now be discussed in detail and the results compared, whenever possible, to other existing results. Particular attention will be paid to the problem of modelling the various case histories.

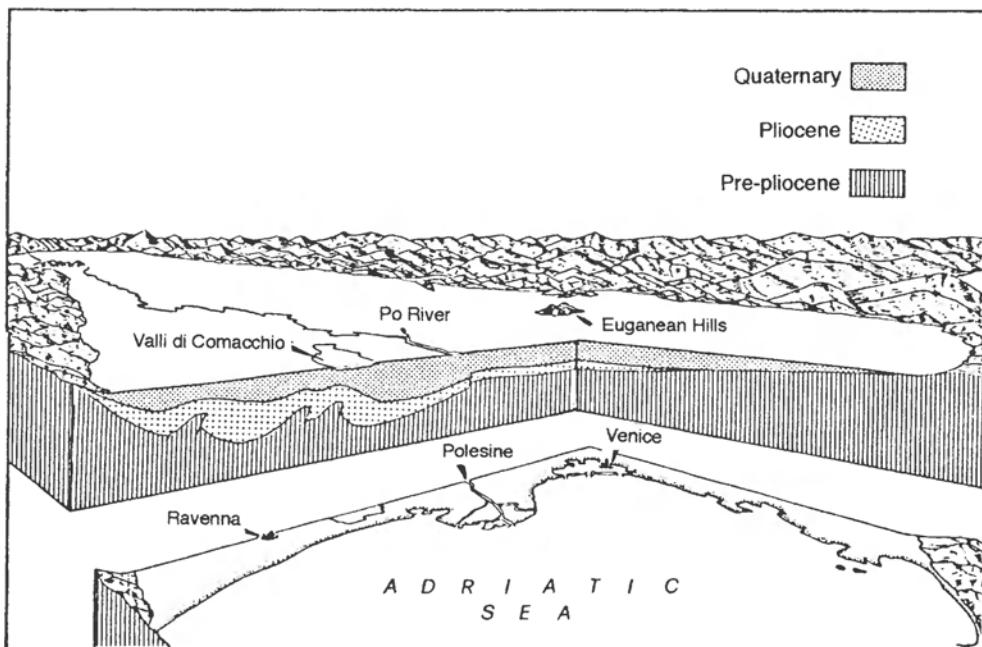


Figure 7.1 Cross-section of the Po valley between Ravenna and Venice

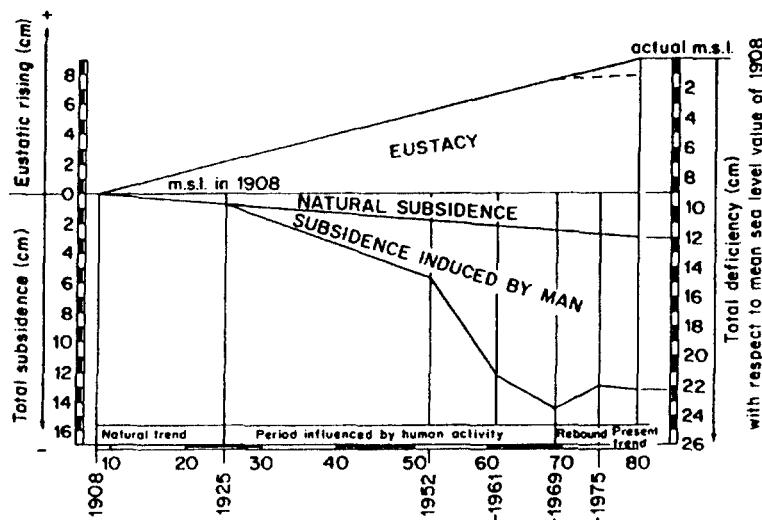
## 7.2 THE SUBSIDENCE OF VENICE

### 7.2.1 Background

Three principal factors have contributed to the lowering of Venice and its surrounding regions with respect to the mean sea level. Two of these factors are of natural origin: (a) eustatic rising of the mean sea level, and (b) land sinking due to the natural consolidation of fine-grained deposits in the lagoon substratum. The third factor is man-induced subsidence due to groundwater withdrawal.

The glacial-eustatic rise depends upon world climatic changes. The eustatic rise from the beginning of the century has an average value of 1.27 mm per annum [1], resulting in a loss of elevation of approximately 9 cm during the period 1908–1980. The effects of natural consolidation are estimated to have produced a subsidence of roughly 0.42 mm per annum [2]. A substantial contribution to this subsidence was made by an increased salt concentration in the saturation water of the clay sediments. This was due to a change of the brackish characteristics of the lagoon to a saline environment during the evolution of the basin. The actual rate of natural subsidence is about 0.4 mm per annum, contributing by 1980 about 3 cm to the total land elevation loss since the beginning of this century.

The man-induced subsidence commenced around 1925 with the development of the industrial zone of Marghera on the mainland. In the period 1952–1969 the water withdrawn for this region resulted in a settlement of 10 cm. Later, the rate of land



**Figure 7.2** The combined lowering of the surface level of Venice from 1908 to 1980 and the contribution of the three responsible factors. Trend value of natural components: eustatic rising = 1.27 mm per year (1896–1967); subsidence = 0.4 mm per year (before 1925) (Reprinted, by permission, from Gatto, P. and Carbognin, L., 1981, The lagoon of Venice: natural environmental trend and man-induced modification, *Hydr. Sci. Bull.*, **26**, 379–91)

subsidence decreased due to a reduction in the withdrawal rate of groundwater. The aquifers recovered and finally in 1975 a land rebound of approximately 2 cm was recorded.

The overall effects of these three factors for the period 1908–1980 are schematically represented in Figure 7.2 [3]. Notice that the industrially induced subsidence contributed 47 percent of the total loss in elevation, taking into account the rebound. The joint effect of the three factors is a rising of the sea level by a total of 22.4 cm. This is a serious situation for Venice, which has a mean elevation of only 1 m above mean sea level [3]. If we consider that the mean spring high tide is 55 cm over tide gauge zero, then this rise of 22 cm represents 30 percent of the safety margin for ordinary high tides. As a result, flooding of Venice occurs much more frequently than in the past.

This industrially caused subsidence was due to the extensive groundwater withdrawal from the aquifer-aquitard system underlying the lagoon of Venice and the surrounding mainland [4–6].

Particularly heavy pumpage took place until 1972 in Mestre and in its industrial area, Marghera, both situated on the mainland close to the lagoon. Since 1969 a research group of the Consiglio Nazionale delle Ricerche (CNR) has been active in Venice, collecting and processing a large amount of experimental and field data regarding the local and regional stratigraphy. This information includes the material properties obtained from the analysis of core sample from test holes [7], the withdrawal rates of water at Marghera, Venice and Lido, and the records of piezometric decline and subsidence. Comprehensive details of available data can be found in the literature [8,9].

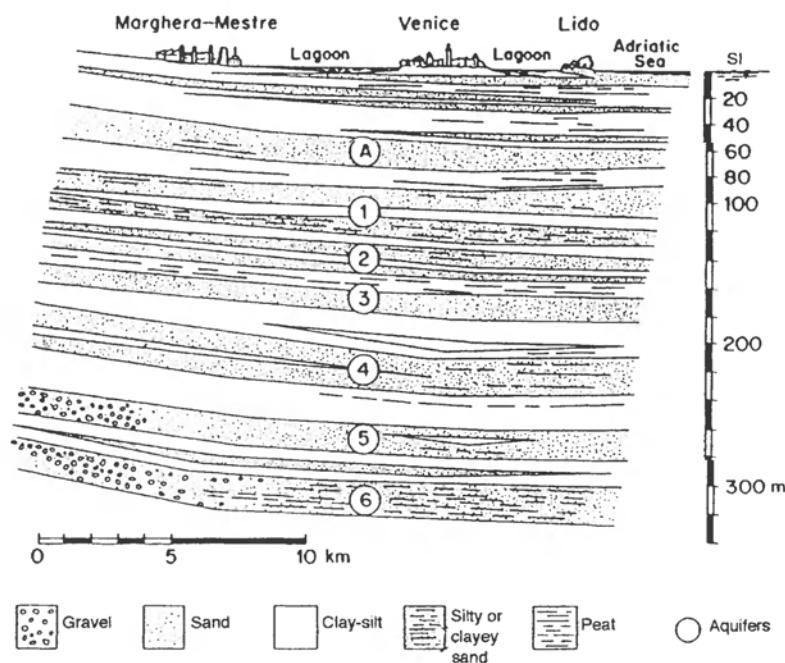


Figure 7.3 Cross-section of the confined aquifer system

Table 7.1 Limits of the aquifers, aquitards, material properties and permeabilities used

	Aquifers			Aquitards	
	Depth(m)	K (m per day)	E(MPa)	K(m per day)	E(MPa)
1	0–94			0.0432	26.4
	94–130	13.82	817.0	0.00432	8.6
2	130–135			$0.864 \times 10^{-2}$	9.5
	135–156	4.147	817.0		
3	156–168			0.0138	12.9
	168–185	2.765	817.0		
4	185–216			0.00173	17.3
	216–238	5.616	817.0		
5	238–263				
	263–310	5.616	817.0		

In Marghera, Mestre and Venice six main aquifers are exploited. They are located in the upper 350 m, as indicated in Figure 7.3. The depths of the aquifers and aquitards are shown in Table 7.1. For the purpose of mathematical modelling, the sixth aquifer, which is exploited only in Marghera, is combined with the fifth. This procedure was also adopted by Gambolati *et al.* [10] in a ‘two-step model’. A detailed description of these

**Table 7.2** Permeability and porosity values as measured in the laboratory on samples of clean sand taken from the five main aquifers

Aquifer	Average permeability (cm s <sup>-1</sup> )	Average porosity
1	10 <sup>-3</sup>	0.43
2	10 <sup>-4</sup>	0.45
3	2 × 10 <sup>-3</sup>	0.38
4	6 × 10 <sup>-4</sup>	0.42
5	2 × 10 <sup>-4</sup>	0.37

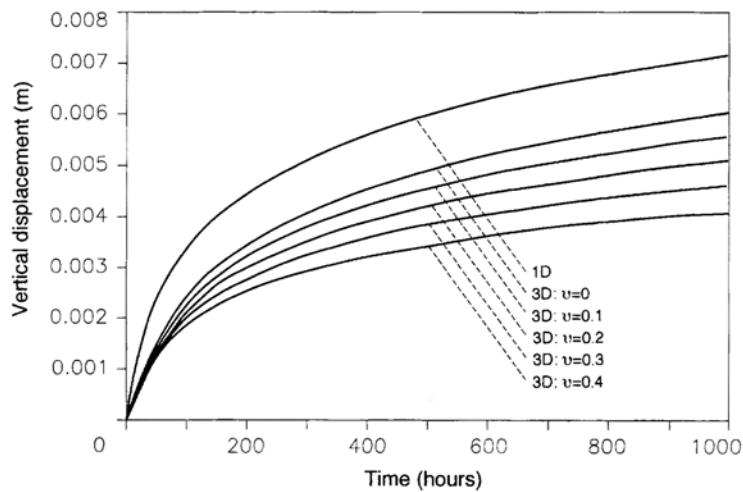
aquifers is given in that article. The results of Gambolati *et al.* [10] will be used for comparison with those obtained via the fully coupled model (6.4).

A linear elastic constitutive relationship has been assumed for the skeleton along with a linear fluid flow regime. This is mainly because there is a lack of information concerning the non-linear behaviour of the porous medium. The material properties and permeabilities used for the simulation are shown in Table 7.1. This data corresponds to the set VE2, as used by Gambolati *et al.* [10] for the fluid flow equations (hydrologic model). This data has been obtained from laboratory tests [7]. The permeabilities in particular were modified by trial-and-error techniques in order to match the known piezometric decline in Venice. The original values are shown for comparison in Table 7.2. Laboratory tests have also shown that the value of soil compressibility (coefficient of volume change) on unloading is 20 percent of the corresponding value in compression.

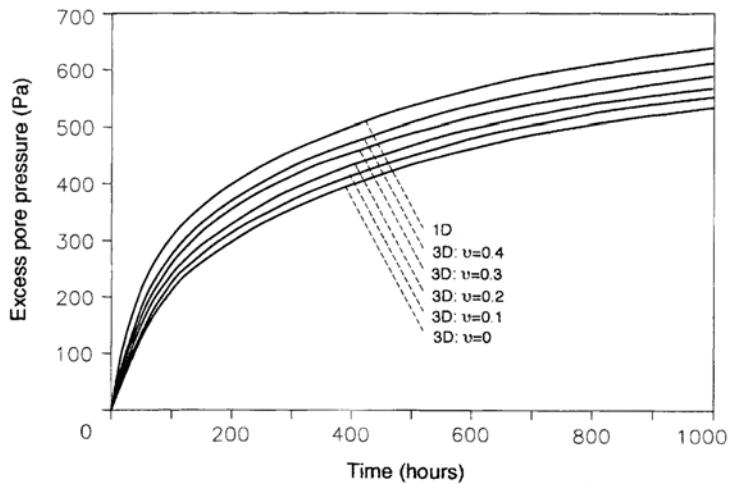
The Young's modulus and Poisson's ratio for the skeleton are obtained from the compressibilities according to

$$m_v = \frac{1}{1+e} \frac{\Delta e}{\Delta \sigma_z} = \frac{(1+v)(1-2v)}{E(1-v)} \quad (7.1)$$

where  $e$  is the void ratio. The solution to this equation is not unique and requires a fixed value of Poisson's ratio. A typical value of  $v = 0.25$  was assumed for the aquifers and  $v = 0.45$  for the aquitards. The effects of such a choice are firstly investigated by analysing the single-aquifer behaviour as a reference solution (Section 6.5). Different sets of  $E$  and  $v$ , for the same compressibility  $m_v$ , have no consequence for one-dimensional consolidation, but they do matter in three-dimensional consolidation. They shift the curves for  $p_w$  onto a  $p_w, t$  plane, but do not change the time transient behaviour (6.75). Figures 7.4 and 7.5 present the time histories of the pore pressure and the compaction for the isolated aquifer at a distance 10  $h$  from the well ( $h$  is the thickness of the aquifer) and were obtained using different mechanical parameters for the same  $m_v$ . The differences between the results are more remarkable for the displacement values than for the pore pressure. For water pressure, the results obtained with  $v = 0.25$  in three-dimensional conditions differ from those of the limiting values ( $v = 0$  and  $v = 0.4$ ) by 5 percent, whereas the difference with one-dimensional solution is of the order 12 percent at  $t = 1000$  hours. Hence the choice of Poisson's ratio does not substantially alter the response of the model.

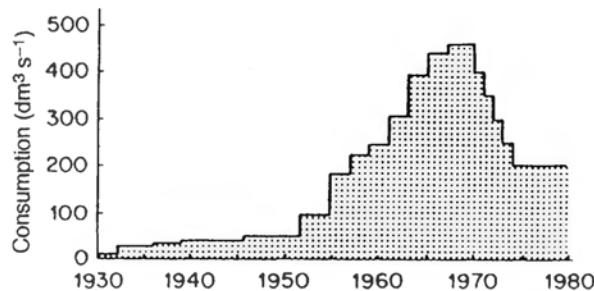


**Figure 7.4** Isolated aquifer: analytical solutions for vertical displacement versus time using one-dimensional and three-dimensional hypotheses with different values of Poisson's ratio  $v$ ;  $r = 10\text{ h}$



**Figure 7.5** Isolated aquifer: analytical solutions for excess pore pressure versus time using one-dimensional and three-dimensional hypotheses with different values of Poisson's ratio  $v$ ;  $r = 10\text{ h}$

The water withdrawal from the Marghera wells since 1930 and the respective percentages from each aquifer are shown in Figure 7.6 and Table 7.3. The data until 1969 has been taken from the work of Serandrei-Barbero [11], and the withdrawal from 1970 to 1976 has been obtained from Carbognin *et al.* [12]. In 1969 the maximum withdrawal in Marghera was about 460 litres per second. After 1970 the industrial pumpage was reduced gradually. A sensible decrease of active wells occurred in the period from 1972 until 1975, where a new industrial aqueduct was put into operation. During the same period industrial activity was also reduced in Marghera due to an

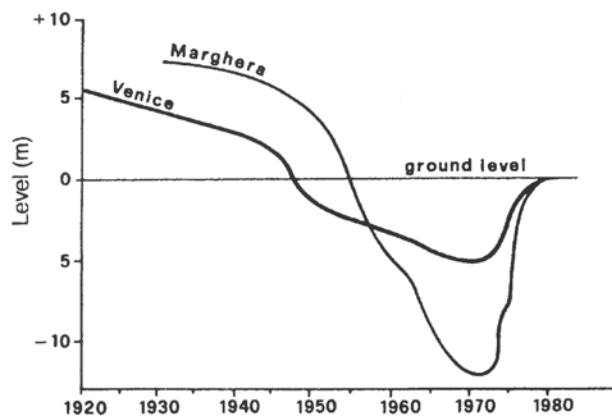


**Figure 7.6** Groundwater consumption in the Marghera industrial area

**Table 7.3** Subdivision of the groundwater consumption between the different aquifers (%)

Period	Aquifer				
	1	2	3	4	5
1930–1935	43	7	21	29	0
1935–1940	22	7.5	19	44	7.5
1940–1945	25	7	18	43	7
1945–1950	24	7	17	45	7
1950–1955	27	4	14	43	12
1955–1960	21	5	17	35	22
1960–1965	19	6	22	31	22
1965–1970 <sup>a</sup>	18	6	22	33	21

<sup>a</sup> Values constant from 1970 to 1980.



**Figure 7.7** Average piezometric level in the industrial area and in Venice from 1920 to 1980

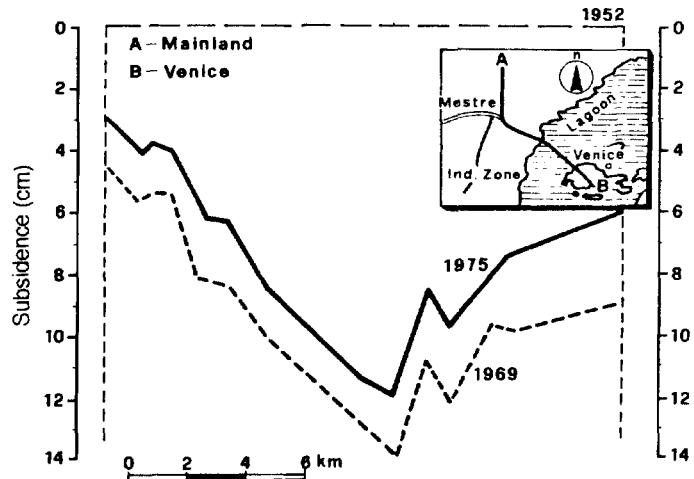


Figure 7.8 Recorded subsidence since 1952

economic crisis. The withdrawal has been estimated to be about 200 litres per second until 1980, subsequently being drastically reduced and eventually stopped completely.

The mean values of the observed piezometric decline from 1920 to 1975 in the same regions are shown in Figure 7.7. A remarkable recovery of the piezometric levels after 1970 can be seen. In 1975 the average recovery in Marghera reached a maximum of over 8 m and in Venice more than 3 m. Data up to 1970 has been taken from Serandrei-Barbero [11] and data for the period 1970–1975 has been taken from Carbognin *et al.* [12]. In Figure 7.8 the recorded subsidence since 1952 is shown for a levelling line A–B from the mainland to Venice itself [12]. The rebound following the piezometric recovery after 1970 may be clearly observed.

### 7.2.2 The Mathematical Model

An axisymmetric model has been used to simulate the available field data which implies a horizontal stratigraphy, central pumping area (Marghera) and radial symmetry of the material properties. A realistic simulation of the well distribution is of great importance for the case of a fully coupled model. It is incorrect to replace a large number of wells by a single pumping site, as is usually done in uncoupled models [10], since this alters the relationship between the area open to flow and the outflow. Hence extremely high body forces result, causing excessively large settlements.

By taking into account the spatial density of active wells, it is more realistic to consider the source points distributed over a certain area. In the model used for the simulation, the distribution of the source points has been assumed to act over a circular area with a diameter of 3.0 km. This assumption was supported by an examination of available piezometric contours of the aquifers considered (Figure 7.9) [9]. It also resulted in a reasonably close matching of the subsidence values in Marghera. This spatial distribution of the source wells has been modelled by specifying the nodal flux at the nodes of each aquifer, contained in the vertical cylindrical surface with a radius

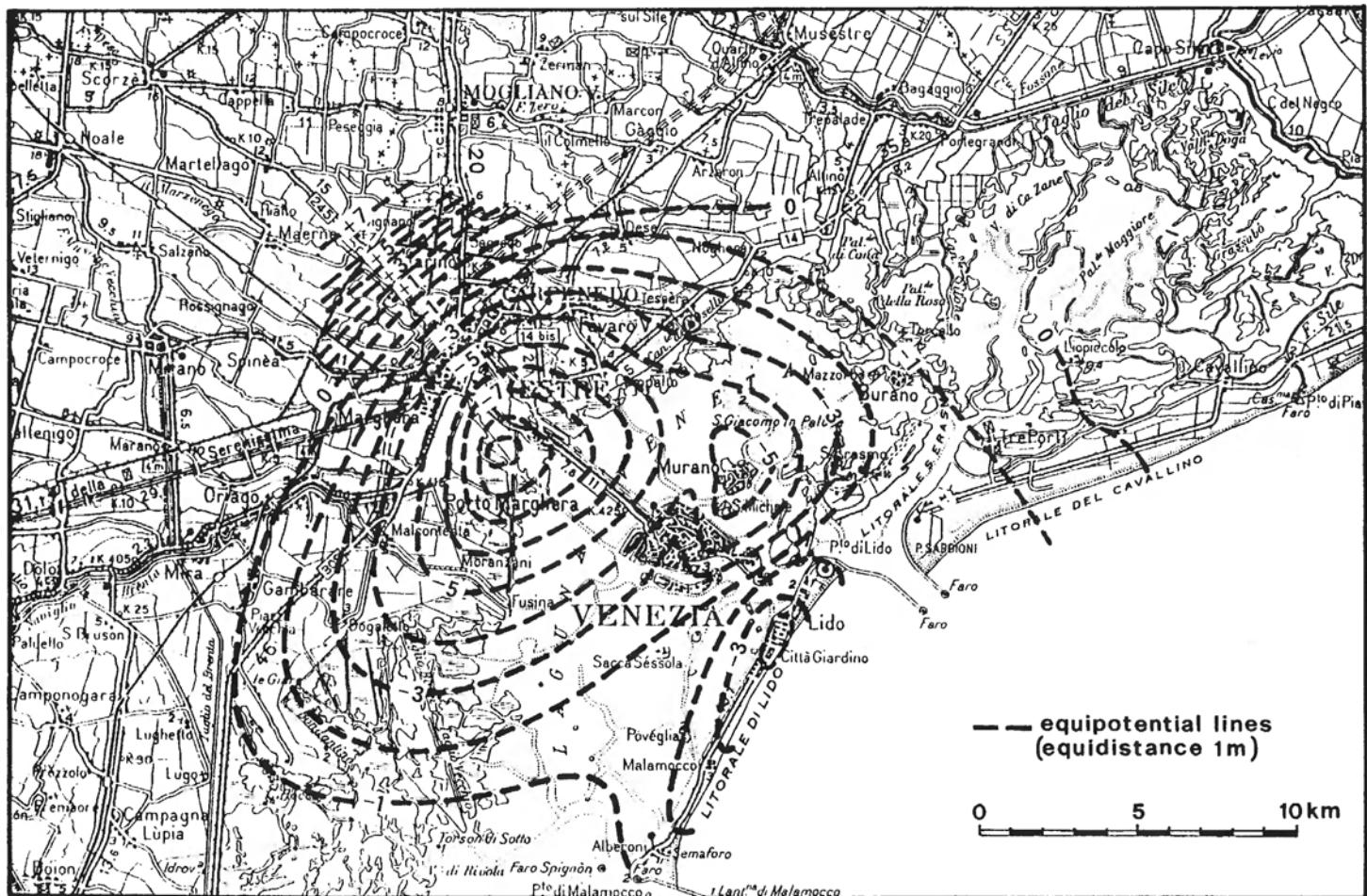


Figure 7.9 Piezometric surface of the fourth aquifer

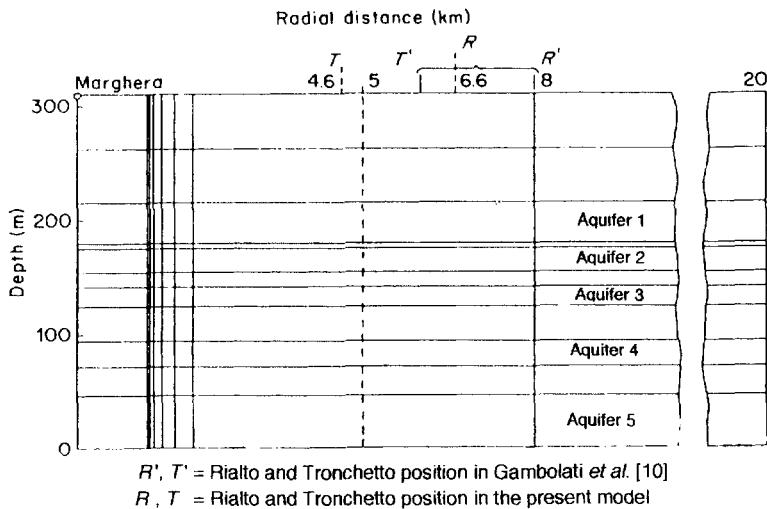


Figure 7.10 Finite element discretisation of the investigated area

$R = 1.5$  km. This assumption yielded satisfactory results both for the vertical displacement and the pore pressure distribution in the central region of the model.

The finite element discretisation is given in Figure 7.10 and clearly shows the idealised geological stratigraphy. Isoparametric parabolic elements have been used, which considerably reduce the number of elements necessary for a realistic representation of the region. The mesh has also been chosen in order to compare answers with results of Gambolati *et al.* [10], except for the position of the reference point in Venice (Rialto). In the present model, this is located at 6.6 km (instead of 8 km) from the centre of the model (Marghera), as suggested on the basis of the piezometric surfaces [9].

The mesh presented in Figure 7.10 is sufficient to give the overall response reported in what follows. For more detailed results, a finer mesh is mandatory with a consequent increase of computing time.

The assumed boundary conditions were a fixed potential with zero horizontal displacements at the external vertical boundary. This assumption appears reasonable as the aquifers are recharged landward. The other possibility was to assume an impermeable outer boundary, which is more likely to prevail seaward. However, trial calculations indicated that the latter case produced excessive settlements and piezometric decline in the period of heavy pumpage. The base of the system was taken to be impermeable with no vertical displacements. An impermeable surface with free displacements was assumed at the top of the model. A more detailed investigation of the influence of the boundary conditions on the behaviour of the model is reported in Section 7.3.2.2. The initial conditions were assumed as being the values in 1930.

In the present model, the pumpage of one additional well on the city limit of Venice (Tronchetto) has been disregarded because a preliminary calculation had shown its influence to be small.

Several trial runs have been carried out to determine the optimal value of time increment necessary. A time step of 91 days showed no oscillation of the pore pressures

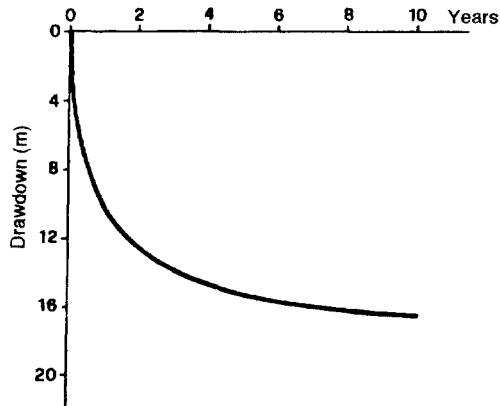


Figure 7.11 Third aquifer: pore pressure decrease versus time for a constant outflow rate

either in time or space, even if an initial outflow rate of 400 liters per second was maintained throughout the simulation. Figure 7.11 plots versus time the drop of pore pressures in the third aquifer at the centre of the model. In order to simulate the problem for the whole period from 1930, a large time step was needed even if a penalty of pore pressure oscillation resulted. For the real case of a variable outflow rate with time, a time increment of one year showed no oscillation of the displacements but some oscillation of the pore pressures at the central part of the model. These oscillations vanished at a radial distance of 2 km from the centre. The third aquifer, which has the lowest permeability,

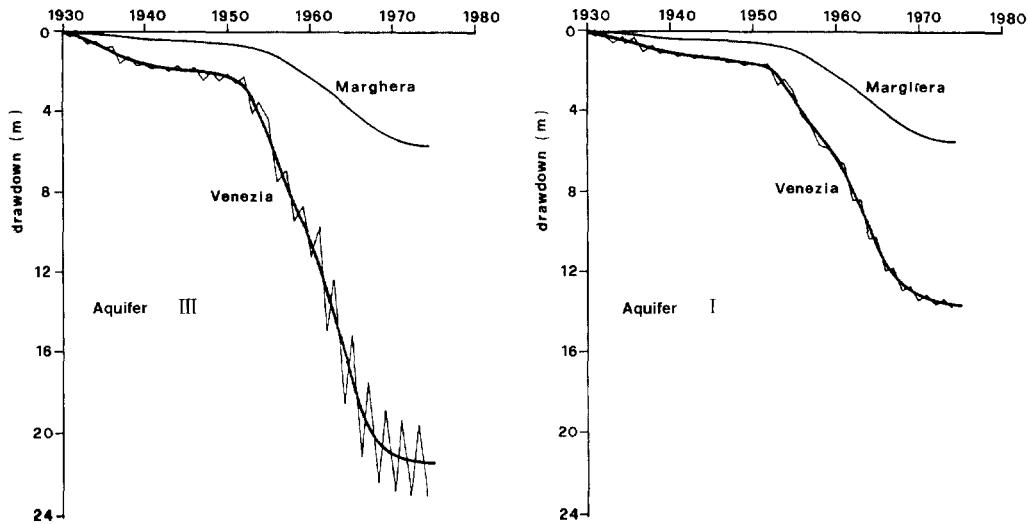


Figure 7.12 Oscillation of the pore pressure in the centre part of the model for variable outflow and a time step length of 1 year: (a) third aquifer, (b) first aquifer

was most affected by these oscillations which are shown in Figure 7.12(a), together with the time transient mean value of the pore pressures. This mean value was sufficiently accurate for the purposes of the simulation. The resulting piezometric decline at 8 km from the centre of the model is also shown.

The first aquifer, which contributes a higher or almost equal pumping rate, when compared to the third aquifer (Table 7.3), shows an almost negligible oscillation for the same time step length, as indicated in Figure 7.12(b). This is due to a higher permeability value, which is five times the value for the third aquifer (Table 7.1).

Finally, this time increment of one year was adopted, as in elastic consolidation analyses the oscillation of the nodal pore pressures do not seriously affect the displacements.

For the calibration of the fully coupled model, the available field data until 1969 has been taken into account. During this period, a satisfactory match for both the subsidence and the piezometric decline in Venice and Marghera was obtained. Results obtained using this model have been published in the literature [13,14]. The same model was subsequently capable of matching the behaviour of the region since 1969 by taking into account the new records available for the period 1969–1976 [15].

### 7.2.3 Results

The final parameters for the model have been determined via a trial-and-error technique by matching the recorded history for the period 1930–1969. Before further data became available, the pumping rate was kept constant after 1969 and equal to its maximum value. The model indicates a tendency towards steady-state conditions once the forcing term (i.e. water pumpage) is kept constant. This behaviour, numerically calculated and experimentally observed [7], is shown in Figure 7.13, where it is compared with the predictions made using the uncoupled referenced model.

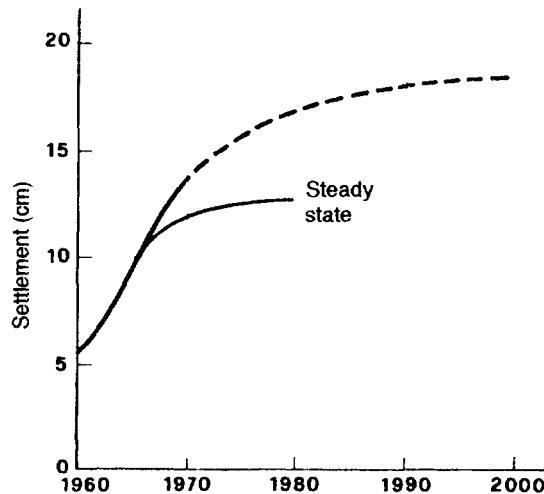
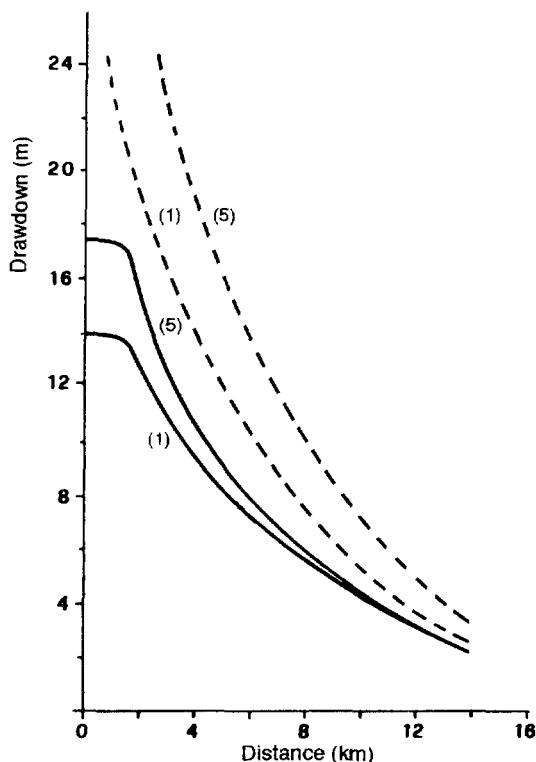


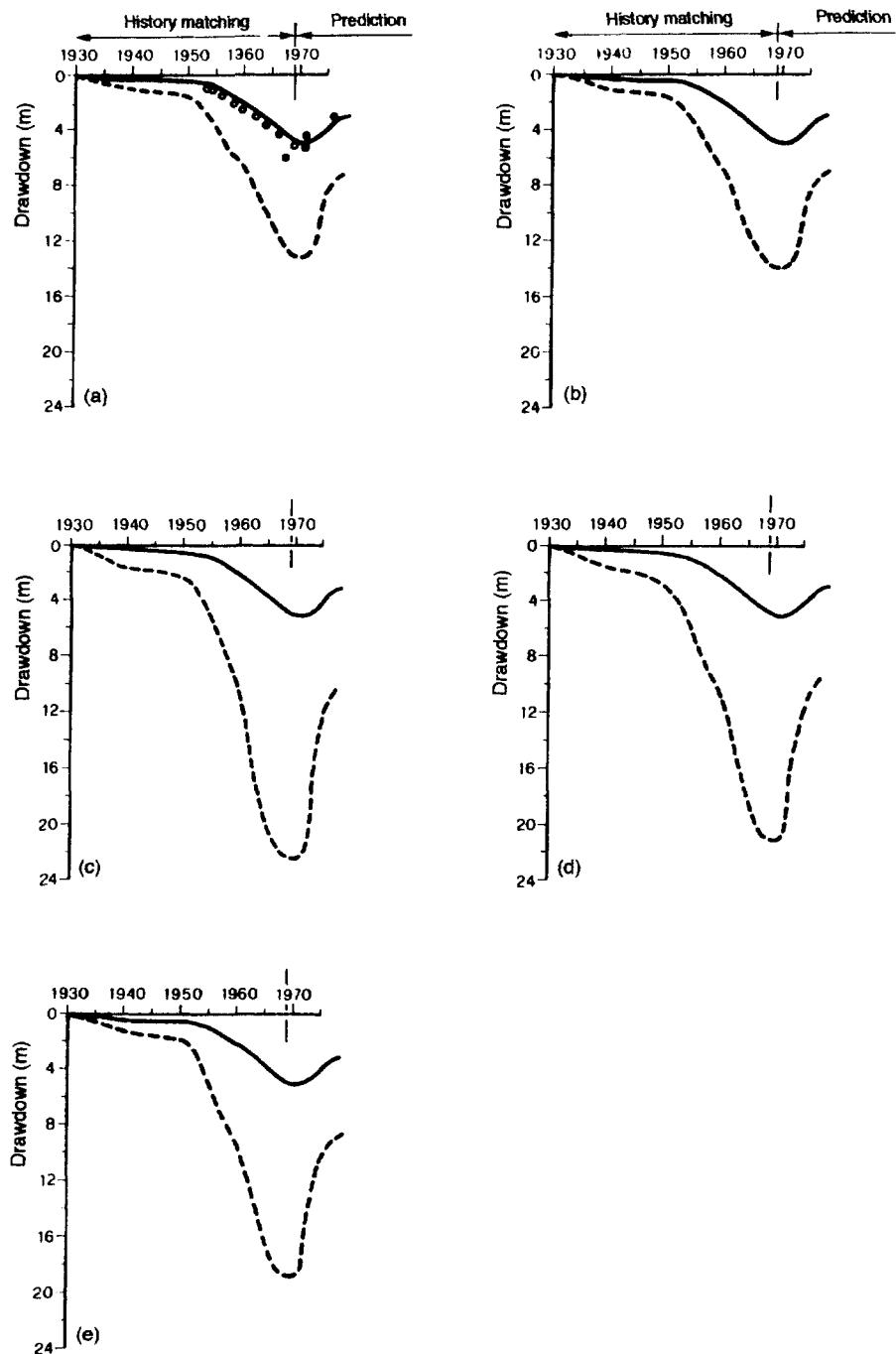
Figure 7.13 Settlement versus time: (—) fully coupled model, (- - -) uncoupled model



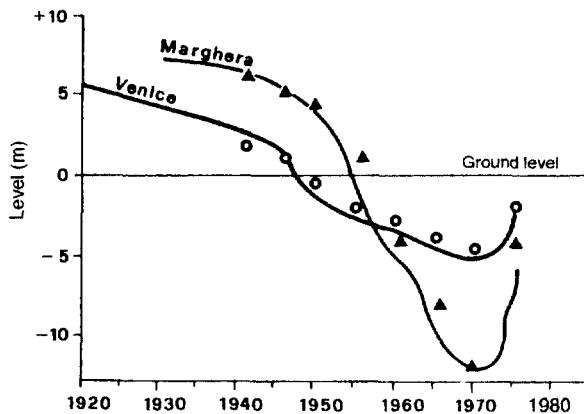
**Figure 7.14** Piezometric decline versus distance from the centre under steady-state conditions for a withdrawal rate of  $454 \text{ dm}^3 \text{ s}^{-1}$ : (1) for the first aquifer and (5) for the fifth aquifer. Key: (---) two-step model, (—) coupled model

The piezometric decline versus distance from Marghera at steady-state conditions for both the first and fifth aquifer is shown in Figure 7.14 and for a withdrawal rate equal to the maximum value in 1969. These results are compared with those obtained from the two-step model [10] and the influence of the spatial distribution of wells in the coupled model can clearly be seen. The curve obtained with the coupled model makes it possible to match the observed values at the central region, but that is impossible with the uncoupled two-step model.

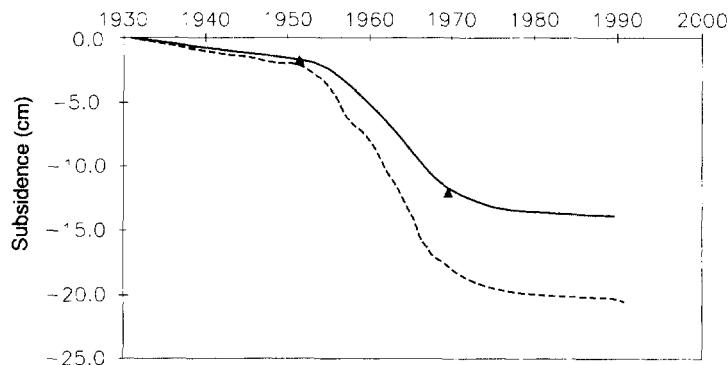
Figure 7.15 shows the piezometric decline versus time for the five aquifers in Venice and in Marghera. The numerical results of the first aquifer are compared with the available experimental records [12], and the agreement is remarkable. The period used for history matching by a trial-and-error technique is shown in Figure 7.15. The results obtained after 1969 are predictive, i.e. only the recorded moduli for unloading-reloading and the new withdrawal rates have been introduced into the model. All other parameters are kept unchanged. Even more remarkable is the agreement between the calculated and recorded mean values of the head drawdown of the five aquifers (Figure 7.16).



**Figure 7.15** Finite element results for drawdown of the five aquifers at Marghera (---) and at 8 km away (—): (a) first aquifer, (b) second aquifer, (c) third aquifer, (d) fourth aquifer, (e) fifth aquifer; (..) experimental results for the first aquifer



**Figure 7.16** Mean values of the piezometric decline: (—) observed, (○, ▲) finite element results



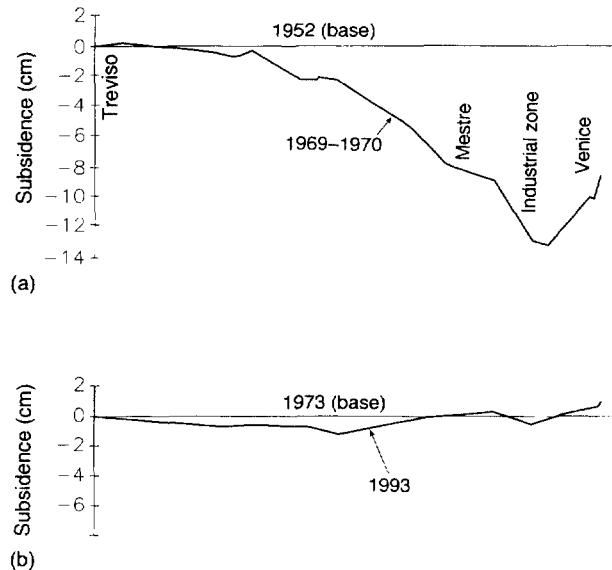
**Figure 7.17** Subsidence versus time: (---) numerical result at 5 km; (—) numerical result at 8 km; (▲) observed subsidence at 6.6 km

The subsidence versus time at a point 8 km from the centre of the model is shown in Figure 7.17 (pumping constant from 1969 onwards). Also shown are the subsidence values at a point situated 5 km from the centre of the models as well as the recorded values at a point 6.6 km from the centre.

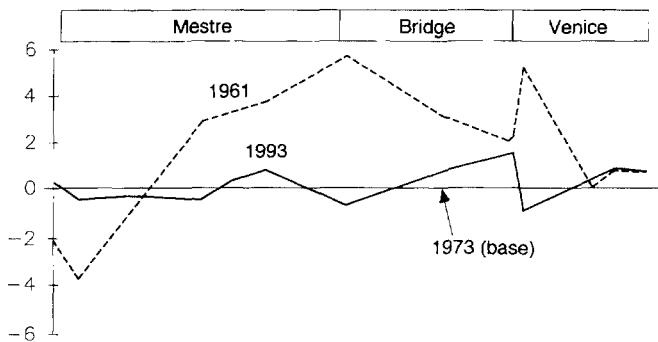
The vertical displacements at the surface create the observed subsidence bowl. No recorded data exists for the horizontal displacements, but these displacements have been observed in other subsidence cases, e.g. in the San Joaquin Valley [16], and have probably also occurred in Venice.

A careful regional levelling exercise was carried out in 1993 along the same path as in Figure 7.8 (Figure 7.18) [17]. The results allow for an evaluation of the present trends of land subsidence after the reduction and the eventual termination of groundwater exploitation.

Compared with 1973, the 1993 levelling exercise generally indicates that the mainland and surrounding areas of Venice and the city are substantially stable, whereas



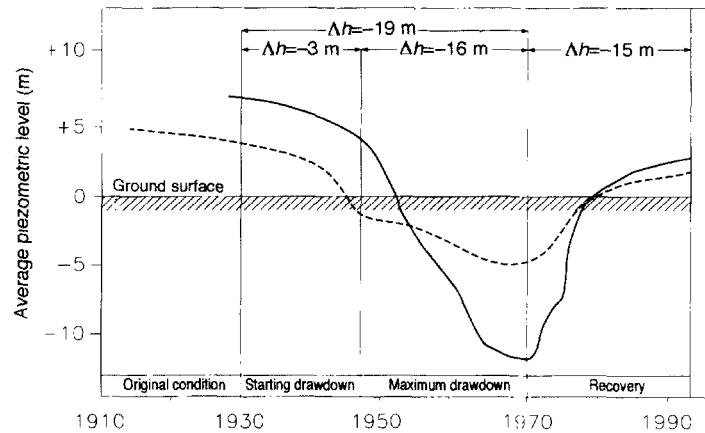
**Figure 7.18** Ground elevation along the levelling line from Treviso to Venice (Pta della Dogana): (a) 1952 (base) to 1969; (b) 1973 (base) to 1993 (Reprinted, by permission of A. A. Balkema, from Barends *et al.* [49])



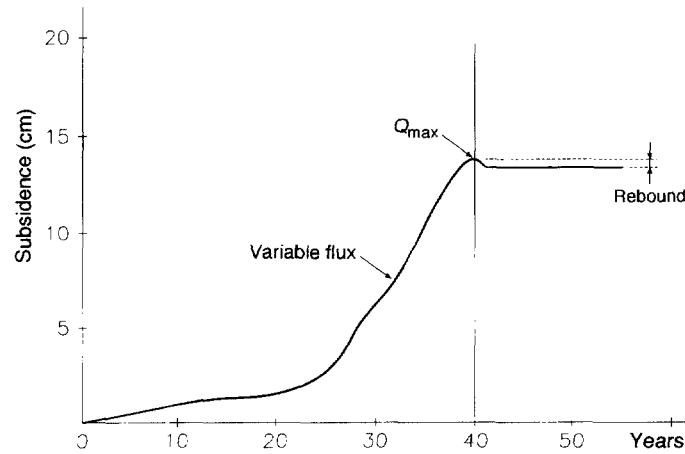
**Figure 7.19** Ground elevation along the levelling line from Mestre to the extreme tip of Venice in the years 1961, 1973 (base) and 1993 (Reprinted, by permission of A. A. Balkema, from Barends *et al.* [49])

at the extreme edges of the lagoon and littoral land, sinking is still going on with a rate of subsidence smaller than was measured in the previous 20-year period. This last phenomenon takes place in areas of more recent formation with respect to the others and close to river outflows. Hence natural compaction is more active here and the relationship with aquifer exploitation is questionable [17].

What is remarkable is that the subsidence, dependent on groundwater pumping, recorded in the period 1952–1970 (Figure 7.18(a)) had practically stopped by 1993 and a small rebound took place in some areas. Figure 7.19(b) presents the ground surface elevation taking the 1973 values as reference. The results presented, which show the



**Figure 7.20** Evolution in time of the average piezometric level in (—) the industrial area and (---) Venice (Reprinted, by permission of A. A. Balkema, from Barends *et al.* [49])



**Figure 7.21** Modelling of rebound after pumping has completely stopped

ground elevation along the levelling line from Mestre to the extreme tip of Venice in the years 1961, 1973 (base) and 1993, confirm the stability of the soil surface. This situation is validated by the analysis of the temporal piezometric surface variations of the once exploited aquifers [18]. Figure 7.20 presents the evolution in time of the average piezometric level in the industrial area and in Venice. The same level as occurred in 1952 had been recovered in the aquifers by the year 1993.

These results confirm the forecasts of the numerical fully coupled model; the tendency to steady-state conditions (Figure 7.13) and the rebound analysis are presented in Figure 7.21.

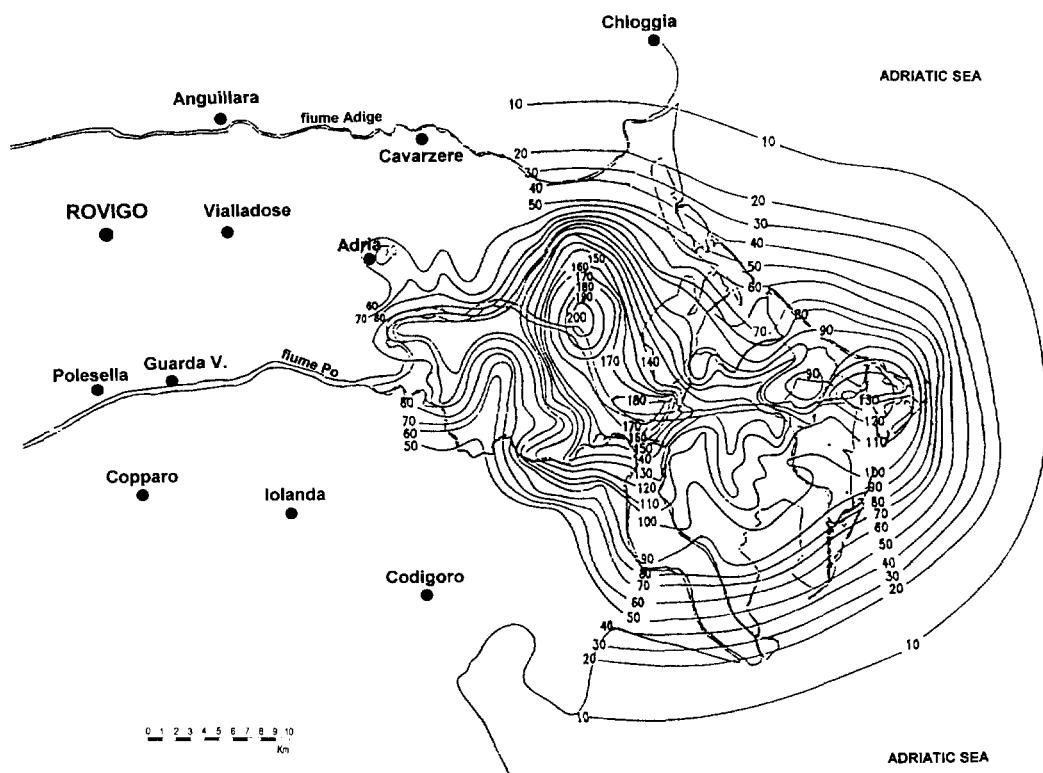
The numerical results presented in this section indicate that the fully coupled consolidation model realistically simulates the piezometric decline and resulting subsidence in Venice, caused by the pumpage of underground water in the industrial area of Marghera. Also, the model clearly indicates the short time delay between changes of the withdrawal rate in Marghera and the resulting effects in Venice.

### 7.3 SUBSIDENCE IN THE PO DELTA AND THE POLESINE

#### 7.3.1 Background

In the 1950s and early 1960s the Po Delta and the Polesine region of northern Italy experienced abnormally large surface settlements due to the production of natural gas (methane) [4,6]. The subsidence was caused by the large quantity of water containing dissolved methane, which was withdrawn from the underlying reservoir sediments. Gas production commenced in the region in 1938 but did not achieve a significant value until the early 1950s, when the demands of an industrial expansion resulted in a sharp rise. During the early period, the gas produced was equal in volume to that of the water, but this decreased during the lifetime of the field until the water exceeded the gas volume by 40 percent during the final stage of production.

The period of greatest production occurred in 1959, when 3860 wells covering an area of  $2600 \text{ km}^2$  produced a total of  $260 \times 10^6 \text{ m}^3$  of gas. The surface settlement of the region resulted in the Po River flooding several times, with resulting damage to property. Therefore most of the wells were closed in 1963. Even after closure the settlement continued until 1966 with an exponential decay, as recorded by Zambon [4]. The maximum recorded subsidence during the whole period was 3.20 m.

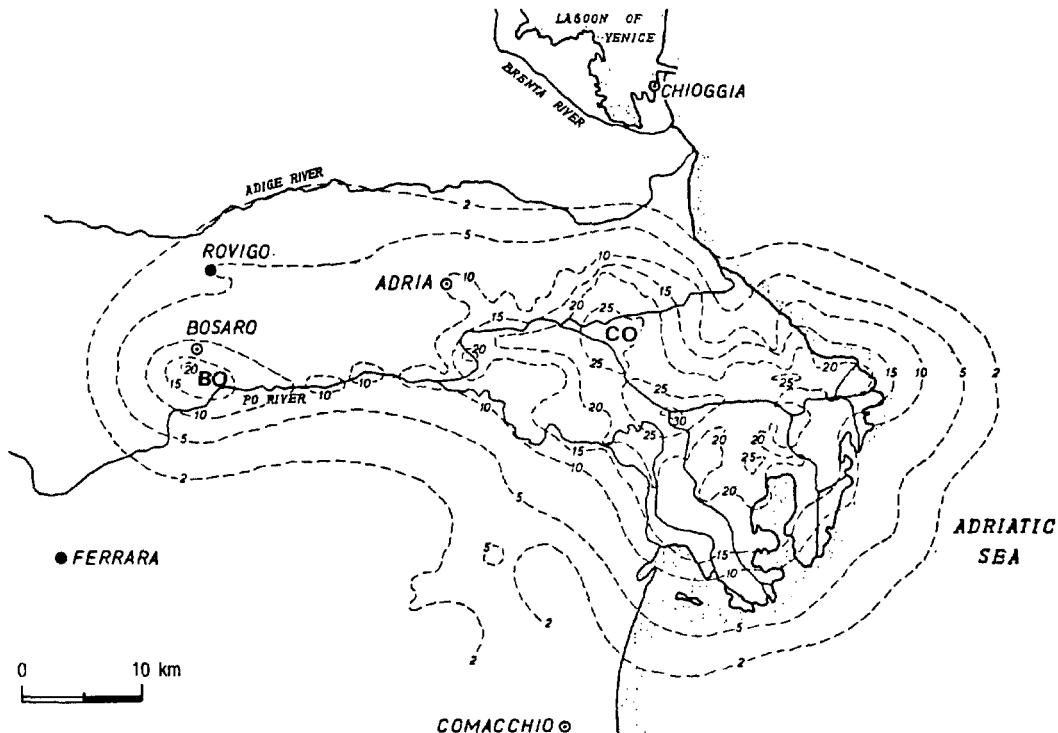


**Figure 7.22** Land subsidence (cm) in the Po Delta from 1951 to 1960

The reservoir consists of layers of sand separated by clay, and Selli [19] indicates that five different layers in the area were drained for gas production. These five layers are mainly contained in the upper 600 m of the Quaternary sediments, which attain a thickness of 2000 m in some regions of the Po River delta. In only a few instances does the depth of a well exceed 600 m. Field measurements made in such a deep well indicate that the greatest settlements occur in the upper regions of the sediments and are negligible in deeper layers [20].

The area under investigation was extensively studied by Zambon [4] and by Caputo *et al.* [21], who produced contour maps of surface settlements (Figures 7.22 and 7.23). Inspection of these maps reveals that certain areas have undergone especially large settlements and the subsidence contours are almost circular in form where there is a concentration of wells. The complete region is not sufficiently documented for comparison with a mathematical model, but two areas in particular have sufficient field data to warrant attention.

The first area has its centre near Bosaro, a small village close to Rovigo, and is indicated as BO in Figure 7.23. This area was extensively studied by Schrefler *et al.* [22]. Attention is here focused on the second area, which has its centre near the village of Contarina, indicated as CO in Figure 7.23. In the case of Contarina, subsidence started as early as 1951, but the phenomenon was only sufficiently observed and documented from 1958 onwards, after the damage caused by subsidence became evident.



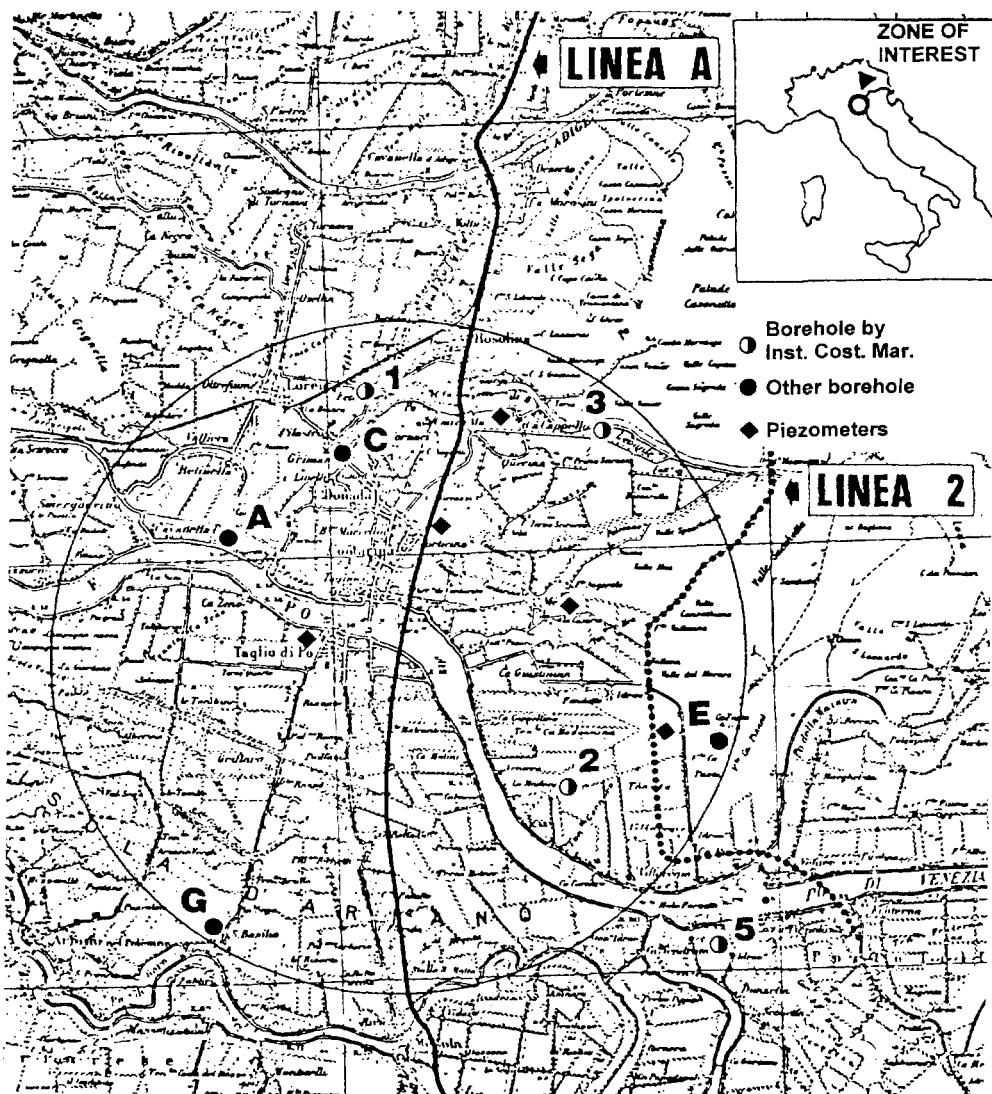
**Figure 7.23** Settlement contours (cm) for 1958 in the Polesine region: BO and CO are the investigated areas

The available material data allowed the application of variable elastic and elasto-plastic constitutive relationships, and the problem of rebound was investigated in detail.

### 7.3.2 The Contarina Model

#### 7.3.2.1 Available Data

The region investigated in the Contarina model is shown in Figure 7.24. It covers a circular area of approximately  $200 \text{ km}^2$  with a diameter of 16 km. The area was chosen



**Figure 7.24** The region investigated in the Contarina model

due to the relatively large amount of data available. This included material properties, settlement histories along two lines crossing the area (indicated by A and 2 on Figure 7.24), piezometric levels showing their recovery after the shutdown of the wells, and also information regarding the pumping rates. Unfortunately, the data on pumping rates was not sufficiently accurate for a complete simulation of the case history.

Nevertheless, the available material properties allowed a non-linear constitutive relationship to be used, hence an extensive parametric investigation was carried out to assess the influence of this non-linearity as well as that of the boundary conditions [23]. Particular emphasis was given to the subsidence rebound problem, where a variable elastic model with capability for rebound analysis was compared with an elastoplastic model.

The material properties were obtained from the soil samples of three test holes of depth 600 m, whose locations are shown in Figure 7.24. The laboratory tests were carried out in the laboratory of the Maritime Construction and Geotechnical Institute of the University of Padua [24]. As in the Bosaro case, five aquifers in the upper 600 m were exploited. The stratigraphy was obtained from test holes and from the production wells (Figure 7.25).

The permeabilities are shown in Table 7.5. These vary with the depth and range in values from  $5 \times 10^{-7}$  to  $5 \times 10^{-8} \text{ cm s}^{-1}$  for the aquitard clays and from  $5 \times 10^{-3}$  to  $3.5 \times 10^{-3} \text{ cm s}^{-1}$  for the aquifers. The permeability values for the aquitards are slightly lower than the corresponding values for the Venice VE2 model, due to the larger plasticity of the clays (Figure 7.26). The range of the particle size curves for the aquifer from a depth of 230–270 m is shown in Figure 7.27. The hyperbolic model described in Section 4.4.2.1 is assumed as the constitutive relationship for the material. The material properties resulting from the laboratory tests are given in Table 7.4. The value of the constant  $R_f$  was assumed to be 0.90.

The values of some of the constants of Table 7.4 are varied in order to determine their influence upon the subsidence. The properties for the elastoplastic model are given in Section 7.3.2.3 where the rebound will also be simulated.

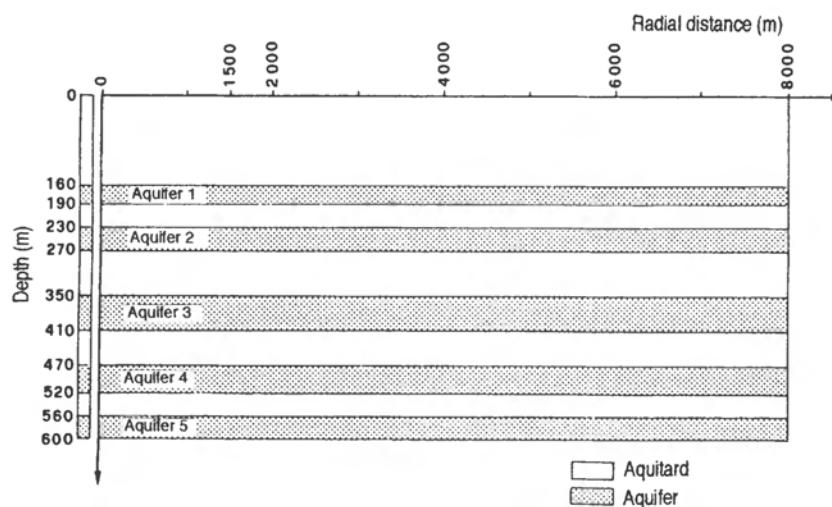
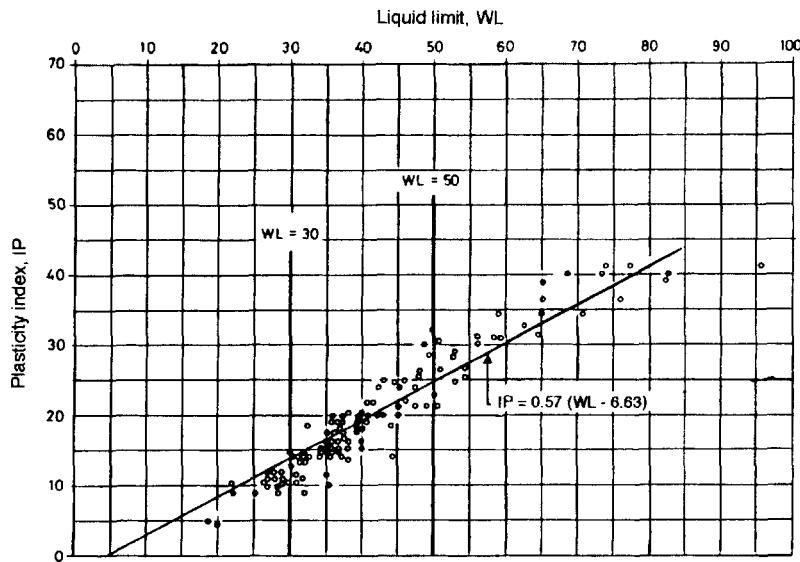
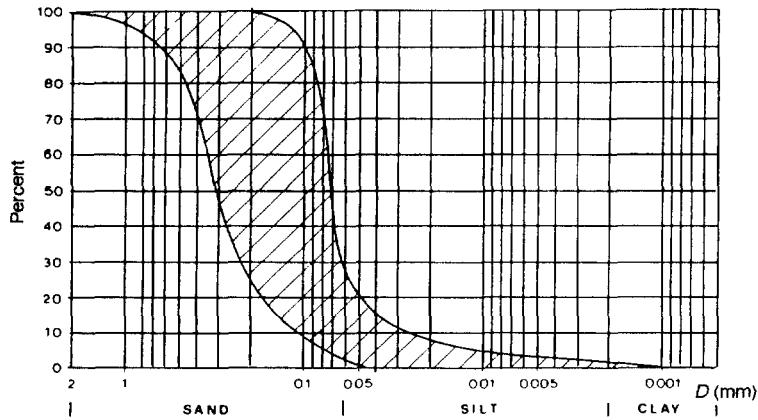


Figure 7.25 Cross-section of the confined aquifer system



**Figure 7.26** Plasticity index versus liquid limit for aquitard clays



**Figure 7.27** Granulometric curves for the second aquifer

The following pumping schedule was assumed for the parametric investigations [24]:

From 1958 to 1960	$15 \times 10^6 \text{ m}^3 \text{ per year}$
From 1961 to 1962	$7.5 \times 10^6 \text{ m}^3 \text{ per year}$
From 1963 to mid 1964	$3.75 \times 10^6 \text{ m}^3 \text{ per year}$
From mid 1964 onwards	zero

The contributions to the total flow of the five aquifers were 5, 10, 60, 15 and 10 percent respectively.

**Table 7.4** Soil layer depth, material properties (non-linear elastic constitutive law) and permeabilities for the Contarina model

	Depth(m)	<i>M</i>	<i>n</i>	<i>C</i> (Pa)	$\phi$ (deg)	<i>v</i>	Permeabilities (m per day)
1	0–160	70	0.0	0.0	27	0.45	$0.5 \times 10^{-3}$
	160–190	650	0.77	10.0	38	0.25	4.32
	190–230	10	0.90	0.0	26	0.45	$0.37 \times 10^{-3}$
2	230–270	650	0.77	10.0	38	0.25	4.00
	270–350	10	0.90	0.0	26	0.45	$0.27 \times 10^{-3}$
3	350–410	650	0.77	10.0	38	0.25	4.00
	410–470	10	0.90	0.0	26	0.45	$0.15 \times 10^{-3}$
	470–520	650	0.77	10.0	38	0.25	3.5
4	520–560	10	0.90	0.0	26	0.45	$0.5 \times 10^{-4}$
	560–600	650	0.77	10.0	38	0.25	3.0

An inspection of the stratigraphy of the region, as well as the form of the subsidence bowl shown in Figure 7.22 suggested an axisymmetric model. The radius of the central circular pumping site was chosen to be 1500 m.

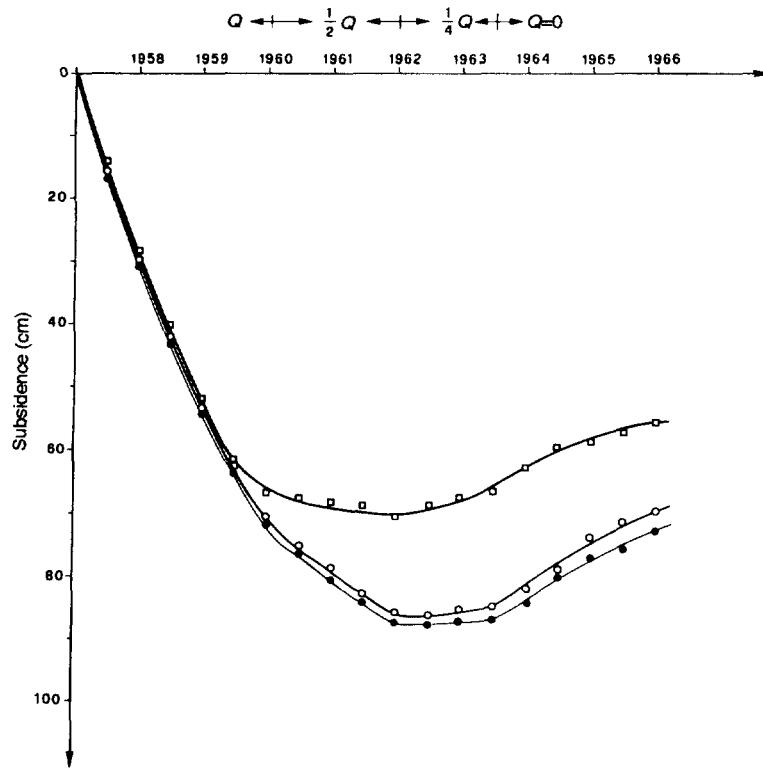
### 7.3.2.2 Parametric Investigations

The boundary conditions investigated are shown in Table 7.5. Most of the analyses were limited to the period of compaction only, as the rebound problem is investigated in detail in the next section. The time step in the present case was assumed to be 180 days. This ensured that the pore pressure oscillations were maintained within reasonable limits as well as minimising the required number of time steps for the whole period involved.

The influence of assuming a non-linear material behaviour on the surface settlement, as compared to a linear model, was investigated for boundary condition E (Table 7.5). The transient subsidence at the limit of the pumping zone ( $R = 1500$  m) is compared in Figure 7.28 for three values of the tangent modulus.

**Table 7.5** Contarina model: boundary conditions for parametric investigations

	Top limiting surface			Bottom limiting surface			Lateral limiting surface		
	Flow		Flow	Displacements		Flow		Displacements	
	Fixed pot.	Imperm.	Imperm.	Vert.	Horiz.	Fixed pot.	Imperm.	Vert.	Horiz.
A		X	X	= 0	$\neq 0$	X		$\neq 0$	= 0
B		X	X	= 0	$\neq 0$		X	$\neq 0$	= 0
C	X		X	= 0	$\neq 0$	X		$\neq 0$	= 0
D	X		X	= 0	= 0	X		$\neq 0$	= 0
E	X		X	= 0	$\neq 0$		X	$\neq 0$	= 0

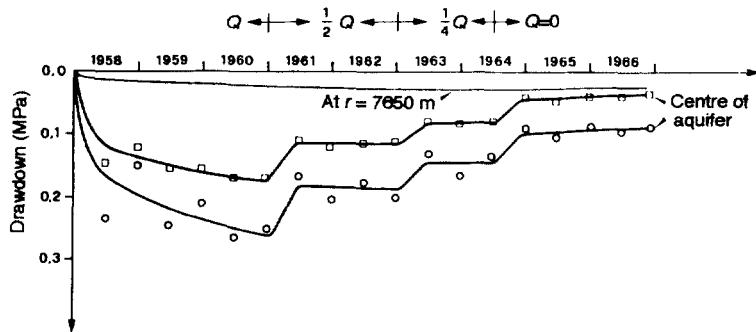


**Figure 7.28** Effects of the non-linear stress-strain constitutive relationship upon subsidence for boundary condition E at  $R = 1500$  m: (□)  $E = E_i$ ; (○)  $E = E_t$ ,  $t = 0$ ; (●)  $E = E_t$ ,  $t = \text{variable}$

1.  $E = E_i$  shows the results for the constant tangent modulus equal to its initial value, obtained through equation (4.16).
2. The curve  $E_t(t = 0)$  indicates the results for the constant tangent modulus, obtained with equation (4.17) for the initial values of the stresses.
3. Finally,  $E_t(t = \text{variable})$  shows the true non-linear behaviour, due to the variation of the stress states as consolidation proceeds.

In the case 3 the influence of the non-linearity is less than 5 percent, if no correction for rebound is carried out. Hence, if only the compaction problem is investigated, then a linear analysis with a proper choice of Young's modulus will give perfectly acceptable results. Notice that in this case a simple reduction of the outflow rate does not immediately produce a rebound of the settlement. This is true for all of the five boundary conditions investigated, irrespective of whether the system is open or closed (Figure 7.30). Only a reduction of the outflow to one-quarter of its initial value will terminate the downward trend of the settlements.

In Figure 7.29 the drawdown in the third aquifer at its centre and at a radial distance of 7650 m is shown for boundary condition E. Also given are the drawdown values at the centre, as determined by a linear elastic analysis with  $E = E_i$ . Each reduction in the



**Figure 7.29** Pore pressure decrease in the third aquifer for boundary condition E: ( $\square$ )  $E_t$ , ( $\circ$ )  $E = E_f$ ; the values at the centre of the aquifer are compared with those obtained using a particular linear elastic analysis

outflow rate produces an immediate sharp decrease in the drawdown. No such behaviour is observed for the displacements (Figure 7.28).

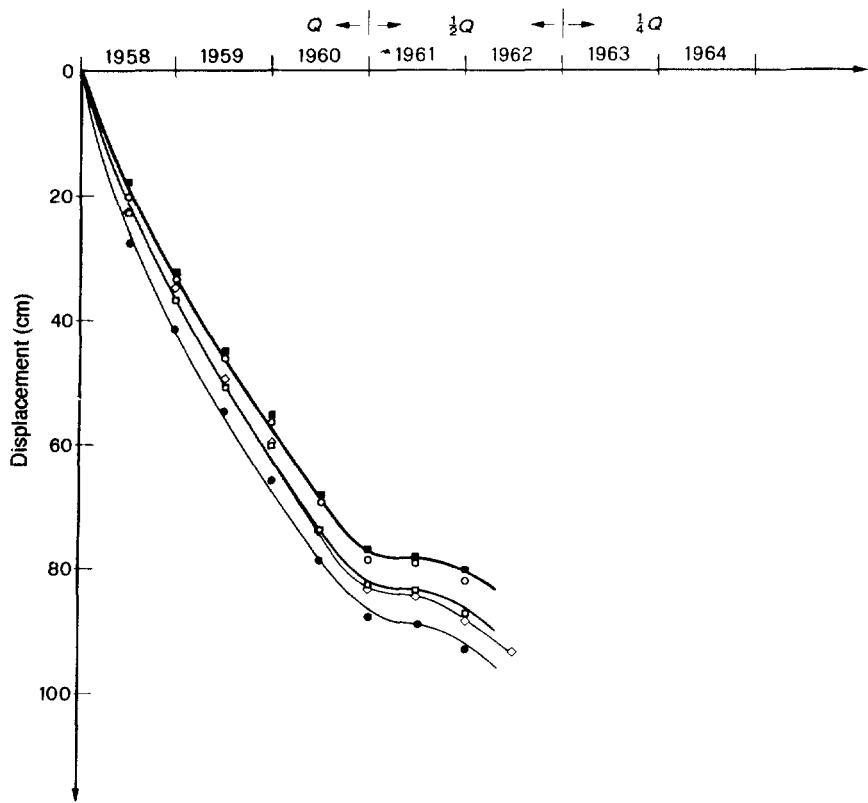
Boundary condition E represents a quasi-closed system, as only a limited quantity of water passes through the thick top layer of clay. Hence, once the right-hand term is zero, the drawdown in the whole reservoir can be observed to level off.

Figures 7.28 and 7.29 also show the influence of Young's modulus on the subsidence and drawdown values. For a given outflow rate, a stiffer material requires a much larger drawdown but a lower subsidence value.

The influence of the boundary conditions on the central surface subsidence during compaction is shown in Figure 7.30. The largest settlement is obviously obtained with boundary condition B, which represents a system with no possibility of recharge. Note that the constraint on the horizontal displacements at the bottom has no influence, as indicated by the results for boundary conditions C and D. The scenario of the boundary condition influence is completed with Figure 7.31, where a section is shown through the subsidence bowl at the end of 1961. A fixed potential at the top surface produces a slight swelling at the centre, which occurs only in the top clay layer. This local swelling is due to the way the nodal flux is imposed (Section 7.2.2). A fixed potential at the top surface permits a small influx of water, even though a clay layer exists above the first aquifer. Hence not all of the water withdrawn from the first aquifer is supplied by this aquifer and the confining aquitard, therefore a smaller drawdown is required at the centre. The maximum value of subsidence is nevertheless equal for all cases where the same flow condition exists at the outer border.

The surface horizontal displacements in 1961 for boundary conditions A and B are shown in Figure 7.32. The difference in shape is due to the flow condition at the outer boundary.

The influence of the clay friction angle on the subsidence values was also investigated for boundary condition D. The results obtained for  $\phi = 26^\circ$  are compared in Figure 7.33 with those for  $\phi = 22^\circ$ . The total subsidence, as well as the values for the sand and clay layers, can clearly be seen. The total thicknesses for the sand and clay layers are 220 m and 350 m respectively. In this case the contribution of the sand layers to the total subsidence is almost negligible when compared with the aquitards. The two friction



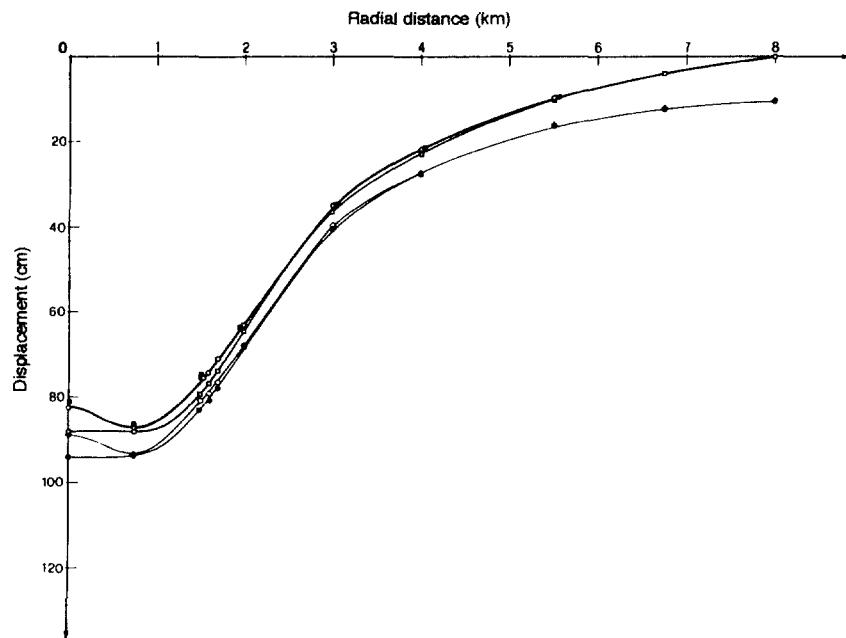
**Figure 7.30** Surface subsidence versus time at  $R = 1500$  m: (○) boundary condition A; (●) boundary condition B; (◇) boundary condition C; (■) boundary condition D; (□) boundary condition E

angle values produce not only a different magnitude of the subsidence, but also an intersection of the two profiles initially. During the first few time steps, the drawdown in those aquifers confined by a stiffer material is greater than in the aquifers confined by a softer material. This produces a larger settlement. Later, this difference in drawdown diminishes, while the softer material undergoes a larger deformation.

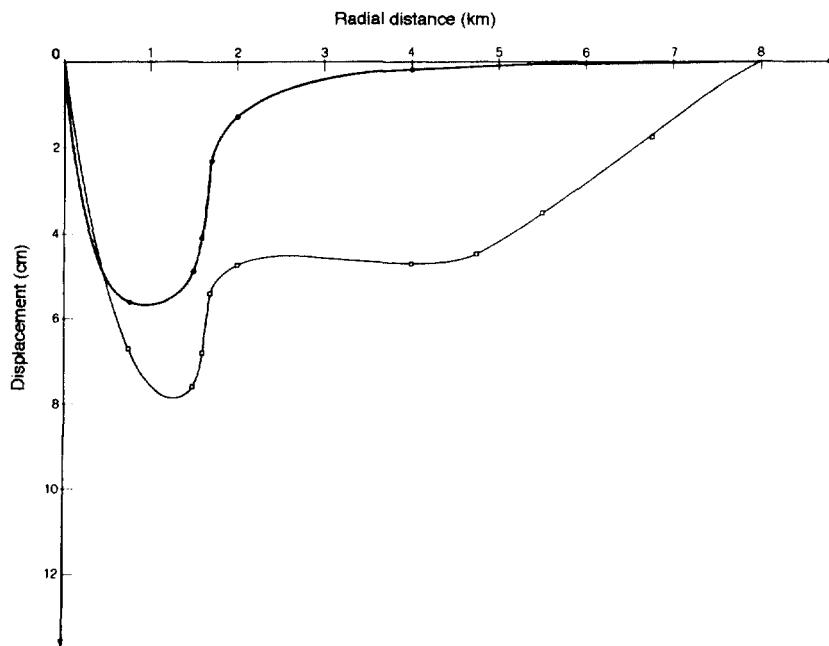
The different shapes of the subsidence bowl in 1963 obtained with the two values of the friction angle are shown in Figure 7.34.

In all cases investigated the temporal variation of the effective stresses was very small. The initial effective stresses (solid line) are compared in Figure 7.35 with the effective stresses (dots) at the Gauss points close to the centre of the model at the instant of the largest drawdown (1960) for boundary condition E; see Figure 7.29. The largest increment (equal to 4 percent), occurred the third aquifer, which produced 60 percent of the total outflow. The increment in the aquitard between the third and fourth aquifers was 2 percent, whereas all other variations were less than or equal to 1 percent.

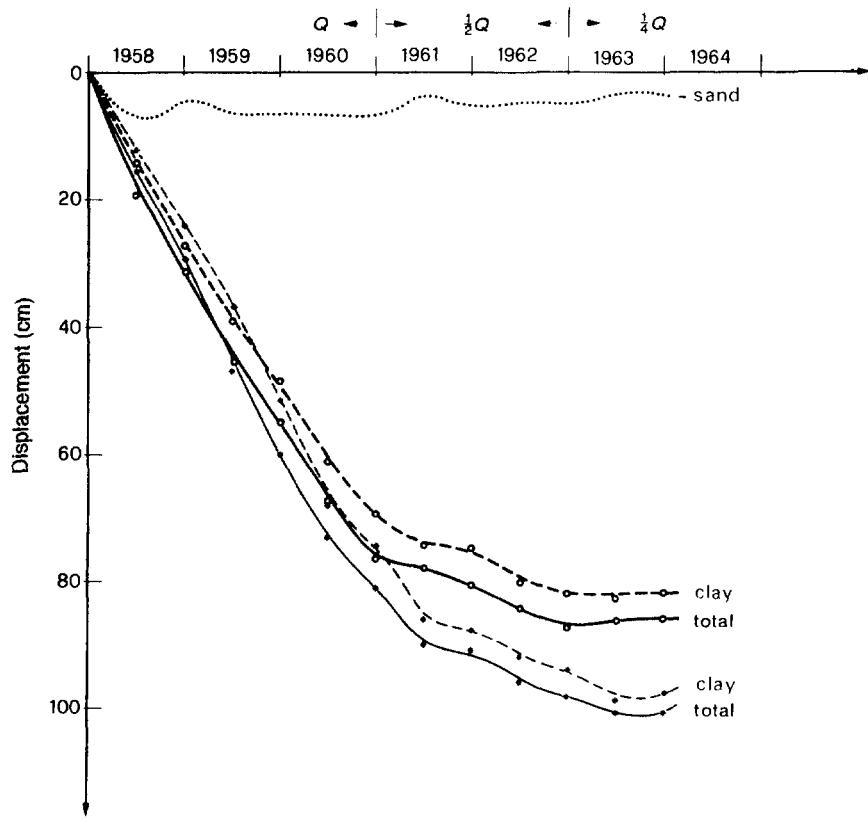
The drawdown along the centreline at the same instant is shown in Figure 7.36. The variations of the effective stresses are of the same order as for the drawdown, hence the



**Figure 7.31** Surface subsidence in 1961 versus radial distance: (○) boundary condition A; (●) boundary condition B; (△) boundary condition C; (■) boundary condition D; (◇) boundary condition E



**Figure 7.32** Horizontal surface displacements in 1961 versus radial distance: (●) boundary condition A; (○) boundary condition B



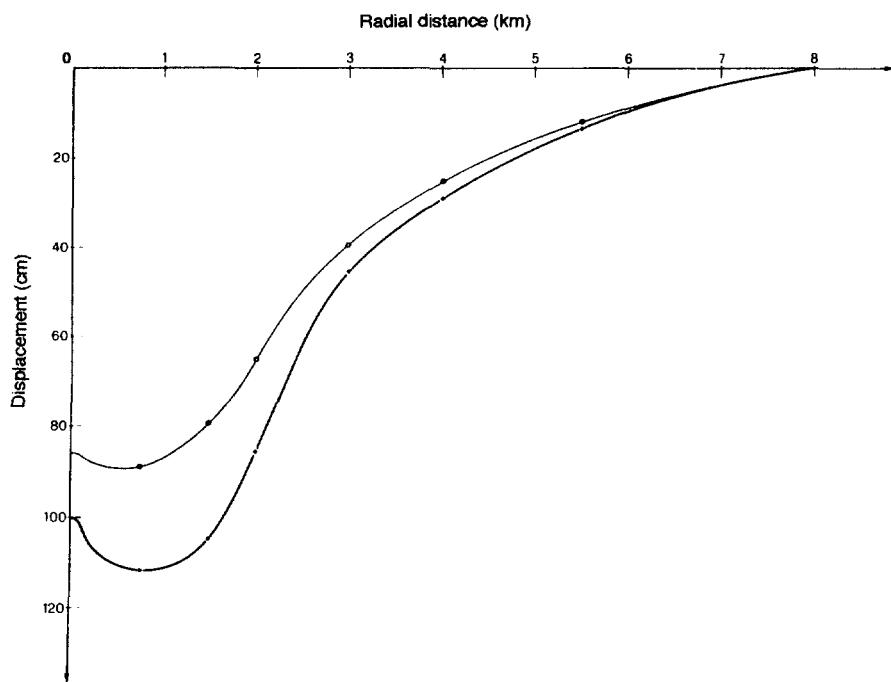
**Figure 7.33** Settlements close to the centre of the model versus time for boundary condition D:  
 (○)  $\phi_S = 38^\circ$  for sand,  $\phi_A = 26^\circ$  for clay; (●)  $\phi_S = 38^\circ$  for sand,  $\phi_A = 22^\circ$  for clay

condition  $\partial\sigma_{zz}/\partial t = 0$ , which is required for the uncoupling of the flow and equilibrium equations (Section 6.1.1) can be regarded as being reasonably well satisfied in subsidence problems. However, the condition regarding the horizontal displacements is not satisfied, as can easily be seen in the Figure 7.32.

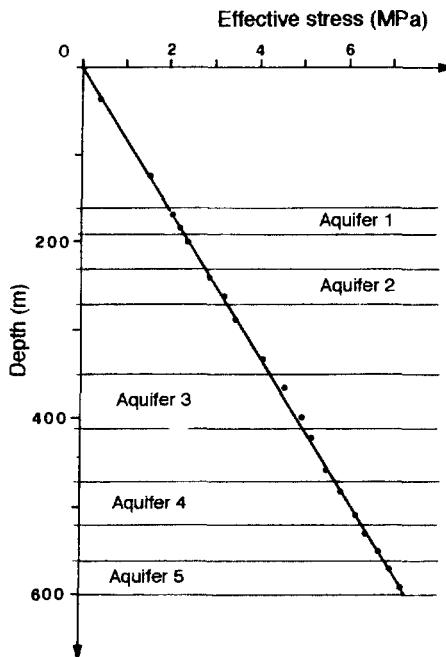
The hyperbolic stress-strain curve for the aquitard between the third and fourth aquifers is given in Figure 7.37. This is the aquitard which has the greatest variation in stress.

Figures 7.38 shows the same curve for the sand of the third aquifer. In this case the behaviour can be assumed linearly elastic.

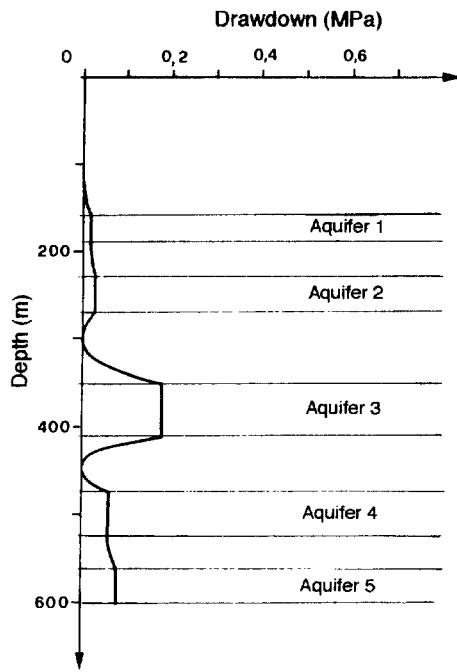
The transient behaviour of a system composed of a fictitious top layer, and the second and third aquifers, along with the aquitard between them, was studied with the refined mesh shown in Figure 7.39. This was carried out to check if the rather coarse mesh used in all the subsidence models gave adequate results, particularly for the aquitards. The boundary conditions were of type A and the assumed pumping schedule was the same as that used for the second and third aquifers in the complete model. The material properties for the top layer, the two aquifers and the aquitard are shown in Table 7.4 and



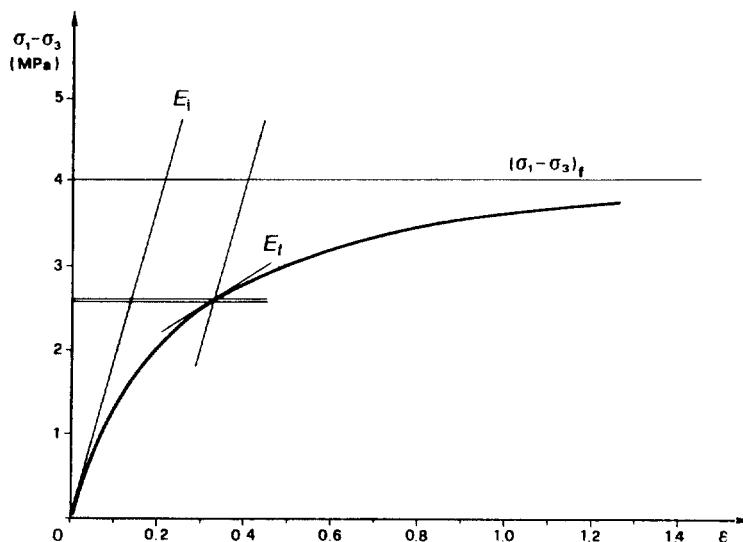
**Figure 7.34** Surface subsidence in 1961 versus radial distance for boundary condition D: (---)  $\phi_S = 38^\circ$  for sand,  $\phi_A = 26^\circ$  for clay; (●)  $\phi_S = 38^\circ$  for sand,  $\phi_A = 22^\circ$  for clay



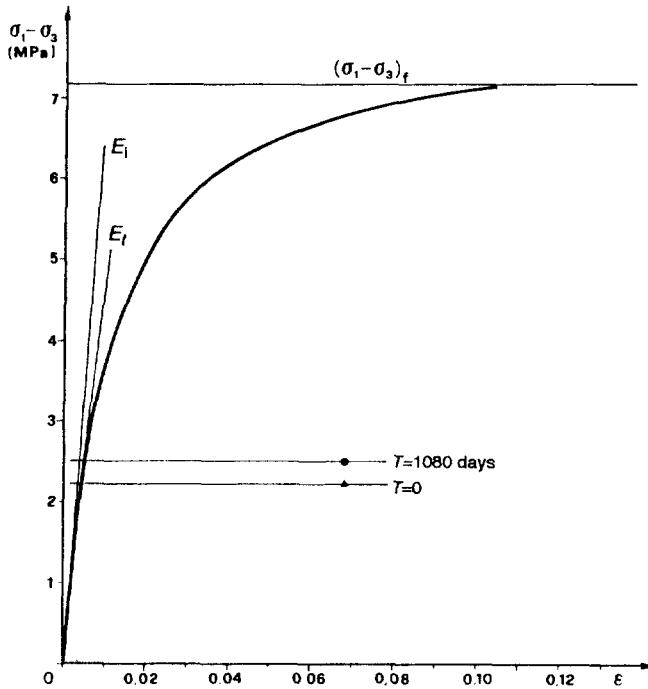
**Figure 7.35** Initial effective stresses (—) and effective stresses (●) in 1960 close to the centre of the model



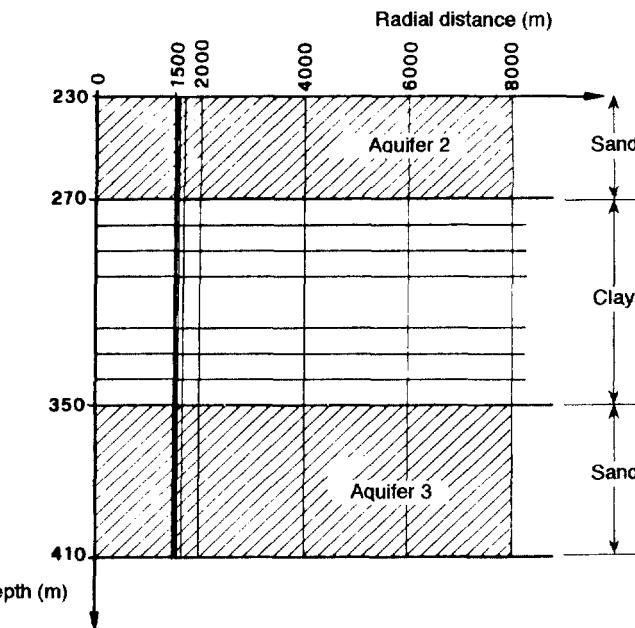
**Figure 7.36** Pore pressure decrease in 1960 along the centreline of the model



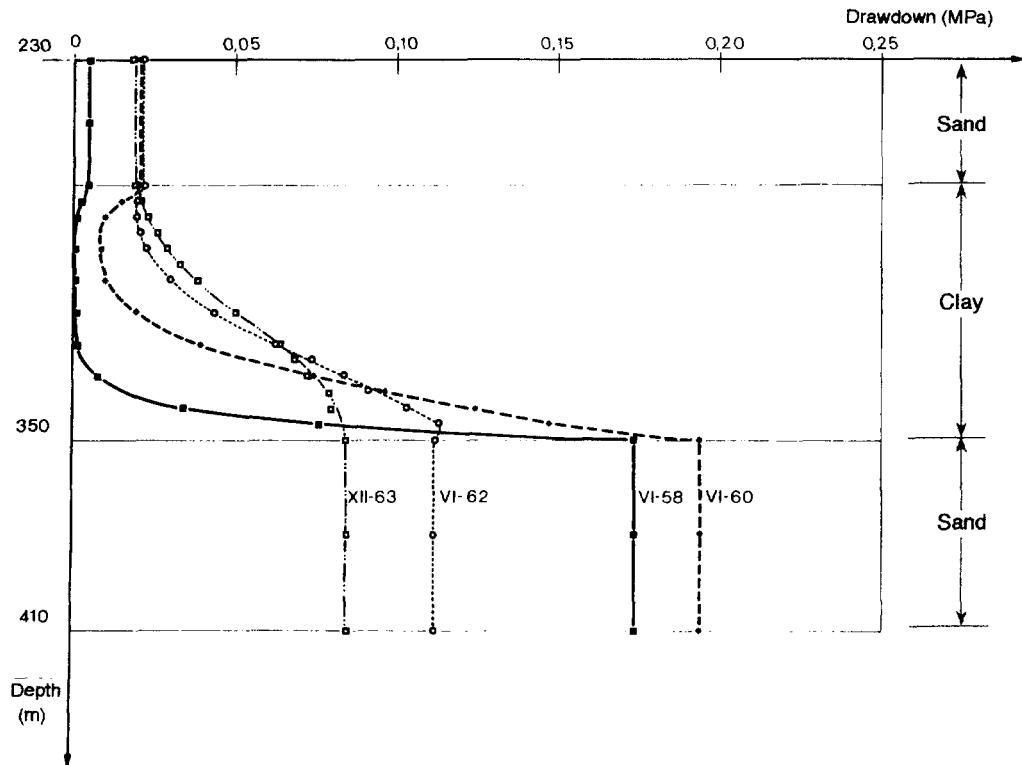
**Figure 7.37** Hyperbolic stress-strain curve for clay between the third and fourth aquifers:  
 $E_i = 18.72 \text{ MPa}$ ,  $E_r = 2.58 \text{ MPa}$ ,  $\Delta(\sigma_1 - \sigma_2) = 0.73 \text{ percent}$



**Figure 7.38** Hyperbolic stress-strain curve for the sand of the third aquifer: (●)  $E_i = 316.03$  MPa, (▲)  $E_r = 367.70$  MPa;  $E_f = 707.09$  MPa,  $\Delta(\sigma_1 - \sigma_2) = 13.1$  percent



**Figure 7.39** Refined finite element discretisation between the second and third aquifers



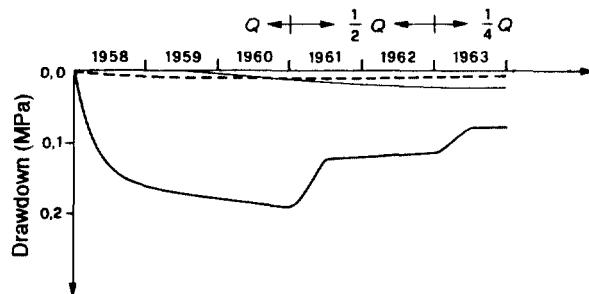
**Figure 7.40** Pore pressure decrease at the centreline for different times

the drawdown values at different times are given in Figure 7.40. The shape shown in Figure 7.36, where only one element is used per aquitard and aquifer, compares well with that obtained using the refined solution.

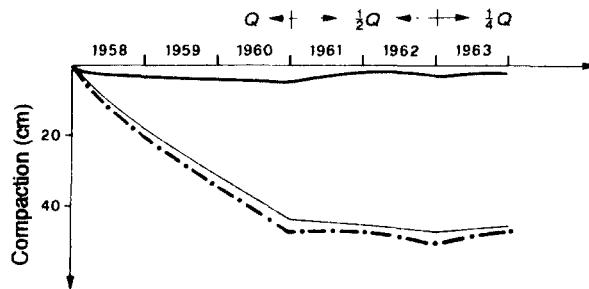
The results shown in Figure 7.36 seem to favour a parabolic pore pressure interpolation as against the suggestions of several authors, e.g. Sandhu *et al.* [25], of using a lower-order interpolation than was used for the displacements. A full discussion on the interpolations to be used in consolidation problems can be found in the literature [26–28]. The present case involves a far from incompressible or nearly incompressible material, hence it is possible to use the same interpolation for both variables (Chapter 3), which results in obvious advantages. A linear interpolation of the pore pressure would require at least three elements along each aquitard thickness, which would render the simulation much more expensive.

The drawdown versus time in the second and third aquifers and in the centre of the aquitard for the refined model is shown in Figure 7.41. The drawdown in the third aquifer is greater than obtained with the complete model, as the reduced model ignores the contribution of the lower confining aquitard.

The compaction behaviour versus time for the two aquifers and the intervening aquitard is shown in Figure 7.42. Notice that the contribution of the aquifers is again low in comparison with the aquitards. These simulations were also carried out with



**Figure 7.41** Pore pressure decrease versus time: (- - -) in the second aquifer, (—) in the third aquifer, (—) in the centre of the aquitard

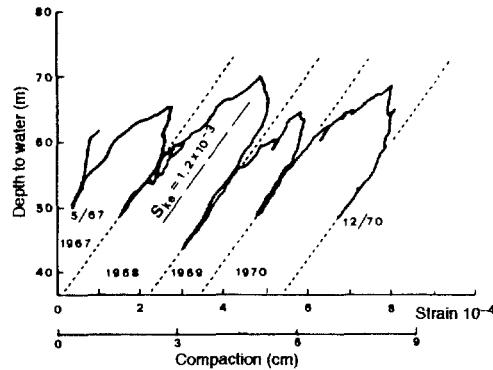


**Figure 7.42** Compaction versus time: (—) in the second and third aquifers, (—) in the aquitard, (- · -) total for the second and third aquifers

anisotropic permeability values for the aquitards. The same results as above were obtained by increasing the horizontal permeability by a factor of 10. A similar reduction of the vertical permeability produced completely different results, e.g. the vertical displacements were reduced by 66 percent. The model is therefore sensitive to the vertical and not the horizontal permeability values of the aquitards. Simplified models for flow in porous media consider in fact only vertical flow in the aquitards and horizontal flow in the aquifers [29].

### 7.3.2.3 Subsidence Rebound

The linear and non-linear elastic models are unable to simulate the expanding aquifer system when the withdrawal rate is significantly reduced, unless a correction is made. This occurs because the problem is essentially hysteretic, i.e. the material behaviour for the solid and the liquid depends on the sign of the forcing functions. The assumption of an elastoplastic model for the solid, even if the effective stress variations involved are usually very limited, produces acceptable answers. The applicability of such an approach is quite evident by inspecting the stress-strain curves in Figure 7.43 [30]. These curves were obtained from field measurements of sediment compaction and expansion and the correlative changes in fluid pressure. The change in water depth represents a change in the effective stresses within the aquifers of the confined system. It



**Figure 7.43** Stress change, compaction and strain for a well in Western Fresno County (Reprinted, by permission, from Poland, J.F., 1981, Subsidence in United States due to groundwater withdrawal, *J. Irrig. Drain. Div. ASCE*, **107**, 115–35)

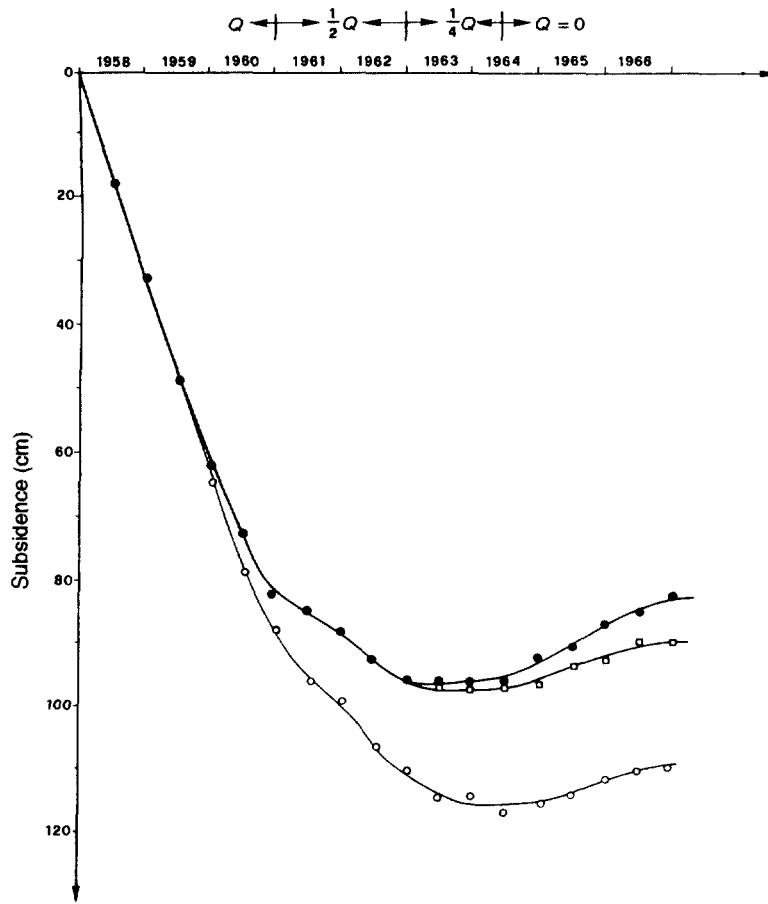
appears in this case that the response is essentially elastic in the aquifers and aquitards when the depth to water is less than about 55 m. For an effective stress increase in the loading, which exceeds the preconsolidation stress, the virgin compaction of the aquitards is chiefly inelastic and irrecoverable upon a decrease in stress [30]. In this section rebound is simulated in two ways:

1. A hyperbolic stress–strain law together with a restart procedure allows the replacement of the current tangent stiffness modulus by the modulus for unloading–reloading; this procedure applies some ingredients of generalised plasticity and was devised [31] before generalised plasticity had been introduced [32].
2. An elastoplastic solution (classical plasticity).

In the first case the change of sign of the variation of the stress invariants with time may be used as a signal for unloading [28], since in subsidence problems these invariants move approximately up and down along a line. However, the actual value of the changes between any two time steps is usually much less than 1 percent; hence this criterion is not sensitive enough from an operational viewpoint.

It is preferable to obtain a better insight by inspection of the cross-sectional behaviour rather than use the model merely as a ‘black box’. The settlement history obtained from a trial run, using the compaction tangent modulus, gives a sufficiently accurate indication when the modulus for unloading–reloading should be applied. This is clearly seen in Figure 7.28, where a change in modulus seems appropriate the end of 1962, when the withdrawal rate was reduced to one-quarter of its initial value. This was also indicated by the stress invariants. An inspection of Figures 7.29 and 7.41 excludes drawdown as a reliable indicator. The central subsidence obtained from the non-linear elastic model, with and without the repeat solution, is shown in Figure 7.44. In this case the initial modulus was taken as the modulus for unloading–reloading since no other data was available.

An elastoplastic analysis, based on classical elastoplasticity, avoids the need to repeat the solution. But due to the modelling effort, this is no panacea for subsidence problems



**Figure 7.44** Subsidence versus time: (●) variable elastic, (○) variable elastic with correction for rebound, (◐) elastoplastic

[31]. The small variations of stress state in the aquitards require that the parameters have to be chosen in such a way that all stress points must lie on the yield surface from the onset and for the complete solution up to the rebound analysis. Otherwise it is practically impossible to simulate elastoplastic behaviour.

A similar test case was analysed with an elastoplastic model. A Mohr–Coulomb model was used for the sand layers and the modified Cam clay model for the clay aquitards. A time step of 180 days was assumed as in the previous non-linear models.

The parameters for the critical state model chosen for the clays of the aquitards are

$$M = 1.05$$

$$\chi = 20$$

$$p_{c0} = 0.48, 1.26, 1.86, 2.64 \text{ and } 3.24 \text{ MPa}$$

for the respective aquitard depths. All the other parameters have the same values as used

in the variable elastic model, comprising the Young's modulus for elasticity, which was taken as being equal to the initial tangent modulus  $E_i$  and the parameters necessary for the Mohr–Coulomb model for the aquifers [31].

With this choice the sand of the aquifers behaved elastically during the whole simulation. All aquitards, with the exception of the top one, were assumed to be in the plastic range. This was ensured numerically by the choice of preconsolidation pressure. The resulting surface settlements are shown in Figure 7.44. The results show the same behaviour during rebound as in the repeat solution of the non-linear elastic model.

The difference between the two models lies in the differing modes of compaction modelling. In the non-linear elastic model this is governed by the current tangent modulus [33]; in the elastoplastic model the compaction of the aquifers, which are elastic, is negligible. The compaction is therefore governed by the strain-hardening law of the aquitards (Figure 4.11). The mechanisms are different, hence the different response.

A straightforward use of generalised plasticity (Section 4.5.5), which already incorporates the criterion for unloading, now seems to be the most appropriate approach for handling the rebound problem.

## 7.4 SUBSIDENCE ABOVE GAS RESERVOIRS: THE RAVENNA CASE

### 7.4.1 Background

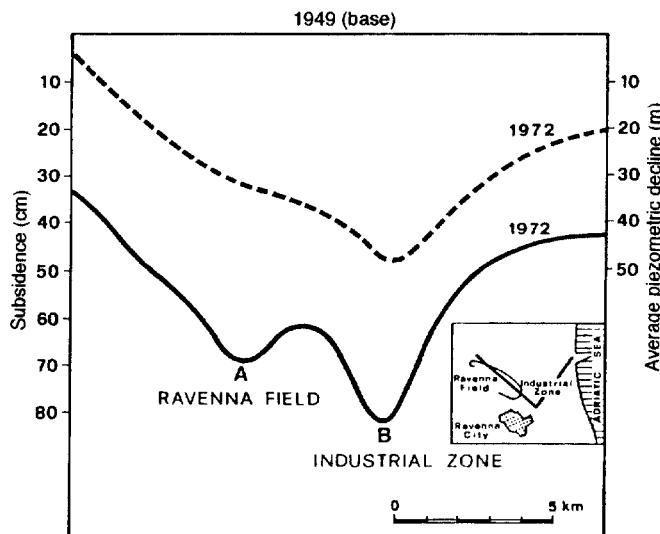
The subsidence of the Ravenna region is caused mainly by water pumpage in the upper part of the underlying sediments for industrial and agricultural purposes. However, the removal of gas from a local gas reservoir cannot be ignored. A possible correlation between the intensive gas production and the surface subsidence was first shown by Zambon [34].

Heavy water pumpage started in the region at the beginning of the 1950s. The maximum recorded head drawdown in the aquifers of the upper layers is more than 40 m and the total subsidence since 1949 is almost 1 m. A large part of the region is less than 1 m above mean sea level, so it is clear that disastrous consequences could occur. Important historical monuments, port structures and existing geographical boundaries may be affected by the land settlement.

Many similarities with the Venice case become evident on inspection of the physical environment and geological history. A comparison between the subsidence and fluid withdrawal history of Ravenna and Venice was made by Carbognin *et al.* [35]. Their work gives a broad description of the available field data and the existing stratigraphy of the upper layers.

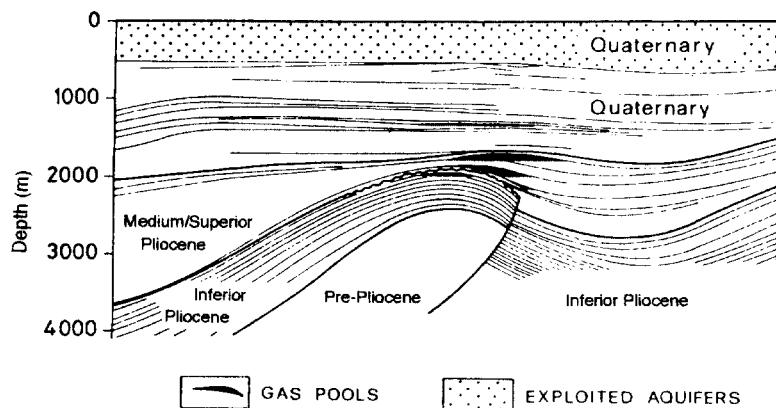
The events in the Ravenna region are more intricate due to the fact that one gas field, the Ravenna field, was exploited onshore, and offshore gasfields are still being exploited, which might possibly affect the coastal area. A local increase of the land settlement above the Ravenna field has in fact been monitored. Figure 7.45 [35] shows a secondary minimum (A) of the subsidence located above the Ravenna field. No such minimum exists at this point for the piezometric level.

The solution presented in Section 6.4 applies to reservoirs of limited lateral extension when compared to the depth of burial. A typical case of this kind is the Ravenna field,



**Figure 7.45** Ravenna region: (—) observed subsidence and (---) average piezometric decline (Reprinted, by permission, from Carbognin, L., Gatto, P., Mozzi, G. and Gambolati, G., 1978, Land subsidence in Ravenna and its similarities with the Venice case, in *Evaluation and Prediction of Land Subsidence*, S.K. Saxena (ed.), ASCE, New York, pp. 254–66)

where the production zone covers seven intervals distributed over a vertical section from 1270 m to 2048 m below the surface. The most important reservoir in this field is located between 1678 m and 1803 m, has a maximum formation thickness of 65 m and covers an area of 15.6 km<sup>2</sup>. The porosity of the reservoir sand is 26.6 percent, the mean value of the permeability is 168 mD, the temperature is 38 °C and the initial pressure is 19.7 MPa.



**Figure 7.46** Cross-section with gas pools and aquifers exploited at different levels (Reprinted, by permission, from Evangelisti, G. and Poggi, B., 1970, Sopra I fenomeni di deformazione dei terreni da variazione della pressione di strato, in *Proc. Acc. delle Scienze dell'Istituto di Bologna*, Series 2, 6)

A cross-section through the Ravenna field is shown in Figure 7.46 [36]. The exploited aquifers are located in the upper part of the quaternary sediments (dotted area) and the gas pools are shown in black. It is difficult to obtain information regarding the stratigraphy, material properties of the deeper soil layers, gas production, and the drop of the reservoir pressure for successful modelling of the area.

It has already been stated [37] that in the Ravenna case the exploitation of deep gas reservoirs does not produce any significant surface settlement. This statement was based on calculations made by Evangelisti and Poggi [38], who used an uncoupled poroelastic model. They applied the material balance equation in the reservoir and the uncoupled one-phase flow equation in the adjacent aquifer for the determination of the water inflow in the reservoir, thus neglecting the horizontal displacements. The contribution to fluid accumulation of the rate of change of the total strain only was taken into account and the fluid flow equation was solved by direct integration for a geometrically simple case. The deformation analysis was then carried out in a homogeneous linearly elastic half-space using the poroelastic theory. Their results are not in accordance with the field observations shown in Figure 7.45. Hence a hypothetical axisymmetric cross-section with a waterdrive gas reservoir at 1800 m below the surface was studied using the solution procedure outlined in Section 6.4. The example is similar to that investigated by Evangelisti and Poggi [38], and has the same reservoir, aquifer flow and deformation properties, so that useful comparisons can be made between the two techniques. However, in this case the overburden was not assumed homogeneous, but this would have no influence on the flow regime as solved by Evangelisti and Poggi.

The reservoir thickness is approximately 80 m, and the radius  $R_w$  was assumed to vary between 1000 m and 2000 m. The adjacent aquifer was assumed to have the same thickness as the reservoir itself. Both reservoir and aquifer are overlaid by a clay layer of 80 m thickness with the remaining depth comprising sand. In the work of Evangelisti and Poggi [38] the overburden was sandstone and in the application of the poroelastic theory the whole cross-section, comprising reservoir and aquifer, was assumed to be homogeneous with respect to the material properties.

The base of the model system was assumed impermeable with no vertical displacements. The horizontal constraints at the base were assumed free in one case and then as being fixed. Both assumptions gave the same overall behaviour, which is in accordance with the observations of Section 7.3.2.2. The top of the model was taken to be impermeable with no constraints for the displacements. At the outer boundary, horizontal displacements were fixed, vertical displacements were free and a fixed flow potential was assumed.

The initial reservoir pressure was 19.7 MPa and the volume of the gas initially contained in the reservoir was  $13 \times 10^9 \text{ m}^3$  at a temperature of 40 °C [38]. The modulus of elasticity for the sand,  $E = 990 \text{ MPa}$ , was taken from the work of Evangelisti and Poggi [38], and the modulus of elasticity for the clay,  $E = 392 \text{ MPa}$ , was obtained from samples of a deep borehole in Venice [13]. Poisson's ratio was assumed to be 0.20 for the sand and 0.45 for the clay [13]. The permeability for the clay was  $8.65 \times 10^{-5} \text{ m per day}$  and the permeability for the sand on top of the model was 4.32 m per day. Three values were investigated for the aquifer that limited the extent of the reservoir, 0.0471, 0.7358 and 1.8496 m per day (respectively 12.8, 200 and 500 mD).

In general, the values assumed in the parametric study for the geometrical parameters, temperature and initial pressure are within the range of those observed in the Ravenna

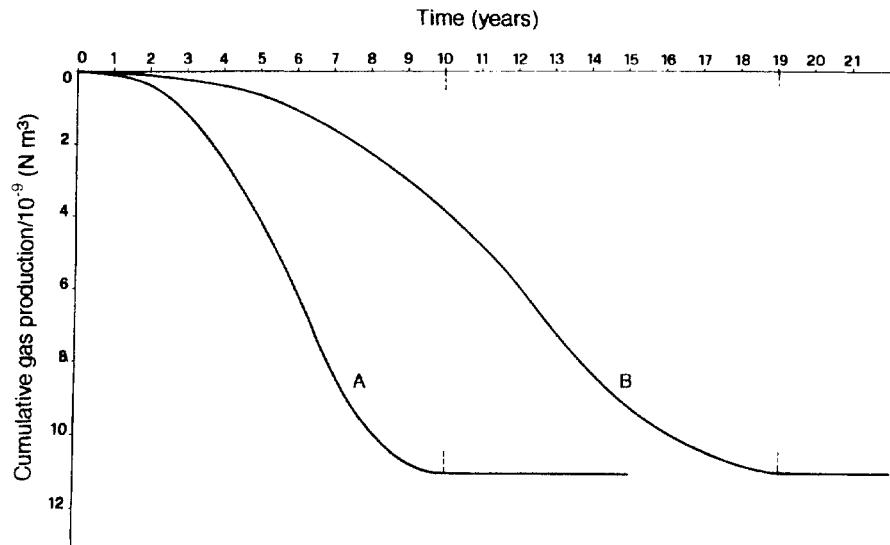


Figure 7.47 Production histories assumed for the examples

field, whereas the stress–strain properties were obtained from a deep borehole in Venice. Both areas have a similar geological history and physical environment.

Two different production histories were taken into account, as shown in Figure 7.47. For graph A the time of exploitation was 10 years and in graph B it was 19 years. In both instances the cumulative gas production was  $11.076 \times 10^9 \text{ N m}^3$ .

Extensive parametric investigations were carried out with this test case, especially the importance of various parameters such as time of exploitation, lateral extent of the reservoir, aquifer permeability and the influence of horizontal displacements in the flow equations. The results show that the exploitation of a deep gas reservoir can influence surface settlements and further investigations should be made into the problem. Finally, a comparison was made between the behaviour of a waterdrive and a volumetric gas reservoir.

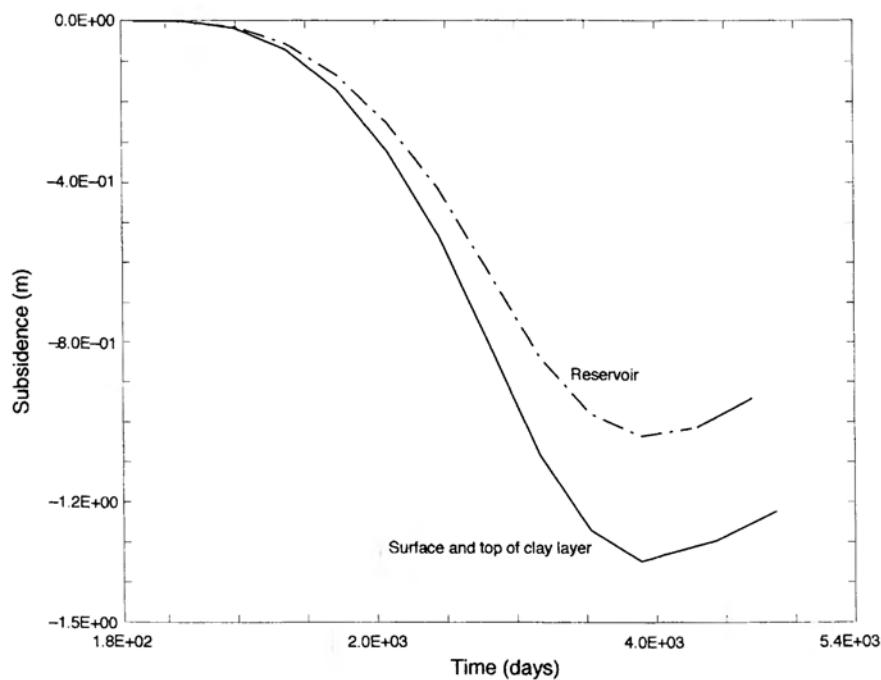
Note that the numerical simulator takes into account both the compaction of the reservoir itself and the possible consolidation of the material above and/or below the reservoir. This is often neglected in other solutions quoted in the literature.

#### 7.4.2 Results

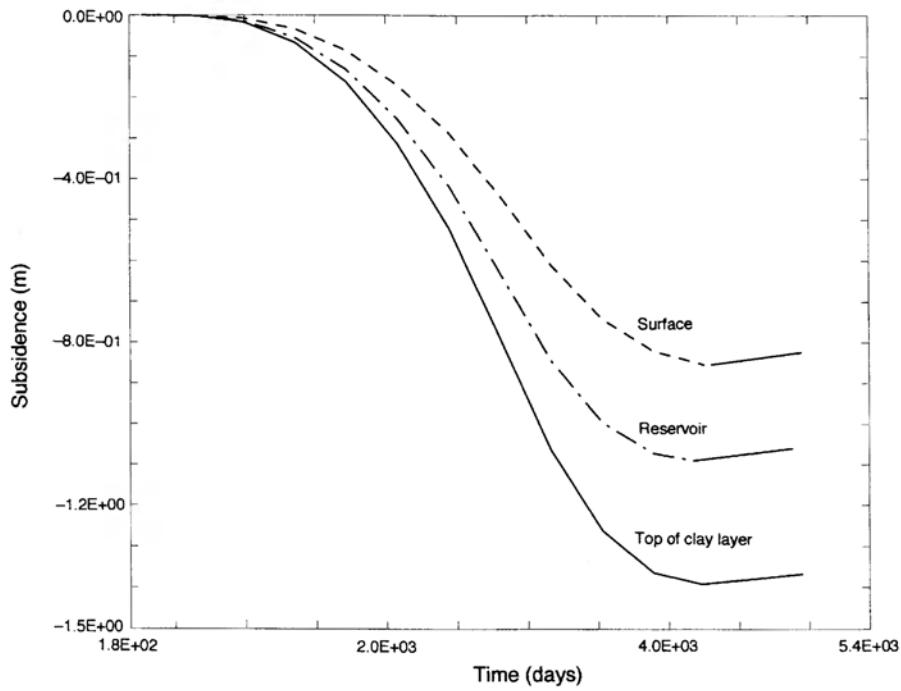
The application of the computational scheme described in Section 6.4 required a time-step length of six months to prevent divergence of the solution.

First, the 10 year production history was investigated with different values of reservoir radius and permeability [39,40].

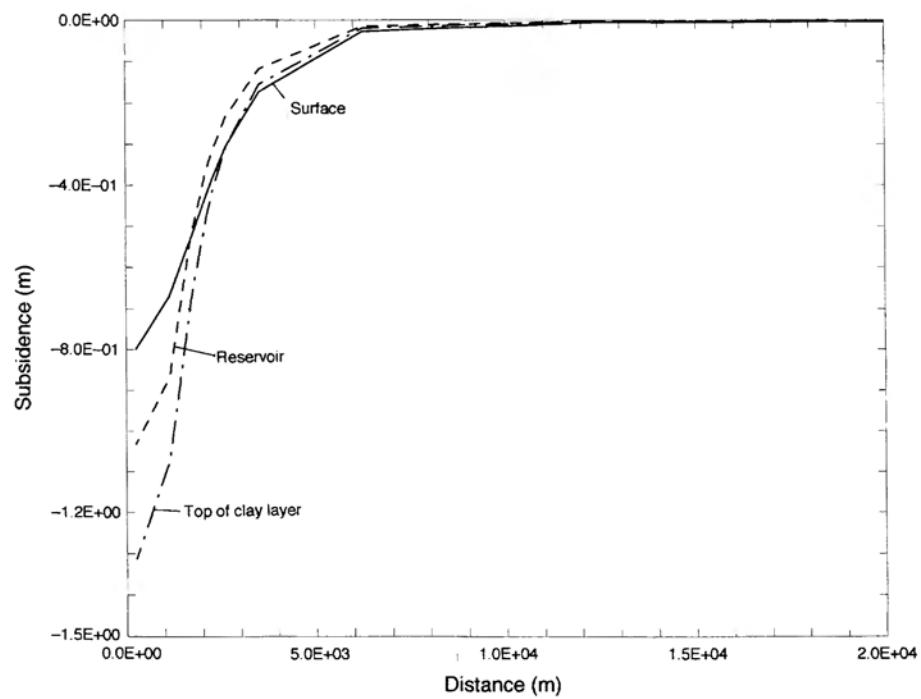
The transient settlement of the top of the reservoir, at the surface and the top of the clay layer is shown in Figure 7.48 for  $R_w = 2000 \text{ m}$  and  $k = 0.0471 \text{ m per day}$ . The reservoir compaction accounts for only about half the surface subsidence; the rest of



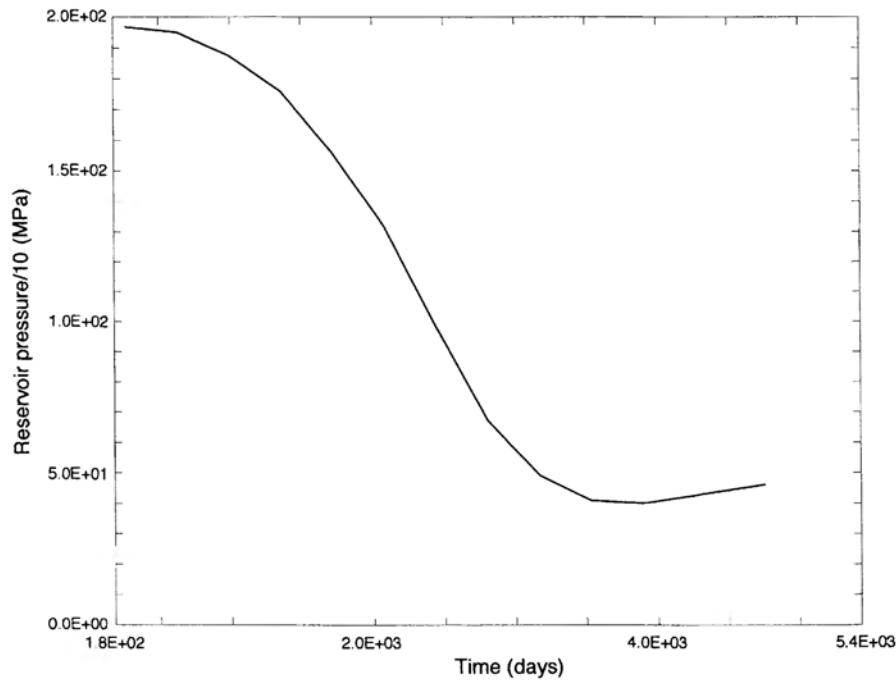
**Figure 7.48** Settlement versus time for  $R_w = 2000$  m,  $k = 12.8$  mD



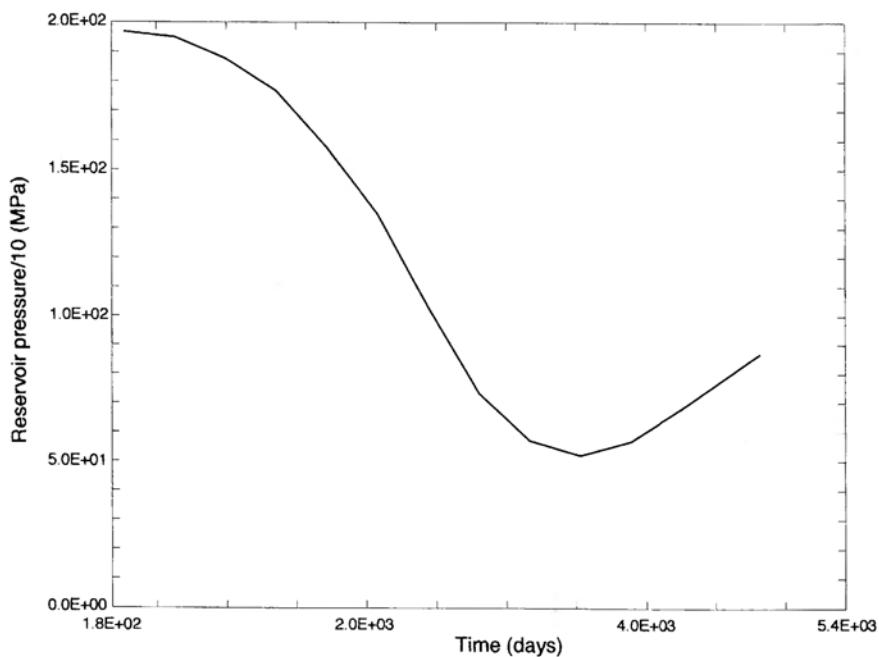
**Figure 7.49** Settlement versus time for  $R_w = 1000$  m,  $k = 12.8$  mD



**Figure 7.50** Settlement versus radial distance at shutdown of the wells:  $R_w = 1000$  m,  $k = 12.8$  mD



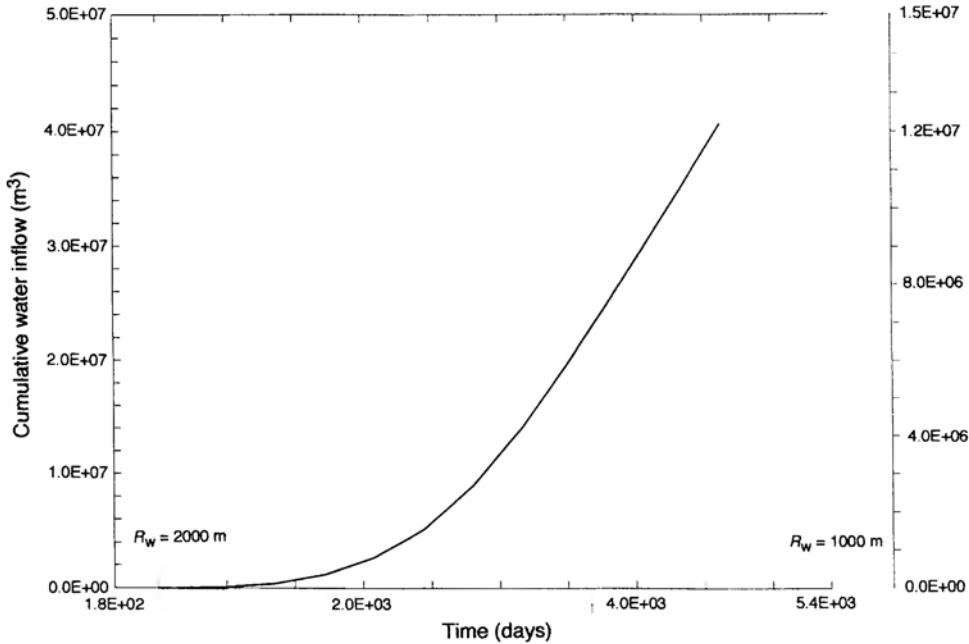
**Figure 7.51** Reservoir pressure versus time for  $R_w = 1000$  m,  $k = 12.8$  mD



**Figure 7.52** Reservoir pressure versus time for  $R_w = 2000$  m,  $k = 12.8$  mD

the surface settlement is due to the clay layer. For comparison the same quantities are shown in Figure 7.49, assuming  $R_w = 1000$  m. The reservoir compaction and the settlement of the top of the clay layer exceed the values of the previous case but the huge sand layer at the top of the model is slightly expanding. The recovery of the deformation after the end of the exploitation is substantially slower than in the previous case. Figure 7.50 shows the settlement versus the radial distance at the shutdown of the wells for  $R_w = 1000$  m. The variation of the reservoir pressure for  $R_w = 1000$  m and  $R_w = 2000$  m is shown versus time in Figures 7.51 and 7.52. The recovery of the pressure after the closedown of the wells due to the encroaching water can be clearly seen. Recovery is faster for a reservoir with a larger radius. This is consistent with the higher water inflow shown in Figure 7.53. Note that the water inflow versus time has a similar shape for  $R_w = 1000$  m and  $R_w = 2000$  m.

Figure 7.54 shows the variation of excess pore pressure with depth at the end of the pumpage for two values of the radial distance. The excess pore pressure disappears above the clay layer. Figure 7.55 shows the variation of excess pore pressure with radial distance in the pool and the aquifer after 10 years of pumpage. The influence of the exploitation vanishes at a distance of 6 km from the centre of the model. The subsidence versus time for different values of the aquifer permeability is shown in Figure 7.56. The increase of permeability reduces the maximum observed subsidence because the major quantity of the penetrating edge water causes the drop in the reservoir pressure to be smaller. When the permeability increases, the peak value of subsidence is reached earlier. The radial extension of the resulting subsidence bowl increases with the permeability because the flow in the aquifer affects zones with greater radial distance.



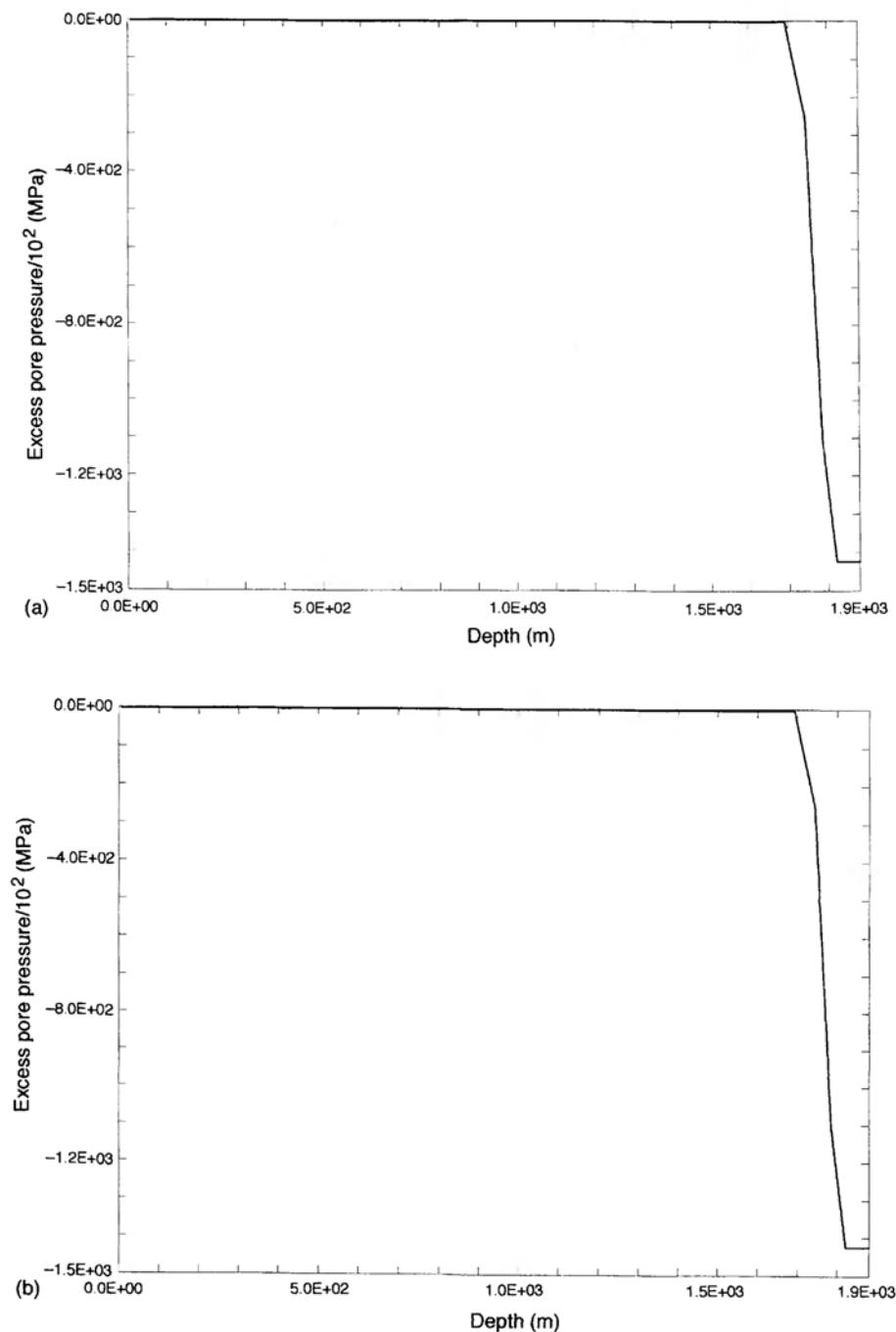
**Figure 7.53** Cumulative water inflow for  $R_w = 1000\text{ m}$  and  $R_w = 2000\text{ m}$ ,  $k = 12.8\text{ mD}$

The influence of the formation thickness upon the reservoir pressure is indicated in Figure 7.57.

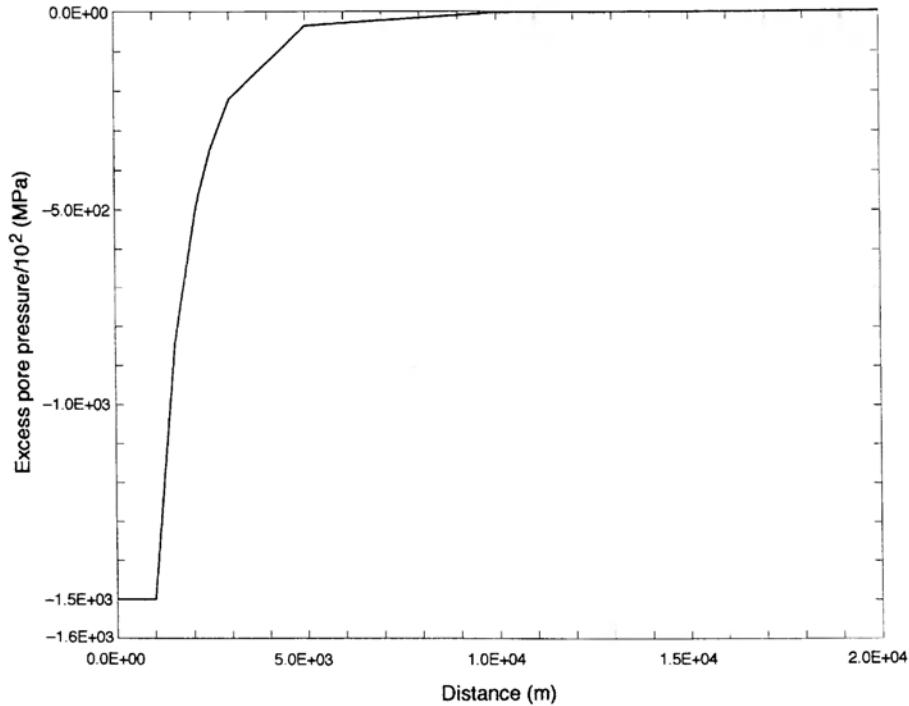
The production history over a withdrawal period of 19 years has been studied in detail for a reservoir with  $R_w = 2000\text{ m}$  and  $k = 0.0471\text{ m per day}$ . The transient reservoir compaction at the top of the clay layer and at the surface is depicted in Figure 7.58. The horizontal and vertical displacements after 19 years of production along a vertical line passing through the reservoir boundary are shown in Figure 7.59 for a base with fixed and free horizontal displacements. The effects on the surface are similar for both boundary conditions. The same is true for the transient reservoir pressure (Figure 7.63). The horizontal and vertical displacements after 19 years of production along a radius of half the reservoir thickness are shown respectively in Figures 7.60 and 7.61 for free horizontal displacements at the base. Notice that the reservoir is compacting vertically and contracting laterally. The implication is that the resulting cumulative water inflow at a particular instant is much less in the fully coupled consolidation model than in the uncoupled poroelastic model [38], where horizontal displacements are disregarded in the flow equation.

This appears clearly in Figure 7.62, where the cumulative water inflow for both production histories ( $k = 1.28\text{ mD}$ ) have been compared using the coupled and uncoupled models. The final values are obviously identical, but the transient behaviour is considerably different in both cases.

It is obvious that a smaller water inflow at a certain moment would give rise to an increased drop of reservoir pressure, and this is borne out in Figure 7.63, where the two solutions have been compared for both production histories. Again, as in the Venice



**Figure 7.54** Pore pressure decrease versus depth at shutdown of the wells: (a)  $R_w = 2000$  m,  $x = 0$ ; (b)  $R_w = 1000$  m,  $x = 2000$  m



**Figure 7.55** Pore pressure decrease versus radial distance at shutdown of the wells,  
 $R_w = 2000$  m

example (Figure 7.13), a stronger trend towards stabilisation can be seen once the forcing terms have reached equilibrium.

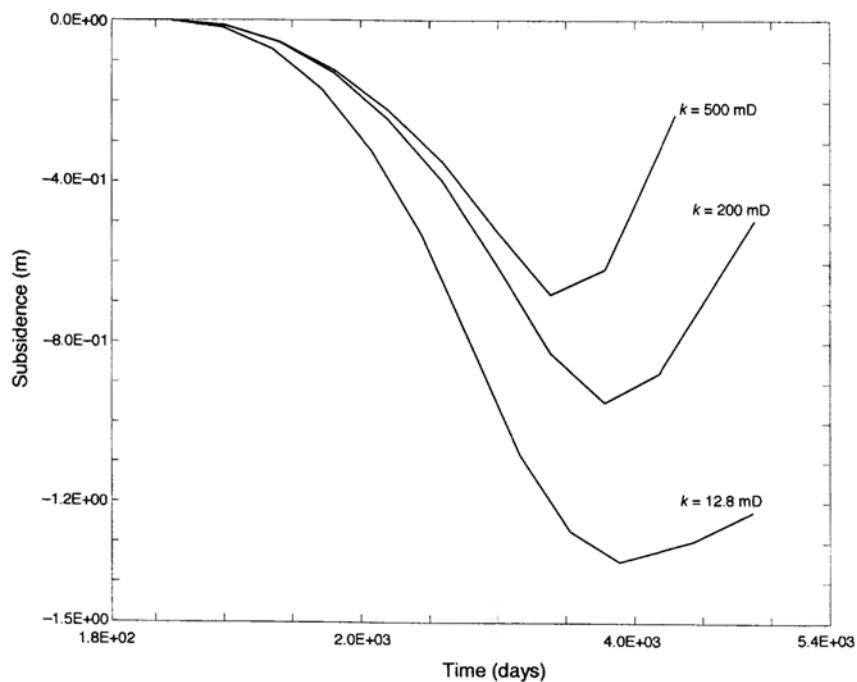
Once the initial conditions are restored in the reservoir by water inflow (in an elastic solution), the cumulative water inflow is equal to the loss of gas through the wells. This follows from the material balance equation (6.66), by assuming  $B_{gi} = B_{g\infty}$ , and was also achieved numerically (Figure 7.62). The total gas production at downhole conditions is in fact  $11.076 \times 10^9 \times B_{gi} = 53.5 \times 10^6 \text{ m}^3$ .

#### 7.4.3 Comparison with a Volumetric Reservoir

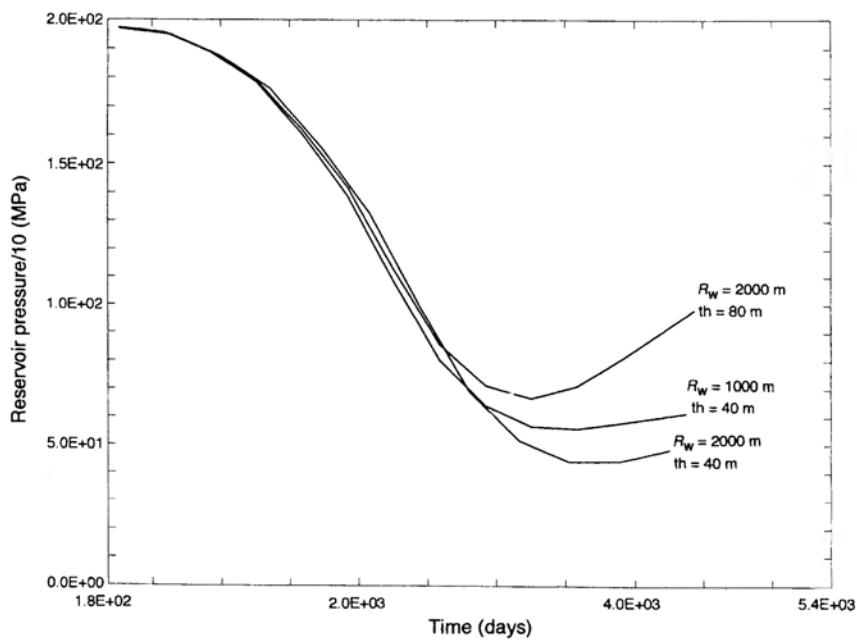
A volumetric reservoir was considered for comparison [41]. The cross-section is almost identical to the previous case, the only difference being that the pool is also laterally bounded by clay. The production history taken into account is given in Figure 7.47.

The same quantities as in Figures 7.48 and 7.49 are shown in Figure 7.64 for volumetric reservoir. Again, the reservoir compaction and settlement at the top of the clay layer is greater for  $R_w = 1000$  m than for  $R_w = 2000$  m. This time the expansion of the huge sand layer on the top of the limiting clay layer can be observed in both cases.

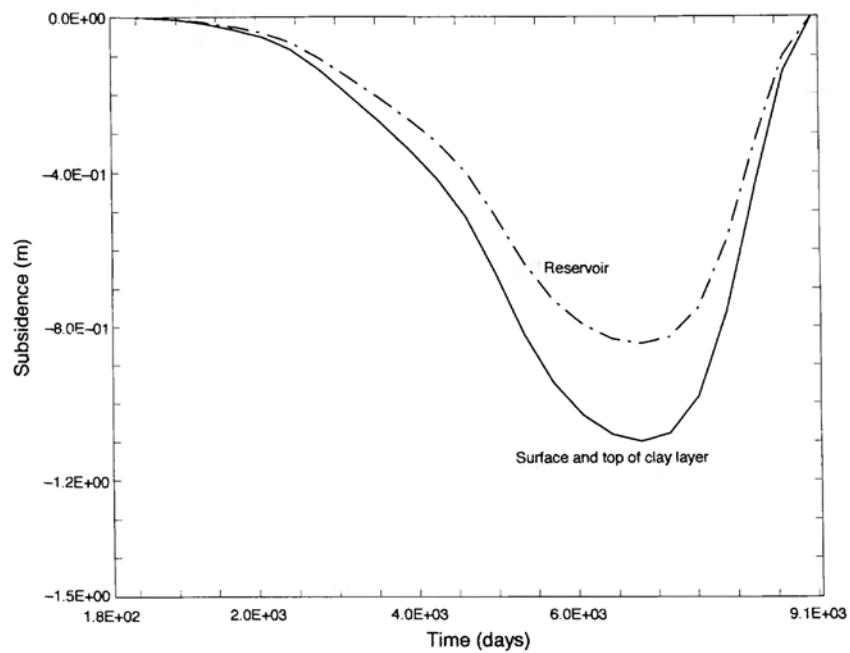
On comparing the diagrams for  $R_w = 2000$  m in Figures 7.48 and 7.64 it can be seen that the maximum compactations of the reservoir and the clay layer are similar but



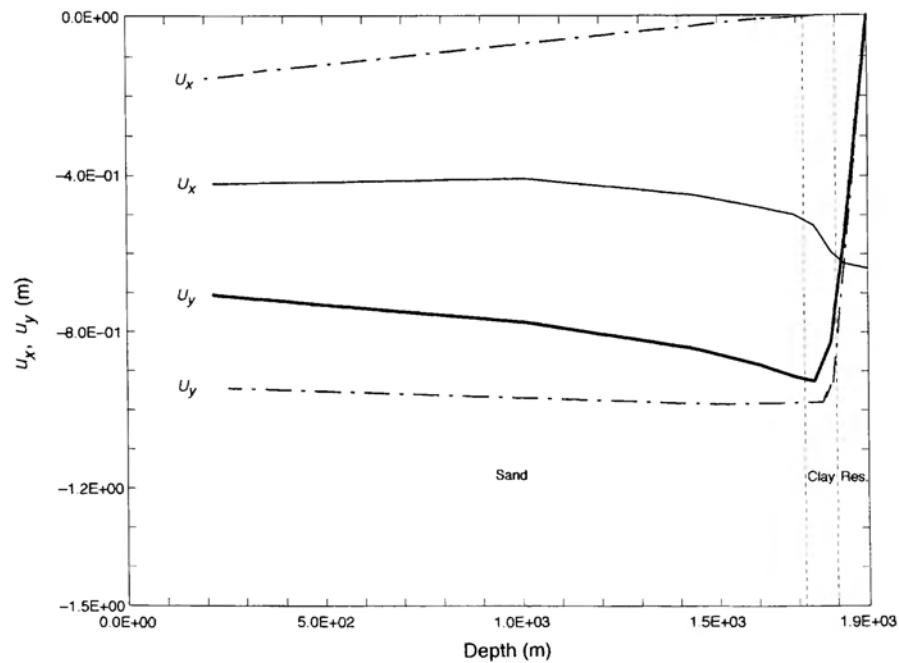
**Figure 7.56** Surface subsidence versus time for different values of aquifer permeability,  $R_w = 2000 \text{ m}$



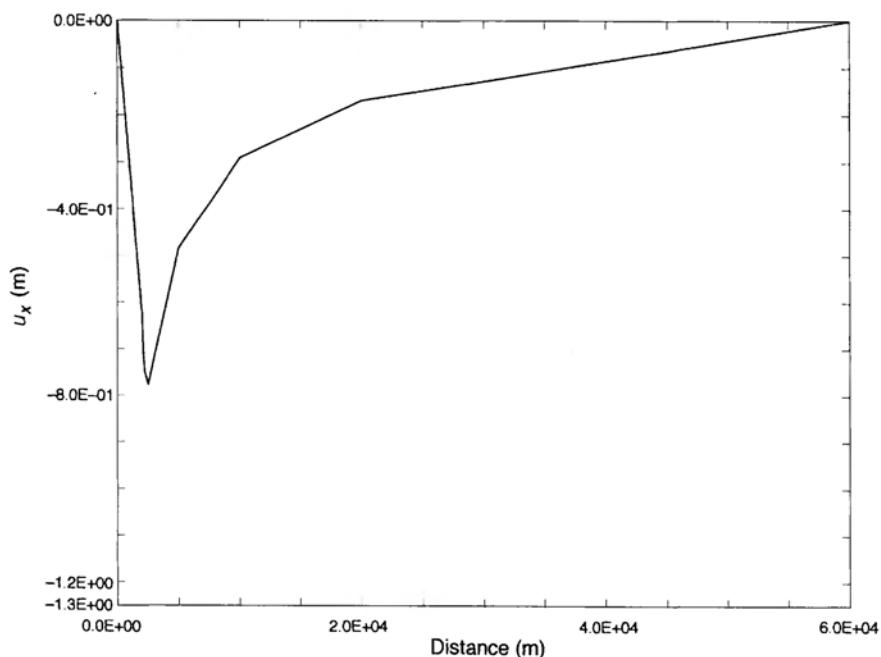
**Figure 7.57** Waterdrive reservoir: effects of the reservoir radius ( $R_w$ ) and formation thickness (th) on the reservoir pressure versus time,  $k = 12.8 \text{ mD}$



**Figure 7.58** Settlement at the centre versus time for production history B



**Figure 7.59** Horizontal and vertical displacements at shutdown along a vertical line on the outer boundary of the reservoir for a base with (---) fixed and (—) free horizontal displacements.

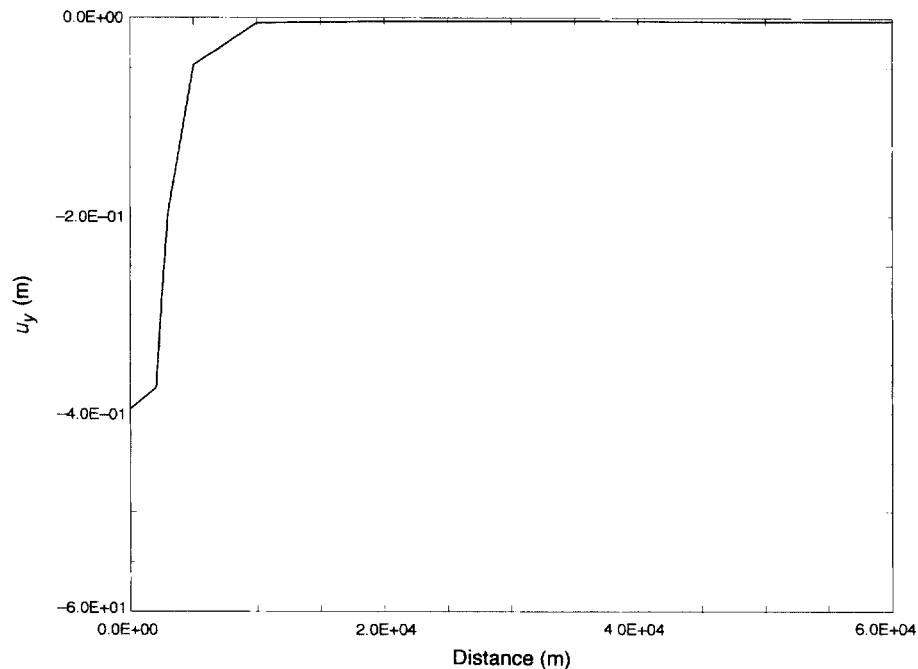


**Figure 7.60** Horizontal displacement versus radial distance plotted after 19 years of production and at half the reservoir thickness

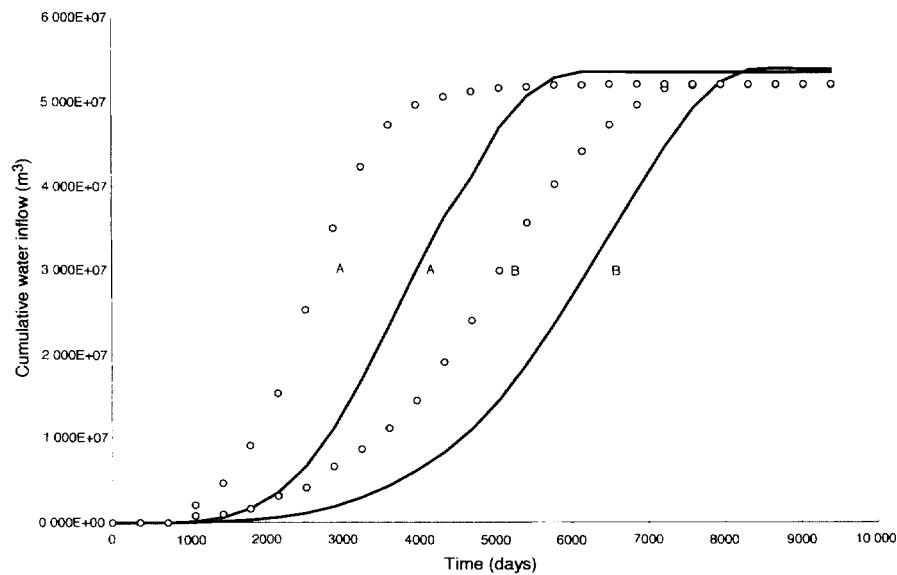
recovery after abandonment is obviously quicker for a waterdrive reservoir. The similarity can be explained by the fact that, even in a volumetric reservoir bound by clay, a significant amount of water is drained from the clay (Figure 7.65). For instance, at the shutdown of the wells (10 years after the start of production) the cumulative water inflow is  $15.4 \times 10^6 \text{ m}^3$  for the volumetric reservoir and  $20.5 \times 10^6 \text{ m}^3$  for the waterdrive reservoir. In this case (reservoir permeability  $k = 12.8 \text{ mD}$ ) the quantity of water drained from the bounding clay is greater than the inflow from the adjacent aquifer. The difference between the cumulative water inflow for the volumetric and waterdrive reservoirs increases with time and therefore causes differing behaviour of the settlement during recovery. Similar conclusions can be made when comparing the diagrams for  $R_w = 1000 \text{ m}$  of Figures 7.49 and 7.64. Note that the quantity of water drained from the surrounding clay is disregarded in the poroelastic models. The enormous quantity of water drained from the clay compared with the encroaching edge water also explains the almost negligible difference between the maximum values of the reservoir pressure drop in the two reservoir types (Figure 7.66).

The curves for reservoir pressure versus time and the cumulative water inflow for a volumetric reservoir, with radius  $R_w = 1000 \text{ m}$  and  $2000 \text{ m}$ , are shown respectively in Figures 7.67 and 7.68. Again, the recovery of the reservoir pressure for  $R_w = 1000 \text{ m}$  is significantly smaller than for  $R_w = 2000 \text{ m}$ . Ten years after the onset of production the cumulative water inflow for  $R_w = 2000 \text{ m}$  is 3.6 times that for  $R_w = 1000 \text{ m}$ .

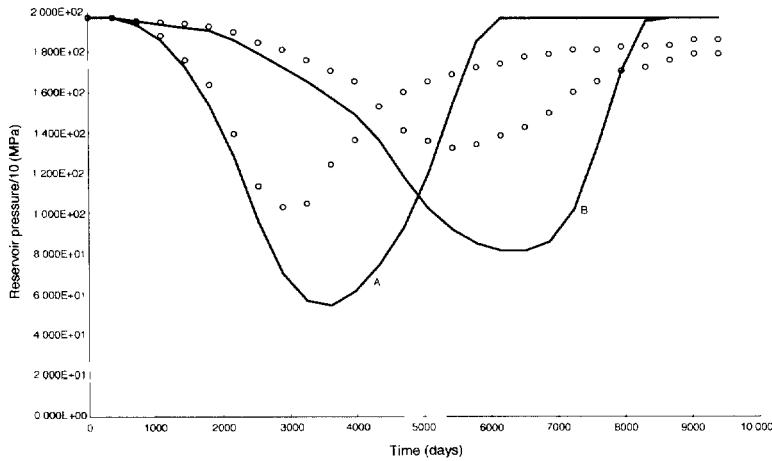
The subsidence bowl profiles at closure of the wells for a waterdrive reservoir and a volumetric reservoir are compared in Figure 7.69. The areal extent of the bowl is



**Figure 7.61** Vertical displacement versus radial distance plotted after 19 years of production and at half the reservoir thickness



**Figure 7.62** Cumulative water inflow in the reservoir,  $R_w = 2000$  m: (—) fully coupled consolidation model and (○) uncoupled poroelastic model



**Figure 7.63** Reservoir pressure using two different models and two production histories;  $R_w = 2000$  m,  $k = 12.8$  mD: (—) fully coupled consolidation model and (○) uncoupled poroelastic model

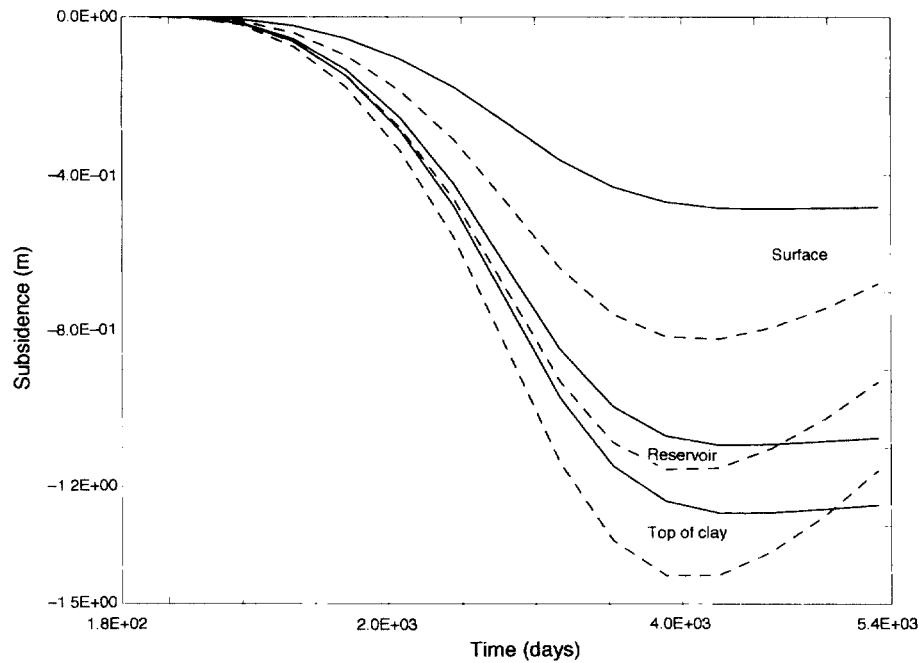
significantly larger for the waterdrive reservoir than for the volumetric reservoir. This is because the lateral aquifer extends the influence zone of the pressure gradient, and the clay above the aquifer is also consolidating. This fact is again disregarded in the poroelastic models. The subsidence bowl profiles, the settlement at the top of the clay layer and the reservoir compaction versus radial distance for both types of reservoir ( $R_w = 1000$  m) are compared in Figure 7.70. The conclusions are the same as for Figure 7.69. In the case of the volumetric reservoir the effects of a large settlement distribution over a small area at the top of the clay layer are not completely transmitted to the ground surface.

The following conclusions can be derived from the parametric investigations regarding the ground subsidence above volumetric and waterdrive gas reservoirs which were made with a fully coupled consolidation model. The existence of clay layers bounding an exploited reservoir can substantially influence the surface subsidence. The maximum subsidence and the diameter of the subsidence bowl above a waterdrive reservoir is considerably greater than above a similar volumetric reservoir. The quantity of water drained from the bounding clay layers can become significant and produce a waterdrive effect in volumetric reservoirs.

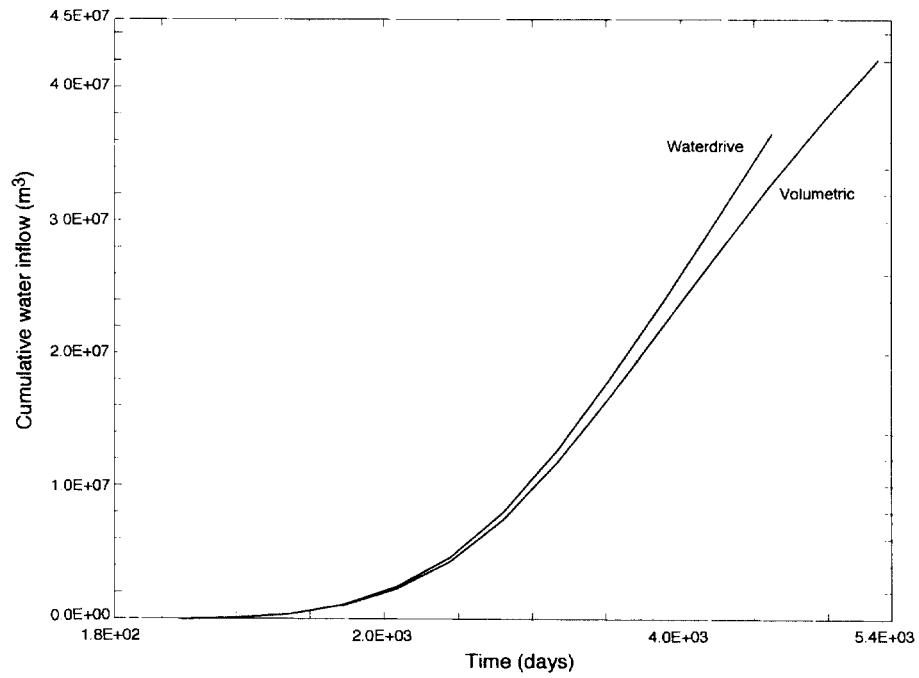
Finally, in similar situations, a fully coupled consolidation model simulates the true subsidence behaviour more realistically than a poroelastic model.

#### 7.4.4 New results for the Ravenna field

The production history at the Ravenna field (more precisely called the Ravenna Terra field) lasted from 1952 up to May 1982. Afterwards AGIP, the exploiting company, released the information regarding this field and the production [42]. In particular, the transient pressure drop was given, with a maximum value of about 10 MPa. The



**Figure 7.64** Settlement versus time in a volumetric reservoir,  $k = 12.8 \text{ mD}$ : (—)  $R_w = 1000 \text{ m}$  and (---)  $R_w = 2000 \text{ m}$



**Figure 7.65** Cumulative water inflow versus time:  $R_w = 2000 \text{ m}$ ,  $k = 12.8 \text{ mD}$

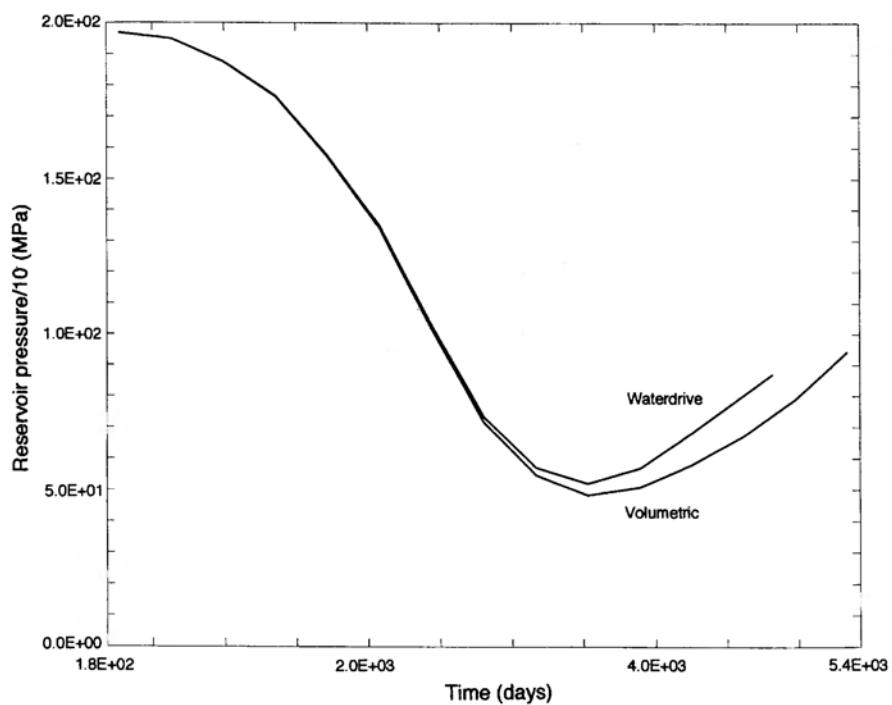


Figure 7.66 Reservoir pressure versus time:  $R_w = 2000$  m,  $k = 12.8$  mD

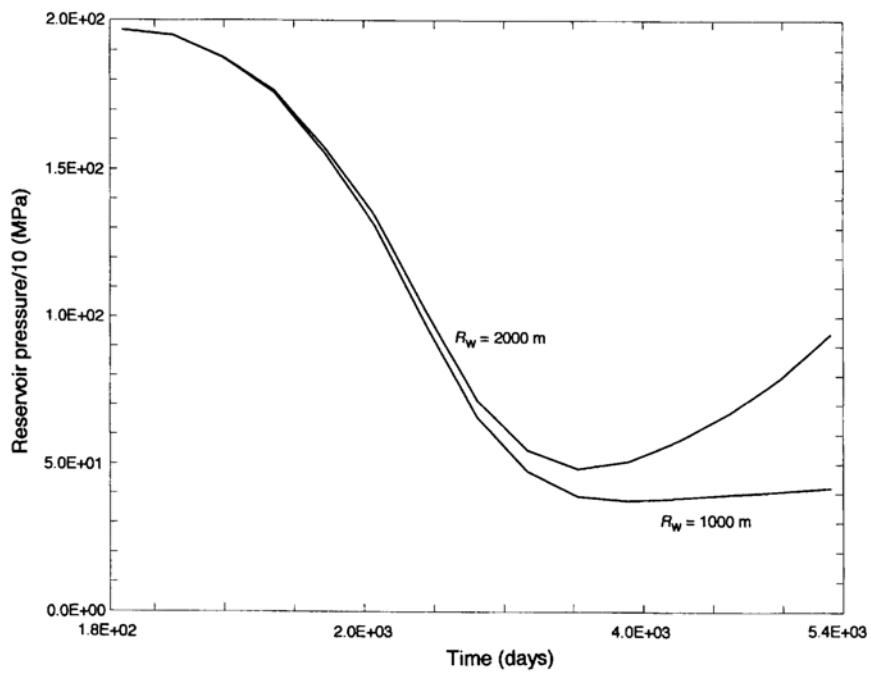
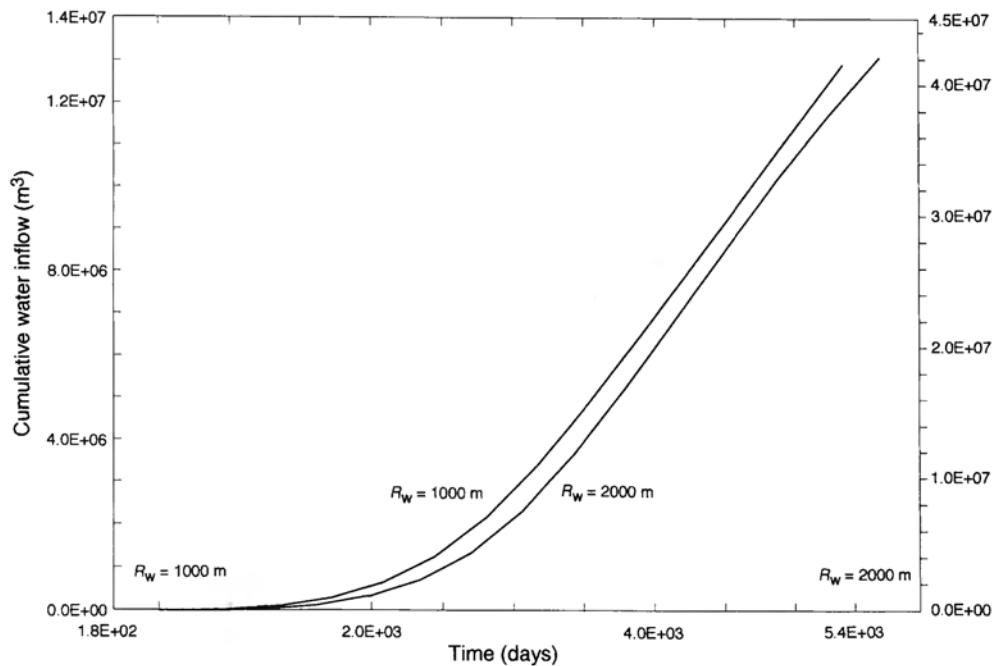
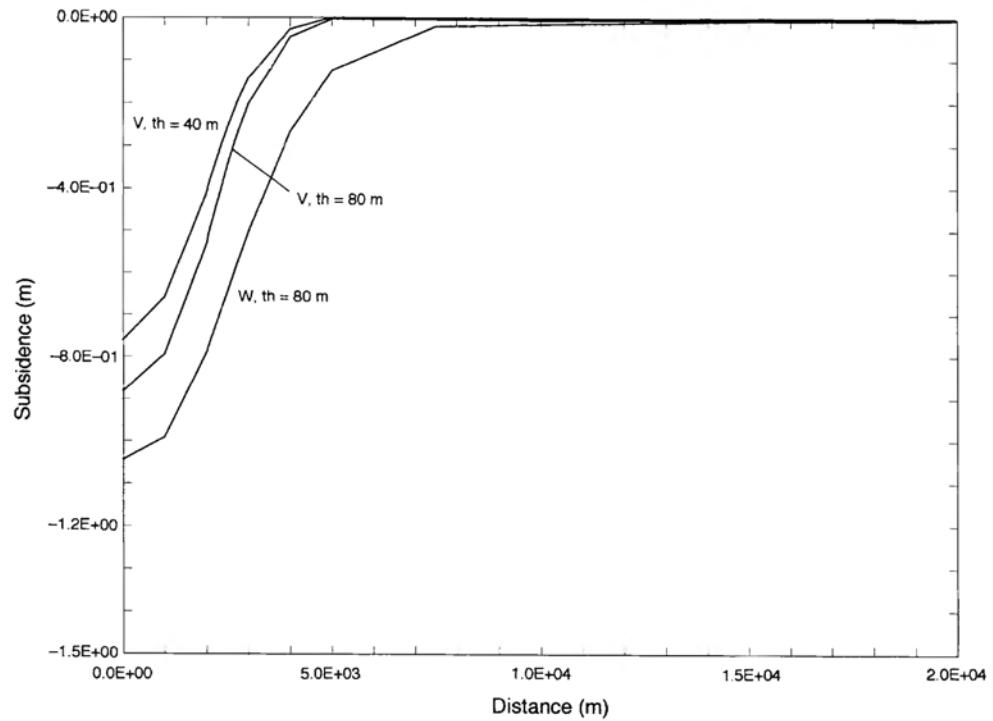


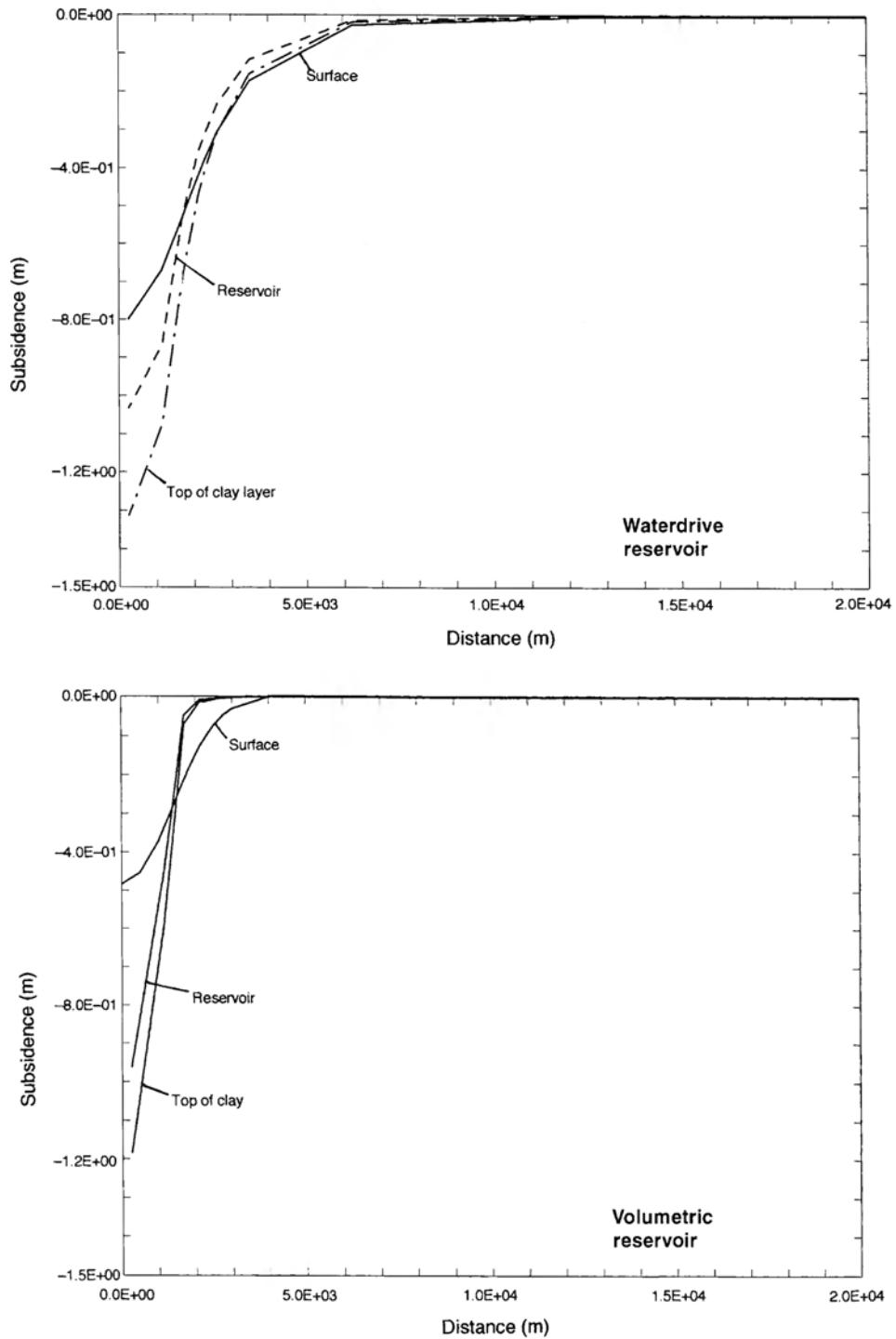
Figure 7.67 Volumetric reservoir: reservoir pressure versus time,  $k = 12.8$  mD



**Figure 7.68** Volumetric reservoir: cumulative water inflow versus time,  $k = 12.8\text{ mD}$



**Figure 7.69** Subsidence bowl profile at closedown of the wells,  $R_w = 2000\text{ m}$ ,  $k = 12.8\text{ mD}$ : th = thickness, V = volumetric reservoir, W = waterdrive reservoir

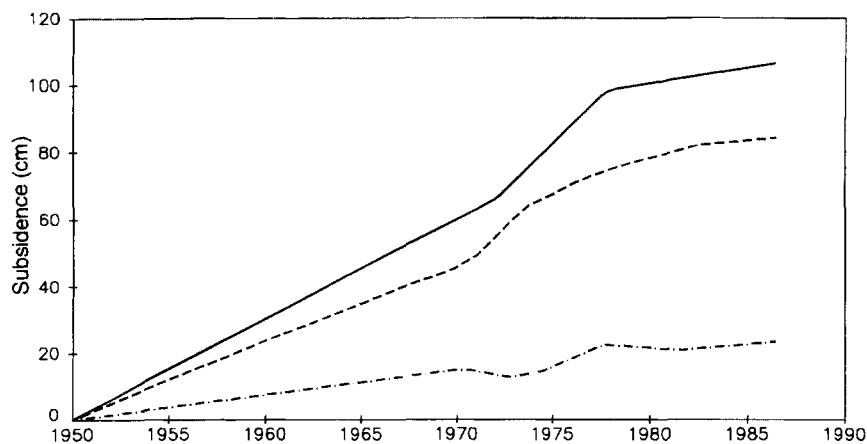


**Figure 7.70** Settlement versus radial distance at closedown of the wells:  $R_w = 1000$  m,  $k = 12.8$  mD

subsidence due to groundwater withdrawal from the superficial aquifers and to the exploitation of the gas reservoir was then extensively studied by Gambolati and coworkers [42]. In particular, for the analysis of the reservoir compaction and the ensuing subsidence, these authors used a time-independent solid-phase analyser [43], where the reservoir pressure was taken into account as a volume force and a linear elastic solid-phase behaviour was assumed.

The pressure drop in the reservoir, reported by AGIP, was almost constant from 1970 onward and the Gambolati model also gave an almost constant surface subsidence profile from 1970 to 1982. Careful inspection of data on surface lowering reported in the literature [42,44] show that, at a benchmark (Porta Adriana) close to the Ravenna Terra reservoir, the surface continued to settle at a constant rate from 1977 up to 1989 (last available data); see Figure 7.71. The same is true for the area directly above the reservoir. The water pressure in the superficial aquifers in that area also increases, at least since 1982, and the corresponding subsidence shows a rebound. Given this data and excluding further effects of the superficial aquifers, the observed subsidence behaviour cannot be explained by a pressure drop in the reservoir alone. The mechanics of partially saturated porous media, where the capillary effects are also taken into account according to the model of Section 4.6, results in a further subsidence mechanism. In fact, from equation (4.97) it follows that changes in the effective stress (the Bishop stress) at constant overburden load are not only obtained via pressure changes, but also by changes in suction (capillary pressure) and in degree of saturation. The last two are linked by the capillary pressure–saturation relationship, e.g. equation (4.95).

Water inflow causes the changes in degree of saturation within the reservoir during and after the exploitation (Figure 7.62). Hence, from such a model, ongoing subsidence could be expected both during the period of exploitation, when the reservoir pressure remained almost constant, and after its cessation. Such a model was successfully applied by Bolzon and Schrefler [45], showing the benefit of an unsaturated soil approach in subsidence analysis above gas reservoirs.



**Figure 7.71** Recorded subsidence at (—) Porta Adriana and (---) Canale Candiano; 1975 maximum subsidence due to gas extraction [6]; 1980 maximum subsidence in Ravenna due to groundwater pumping [6]; 1982 end of gas production; 1975 reduction of groundwater pumping. Also plotted is the difference between the two curves (- · -)

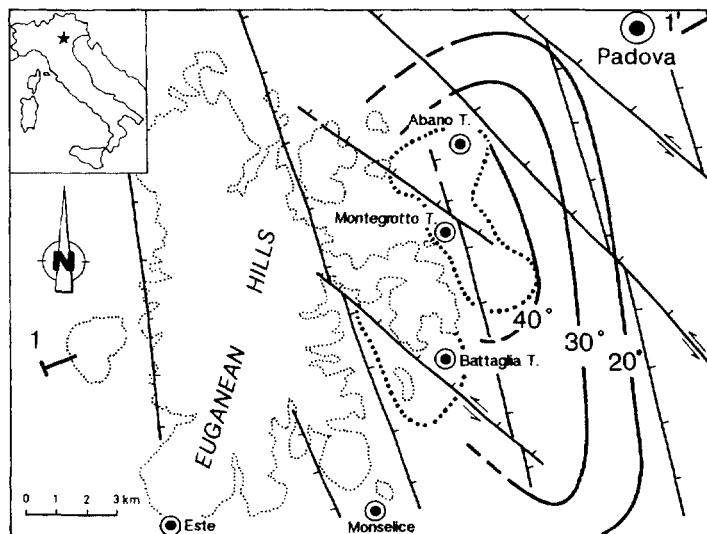
## 7.5 SUBSIDENCE OF ABANO TERME

### 7.5.1 Background

Systematic studies on the Euganean thermal field (north-eastern Italy) began in 1970 [46]. Geological and chemical-isotopic investigations resulted in the conclusion that the Euganean underground water system is a deep circuit with remote feeding (alpine and prealpine zones). In 1986, a research contract between Regione Veneto (local government) and the University of Padua boosted the investigation and involved multi-disciplinary capabilities. More historical and technical information on this test case can be found in the literature [47].

The Euganean Hills extend over an area of 100 km<sup>2</sup> and are located at the south-east of Padua. They are mainly composed of volcanic rocks, which are surrounded by alluvial plains with clay and silt deposits. Figure 7.72 presents a tectonic map of the hydrothermal reservoir with isotherm lines at a depth of 150 m. Figure 7.73 shows the geological cross-section of the region: several faults, also of regional importance, characterise the area and are clearly seen. The thickness of the Quaternary cover ranges from some 10 m to a maximum of about 230 m and increases north-east toward Padua, where it is more than 500 m thick.

The hydrothermal field occupies an area of about 20 km<sup>2</sup> immediately east of the Euganean Hills and can be classified as a convection system, dominated by the water phase. The heated fluids, at depths greater than 4000 m, rise along a fracture system due to their lighter density. They are replaced by colder water coming from the recharged zones. Hence a great heat flux is rapidly transferred by convection to the carbonatic reservoir, which is situated at relatively low depth and exploited from about 250 wells. In deep layers there was detected a more relevant vertical permeability with respect to



**Figure 7.72** Tectonic map of the Euganean reservoir with isothermal lines at a depth of 150 m

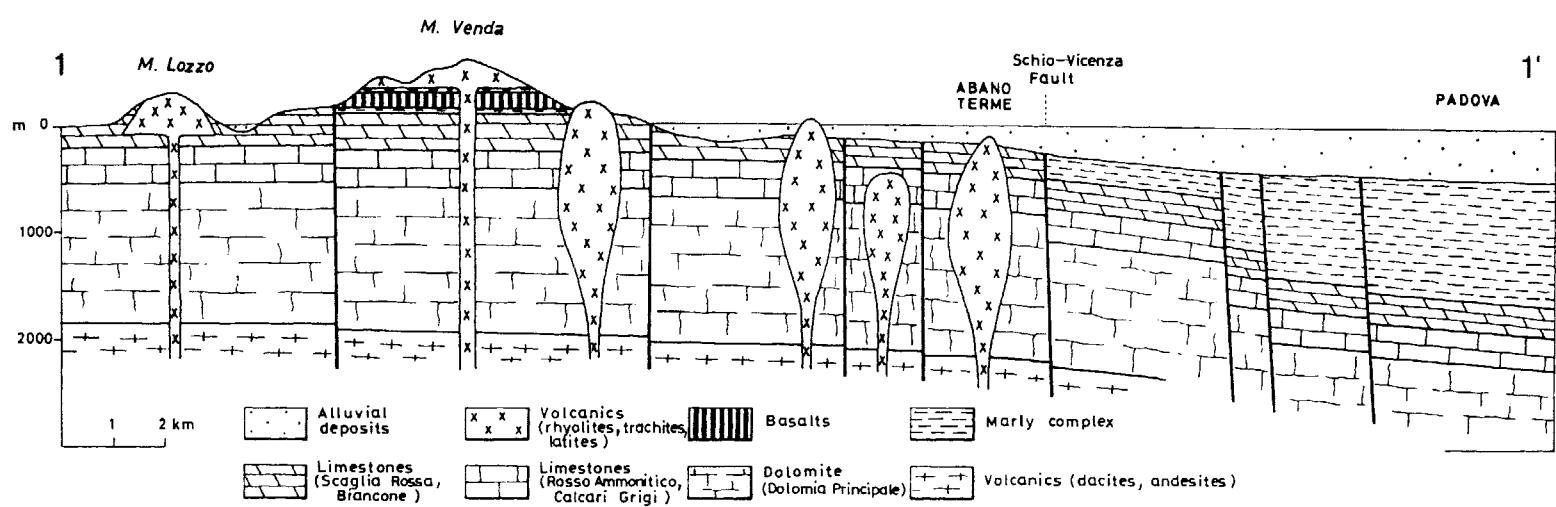


Figure 7.73 Cross-section of the hydrogeological system

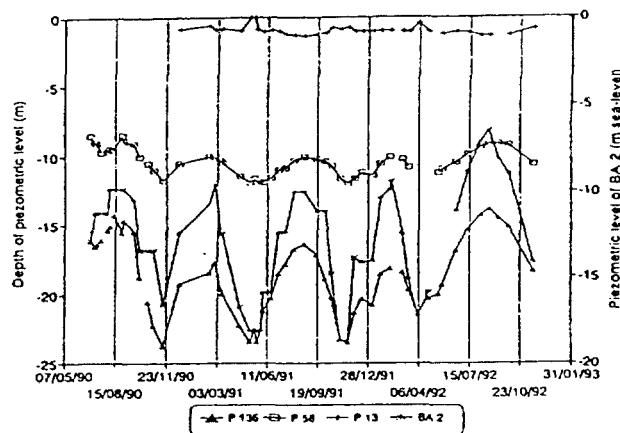
the horizontal. This is due to the presence of fractured rocks. A lateral expansion takes place when thermal water reaches the shallow zone of the reservoir. The water is partly stored inside the fractures and partly moves sideways or rises again within the alluvial cover to within 10 m of the surface. When a vertical flux is more pronounced, the temperature of the thermal water ranges between 80 °C and 87 °C, remaining practically at the same temperature as at the bottom holes.

The system exploitation occurs in the alluvial cover and, more recently, in the bedrock to a large extent. The maximum density of the well spacing is 10–15 wells per hectare and the total average production in the two principal areas of utilisation is about 0.7–0.8 m<sup>3</sup> s<sup>-1</sup>, with varying pumpage during the year and a maximum in the thermal health season.

An artificial piezometric regime in the thermal aquifer is produced by high pumping rates during the health season (Figure 7.74): rapid drops were recorded and a levelling of the subsequent retrieval. This was obtained during the two or three months of exploitation, but during the past few years a general permanent piezometric decrease was observed, with a maximum of 2 m per year. A continuously sampled well allowed an accurate determination of the aquifer in the cover, with a permeability ranging from  $1.12 \times 10^{-5}$  to  $7.7 \times 10^{-6}$  m s<sup>-1</sup>.

Two different vertical movements were observed in the Euganean Hills area: the first was connected to geological factors (elevation of alpine sector and lowering of the Padana plain) and the second related to aquifer exploitation. The first elevation measures were started in 1959 and observed a maximum subsidence in Abano Terme of several decimetres per year. This situation improved during the 1980s and today the subsidence is reduced to a few centimetres per year (Figure 7.75). This is mainly due to a different exploitation technique for the reservoir. In recent years the pumping has been principally from the calcareous bedrock, not from the alluvial cover. Maximum settlement is always localised near Abano, where the density of wells is higher.

To investigate the geotechnical properties of soils, a borehole about 300 m deep was drilled and continuously sampled. A description of the soil profile is presented in



**Figure 7.74** Piezometric levels at the Barillari wells during the period 1990–1993

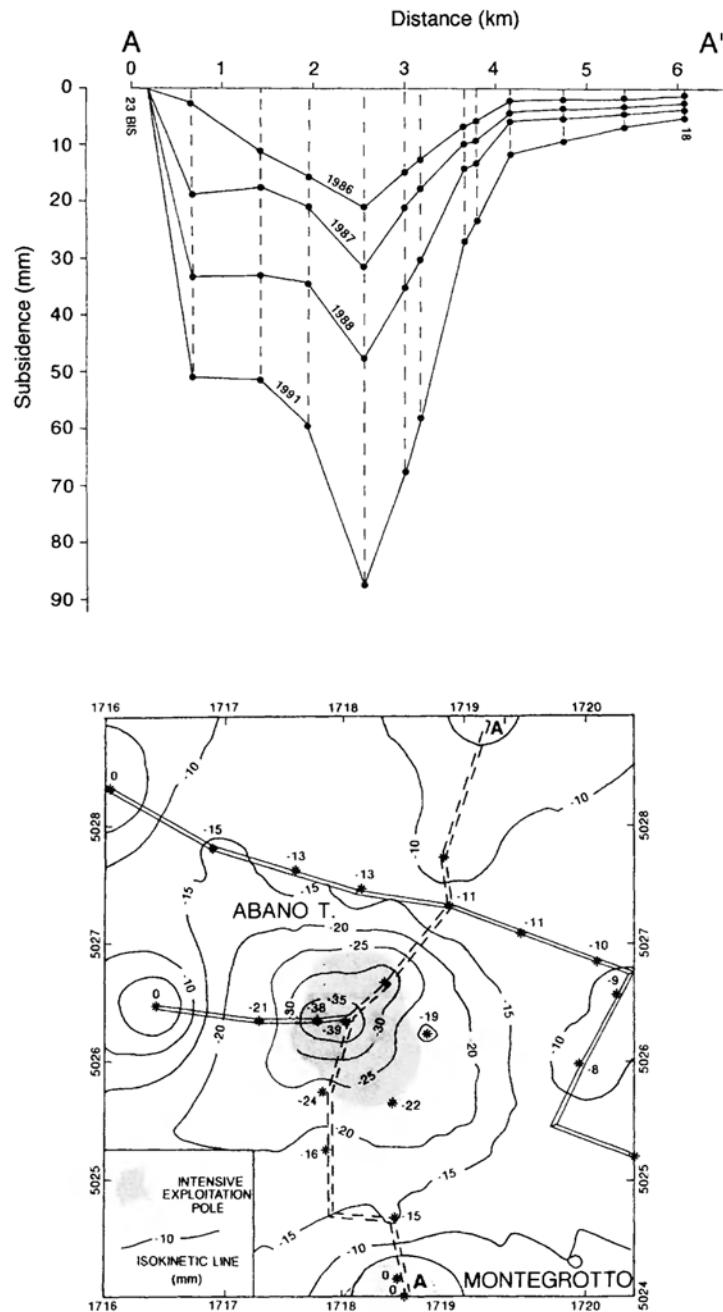


Figure 7.75 Recorded subsidence during the period 1985–1991 and an isokinetic map

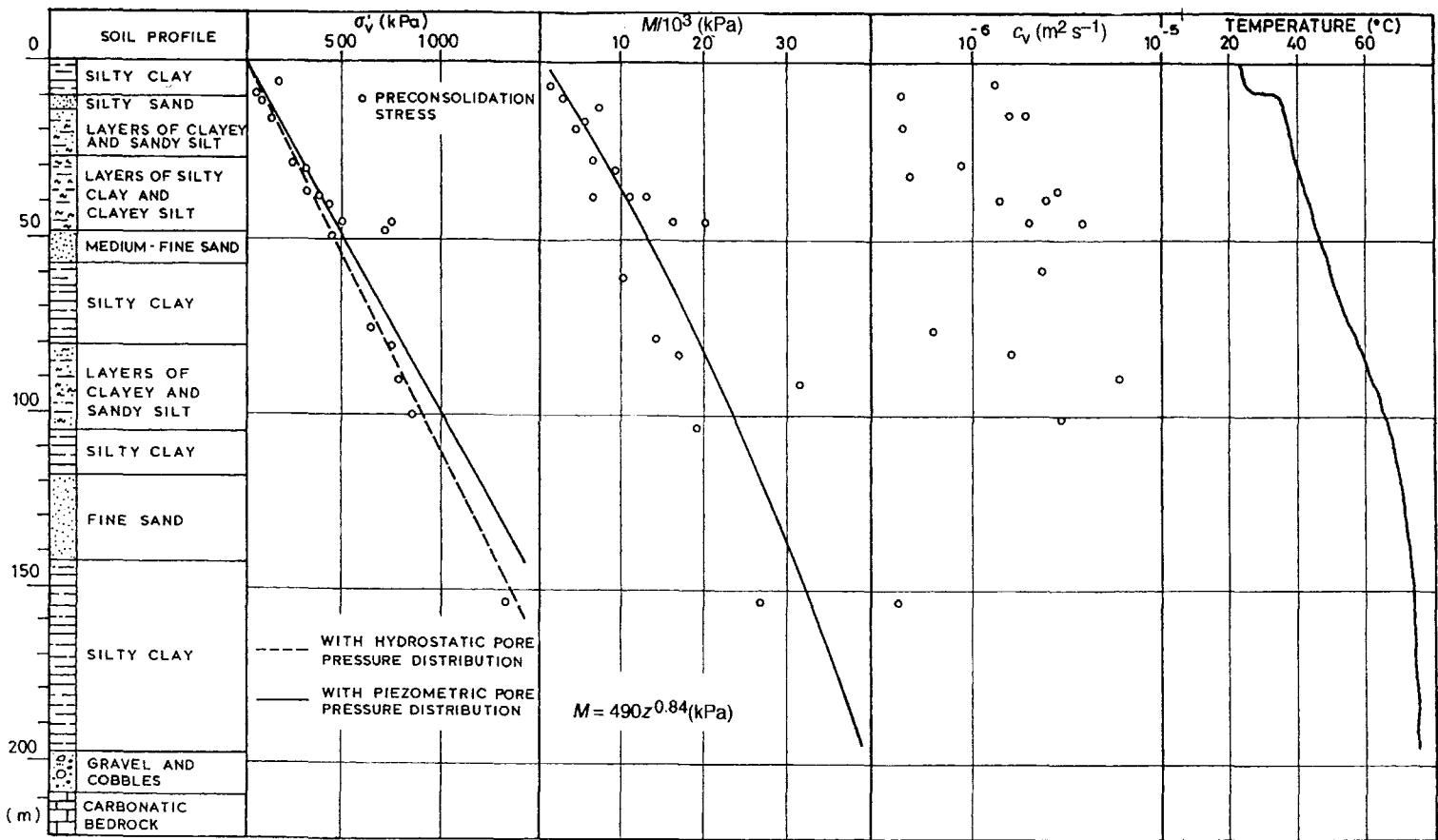
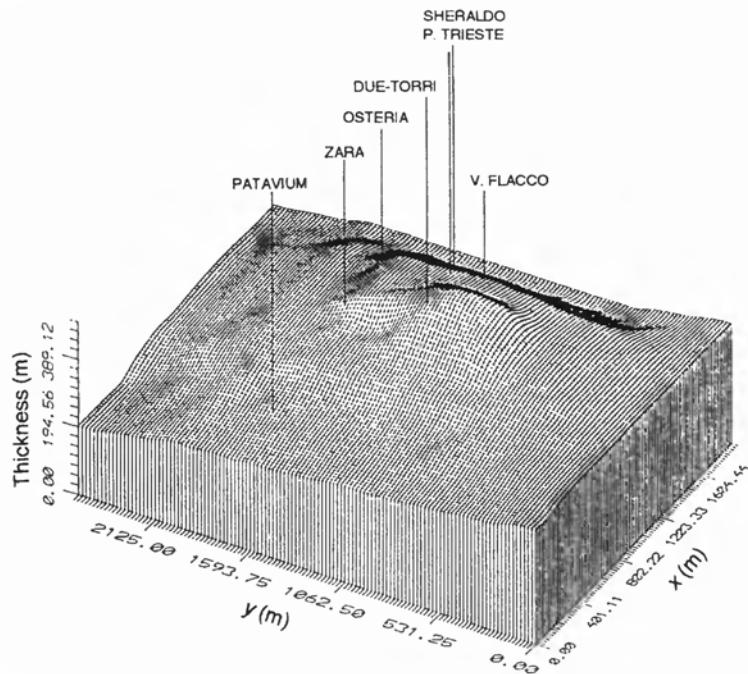


Figure 7.76 Stratigraphy and geotechnical properties at the borehole



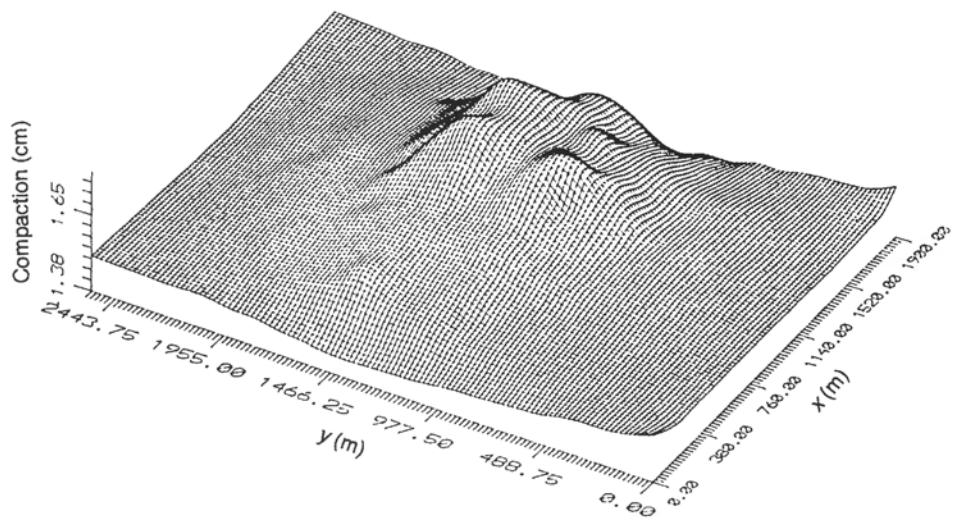
**Figure 7.77** Case study for the Abano Terme reservoir: thickness and well distribution (reduced mesh) (Reprinted, by permission, from Simoni, L. and Schrefler, B.A., 1989, FE solution of a vertically averaged model for regional land subsidence, *Int J. Num. Meth. Engng*, **27**, 215–30)

Figure 7.76, where the preconsolidation stress  $\sigma'_v$ , oedometric modulus, one-dimensional coefficient of consolidation  $c_v$  and temperature also appear as functions of depth.

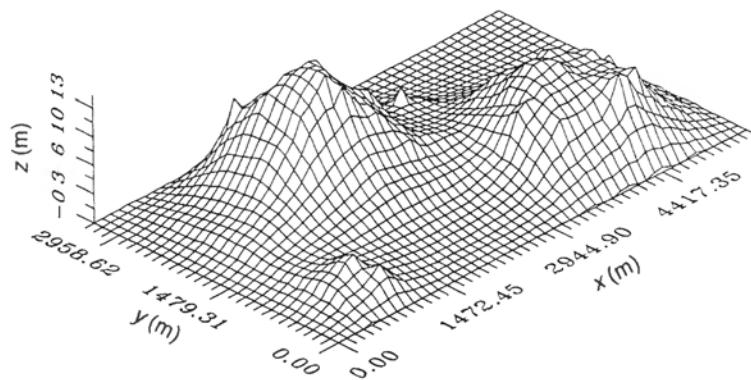
### 7.5.2 The Mathematical Model

The numerical simulator described in Section 6.2 was used in modelling the behaviour of the alluvial cover where subsidence takes place. This technique enables the problem to be studied without expensive three-dimensional models or without oversimplified two-dimensional solutions, e.g. axisymmetric (there is no axial symmetry in the geological setting), plane stress or plane strain. Measurements performed over a time span of more than 30 years indicate an almost constant value of the temperature distribution in the reservoir, so a fully saturated isothermal model was chosen, which employs one equilibrium equation for the multiphase medium and one balance equation for water.

The areal extent of the numerical model (Figure 7.77) was assumed from the presence of vertical fractures. As indicated in Section 6.2, the horizontal displacements, pore pressure and compaction values are the unknowns. The model comprised 432 parabolic finite elements in a horizontal plane, involving 1381 nodal points and 5524 d.o.f. The reservoir contained 63 boreholes, with pumping rates ranging between 150 and



**Figure 7.78** Spatial distribution of the aquifer compaction at 6 months after the start of pumping (reduced mesh) (Reprinted, by permission, from Simoni, L. and Schrefler, B.A., 1989, FE solution of a vertically averaged model for regional land subsidence, *Int J. Num. Meth. Engng*, 27, 215–30)

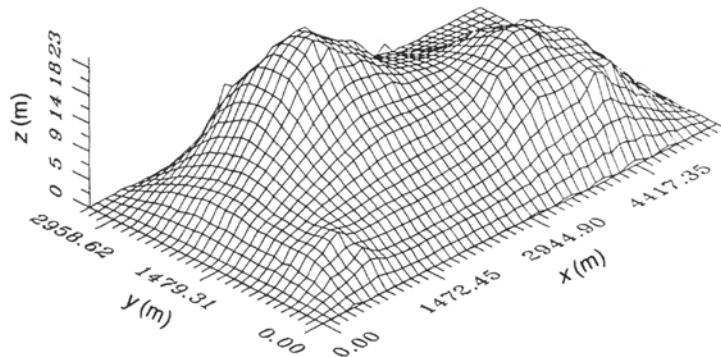


**Figure 7.79** Spatial distribution of the excess pore pressure after 2 years of pumping

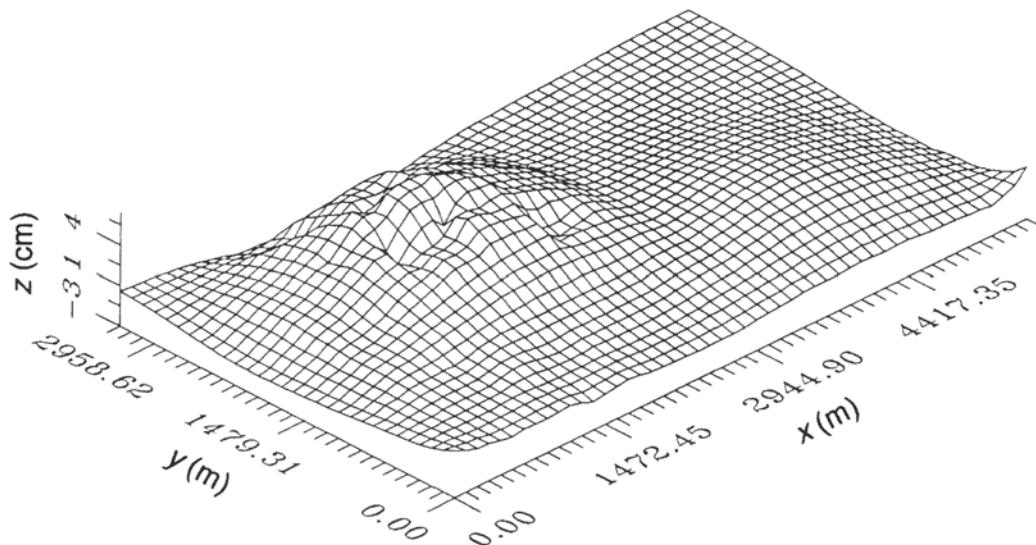
600  $\text{dm}^3 \text{ min}^{-1}$ . The following data was used (average values for the alluvial cover, which comprised aquitards and aquifers):

Young's modulus of the skeleton	2.200 MPa
Poisson's ratio	0.4
Permeability coefficients	$10^{-9} \text{ m s}^{-1}$
Porosity	0.5

The variable aquifer thickness is shown in Figure 7.77 along with the irregular distribution of the pumping stations; every well has a different time history for the pumping rate.

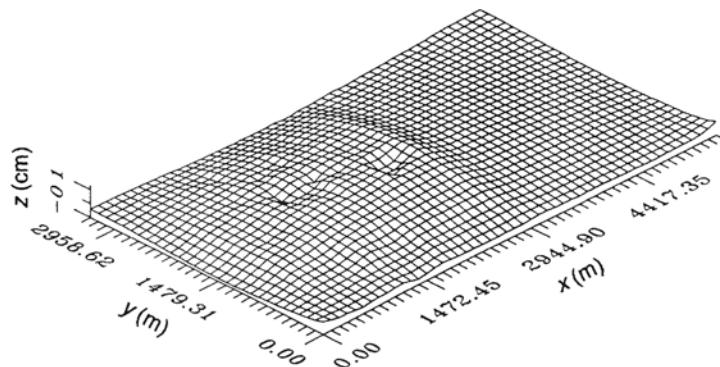


**Figure 7.80** Spatial distribution of the excess pore pressure after 9 years of pumping



**Figure 7.81** Spatial distribution of the compaction increment between the first year and the second year

The pumping rate was assumed to be constant during the year and a trial-and-error procedure was used in order to determine the amount of water drawn from the cover as compared to the total water withdrawn from the system. Figure 7.78 presents the calculated spatial distribution of compaction after 6 months of pumping, whereas Figures 7.79 and 7.80 show the spatial distribution of the excess pore pressure after 2 and 9 years, respectively. Notice the decrease in the rate of compaction change in a year: for the period between the first and second year (Figure 7.81) the compaction is about 5 cm, whereas between the eighth and ninth year. (Figure 7.82), the recorded and calculated increments of compaction are both about 1 cm. The calculated maxima for the pore pressure changes compare reasonably well with the recorded maxima (Figure 7.74).



**Figure 7.82** Spatial distribution of the compaction increment between the eighth year and the ninth year

## 7.6 CONCLUSIONS

The case studies and parametric studies discussed in this chapter give valuable insight into the problems of surface subsidence and related phenomena.

In Venice and Abano Terme, an in-depth study and successful simulation of the observed phenomena was possible because of the great amount of data on stratigraphy, material properties, withdrawal and subsidence.

An axisymmetric model was applied to the Venice problem and a vertically averaged model was applied to the Abano Terme problem, which enables an accurate three-dimensional analysis using a simplified model. The spatial distribution of the source wells in the Venice problem was taken into account via an all-embracing pumping zone. The diameter of this central pumping zone was obtained from piezometric contour lines. The finite element mesh in the calculation was designed to account for the steep gradients of pore pressure close to the boundary of this extraction area. In the Abano Terme problem the actual distribution of wells was accounted for.

Investigations were performed to calculate a time step that would keep the oscillations of the pore pressures within reasonable limits and would allow the economic simulation of long time histories (over 40 years in Venice). In the different case histories this time step varied between 3 months and 1 year.

Whenever possible, the quantitative results were checked against field measurements. In Venice this was possible both for drawdown values and vertical displacements. Unfortunately, there was a lack of recordings for the horizontal displacement.

A linear elastic constitutive relationship was used in both cases. Due to the peculiar nature of subsidence problems, simple linear and non-linear elastic models perform adequately, providing that only the compaction state is studied. The small rebound observed in the Venice simulation was modelled using an incremental procedure by changing the value of the Young's modulus.

The Venice model was calibrated with available data only until 1970; in theory it should then be capable of simulating the subsequent behaviour of the area. Further experimental data was received after the calibration, and it proved that the model can

perform predictive simulations. The superiority of coupled subsidence models over the uncoupled variety has been clearly demonstrated.

In the other two investigations a lack of sufficient field data prevented an in-depth study of the real situation. This was the case for the Contarina and Ravenna models. Extensive parametric studies have been carried out with the existing data to assess the sensitivity of the models to some of the important parameters. For Contarina the effects of different boundary conditions and material properties were investigated in detail, assuming that a non-linear elastic constitutive relationship was appropriate.

Land subsidence due to fluid withdrawal is usually permanent and only a very small recovery can be achieved by termination of pumping or by injection. The elastoplasticity model is the most appropriate constitutive relationship for the soil skeleton when such recovery mechanisms are investigated. Rebound after the cessation of pumping has therefore been studied with an elastoplastic constitutive law and a generalised plasticity framework seems to be the most appropriate procedure for this purpose. The difficulties encountered during the modelling procedure have been pointed out.

In the Ravenna simulation, where gas is being produced from deep reservoirs, the parametric studies were carried out on a hypothetical cross-section taken from the literature [38], using data related to this area. In this case it was only possible to compare the coupled and uncoupled solution [40]. The calculations have shown, however, that even if production occurs from a deep gas reservoir, the surface subsidence can still be important.

The resulting pattern of the horizontal displacements indicates movement towards the production area on the surface and at reservoir level. These horizontal displacements at reservoir level are the cause of the different results obtained with the two approaches. Bear and Corapcioglu [48] observed that approximately half the volumetric strain is produced by vertical subsidence and the other half by horizontal displacements. Hence, if the influence of the horizontal strains is disregarded, the water inflow into the reservoir is overevaluated with a resulting overestimation of the reservoir pressure. The resulting vertical subsidence is therefore much smaller when using the uncoupled model [40], as it neglects the horizontal displacements generated by the flow model.

The following conclusions can be drawn from the results presented. A successful simulation of the subsidence problem requires an enormous amount of data, hence extensive field-work by geologists, hydrologists and soil mechanicians. When data is available, the models presented here are capable of simulating both the observed and future subsidence behaviour of the reservoir in isothermal conditions. The problem is further exacerbated when dealing with subsidence related to petroleum reservoirs as they are usually deep-seated and could be many kilometres beneath the earth's surface. This is discussed in Chapter 8. And to investigate subsidence above geothermal reservoirs, an appropriate model is presented in Chapter 10.

## REFERENCES

1. Morsetti, F. (1969) Le variazioni relative al livello marino nell'Adriatico dal 1896 al 1967 ed il problema dello sprofondamento di Venezia. *Boll. Geofis. Teorica ed Applicata*, **II**(43/44), 243–254.

2. Gatto, P. (1981) *Il sottosuolo del litorale veneziano*, Tech. Rep. 108, Ist. Stud. Din. Grandi Masse, Venice.
3. Gatto, P. and Carbognin, L. (1981) The lagoon of Venice: natural environmental trend and man-induced modification. *Hydr. Sci. Bull.*, **26**(4), 379–91.
4. Zambon, M. (1967) Abbassamento del suolo per estrazioni di acqua e gas: Deduzioni ed indirizzi logicamente conseguenti per la sistemazione del Delta del fiume Po, in *Proceedings of the XIIIth National Conference on Reclamation*, Rome, pp. 345–70.
5. Poland, J. F. and Mosterman, L. J. (1969) *Reconnaissance Investigation of the Subsidence of Venice and Suggested Steps Towards its Control*, Report, Unesco, Paris.
6. Benini, G. (ed.) (1971) *Relazione fra emungimenti ed abbassamenti del suolo*, Rep. 1, Com. per lo studio dei provvedimenti a difesa della citta di Venezia ed a salvaguardia dei suoi caratteri ambientali e monumentali, III gruppo di lav., Min. LI PP, Rome.
7. Ricceri, G. and Butterfield, R. (1974) An analysis of compressibility data from a deep borehole in Venice. *Géotechnique*, **24**(2), 175–91.
8. Carbognin, L. and Gatto, P. (1976) A methodology for hydrological data collection in the Venetian plain, in *Proceedings of the IBM Seminar on Regional Groundwater Hydrology and Modelling*, Venice.
9. Mozzi, G., Benini, G., Carbognin, L., Gatto, P. and Masutti, M. (1975) *Situazione idrogeologica nel sottosuolo di Venezia: Evoluzione delle pressioni di strato negli acquiferi artesiani*. Tech. Rep. 66, CNR, Lab. Stud. Din. Grandi Masse, Venice.
10. Gambolati, G., Gatto, P. and Freeze, A. (1974) Mathematical simulation of the subsidence of Venice, 2, results. *Water Resources Research*, **10**(3), 563–77.
11. Serandrei-Barbero, R. (1973) *Indagine sullo sfruttamento artesiano nel comune di Venezia 1846–1970*, Tech. Rep. 31, CNR, Lab. Stud. Din. Grandi Masse, Venice.
12. Carbognin, L., Gatto, P., Mozzi, G., Gambolati, G. and Ricceri, G. (1976) New trend in the subsidence of Venice, in *Proceedings of the 2nd International Symposium on Land Subsidence*, Anaheim CA, IAHS (International Association of Hydrological Sciences), 121, 65–81.
13. Lewis, R. W. and Schrefler, B. A. (1978) A fully coupled consolidation model of the subsidence of Venice. *Water Resources Research*, **14**(2), 223–30.
14. Lewis, R. W. and Schrefler, B. A. (1978) A finite element analysis of surface subsidence, in *Evaluation and Prediction of Subsidence*, S. K. Saxena (ed.), ASCE, New York, pp. 400–16.
15. Schrefler, B. A. (1978) Investigation of land subsidence in the lower Po valley using a finite element consolidation model, in *Proceedings of the XXth ICES UGW Conference*, Vol. 1, EIDA, EUROCOPY, ISPRA; see also *ICES J.*, **10**, 38–57.
16. Guacci, G. (1978) The Pixley fissure, San Joaquin Valley, California, in *Evaluation and Prediction of Subsidence*, S. K. Saxena (ed.), ASCE, New York, pp. 303–19.
17. Carbognin, L., Tosi, L. and Teatini, P. (1995) Analysis of actual land subsidence in Venice and its hinterland (Italy), in *Land Subsidence*, F. B. J. Barends, F. J. J. Brouwer and F. H. Schröder (eds), Balkema, Rotterdam, pp. 129–37.
18. Carbognin, L., Marabini, F. and Tosi, L. (1995) Land subsidence and degradation of the Venetian littoral, in *Land Subsidence*, F. B. J. Barends, F. J. J. Brouwer and F. H. Schröder (eds) IAHS, The Hague, Netherlands, pp. 391–402.
19. Sellì, R. (1943) Le conoscenze geologiche sul quaternario gassifero del Polesine e del Ferrarese settentrionale in *Proceedings of the VIth Convegno nazionale del Metano*, 515–35.
20. Ministero Del Lavori Pubblici (1963) *Anormali abbassamenti dei terreni del Polesine e del Delta Padano*, Comitato esecutivo per le ricerche ed esperienze relative all'abbassamento del suolo del Polesine e del Delta Padano, Parma.
21. Caputo, M., Pieri, L. and Uguendoli, M. (1976) *Geometric investigation of the subsidence in the Po Delta*, CNR-Lab. Stud. Din. Grandi Masse, Venice.
22. Schrefler, B. A., Lewis, R. W. and Norris, V. A. (1977) A case study of the surface subsidence of the Polesine area. *Int. J. Num. Anal. Meth. Geomech.*, **1**, 377–86.

23. Ricceri, G. and Schrefler, B. A. (1978) Condizioni di contorno, leggi costitutive e parametri geotecnici nella simulazione dei processi di subsidenza. *Proc. Acc. Padavina di Scienze, Lettere, Arti XCI*, II, 133–150.
24. Mazzalai, P., Ricceri, G. and Schrefler, B. A. (1978) Studio della subsidenza nel Delta Padano, in *Proceedings of the Conference I problemi della subsidenza nella politica del territorio e della difesa del suolo*, Pisa, Vol. 4, 50–64.
25. Sandhu, R. S., Liu, H. and Singh, K. J. (1977) Numerical performance of some finite element schemes for analysis of seepage in porous elastic media. *Int. J. Num. Anal. Meth. Geomech.*, **1**, 177–94.
26. Zienkiewicz, O. C. and Taylor, R. L. (1989) *The Finite Element Method*, Vol. 1, 4th ed, McGraw-Hill, New York.
27. Zienkiewicz O. C. (1982) Basic formulation of static and dynamic behaviour of soil and other material, in *Numerical Methods in Geomechanics*, J. B. Martins (ed.), D. Riedel, London.
28. Zienkiewicz, O. C., Qu, S., Taylor, R. L. and Nakazawa, S. (1986) The patch test for mixed formulation. *Int. J. Num. Meth. Engng.*, **22**, 39–62.
29. Bredehoeft, J. D. and Pinder, G. F. (1970) Digital analysis of areal flow in multi-aquifer groundwater systems: a quasi three-dimensional model. *Water Resources Research*, **6**, 883–88.
30. Poland, J. F. (1981) Subsidence in United States due to groundwater withdrawal. *J. Irr. Drain. Div. ASCE*, **107**(IR2), 115–35.
31. Schrefler, B. A., Lewis, R. W. and Majorana, C. E. (1982) The problem of rebound in subsidence models, in *Finite Element Flow Analysis*, T. Kawai (ed.), University of Tokyo Press, **Tokyo**, pp. 689–96.
32. Zienkiewicz, O. C. and Mroz, Z. (1984) Generalized plasticity formulation and application to geomechanics, in *Mechanics of Engineering Materials*, C. S. Desai and R. H. Gallagher (eds), Wiley, Chichester, pp. 655–79.
33. Naylor, D. J., Pande, G. N., Simpson, B. and Tabb, R. (1981) Variable elastic stress-strain laws, in *Finite Elements in Geotechnical Engineering*, Pineridge Press, Swansea, Ch. 4.
34. Zambon, M. (1968) Abbassamenti anormali del suolo nel Ravennate. *La Bonifica*, **1/2**, 58–64.
35. Carbognin, L., Gatto, P., Mozzi, G. and Gambolati, G. (1978) Land subsidence of Ravenna and its similarities with the Venice case, in *Evaluation and Prediction of Subsidence*, S. K. Saxena (ed.), ASCE, New York, pp. 254–66.
36. Ministero Dell' Agricoltura e Delle Foreste (1976) *L'abbassamento del suolo della zona litorale Ravennate*, Concessione studio DM 9-11-71, No. 481.
37. Ippolito, T. (1976) La subsidenza di Ravenna. *Le Scienze*, **99**, 104–11.
38. Evangelisti, G. and Poggi, B. (1970) Sopra i fenomeni di deformazione dei terreni da variazione della pressione di strato. *Proc. Acc. delle Scienze dell'Istituto di Bologna*, Serie II, 6.
39. Schrefler, B. A. and Lewis, R. W. (1980) A consolidation analysis of a waterdrive gas reservoir, in *Proceedings of the Third International Conference on Finite Elements in Flow Problems*, Banff, Alberta, pp. 237–47.
40. Lewis, R. W. and Schrefler, B. A. (1982) A finite element simulation of the subsidence of a gas reservoir undergoing a waterdrive, in *Finite Elements in Fluids*, Vol. 4 R. H. Gallagher, D. H. Norrie, J. T. Oden and O. C. Zienkiewicz (eds), Wiley, Chichester, pp. 179–98.
41. Schrefler, B. A., Lewis, R. W. and Majorana, C. E. (1981) Subsidence above volumetric and waterdrive gas reservoirs. *Int. J. Num. Meth. Fluids*, **1**(2), 101–15.
42. Gambolati, G., Ricceri, G., Bertoni, W., Brighenti, G. and Vuillermin, E. (1991) Mathematical simulation of the subsidence of Ravenna. *Water Resources Research*, **27**(11), 2899–2918.
43. Gambolati, G., Sartoretto, F., Rinaldo, A. and Ricceri, G. (1987) A boundary element solution to land subsidence above 3-D gas-oil reservoirs. *Int. J. Num. Anal. Meth. Geomech.*, **11**, 489–502.

44. AGIP (1992) *Subsidenza naturale e da estrazione di fluidi dal sottosuolo*, AGIP, San Donato, Italy.
45. Bolzon, G. and Schrefler, B. A. (1997) Compaction in gas reservoirs due to capillary effects, in *Computational Plasticity*, D. R. J. Owen, E. Onate and E. Hinton (eds), CIMNE, Barcelona, pp. 1625–30.
46. Piccoli, G., Bellati, R., Binotti, C., Di Lallo, E., Sedeà, R., Dal Pra, A., Cataldi, R., Gatto, G. O., Ghezzi, G., Marchetti, M., Bulgarelli, G., Schiesaro, G., Panichi, C., Tongiorgi, E., Baldi, P., Ferrara, G. C., Massari, F., Medizza, F., Iliceto, V., Norinelli, A., De Vecchi, G. P., Gregnanin, A., Piccirillo, E. M. and Sbettega, G. (1976) Il sistema idro-termale euganeo-berico e la Geologia dei Colli Euganei. Mem. Ist. Geol. Miner. Univ. Padova, 1–256.
47. Antonelli, R., Fabbri, P., Iliceto, V., Majorana, C., Previtallo, P., Schrefler, B. A. and Sedeà, R. (1995) The geothermal Euganean field: a subsidence modelling approach, in *Proceedings of the World Geothermal Congress*, Florence, International Geothermal Association, pp. 1263–68.
48. Bear, J. and Corapcioglu, M. Y. (1981) Mathematical model for regional land subsidence due to pumping. 2. Integrated aquifer subsidence equations for vertical and horizontal displacements. *Water Resources Research*, **17**(4), 947–58.

# 8

## Modelling Three-Phase Flow in Deforming Saturated Oil Reservoirs

### 8.1 INTRODUCTION

The application of the finite element method to soil consolidation problems is becoming increasingly relevant in petroleum reservoir analysis, in particular with problems involving fluid flow–stress analysis. The fluid flow–stress coupling problem was first introduced by Terzaghi [1] in 1924 as a consolidation phenomenon. This one-dimensional consolidation theory has since been widely used in practice to calculate ground settlements. Subsequently, Biot [2] extended this consolidation theory into a more general three-dimensional case based on a linear stress–strain constitutive relationship and a linear form of Darcy’s flow law.

In petroleum reservoir engineering, the application of stress–fluid flow analyses have been limited mainly to the simulation of hydraulic fracturing processes [3,4] and compaction or subsidence problems [5–8] in various situations. The reservoir rock deformation, also known as reservoir compaction, is normally neglected except in a few cases where the deformation of unconsolidated formations can cause a considerable effect on reservoir performance. In this chapter, the one-phase flow equation coupled with the consolidation theory described in previous chapters is applied to simulate the problems of three-phase flow in deforming saturated oil reservoirs, e.g. the prediction of surface subsidence and the influence of compaction on the reservoir performance.

The phenomena of surface subsidence associated with the production of oil or gas from an underground reservoir is not common, but it can often cause environmental and technical problems. Field studies related to subsidence problems have been reported by many authors; examples are the Wilmington oil field in California [9], oil reservoirs under Lake Maracaibo in Venezuela [10,11] and the Groningen gas field in Holland [12]. Numerical models have often been applied to investigate the problems. Most of the simulators were based on three different features:

1. Coupled or uncoupled models
2. Type of constitutive model (e.g. elastic or elastoplastic laws)
3. Solution techniques used

In petroleum reservoir simulation, two solution techniques are commonly used to solve the unknowns, i.e. the finite difference and finite element methods. One might ask which

method provides the most accurate and economical solution. The answer depends upon a number of factors. The finite difference approximation has been applied to many reservoir problems and is still widely used. However, the versatile finite element method is becoming increasingly popular in petroleum simulation studies partly due to its natural flexibility in being able to handle the boundary conditions for simulating complex geometric problems. Our present ‘fully coupled’ finite element model incorporates non-linear stress–strain behaviour for the consolidation phenomenon and is capable of simulating three-phase flow having relative permeability contrasts, complex rock and fluid properties and the effects of capillary pressure. The model is enhanced further by its capability in simulating reservoirs with differing types of boundary conditions.

## 8.2 DEVELOPMENT OF THE GOVERNING EQUATIONS

The four fully coupled partial differential equations, including the three immiscible and compressible fluid flow equations along with the equilibrium equation, are derived using the equations of Chapter 2. The continuity equations in a deforming porous medium are modelled in a similar manner to those applied for a non-compacting case, i.e. by combining the continuity equation for each phase with Darcy’s equation and by using appropriate equations of state. The only difference is the inclusion of several factors involving the rock matrix displacements in the calculation of the rate of accumulation for each flowing fluid. Only the summary of the derivation is presented here.

### 8.2.1 The Equilibrium Equation for a Three-Phase System

For water-wet saturated oil reservoirs, it is assumed that the average effective pore pressure is a function of the respective fluid saturations. On differentiating an extended form of equation (2.175) w.r.t time and simplifying, the following equation is obtained:

$$\frac{dp}{dt} = S_w'' \frac{dp_w}{dt} + S_o'' \frac{dp_o}{dt} + S_g'' \frac{dp_g}{dt} \quad (8.1)$$

where

$$\begin{aligned} S_w'' &= S_w^n + \delta p_{cw} \left( \frac{dS_w}{dp_{cw}} \right)^{n+1} \\ S_o'' &= S_o^n - \delta p_{cw} \left( \frac{dS_w}{dp_{cw}} \right)^{n+1} - \delta p_{cg} \left( \frac{dS_g}{dp_{cg}} \right)^{n+1} \\ S_g'' &= S_g^n + \delta p_{cg} \left( \frac{dS_g}{dp_{cg}} \right)^{n+1} \\ \delta p_{cw} &= p_{cw}^{n+1} - p_{cw}^n \\ \delta p_{cg} &= p_{cg}^{n+1} - p_{cg}^n \end{aligned}$$

Substituting (2.270) into (3.118) and recalling (2.300) gives the final form of the equilibrium equation for a three-phase flow system:

$$\begin{aligned} & \int_{\Omega} \delta \boldsymbol{\varepsilon}^T \mathbf{D}^T \frac{\partial \boldsymbol{\varepsilon}}{\partial t} d\Omega - \int_{\Omega} \delta \boldsymbol{\varepsilon}^T S_w'' \left( \mathbf{m} - \frac{\mathbf{D}_T \mathbf{m}}{3K_s} \right) \frac{\partial p_w}{\partial t} d\Omega \\ & - \int_{\Omega} \delta \boldsymbol{\varepsilon}^T S_o'' \left( \mathbf{m} - \frac{\mathbf{D}_T \mathbf{m}}{3K_s} \right) \frac{\partial p_o}{\partial t} d\Omega - \int_{\Omega} \delta \boldsymbol{\varepsilon}^T S_g'' \left( \mathbf{m} - \frac{\mathbf{D}_T \mathbf{m}}{3K_s} \right) \frac{\partial p_g}{\partial t} d\Omega \quad (8.2) \\ & - \int_{\Omega} \delta \boldsymbol{\varepsilon}^T \mathbf{D}^T \mathbf{c} d\Omega - \int_{\Omega} \delta \boldsymbol{\varepsilon}^T \mathbf{D}^T \frac{\partial \boldsymbol{\varepsilon}_o}{\partial t} d\Omega - \frac{\partial \hat{\mathbf{f}}}{\partial t} = 0 \end{aligned}$$

Note that (3.118) can be used for simulating reservoir problems in various conditions. For example, the term  $S_g''$  is equal to zero (free gas is not present) when simulating an immiscible two-phase flow (oil and water phases) in a deforming porous medium.

### 8.2.2 Three-Phase Flow Equations

The main assumptions made in deriving the three-phase flow equations are mainly concerned with the fluid and rock properties. For example, three-phase relative permeability functions can be approximated from two-phase data which consists of a set of oil–water and oil–gas relative permeability data. The reservoir is assumed to be saturated with hydrocarbon fluids, oil and gas phases, and contains water initially. The gaseous hydrocarbons usually occur in the form of free gas which is dissolved in the oil phase. Gas solubility in the water phase is assumed to be zero. On taking into account several factors which contribute to the rate of fluid accumulation, the general form of the continuity equation for each flowing phase may be expressed as follows:

$$\begin{aligned} & -\nabla^T [T_m \nabla (P_l + \rho_l gh)] + \frac{\phi}{B_l} \frac{\partial S_l}{\partial t} + \phi S_l \frac{\partial}{\partial t} \left( \frac{1}{B_l} \right) + \phi \frac{\partial}{\partial t} \left( R_{sl} S_l \right) \\ & + \lambda_f \left[ \left( \mathbf{m}^T - \frac{\mathbf{m}^T \mathbf{D}_T}{3K_s} \right) \frac{\partial \boldsymbol{\varepsilon}}{\partial t} + \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{c}}{3K_s} \right. \\ & \left. + \left( \frac{(1-\phi)}{K_s} - \frac{1}{(3K_s)^2} \mathbf{m}^T \mathbf{D}_T \mathbf{m} \right) \frac{\partial p}{\partial t} \right] = 0 \quad (8.3) \end{aligned}$$

for  $l = o, w, g$

where

$$\begin{aligned} T_m &= \mathbf{k} \left( \frac{k_{rl}}{\mu_l B_l} + R_{sl} \frac{k_{rl}}{\mu_l B_l} \right) \\ \lambda_f &= \frac{S_l}{B_l} + \frac{R_{sl} S_l}{B_l} \quad (8.4) \end{aligned}$$

$B_l$  is the formation volume factor (Section 6.4), and

$\mathbf{c}$  is a creep term (Chapter 11)

The term  $T_m$  represents the total fluid mobility of each flowing phase and  $R_{sl}$  is the gas solubility term in phase l. In this text,  $R_{sw}$  and  $R_{sg}$  were assumed to be negligible.

This mobility term is strongly dependent on the unknowns, e.g. the relative permeabilities depend on the fluid saturations. For the initial conditions it is assumed that the reservoir was in capillary pressure equilibrium, with the initial water and gas saturations above their critical values, i.e. all phases were mobile. Equations (8.3) are subject to various boundary conditions: namely, open and closed boundaries (Section 3.4).

### 8.3 APPLICATION OF THE FINITE ELEMENT METHOD

The finite element discretisation of the equilibrium and flow equations may now be expressed in terms of the nodal displacements  $\bar{\mathbf{u}}$  and nodal fluid pressures, i.e.  $\bar{\mathbf{p}}_o$ ,  $\bar{\mathbf{p}}_w$  and  $\bar{\mathbf{p}}_g$  by using the Galerkin method. The unknowns are related to their nodal values by the following expressions:

$$\mathbf{P} = \mathbf{N}\bar{\mathbf{p}} \quad \boldsymbol{\varepsilon} = \mathbf{B}\bar{\mathbf{u}} \quad \mathbf{u} = \mathbf{N}\bar{\mathbf{u}} \quad (8.5)$$

where  $\mathbf{N}$  and  $\mathbf{B}$  are the shape function and linear operator, respectively. Upon substitution of equations (8.5), we obtain the discretised form for a three-phase flow system.

For the equilibrium equation:

$$\mathbf{K} \frac{d\bar{\mathbf{u}}}{dt} + \mathbf{L}_w \frac{d\bar{\mathbf{p}}_w}{dt} + \mathbf{L}_o \frac{d\bar{\mathbf{p}}_o}{dt} + \mathbf{L}_g \frac{d\bar{\mathbf{p}}_g}{dt} - \mathbf{C} - \frac{d\mathbf{f}}{dt} = 0 \quad (8.6)$$

where

$$\begin{aligned} \mathbf{K} &= \int_{\Omega} \mathbf{B}^T \mathbf{D}_T \mathbf{B} d\Omega \\ \mathbf{L}_w &= \int_{\Omega} S'_w \left( \mathbf{B}^T \mathbf{m} \bar{\mathbf{N}} - \mathbf{B}^T \mathbf{D}_T \frac{\mathbf{m}}{3K_s} \bar{\mathbf{N}} \right) d\Omega \\ \mathbf{L}_o &= \int_{\Omega} S'_o \left( \mathbf{B}^T \mathbf{m} \bar{\mathbf{N}} - \mathbf{B}^T \mathbf{D}_T \frac{\mathbf{m}}{3K_s} \bar{\mathbf{N}} \right) d\Omega \\ \mathbf{L}_g &= \int_{\Omega} S'_g \left( \mathbf{B}^T \mathbf{m} \bar{\mathbf{N}} - \mathbf{B}^T \mathbf{D}_T \frac{\mathbf{m}}{3K_s} \bar{\mathbf{N}} \right) d\Omega \\ \mathbf{C} &= - \int_{\Omega} \mathbf{B}^T \mathbf{D}_T c d\Omega \\ \mathbf{d}\mathbf{f} &= - \int_{\Omega} \mathbf{B}^T \mathbf{D}_T \frac{d\varepsilon_o}{dt} d\Omega - \int_{\Omega} \bar{\mathbf{N}}^T \frac{db}{dt} d\Omega - \int_{\Gamma} \mathbf{N}^T \frac{d\hat{\mathbf{t}}}{dt} d\Gamma \end{aligned}$$

For water phase:

$$\mathbf{W}_p \bar{\mathbf{p}}_w + \mathbf{W}_w \frac{d\bar{\mathbf{p}}_w}{dt} + \mathbf{W}_o \frac{d\bar{\mathbf{p}}_o}{dt} + \mathbf{W}_g \frac{d\bar{\mathbf{p}}_g}{dt} + \mathbf{W}_u \frac{d\bar{\mathbf{u}}}{dt} + \bar{\mathbf{F}}_w = 0 \quad (8.7)$$

where

$$\begin{aligned} \mathbf{W}_p &= \int_{\Omega} \nabla \bar{\mathbf{N}}^T \frac{\mathbf{k} k_{rw}}{\mu B_w} \nabla \bar{\mathbf{N}} d\Omega \\ \mathbf{W}_w &= \int_{\Omega} \left[ \frac{S_w}{B_w} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_w - \frac{\phi}{B_w} \frac{dS_w}{dP_{co}} + S_w \phi \frac{d}{dt} \left( \frac{1}{B_w} \right) \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{W}_o &= \int_{\Omega} \bar{\mathbf{N}}^T \left[ \frac{\phi}{B_w} \frac{dS_w}{dP_{co}} + \frac{S_w}{B_w} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_o \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{W}_g &= \int_{\Omega} \frac{S_w}{B_w} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_g \bar{\mathbf{N}} d\Omega \\ \mathbf{W}_u &= \int_{\Omega} \bar{\mathbf{N}} \frac{S_w}{B_w} \left( \mathbf{m}^T - \frac{\mathbf{m}^T \mathbf{D}_T}{3K_s} \right) \mathbf{B} d\Omega \\ \bar{\mathbf{F}}_w &= \int_{\Omega} Q_{net} d\Omega + \int_{\Gamma} \mathbf{N}^T d\Gamma \end{aligned}$$

For oil phase:

$$\mathbf{H}_p \bar{\mathbf{p}}_o + \mathbf{H}_w \frac{d\bar{\mathbf{p}}_w}{dt} + \mathbf{H}_o \frac{d\bar{\mathbf{p}}_o}{dt} + \mathbf{H}_g \frac{d\bar{\mathbf{p}}_g}{dt} + \mathbf{H}_u \frac{d\bar{\mathbf{u}}}{dt} + \bar{\mathbf{F}}_o = 0 \quad (8.8)$$

where

$$\begin{aligned} \mathbf{H}_p &= \int_{\Omega} \frac{\mathbf{k} k_{ro}}{\mu_o B_o} \nabla \bar{\mathbf{N}} d\Omega \\ \mathbf{H}_o &= \int_{\Omega} \bar{\mathbf{N}}^T \left[ \left( \frac{dS_g}{dP_{cg}} - \frac{dS_w}{dP_{co}} + B_o S_o \frac{d}{dP_o} \left( \frac{1}{B_o} \right) \right) + \frac{S_o}{\phi} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_o \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{H}_w &= \int_{\Omega} \bar{\mathbf{N}}^T \left[ \frac{\phi}{B_o} \left( \frac{dS_w}{dP_{co}} + \frac{S_o}{\phi} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_w \right) \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{H}_g &= \int_{\Omega} \bar{\mathbf{N}}^T \left[ \frac{\phi}{B_o} \left( - \frac{dS_g}{dP_{cg}} + \frac{S_o}{\phi} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_g \right) \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{H}_u &= \int_{\Omega} \bar{\mathbf{N}}^T \frac{S_o}{B_o} \left( \mathbf{m}^T - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{3K_s} \right) \mathbf{B} d\Omega \\ \bar{\mathbf{F}}_o &= \int_{\Omega} \nabla \bar{\mathbf{N}}^T \frac{\mathbf{k} k_{ro}}{\mu_o B_o} \nabla p_o g h d\Omega + \int_{\Omega} \mathbf{N}^T \frac{S_o}{B_o} \frac{\mathbf{m}^T \mathbf{D}_T c}{3K_s} d\Omega \\ &\quad + \int_{\Omega} \mathbf{N}^T Q_{net} d\Omega + \int_{\Gamma} \mathbf{N}^T q_{ob} d\Gamma \end{aligned}$$

and for gas-phase flow equation:

$$\mathbf{G}_p \bar{\mathbf{p}}_g + \mathbf{G}_w \frac{d\bar{\mathbf{p}}_w}{dt} + \mathbf{G}_o \frac{d\bar{\mathbf{p}}_o}{dt} + \mathbf{G}_g \frac{d\bar{\mathbf{p}}_g}{dt} + \mathbf{G}_u \frac{d\bar{\mathbf{u}}}{dt} + \bar{\mathbf{F}}_g = 0 \quad (8.9)$$

where

$$\begin{aligned} \mathbf{G}_p &= \int_{\Omega} \nabla \mathbf{N}^T \frac{\mathbf{k} k_{rg}}{\mu_g B_g} \nabla \mathbf{N} d\Omega + \int_{\Omega} \nabla \mathbf{N}^T R_{so} \frac{\mathbf{k} k_{ro}}{\mu_o B_o} \nabla \mathbf{N} d\Omega \\ \mathbf{G}_w &= \int_{\Omega} \bar{\mathbf{N}}^T \left[ \frac{R_{so}}{B_o} \phi \frac{dS_w}{dP_{co}} + CG_{ow} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_w \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{G}_o &= \int_{\Omega} \bar{\mathbf{N}}^T \left[ \phi R_{so} S_o \frac{d}{dP_o} \left( \frac{1}{B_o} \right) - \frac{\phi}{B_g} \frac{dS_g}{dP_{cg}} + \phi \frac{R_{so}}{B_o} \left( \frac{dS_g}{dP_{cg}} - \frac{dS_w}{dP_{co}} \right) \right. \\ &\quad \left. + \phi \frac{S_o}{B_o} \frac{dR_{so}}{dP_o} + CG_{ow} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_o \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{G}_g &= \int_{\Omega} \bar{\mathbf{N}}^T \left[ \phi S_g \frac{d}{dP_g} \left( \frac{1}{B_g} \right) + \phi \left( \frac{1}{B_g} - \frac{R_{so}}{B_o} \right) \frac{dS_g}{dP_{cg}} \right. \\ &\quad \left. + CG_{ow} \left( \frac{1-\phi}{K_s} - \frac{\mathbf{m}^T \mathbf{D}_T \mathbf{m}}{(3K_s)^2} \right) S'_g \right] \bar{\mathbf{N}} d\Omega \\ \mathbf{G}_u &= \int_{\Omega} \bar{\mathbf{N}}^T CG_{ow} \left( \mathbf{m}^T - \frac{\mathbf{m}^T \mathbf{D}_T}{3K_s} \right) \mathbf{B} d\Omega \\ \bar{\mathbf{F}}_g &= \int_{\Omega} \bar{\mathbf{N}}^T CG_{ow} \frac{\mathbf{m}^T \mathbf{D}_T c}{3K_s} d\Omega + \int_{\Omega} \bar{\mathbf{N}}^T Q_{G_{net}} d\Omega + \int_{\Gamma} \bar{q}_{gb} d\Gamma \\ &\quad + \int_{\Omega} \nabla \bar{\mathbf{N}}^T \left( \frac{\mathbf{k} k_{rg}}{\mu_g B_g} + R_{so} \frac{\mathbf{k} k_{ro}}{\mu_o B_o} \right) \nabla p_o g h d\Omega \end{aligned}$$

where

$$\mathbf{C}\mathbf{G}_{ow} = \frac{S_g}{B_g} + R_{so} \frac{S_o}{B_o}$$

These equations represent a set of ordinary differential equations in time. For convenience, the equations are written in the following form:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \mathbf{W}_p & 0 & 0 \\ 0 & 0 & \mathbf{H}_p & 0 \\ 0 & 0 & 0 & \mathbf{G}_g \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}_w \\ \bar{\mathbf{p}}_o \\ \bar{\mathbf{p}}_g \end{bmatrix} + \begin{bmatrix} \mathbf{K} & \mathbf{L}_w & \mathbf{L}_o & \mathbf{L}_g \\ \mathbf{W}_u & \mathbf{W}_w & \mathbf{W}_o & \mathbf{W}_g \\ \mathbf{H}_u & \mathbf{H}_w & \mathbf{H}_o & \mathbf{H}_g \\ \mathbf{G}_u & \mathbf{G}_w & \mathbf{G}_o & \mathbf{G}_g \end{bmatrix} \frac{d}{dt} \begin{bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}_w \\ \bar{\mathbf{p}}_o \\ \bar{\mathbf{p}}_g \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{F}}_u \\ \bar{\mathbf{F}}_w \\ \bar{\mathbf{F}}_o \\ \bar{\mathbf{F}}_g \end{bmatrix} \quad (8.10)$$

The time discretisation method used in this text is based on a Kantorovich approach (Section 3.6). The time integration of (8.10) gives the following result:

$$\begin{aligned} & \left[ \begin{array}{cccc} \mathbf{K} & \mathbf{L}_w & \mathbf{L}_o & \mathbf{L}_g \\ \mathbf{W}_u & \mathbf{W}'_p & \mathbf{W}_o & \mathbf{W}_g \\ \mathbf{H}_u & \mathbf{H}_w & \mathbf{H}'_p & \mathbf{H}_g \\ \mathbf{G}_u & \mathbf{G}_w & \mathbf{G}_o & \mathbf{G}'_p \end{array} \right] \left[ \begin{array}{c} \Delta \bar{\mathbf{u}} \\ \bar{\mathbf{p}}_w \\ \bar{\mathbf{p}}_o \\ \bar{\mathbf{p}}_g \end{array} \right]_{t_n + \Delta t} \\ &= \left[ \begin{array}{ccc} \mathbf{L}_w & \mathbf{L}_o & \mathbf{L}_g \\ \mathbf{W}''_p & \mathbf{W}_o & \mathbf{W}_g \\ \mathbf{H}_w & \mathbf{H}''_p & \mathbf{H}_g \\ \mathbf{G}_w & \mathbf{G}_o & \mathbf{G}''_p \end{array} \right] \left[ \begin{array}{c} \bar{\mathbf{p}}_w \\ \bar{\mathbf{p}}_o \\ \bar{\mathbf{p}}_g \end{array} \right]_{t_n} + \left[ \begin{array}{c} d\mathbf{f}/dt - \mathbf{C} \\ \mathbf{F}_w \\ \mathbf{F}_o \\ \mathbf{F}_g \end{array} \right] \Delta t \end{aligned} \quad (8.11)$$

where

$$\begin{aligned} \mathbf{W}'_p &= (\mathbf{W}_w + \mathbf{W}_p \alpha \Delta t) \\ \mathbf{W}''_p &= (\mathbf{W}_w + \mathbf{W}_p (1 - \alpha) \Delta t) \\ \mathbf{H}'_p &= (\mathbf{H}_o + \mathbf{H}_p \alpha \Delta t) \\ \mathbf{H}''_p &= (\mathbf{H}_o + \mathbf{H}_p (1 - \alpha) \Delta t) \\ \mathbf{G}'_p &= (\mathbf{G}_g + \mathbf{G}_p \alpha \Delta t) \\ \mathbf{G}''_p &= (\mathbf{G}_g + \mathbf{G}_p (1 - \alpha) \Delta t) \end{aligned}$$

## 8.4 NUMERICAL PROCEDURES

Equations (8.11) represent a fully coupled and highly non-linear system for three-phase flow in a deforming porous medium. Since all the coefficients depend on the unknowns, iterative procedures are performed within each time step to obtain the final solutions. For this purpose, an implicit formulation was applied in programming the simulation code. The non-linearities of the solid-phase behaviour have been discussed in Chapter 4, hence only the fluid non-linearities and their treatment are discussed in the following section.

### 8.4.1 Treatment of Fluid Non-linear Terms

Note that the non-linear flow coefficients in (8.11) usually represent the mobility terms which depend on the unknowns. They can be written as follows:

$$\mathbf{T}^g = \mathbf{T}^g[f_1(P^i), f_2(S_w^i)] = T_c f_1 f_2 \quad (8.12)$$

where  $T_c$  is the constant part of the flow coefficient,  $f_1 = 1/\mu_1 B_1$  and  $f_2 = k_{rl}$ . These functions represent two types of non-linear variables, weak and strong non-linearities,

which determine the mobility of an l-phase at a Gauss point g. In the present simulation, the pressure and saturation of the fluids are the unknowns at i, which implicitly determine the non-linear parameters.

In this text, all parameters which are functions of pressure of a single phase can be classified as weak non-linearities. The weak non-linear terms include  $B_l^{n+1}$ ,  $(1/B_l)'$ ,  $\mu_l^{n+1}$  and  $R'_s$ , which depend on the degree of pressure change. Some of derivative terms, which involve the weak non-linearities, are defined as follows.

For the term  $B'_l$ :

$$\frac{\partial}{\partial P_l} \left( \frac{1}{B_l} \right) = \left[ \frac{1}{B_l(P_l^{n+1,k+1})} - \frac{1}{B_l(P_l^n)} \right] / (P_l^{n+1,k+1} - P_l^n) \quad (8.13)$$

and for the term  $R'_{so}$ :

$$\frac{\partial R_{so}}{\partial P_o} = \frac{R_{so}(P_o^{n+1,k+1}) - R_{so}(P_o^n)}{P_o^{n+1,k+1} - P_o^n} \quad (8.14)$$

$l = w, o, g$

On the other hand, all saturation-dependent terms are considered as strong non-linearities. In the present simulation code, the strong non-linearities depend on the capillary and fluid saturation changes, which approximate the values of  $S_l^{n+1}$ ,  $S'_l$  and  $k_{rl}$ . Some of the derivative terms, which involve the strong non-linearities, are given below:

$$S'_w = \frac{S_w^{n+1,k+1} - S_w^n}{P_{cw}^{n+1,k+1} - P_{cg}^n} \quad (8.15)$$

and

$$S'_g = \frac{S_g^{n+1,k+1} - S_g^n}{P_{cg}^{n+1,k+1} - P_{cg}^n} \quad (8.16)$$

The non-linear effects of the capillary-dependent functions on (8.15) and (8.16) disappear if the plots of  $S_l$  versus  $P_c$  are linear. Therefore, in many reservoir simulation problems, the strong non-linearities are caused by the saturation-dependent functions, i.e. the relative permeabilities of the multiphase system.

#### 8.4.2 Stability Analysis

Equations (8.11) must provide a stable solution to be of any practical use. A stable solution implies that the effect of an error made at one stage of the computation will not be propagated at later stages. A detailed stability analysis for the non-linear one-phase flow in a deforming porous medium has already been discussed in Section 3.7. As for the multiphase consolidation solution, the stability condition can also be analysed based on the error vector at every iteration, i.e. by introducing a residual term in (8.10) with

zero forcing term. Therefore, a residual  $\mathbf{R}_m^k$ , corresponding to a vector  $\mathbf{P}^k$ , may be defined as

$$\mathbf{R}_m^k = \mathbf{T}^m \mathbf{P}^k + \mathbf{M}^m (\mathbf{P}^{n+1} - \mathbf{P}^n) + \mathbf{Q} \quad (8.17)$$

where the subscript  $m$  indicates that the elements of the matrices  $\mathbf{T}$  and  $\mathbf{M}$  are evaluated as functions of a vector  $\mathbf{P}^m$ . For the case of implicit schemes [1], the vector  $\mathbf{P}^m$  gives the values obtained from the recent calculations, which in general may be different from  $\mathbf{P}^k$ . Using the same definition for  $\mathbf{R}_{n+1}^n$ , equation (8.17) can be expressed in the final form

$$(\mathbf{T}^{n+1} + \mathbf{A}^{n+1}) (\mathbf{P}^{n+1} - \mathbf{P}^n) = -\mathbf{R}_{n+1}^n + \mathbf{Q} \quad (8.18)$$

or it can be more generally written for the case of an iterative method as follows:

$$(\mathbf{T} + \mathbf{A})^s (\mathbf{P}^{n+1,k+1} - \mathbf{P}^n) = -\mathbf{R}_{n+1,k+1}^n + \mathbf{Q} \quad (8.19)$$

where the superscript  $s$  is the level of computation at which the non-linear coefficients  $\mathbf{T}$  and  $\mathbf{A}$  are approximated. In practice,  $s$  is equal to  $(n)$ ,  $(n + 1/2)$  or  $(n + 1, k + 1)$  which represent explicit, semi-implicit or fully implicit formulations, respectively.

Therefore, the final solution satisfies the conditions

$$\mathbf{R}_{n+1}^n \rightarrow 0 \quad (8.20)$$

Many authors [6,13,14] agree that the stability of (8.19) is strongly influenced by the manner in which the non-linear coefficients are approximated. In a finite difference simulation, the coefficient  $\mathbf{T}$ , which includes the mobility term at one node, can be obtained from the upstream or downstream node values calculated at different levels of computation, as determined by the parameter  $s$ . Aziz and Settari [15] and Peacemen [13] have studied the effects of using different types of mobility term on the stability of a non-linear multiphase problem. In similar work by Lewis and Sukirman [6], a fully implicit formulation using the finite element solution has been applied to calculate the non-linear coefficients. As expected, the authors concur that the fully implicit mobility formulation is unconditionally stable.

### 8.4.3 Mass Balance and Convergence Checks

In practice, condition (8.20) is monitored at every iteration level based on the convergence and the mass balance of the final solution. The balance check requires that the rate of fluid accumulation minus the divergence of velocity must be equal to the net flow of the reservoir system, i.e. total flow from any external source minus the total outflow from the reservoir. Equation (8.3) represents a general mass balance equation (MBE) for each of the flowing phases and may also be written as follows, where the net flow value has been evidenced.

For the incremental MBE:

$$\text{rate of } l\text{-phase accumulation} - \nabla^T \left( \frac{-\mathbf{k} k_{rl}}{\mu_l B_l} \nabla (P_l + \rho_l g h) \right) = \Delta Q_l \quad (8.21)$$

The final solution also requires that

$$\|\text{RHS} - \text{LHS}\| < \varepsilon_m \quad (8.22)$$

where  $\varepsilon_m$  is material balance limit.

For the case of cumulative MBE checks, equation (8.21) is expressed as follows:

$$\sum (\text{LHS of equation (8.21)} - \Delta Q_1) < \varepsilon_t \quad (8.23)$$

where  $\varepsilon_t$  is the cumulative material balance limit.

This check is performed after each iteration within a time step after the unknowns have been calculated. In this case the most recent values of fluid saturation are used.

Similarly, a convergence criterion of the final solution is set, based on the maximum fluid pressure change since the last iteration:

$$\|p_i^{n+1,k+1} - p_i^{n+1,k}\| < \varepsilon_c \quad (8.24)$$

where  $p_i$  is the fluid pressure at node  $i$ ,  $(k + 1)$  and  $k$  are the new and old iteration levels respectively, and  $\varepsilon_c$  is the convergence limit. This check is monitored at the new time level ( $n + 1$ ) and is applied to all nodes, i.e. for both reservoir and overburden layers. Note that the converged solution does not mean the oil material balance has been satisfied.

#### 8.4.4 Computational Procedures

The analysis of three-phase, three-dimensional flow through a deforming saturated oil reservoir deals with a six degree of freedom field problem. Equation (8.11) is the final form of the governing equations where the displacement  $\bar{\mathbf{u}}$ , and the fluid pressures  $\bar{\mathbf{p}}$  are the primary unknowns. This unsymmetric coefficient matrix equation is a fully coupled and highly non-linear system which requires simultaneous solution. For this purpose, a fully implicit formulation was applied in programming the simulation code. This method is suitable for many difficult reservoir problems where stability and the use of larger time steps are the main concern. In this text, a cubic spline interpolation method was used to update all the saturation- and pressure-dependent variables, known as rock and PVT data, respectively.

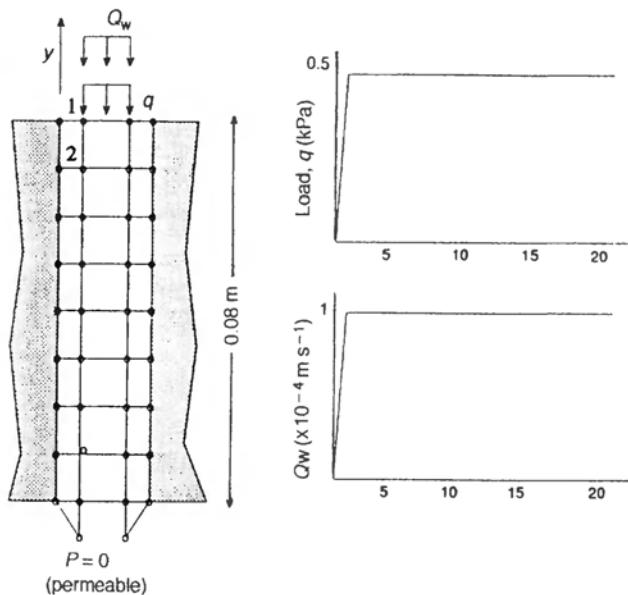
### 8.5 VALIDATION AND APPLICATIONS

A series of simulations were conducted using the developed finite element code in three different stages. First, the validation of the flow consolidation program code was carried out using a non-linear soil column problem presented by Li, *et al.* [16]. Although this test only involves immiscible two-phase flow, it can serve as a partial check for the present model. For this purpose, two simulations were made for both two- and three-phase flows in the soil column. The second stage uses the three-phase coning problem reported by Chappelar and Nolen [17] in order to investigate the effects of

consolidation on the reservoir performance. The last application indicates the utility of the developed code for predicting the surface subsidence in various situations for a hypothetical reservoir–overburden model. Some parametric investigations are discussed in detail.

### 8.5.1 Non-linear Soil Column Analysis

The developed flow consolidation code was verified using a non-linear test presented by Li *et al.* [16] for immiscible two- and three-phase [18] fluid flows. The soil column was subjected to a step loading where a maximum load of 0.5 kPa was applied at time  $t = 1$  s (Figure 8.1). An injection of water phase into the column at the surface was considered. The injection rate reached a constant value of  $0.4 \times 10^{-4} \text{ m}^3 \text{ s}^{-1}$ . In this text the values of  $c = 1.0 \times 10^5 \text{ Pa}$  and  $\phi = 35^\circ$  were used for the determination of the Mohr–Coulomb yield point. On the other hand, Li *et al.* used the Pastor–Zienkiewicz



**Figure 8.1** A non-linear soil column test. The following material properties are relevant:

Modulus of elasticity	$= 30000.0 \text{ Pa}$
Poisson's ratio	$= 0.2$
Solid bulk modulus	$= 0.10 \times 10^{12} \text{ Pa}$
Bulk modulus of liquid phases	$= 0.10 \times 10^6 \text{ Pa}$
Bulk modulus of gas phase	$= 0.10 \times 10^4 \text{ Pa}$
Initial porosity	$= 0.322$
Permeability (for all axes)	$= 0.10 \times 10^{-7} \text{ m}^4 \text{ N}^{-1} \text{ s}^{-1}$
Initial water saturation, $S_{wi}$	$= 0.345$
Soil cohesion, $c$	$= 1.0 \times 10^5 \text{ Pa}$
Friction angle, $\phi$	$= 35^\circ$

**Table 8.1** Fluid and rock properties for the three-phase model<sup>a</sup>

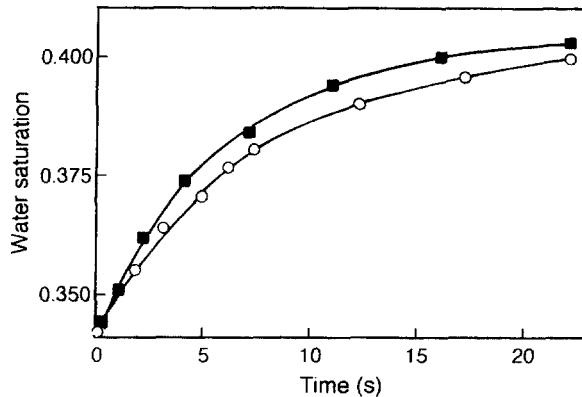
<i>Fluid properties (gas phase)</i>					
Pressure (psi)	$B_g$	$R_{so}$			
2.0	0.027 97	10.00			
3.0	0.023 45	15.0			
4.0	0.020 19	20.0			
5.0	0.017 72	25.0			
6.0	0.015 79	30.0			
7.0	0.014 24	35.0			
8.0	0.012 97	40.0			
9.0	0.011 93	45.0			
10.0	0.011 0	50.0			
<i>Relative permeability (three-phase system)</i>					
$S_g$	$K_{rg}$	$K_{ro}$	$S_w$	$K_{rw}$	$K_{ro}$
0.00	0.0000	1.0000	0.325	0.0000	0.9980
0.12	0.0340	0.2800	0.375	0.0008	0.7500
0.24	0.1240	0.0800	0.425	0.0052	0.5010
0.32	0.1980	0.0100	0.475	0.0130	0.2900
0.40	0.2850	0.0000	0.525	0.0253	0.1350
0.48	0.3590	0.0000	0.625	0.0760	0.0000
<i>Capillary data (three-phase system)</i>					
$P_{ew}$	$S_w$	$P_{cg}$	$S_g$		
0.00	1.000	0.00	0.000		
0.10	0.620	1.00	0.200		
0.20	0.475	3.00	0.600		
0.50	0.214	4.00	0.800		

<sup>a</sup> For all cases the fluid viscosities are  $\mu_o = 1.0 \text{ cP}$ ,  $\mu_w = 1.0 \text{ cP}$ , and  $\mu_g = 0.016 \text{ cP}$ .

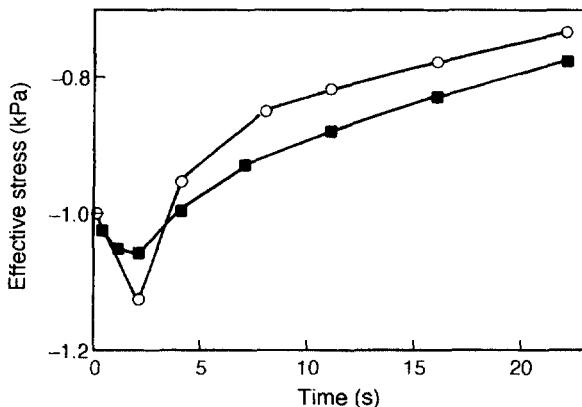
soil model and applied two different numerical solution, the direct and staggered methods. Other material properties are given in the caption to Figure 8.1. Fluid and rock properties for the three-phase model are given in Table 8.1.

The results obtained from the present finite element code are shown in Figures 8.2 to 8.6 for the two-phase and three-phase flow. Figures 8.2 to 8.4 compare non-linear results for the tested soil column obtained from the present finite element model and those presented by Li *et al.* Notice that both predictions are identical, hence they indicate the validity of the present two-phase consolidation program.

The simulation results for the three-phase flow in this simple non-linear soil column test are shown in Figures 8.5 and 8.6. Figure 8.5 illustrates that the three-phase flow model gives a smaller displacement at the surface. This may be because the average pore pressure exerted at the surface of the solid skeleton in the three-phase system is greater than in the two-phase system (Figure 8.6). In other words, the presence of free gas voids



**Figure 8.2** Non-linear solutions for water saturation versus time: (■) at 1 for present FEM model; (○) at 1 for Li and Zienkiewicz model

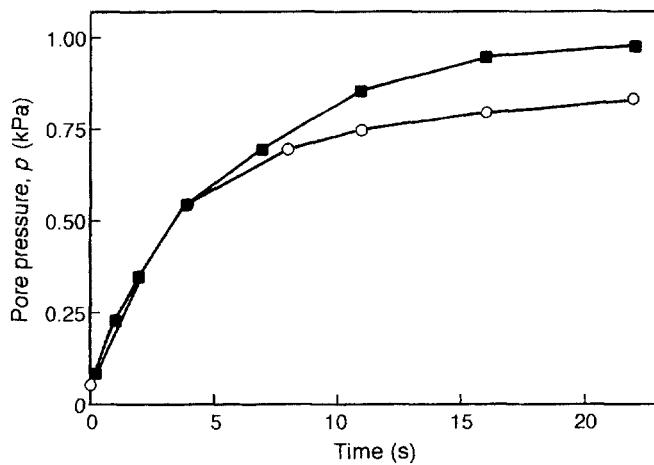


**Figure 8.3** Non-linear solutions for effective stress versus time: (■) at 2 for present two-phase model; (○) at 2 for Li and Zienkiewicz model

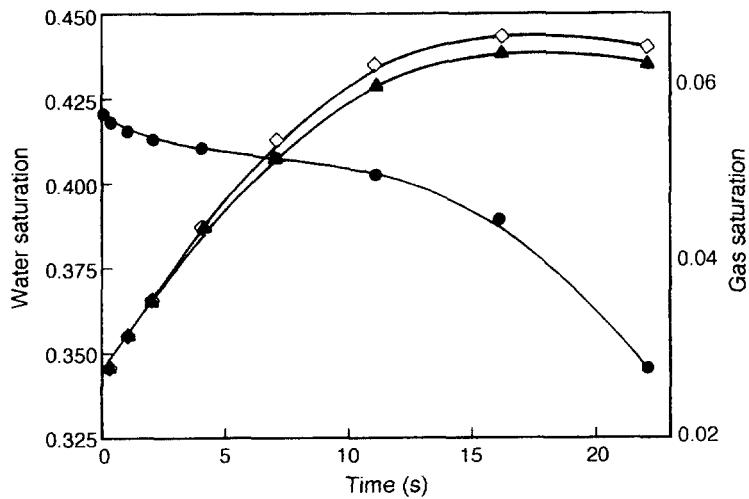
has increased the effectiveness of the pressure maintenance process, i.e. the injected fluid has filled the free voids, reducing the pressure values in the system. Figure 8.6 illustrates the pressure and saturation profiles obtained from the two- and three-phase consolidation simulations using the present code. The results obtained show the same trend for the two models.

### 8.5.2 Reservoir Compaction Problems

This is the second Society of Petroleum Engineers (SPE) comparative solution project on the three-phase coning problem [17]. The actual reservoir model was modified in order to provide a challenging test problem. Table 8.2 summarizes the three different types of reservoir rock used.



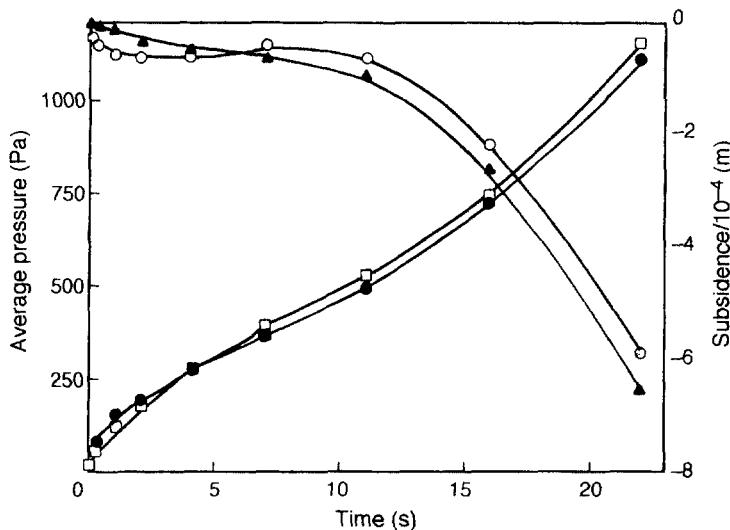
**Figure 8.4** Non-linear solutions for  $p_w$  versus time: (■) at 1 for present FEM model; (○) at 1 for Li and Zienkiewicz model



**Figure 8.5** Fluid saturation for a non-linear soil column: (◊)  $S_w$  at 1 for two-phase model; (▲)  $S_w$  at 1 for three-phase model; (●)  $S_g$  at 1 for three-phase model

In this text the effects of capillary pressure for the water–oil and gas–oil systems were taken into account (Table 8.3). The two-phase relative permeability data required to calculate the three relative permeability curves is shown in Figures 8.7 and 8.8. Figure 8.7 shows the water–oil imbibition data, and Figure 8.8 indicates the gas–oil drainage data. Other data such as the three-phase PVT characteristics were obtained from Chappelar and Nolen [17].

A three-dimensional geometrical representation of the reservoir with Cartesian coordinates was used in all the examples. Two fully sealing faults were placed at the top



**Figure 8.6** Subsidence and pressure profiles for two- and three-phase models: (●) average pressure at 1 for two-phase model; (□) average pressure at 1 for three-phase model; (○) subsidence at 1 for three-phase model; (▲) subsidence at 1 for two-phase model

**Table 8.2** Material properties for the reservoir compaction model

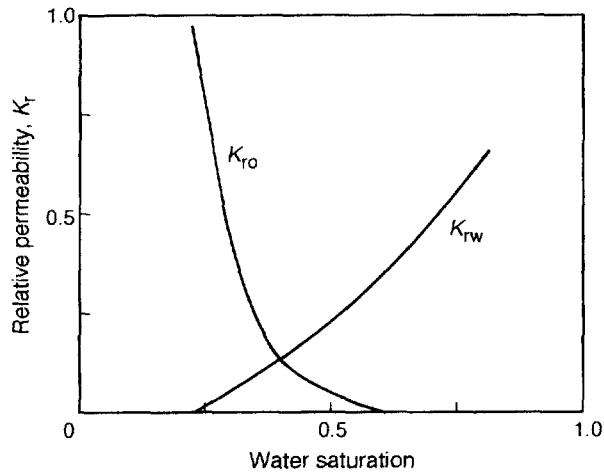
Property	Rock I	Rock II	Rock III
$E(\text{N m}^{-2})$	$0.9631 \times 10^{10}$	$0.9631 \times 10^9$	$0.9631 \times 10^6$
$v$	0.27	0.23	0.20
$C_r(\text{N m}^{-2})$	$0.2000 \times 10^{-9}$	$0.2000 \times 10^{-6}$	$0.2000 \times 10^{-5}$

**Table 8.3** Capillary data used

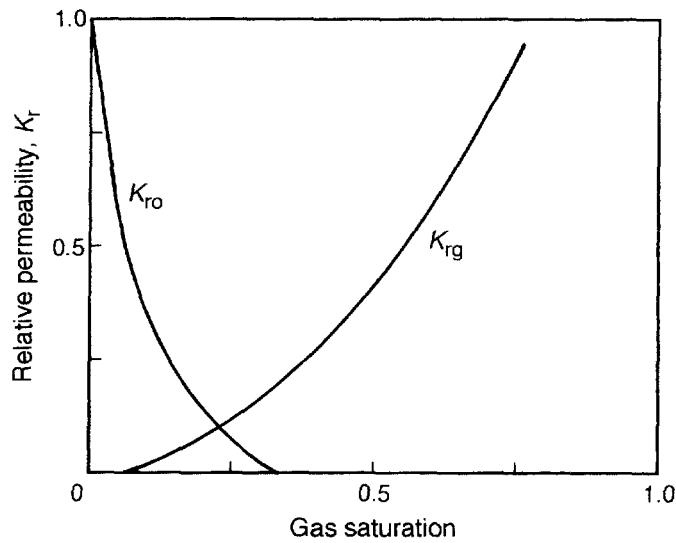
$P_{cw}$	$S_w$	$P_{cg}$	$S_g$
0.50	0.900	0.00	0.000
1.50	0.700	0.50	0.100
3.00	0.400	1.50	0.300
5.00	0.273	3.00	0.600
7.00	0.220	4.00	0.800

and at the base of the reservoir model in order to simulate no-flow boundary conditions. It was assumed that no water was produced at the well, and the produced gas–oil ratio at the surface was calculated by using

$$\text{GOR} = R_s + \frac{k_{rg}\mu_o B_o}{k_{ro}\mu_g B_g} \quad (8.25)$$



**Figure 8.7** Relative permeabilities for water ( $K_{rw}$ ) and oil ( $K_{ro}$ )

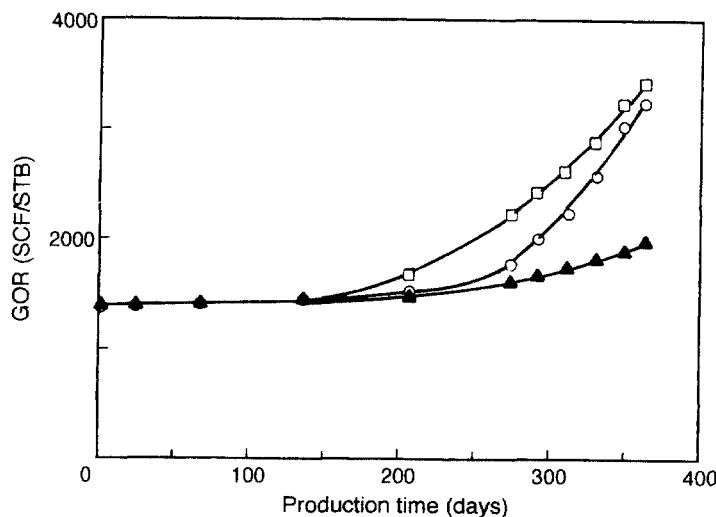


**Figure 8.8** Relative permeabilities for gas ( $K_{rg}$ ) and oil ( $K_{ro}$ )

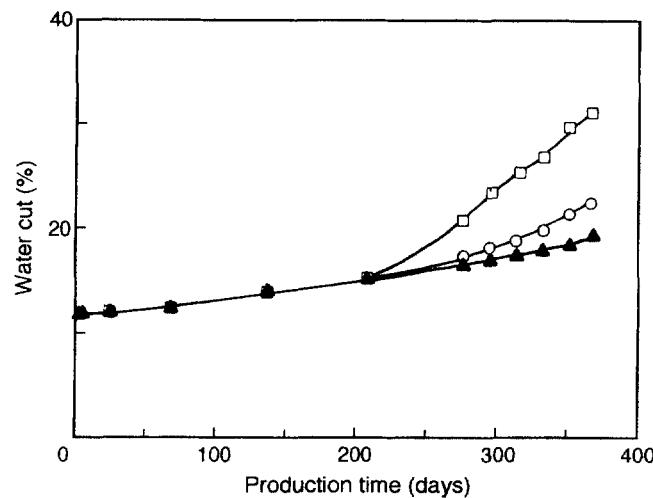
Similarly, the water cut ratio can be calculated from

$$\text{WOR} = \frac{k_{rw}\mu_o B_o}{k_{ro}\mu_w B_w} \quad (8.26)$$

The results obtained from all runs are shown in Figures 8.9 to 8.11. Figures 8.9 and 8.10 indicate that reservoir compaction reduces GOR and WOR at the surface, hence there is

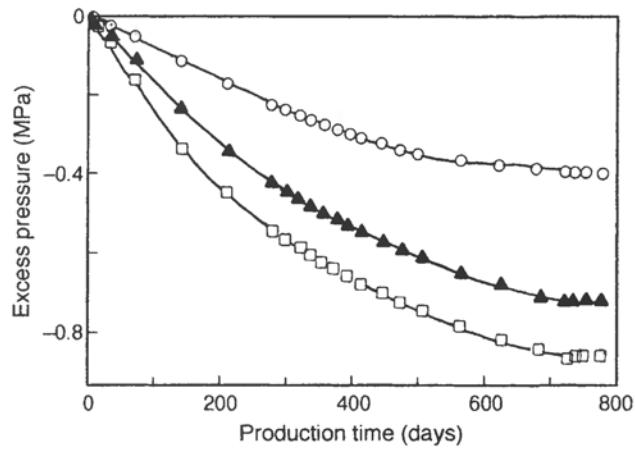


**Figure 8.9** Gas/oil ratio (GOR) versus time for different reservoir rocks: (□)  $E = 0.9631 \times 10^{10}$ , (○)  $E = 0.9631 \times 10^8$ , (▲)  $E = 0.9631 \times 10^6$



**Figure 8.10** Water cut versus time for different reservoir rocks: (□)  $E = 0.9631 \times 10^{10}$ , (○)  $E = 0.9631 \times 10^8$ , (▲)  $E = 0.9631 \times 10^6$

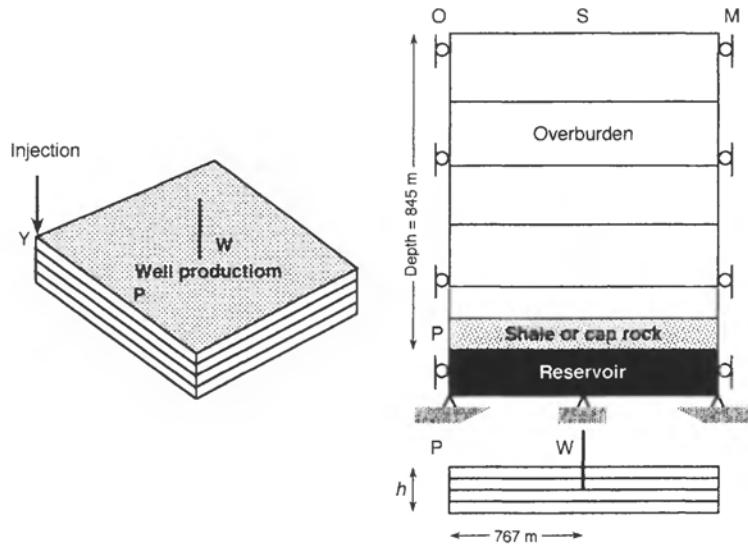
higher oil recovery. Figure 8.11 shows the average pressure drop versus production time for different values of Young's modulus  $E$  (as given in the same figure). This indicates that the rock compaction, which occurs during the oil production, has reduced the average pressure drop in the reservoir system.



**Figure 8.11** Effects of reservoir compaction on reservoir pressure: (◻)  $E = 0.9631 \times 10^9$ , (▲)  $E = 0.9631 \times 10^6$ , (○)  $E = 0.9631 \times 10^5$

### 8.5.3 Surface Subsidence Analyses

The last stage of simulation presented in this chapter is to indicate the utility of the developed finite element code for predicting the surface subsidence above a compacting saturated oil reservoir. To illustrate this application, several simulations were performed using a hypothetical model which consists of a closed reservoir system, cap rock and the overburden layers (Figure 8.12). These simulations are discussed in detail in the



**Figure 8.12** A hypothetical model for subsidence analyses

**Table 8.4** Material properties for the subsidence analyses

Property	Reservoir	Cap rock	Overburden
$E(\text{N m}^{-2})$	$0.9631 \times 10^8$	$0.7631 \times 10^8$	$0.3231 \times 10^7$
$\nu$	0.25	0.23	0.20
$k_{x,y,z}(\text{mD})$	100	0.0	60

**Table 8.5** Time steps for different intervals

Time interval (days)	Number of time steps
0.070	1
0.093	1
5.000	5
10.000	8
25.000	8
50.000	8
75.000	4
100.000	4
125.000	8
150.000	8
200.000	17

**Table 8.6** Well production data

Time interval (days)	Well rate (STB per day)
0–1	1–50
1–20	100–400
21–7000	500

following sections. The fluid data for the saturated oil reservoir was obtained from Lewis and Sukirman [5,6] and the relevant material properties are given in Table 8.4.

In all simulations, cross-sections OP and MN can be either flow or no-flow boundary conditions, and the remaining sections within the system were assumed to be no-flow boundaries (Figure 8.12). It was assumed that the vertical displacement at the base is zero, i.e. a rigid base rock. Four different simulations were conducted to investigate the effect of several modelling parameters on the resulting surface subsidence. These factors are discussed in detail below. Similar results have also been reported by Morgan *et al.* [8], Lewis and Sukirman [6] and Finol *et al.* [10].

The time steps for different intervals and the production data used in all cases are shown in Tables 8.5 and 8.6, respectively. Other reservoir data such as the fluid and the rock properties were obtained from the literature [5,6].

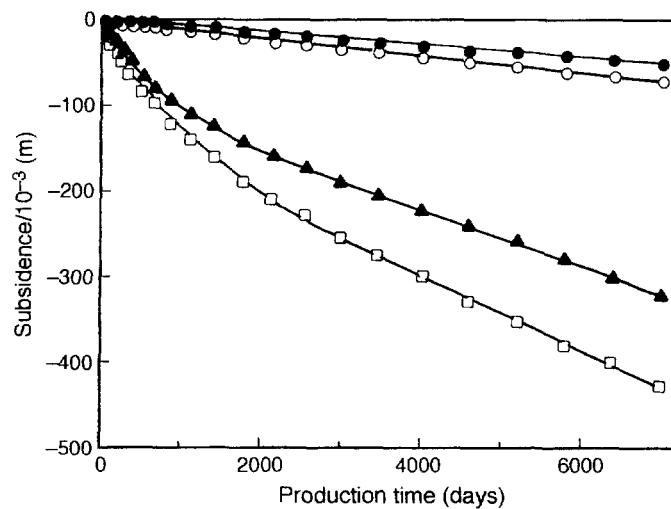
### 8.5.3.1 Effect of Reservoir Parameters

The simulations of reservoir compaction and the consequent surface subsidence for different types of reservoir are presented here. In all cases the reservoirs were simulated using the same fluid properties but with different modelling parameters.

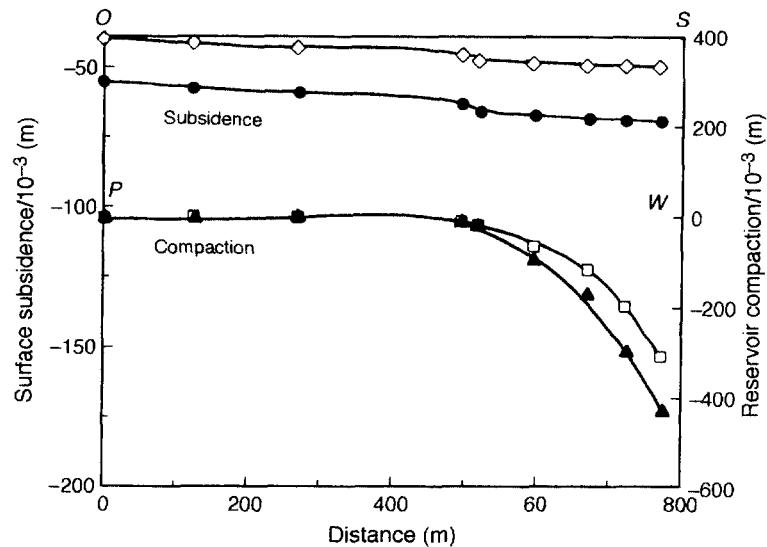
Case 1 is the simulation of two compacting reservoirs with different values of formation thickness. The first reservoir has a formation thickness of 70 m and produced at a constant rate of 1080 STM per day. For the second reservoir, a constant rate of 800 STB per day was produced from a formation thickness of 52 m. The lower production rate for the second reservoir was used to compensate the difference in thickness and to obtain the same production performance. Figure 8.13 illustrates the effect of formation thickness on vertical displacements at two locations, points W and S, respectively representing the reservoir compaction and the consequent surface subsidence. Notice that the subsidence at point S increases by 35 percent when the formation thickness is increased by 38 percent. The two compaction bowls are almost identical in shape (Figure 8.14).

Case 2 simulates the surface subsidence of two similar reservoirs with the same amount of compaction, but located at different depths. This section was used to investigate the effect of the depth of overburden layers on the subsidence predictions. For case 2 the reservoir of case 1 with a formation thickness 70 m, overburden depth of 8.45 m (including the cap rock) and a constant production rate of 1080 STB per day was used as a reference. The studied reservoir is similar to the reference reservoir, except the overburden depth was reduced to 422.5 m.

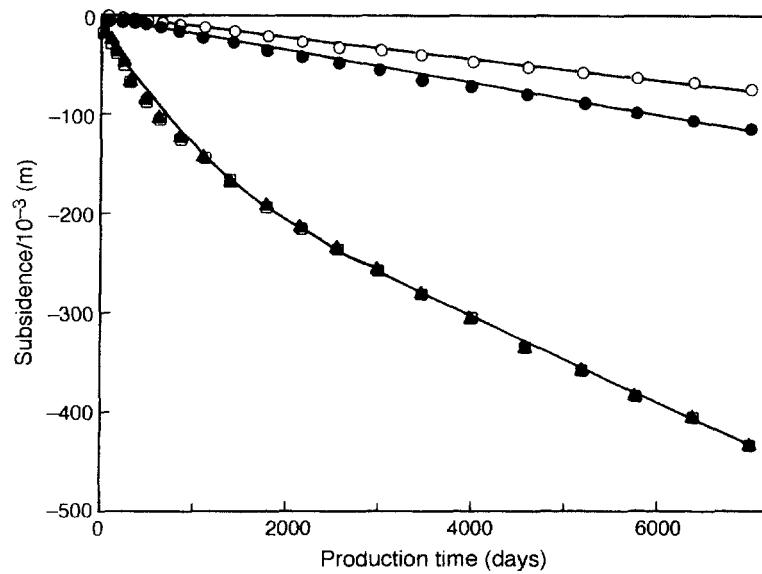
As shown in Figure 8.15, the surface subsidence at S increases by about 50 percent for the same percentage of reduction in the overburden depth. However, the formation



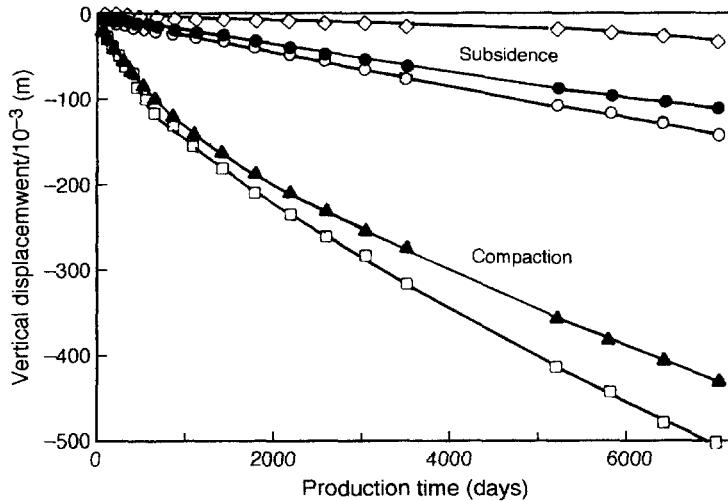
**Figure 8.13** Vertical displacement versus time for different reservoir thicknesses: (●) at S for  $h = 52$  m; (▲) at top of reservoir for  $h = 52$  m; (○) at S for  $h = 70$  m; (□) at top of reservoir for  $h = 70$  m



**Figure 8.14** Vertical displacement profiles at  $t = 19$  years for different reservoir thicknesses: (●) at S for  $h = 70$  m; (◊) at S for  $h = 52$  m; (▲) at top of reservoir for  $h = 70$  m; (□) at top of reservoir for  $h = 52$  m



**Figure 8.15** Vertical displacement versus time for case 2: (●) at S for depth = 1500 ft; (▲) at top of reservoir for depth = 1500 ft; (○) at S for depth = 3000 ft; (△) at top of reservoir for depth = 3000 ft



**Figure 8.16** Vertical displacement versus time for case 3: (●) at S for  $E = 0.96 \times 10^5$  kPa; (▲) at top of reservoir for  $E = 0.96 \times 10^5$  kPa; (○) at S for  $E = 0.15 \times 10^5$  kPa; (□) at top of reservoir for  $E = 0.15 \times 10^5$  kPa; (◊) at S for  $E = 0.96 \times 10^6$  kPa

compaction curves are similar for both reservoirs. This indicates that the overburden depth has no influence on the studied compaction profile.

Case 3 uses three different values of Young's modulus in order to investigate the influence of different degrees of compaction on the predicted subsidence at the surface. The reservoir for case 2, with the material properties in Tables 8.4 and 8.5, was taken as a reference. In case 3 the studied reservoir was similar to the reference reservoir, except it had different values of  $E$ , i.e.  $0.1631 \times 10^8$  and  $0.9631 \times 10^9$  kPa.

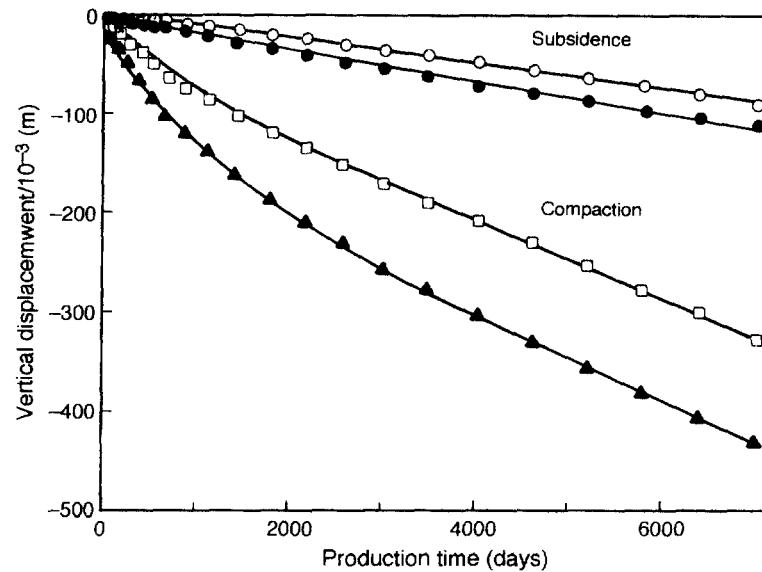
The results are plotted in Figure 8.16. Notice that the subsidence at S increases with the lower value of  $E$  and vice versa. This is an expected result because a higher value of  $E$  implies a more rigid formation rock, which tends to incompressible behaviour, i.e. deformation can be neglected.

Case 4 predicts subsidence at the surface for reservoirs with different areal extents. As presented in case 2, the reservoir with an overburden depth of 422.5 m was used as the reference. The reservoir studied in this case had an area equal to half the reference area (the width and length of the reservoir were reduced to 300 m and 512 m respectively).

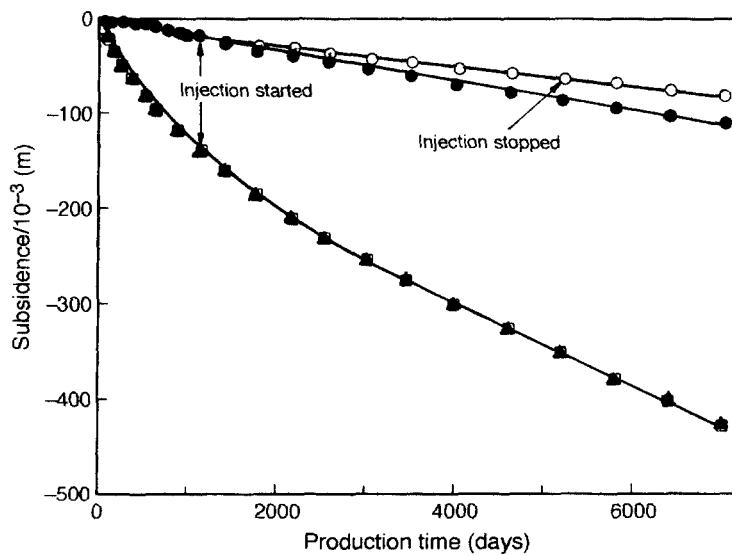
Figure 8.17 compares the vertical displacement profiles at points S and W for the studied reservoirs. The amount of subsidence at point S decreases when the area of the reservoir is reduced by 50 percent. However, the formation compaction curves at point W are identical for both reservoirs.

### 8.5.3.2 Effect of Water Injection Scheme on Subsidence Analysis

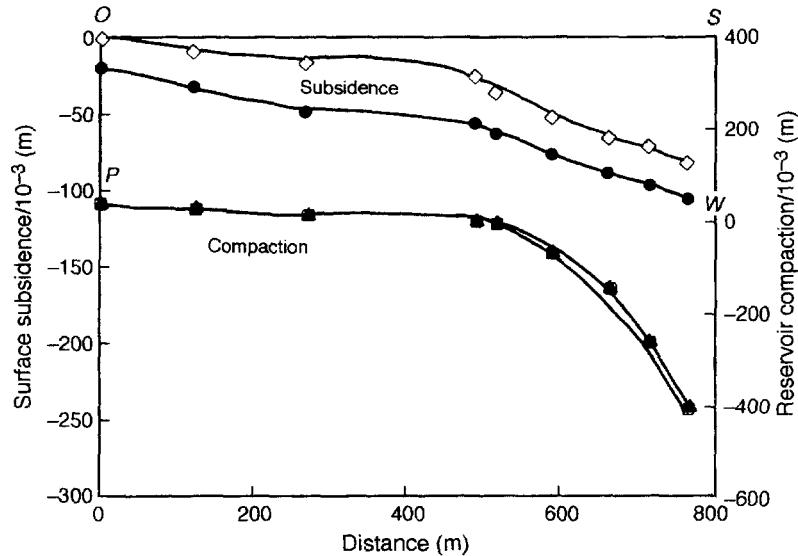
The reservoir studied here is the same as in the previous cases (data in Table 8.4) but it has a water injection well located at point Y. The water was injected at a constant rate of 1000 STB per day after  $t = 3$  years and was terminated at  $t = 16$  years (Figure 8.18).



**Figure 8.17** Vertical displacement versus time for case 4: (●) at S for area =  $A$ ; (▲) at top of reservoir for area =  $A$ ; (○) at S for area =  $\frac{1}{2}A$ ; (□) at top of reservoir for area =  $\frac{1}{2}A$



**Figure 8.18** Vertical displacement versus time for water injection: (●) at S for no injection; (▲) at top of reservoir for no injection; (○) at S for injection at Y; (□) at top of reservoir for injection at Y



**Figure 8.19** Vertical displacement profiles at  $t = 19$  years for water injection: ( $\bullet$ ) at S for no injection; ( $\diamond$ ) at S with injection; ( $\blacktriangle$ ) at top of reservoir for no injection; ( $\square$ ) at top of reservoir with injection

Figures 8.18 and 8.19 show the profiles of reservoir compaction and subsidence at the surface with time. These results indicate that the water injection process can be performed in order to minimise compaction and subsidence problems. In practice the injected fluid reduces the pressure decrease in the reservoir system, hence it produces minimum subsidence at the surface.

## 8.6 CONCLUSIONS

A numerical simulation using a finite element model has been presented for simulating the surface subsidence above a compacting saturated oil reservoir. Numerous studies related to surface subsidence problems have been reported in the literature. In many complex geometrical cases, numerical models have been applied to investigate the problems. The present implementation indicates the versatility and stability of the finite element method when applied to such problems.

In this chapter, a fully coupled implicit formulation using a finite element method has been applied for simulating three immiscible and compressible fluids flowing in a deforming saturated oil reservoir. The derivation of the four governing equations considered the equilibrium equation and the continuity of the fluid flow. These balance equations have taken into account the effects of capillarity and the variation of relative permeability. A Mohr–Coulumb yield surface was used for the elastoplastic soil model.

The application of a finite element spatial discretisation resulted in a semidiscrete form of the equations in time. The developed model has been validated using a nonlinear consolidation soil column problem first presented by Li *et al.* [16]. For both two- and three-phase tests, the present results agreed well with those obtained by Li *et al.*

From the results obtained for the unconsolidated reservoir formation, it appears that the consolidation problem can influence the overall production performance. The last application indicates the utility of the present finite element code in predicting the surface subsidence above a compacting saturated oil reservoir in various situations.

## REFERENCES

1. Terzaghi, K. (1943) *Theoretical Soil Mechanics*, Wiley, New York.
2. Biot, M. A. (1941) General theory of three-dimensional consolidation. *J. Appl. Phys.*, **12**, 155–64.
3. Polillo Filho, A., Maia da Costa A. and Favilla Ebecken, N. F. (1988) An analysis of hydraulic fracturing by the finite element method, in *Numerical Methods in Geomechanics*, G. Swoboda (ed.), Innsbruck, Rotterdam.
4. Vandamme, L. M. and Roegiers, J. C. (1990) Poroelasticity in hydraulic fracturing simulators. *JPT, Sept.*, 1199–1203.
5. Lewis, R. W. and Sukirman, Y. (1993) Finite element modelling of three-phase flow in deforming saturated oil reservoirs. *Int. J. Num. Anal. Meth. Geomech.*, **17**, No. 8, 577–98.
6. Lewis, R. W. and Sukirman, Y. (1994) Finite element modelling for simulating the surface subsidence above a compacting oil reservoir, *Int. J. Anal. Num. Meth. Geom.*, **18**, No. 9, 619–39.
7. Ertekin, T. and Farouq Ali, S. M. (1979) Numerical simulation of the compaction–subsidence phenomena in a reservoir for two-phase non-isothermal flow conditions, in *Proceedings of the Third International Conference on Numerical Methods in Geomechanics*, Aachen, 263–274.
8. Morgan, K., Lewis, R. W. and White, I. R. (1980) The mechanisms of ground surface subsidence above compacting reservoirs and their analysis by finite element method. *Appl. Math. Model.*, **4**, 217–24.
9. Kosloff, D., Scott, R. F. and Scranton, J. Finite element simulation of Wilmington oil field subsidence: II, non linear modelling. *Tectonophysics*, **70**, 159–83.
10. Finol, A. and Farouq Ali, S. M. (1975) Numerical simulation of oil production with simultaneous ground subsidence. *SPE J.*, **15**, 411–24.
11. Merle, H. A., Kentie, C. J. P., van Opstal, G. H. C. and Schreider G. M. G. (1976) The Bachaquero study: a composite analysis of the behaviour of a compaction drive/solution gas drive reservoir. *JPT, Sept.*, 1107–14.
12. Geertsma, J. (1973) Land subsidence above compacting oil and gas reservoirs. *JPT, June*, 734–44.
13. Peacemen, D. W. (1977) A nonlinear stability analysis for difference equations using semi-implicit mobility. *SPE J.*, **Feb.**, 79–91.
14. Crichlow, H. B. (1976) *Modern Reservoir Engineering: A Simulation Approach*. Prentice Hall, Englewood Cliffs NJ.
15. Aziz, K. and Settari, T. (1979) *Petroleum Reservoir Simulation*, Applied Science Publishers, London, pp. 30–38.
16. Li, X. K., Zienkiewicz, O. C. and Xie, Y. M. (1990) A numerical model for immiscible two-phase fluid flow in a porous medium and its time domain. *Int. J. Num. Meth. Engng.*, **30**, No. 6, 1195–1212.
17. Chappellear, J. E. and Nolen, J. S. (1986) Second Comparative Solution project: a three-phase coning study. *JPT, Mar.*, 345–53.
18. Sukirman, Y. (1993) Petroleum reservoir simulation coupling flow and subsidence. PhD thesis, University College of Swansea.

# 9

## Fractured Reservoir Simulation

### 9.1 INTRODUCTION

A large number of existing petroleum reservoirs may be categorized as *naturally fractured*. By *fractured* we usually refer to a reservoir containing a continuous network of fissures and faults which are created by tectonic activities over millions of years of the reservoir-forming process. Examples of fractured reservoirs are the Iranian fields, Ekofisk (Norway), Kirkuk (Iraq), Dukhan field (Qatar) [9], Rospo (Italy) and Emposta (Spain) [33]. The occurrence of this type of reservoir has generated an enormous research effort and a vast amount of money has been devoted to the simulation of naturally fractured reservoirs. However, it is still one of the most challenging fields in the area of simulation.

The behaviour of fractured reservoirs is considerably different from the behaviour of conventional reservoirs because of different constitutive laws that govern the fluid flow within a reservoir. Therefore, fractured reservoir simulation involves different theories and approaches which have made it much more complex and challenging.

Numerical models with different degrees of sophistication have been developed in the past to simulate flow through fractured porous media [6,11,12,17,23]. The initial models were generally very simple and were mainly based on the single-porosity, or continuum, concept. In these models, a fractured system was grossly treated as an equivalent anisotropic continuum. The physical quantities such as fluid pressure, porosity and permeability were averaged over representative blocks of fractured medium, containing a sufficiently large number of fractures. The single-porosity approach is associated with a number of drawbacks [29,48]: the identification of representative blocks and the determination of equivalent permeability values. Moreover, because of the averaging procedure involved, the field variables in these models, e.g. fluid pressure and fluid flux distribution, neither truly represent the situation for fractured or continuum blocks. In fact, the application of single-porosity models could only be justified if one is dealing with large-scale flow problems, and a detailed description of field variables within the fractured media is not of particular concern.

A major departure from the single-porosity approach was made by Barenblatt *et al.* [6] and later by Warren and Root [47]. They modelled a rigid (non-deformable) highly fractured porous medium as an entity having two porosities: one representing a fracture network and the other a continuum porous block. The fracture network was

characterised by very high permeability and very low storage, whereas the porous blocks were characterised by low permeability and high storage. They defined their model in terms of two conservation equations, one for each of the foregoing porosity regions. These equations were coupled by a leakage term expressed as a non-linear source term dependent upon the potential difference between the fractured network and the porous blocks. The solution of these equations at every point within the fractured medium yields two values: one representing the average pressure in the fracture network and the other in the porous blocks.

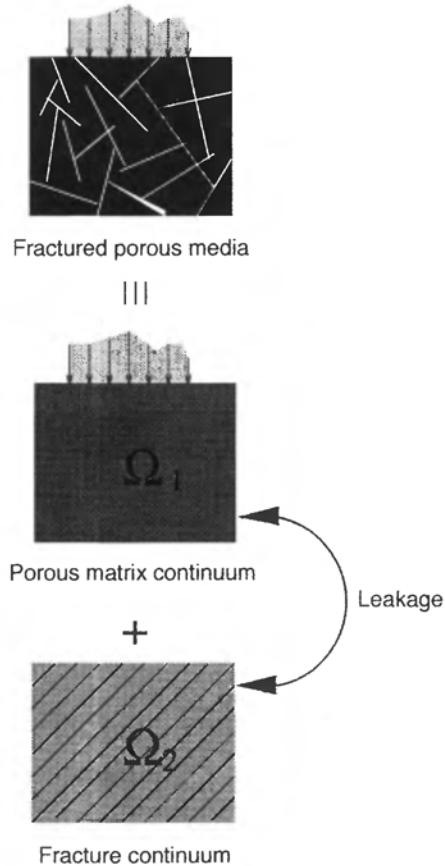
Barenblatt's work was an important step forward in modelling flow through a rigid fractured medium, as unlike single-porosity models, it was able to distinguish between the porous blocks and fractured networks. This work was then followed by other researchers for different aspects. More recently Khalili [25] and Ghafouri *et al.* [20,21,26] incorporated the flow–deformation coupling effects in the double-porosity models. The studies identified the strong coupling between fluid flow and solid deformation.

The simulation of fractured reservoirs is discussed in detail in the following sections which present a highly sophisticated double-porosity model. This model is used to integrate the reservoir description and the analysis of production mechanisms so as to provide production as well as surface subsidence forecasts.

## 9.2 DESCRIPTION OF THE MODEL

The basic mechanism of fluid flow in a fractured porous medium may be explained as follows: the applied external loads and/or well production create a pressure gradient between the fluid within the matrix pores and the fluid in the adjacent fractures. The fluid within the matrix is squeezed out into the fissured continuum due to the produced gradient. Subsequently, flow towards the producing well takes place through the fissured network. In the present model, the fractured porous medium is divided into two overlapping but distinct continua: the first represents flow and deformation in the porous matrix, and the second represents flow in the fissures (Figure 9.1). These two subdomains have the following characteristics:

1. The fluid flow within each subdomain is independent of the flow in the other subdomain and any coupling between the fluid flow in the porous matrix and the fissured network is controlled only via the *leakage* term, i.e. the fluid expelled from the matrix which enters the fissures and vice versa. Hence fluid pressure, saturation, porosity, permeability and the other properties of both the soil and the fluids within the two subdomains are considered separately.
2. Within the first continuum, the fluid flow is assumed to be coupled to the matrix deformation. This coupling is controlled via the rate of change of volumetric strain  $\partial\epsilon_{v1}/\partial t$ .
3. It is assumed that the deformation of the first subdomain is only due to the change of total volumetric strain of the porous matrix and any change of volume of the solid particles is ignored.
4. Two different porosity values should be defined for the porous matrix and fissured region respectively, hence the term *double-porosity model*. These are defined as



**Figure 9.1** Schematic representation of double-porosity model

follows:

$$n_1 = \frac{V_1}{V} \quad (9.1)$$

$$n_2 = \frac{V_2}{V} \quad (9.2)$$

where  $n_1$  and  $n_2$  are the porosity values of the porous matrix and fissured network respectively,  $V_1$  and  $V_2$  are the pore and fissured volume and  $V$  is the total volume in a representative elementary volume (REV).

5. We limit ourselves to the case of porous fractured reservoirs, where the volume of the fractures is normally a small fraction of the total void space and therefore it is assumed that boundary and body forces are carried only by the porous matrix subdomain and the contribution of the fracture continuum is ignored [33,3].
6. Also, for the same reason, it is assumed that the compressibility of the fractured network does not alter the compressibility of the whole porous medium dramatically and may be ignored [33,3].

7. Fracture permeability is determined from the geometry, particularly the aperture. If the fracture compressibility is neglected then the aperture, hence the fracture absolute permeability, do not vary.
8. Imbibition and depletion of the fluid is via the fissured network subdomain only.
9. The two subdomains are assumed to be fully saturated.

### 9.3 DEVELOPMENT OF THE GOVERNING EQUATIONS

In a manner similar to Chapter 8, the equilibrium equation which governs the solid-phase behaviour can be derived. However, bearing in mind the applied loads are carried only by the matrix (assumption 5 in Section 9.2), we develop the equilibrium equation for the matrix continuum only, which is denoted by the subscript 1 throughout the text. This results in the following;

$$\int_{\Omega_1} \delta \boldsymbol{\varepsilon}_1^T \mathbf{D}_T \frac{\partial \boldsymbol{\varepsilon}_1}{\partial t} d\Omega_1 - \int_{\Omega_1} \delta \boldsymbol{\varepsilon}_1^T S''_{w1} \mathbf{m} \frac{\partial p_{w1}}{\partial t} d\Omega_1 - \int_{\Omega_1} \delta \boldsymbol{\varepsilon}_1^T S''_{o1} \mathbf{m} \frac{\partial p_{o1}}{\partial t} d\Omega_1 \\ - \int_{\Omega_1} \delta \boldsymbol{\varepsilon}_1^T S''_{g1} \mathbf{m} \frac{\partial p_{g1}}{\partial t} d\Omega_1 + \frac{\partial \hat{\mathbf{f}}}{\partial t} = 0 \quad (9.3)$$

where

$$d\hat{\mathbf{f}} = - \int_{\Omega} \delta \mathbf{u}^T d\mathbf{b} d\Omega - \int_{\Gamma} \delta \mathbf{u}^T d\hat{\mathbf{t}} d\Gamma + \int_{\Omega} \delta \boldsymbol{\varepsilon}^T d\sigma_o d\Omega \quad (9.4)$$

The various terms used in these equations have the same definition given in the preceeding chapters. Also, continuity equations for various fluid phases could be derived as described in Chapter 8. However, note that the conservation of fluid mass within each subdomain, i.e. the matrix and the fractured network represented by two different porosity values  $n_1$  and  $n_2$ , requires this equation to be written for each subdomain distinctly. The two continuity equations are then combined with each other via the transfer function, or the so-called ‘leakage’ term, as follows:

$$\nabla^T \left( \frac{-\rho_{iz} \mathbf{K}_z}{\mu_{iz}} \nabla p_{iz} \right) + (-1)^{z+1} \rho_{iz} \frac{\bar{\alpha} \mathbf{K}_1}{\mu_{il}} (p_{il} - p_{iz}) + (\alpha - 1) \rho_{stc} (Q_i)_{stc} \\ + \rho_{iz} \phi_z \frac{\partial S_{iz}}{\partial t} + \phi_z S_{iz} \frac{\partial p_{iz}}{\partial t} + (2 - \alpha) \rho_{iz} S_{iz} \mathbf{m}^T \frac{\partial \boldsymbol{\varepsilon}_z}{\partial t} = 0 \quad (9.5)$$

where the various terms have already been defined. An inspection of (9.5) reveals that the leakage term, defined as

$$\pm \rho_{il} \frac{\bar{\alpha} \mathbf{K}_1}{\mu_i} (p_{il} - p_{iz}) \quad (9.6)$$

where  $\bar{\alpha}$ , a coefficient which depends on the fracture width and geometry [47], describes the fluid exchange between the matrix domain and the fractured domain due to the

pressure gradient between the porous matrix and the fractured network. In fact, equation (9.5) exhibits six distinct relationships for  $i = o, w, g$  and  $\alpha = 1, 2$ . On substituting equivalent terms into (9.5) for various variables, the continuity equation for each phase could be written separately, as follows.

For the oil phase:

$$\nabla^T \left( \frac{-\mathbf{k}_o k_{ro\alpha}}{\mu_o B_{o\alpha}} \nabla p_{o\alpha} \right) + (-1)^{\alpha+1} \frac{\bar{\alpha} \mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} (p_{o1} - p_{o2}) + (\alpha - 1)(Q_o)_{stc} \\ + \lambda_{oo\alpha} \frac{\partial p_{o\alpha}}{\partial t} + \lambda_{wo\alpha} \frac{\partial p_{w\alpha}}{\partial t} + \lambda_{go\alpha} \frac{\partial p_{g\alpha}}{\partial t} + (2 - \alpha) \frac{S_{o\alpha}}{B_{o\alpha}} \mathbf{m}^T \frac{\partial \boldsymbol{\varepsilon}_\alpha}{\partial t} = 0 \quad (9.7)$$

where

$$\lambda_{oo\alpha} = \frac{\phi_\alpha}{B_{o\alpha}} \left( S'_{g\alpha} - S'_{w\alpha} + S_{o\alpha} B_{o\alpha} \left( \frac{1}{B_{o\alpha}} \right)' \right) \quad (9.8)$$

$$\lambda_{wo\alpha} = \frac{\phi_\alpha}{B_{o\alpha}} S'_{w\alpha} \quad (9.9)$$

$$\lambda_{go\alpha} = -\frac{\phi_\alpha}{B_{o\alpha}} S'_{g\alpha} \quad (9.10)$$

For the water phase:

$$\nabla^T \left( \frac{-\mathbf{k}_w k_{rw\alpha}}{\mu_w B_{w\alpha}} \nabla p_{w\alpha} \right) + (-1)^{\alpha+1} \frac{\bar{\alpha} \mathbf{k}_1 k_{rw1}}{\mu_w B_{w1}} (p_{w1} - p_{w2}) + (\alpha - 1)(Q_w)_{stc} \\ + \lambda_{ww\alpha} \frac{\partial p_{w\alpha}}{\partial t} + \lambda_{ow\alpha} \frac{\partial p_{o\alpha}}{\partial t} + \lambda_{gw\alpha} \frac{\partial p_{g\alpha}}{\partial t} + (2 - \alpha) \frac{S_{w\alpha}}{B_{w\alpha}} \mathbf{m}^T \frac{\partial \boldsymbol{\varepsilon}_\alpha}{\partial t} = 0 \quad (9.11)$$

where

$$\lambda_{ww\alpha} = -\frac{\phi_\alpha}{B_{w\alpha}} S'_{w\alpha} + \phi_\alpha S_{w\alpha} \left( \frac{1}{B_{w\alpha}} \right)' \quad (9.12)$$

$$\lambda_{ow\alpha} = \frac{\phi_\alpha}{B_{w\alpha}} S'_{w\alpha} \quad (9.13)$$

$$\lambda_{gw\alpha} = 0 \quad (9.14)$$

And finally, for the gas phase:

$$\nabla^T \left( \frac{-\mathbf{k}_g k_{rg\alpha}}{\mu_g B_{g\alpha}} \nabla p_{g\alpha} \right) + \nabla^T \left( R_{so\alpha} \frac{-\mathbf{k}_o k_{ro\alpha}}{\mu_{o\alpha} B_{o\alpha}} \nabla p_{o\alpha} \right) + (-1)^{\alpha+1} \frac{\bar{\alpha} \mathbf{k}_1 k_{rg1}}{\mu_g B_{g1}} \\ (p_{g1} - p_{g2}) + (-1)^{\alpha+1} R_{so\alpha} \frac{\bar{\alpha} K_1 k_{ro1}}{\mu_{o1} B_{o1}} (p_{o1} - p_{o2}) + (\alpha - 1)[(Q_g)_{stc} + R_{so\alpha}(Q_o)_{stc}] \\ + \lambda_{gg\alpha} \frac{\partial p_{g\alpha}}{\partial t} + \lambda_{og\alpha} \frac{\partial p_{o\alpha}}{\partial t} + \lambda_{wg\alpha} \frac{\partial p_{w\alpha}}{\partial t} + (2 - \alpha) C_{go\alpha} \mathbf{m}^T \frac{\partial \boldsymbol{\varepsilon}_\alpha}{\partial t} = 0 \quad (9.15)$$

where

$$\lambda_{gg\alpha} = \phi_\alpha S_{g\alpha} \left( \frac{1}{B_{g\alpha}} \right)' + \phi_\alpha \left( \frac{1}{B_{g\alpha}} - \frac{R_{so\alpha}}{B_{o\alpha}} \right) S'_{g\alpha} \quad (9.16)$$

$$\begin{aligned} \lambda_{og\alpha} &= \phi_\alpha R_{so\alpha} S_{o\alpha} \left( \frac{1}{B_{o\alpha}} \right)' - \frac{\phi_\alpha}{B_{g\alpha}} S'_{g\alpha} \\ &\quad + \frac{R_{so\alpha}}{B_{o\alpha}} (\phi_\alpha S'_{g\alpha} - \phi_\alpha S'_{w\alpha}) + \phi_\alpha \frac{S_{o\alpha}}{B_{o\alpha}} R'_{so\alpha} \end{aligned} \quad (9.17)$$

$$\lambda_{wg\alpha} = \phi_\alpha \frac{R_{so\alpha}}{B_{o\alpha}} S'_{w\alpha} \quad (9.18)$$

$$C_{go\alpha} = \frac{S_{g\alpha}}{B_{g\alpha}} + R_{so\alpha} \frac{S_{o\alpha}}{B_{o\alpha}} \quad (9.19)$$

Equations (9.3), (9.7), (9.11) and (9.15) represent highly non-linear partial differential equations for multiphase flow coupled with the consolidation behaviour occurring in a deforming fractured oil reservoir. The major non-linearities, i.e. the phase saturation  $S_i$ , relative permeability  $k_r$ , formation volume factor  $B_i$ , gas/oil solution ratio  $R_{so}$ , and viscosity  $\mu_i$ , strongly depend on the unknowns and therefore should be updated over time.

#### 9.4 A SPECIAL CASE: SINGLE-PHASE FLOW IN A DEFORMING FRACTURED POROUS MEDIUM

The case of only one fluid, i.e. water, flowing at a saturated condition is considered here. This case is frequently encountered when analysing fractured aquifers and therefore is worth mentioning. In this case the water pressure and displacements are the only unknowns in the formulations, hence the number of continuity equations reduces to one. The equilibrium equation may be derived by recalling (9.3). The equation may be further simplified as the only existing fluid within the porous medium is water, hence subscript  $i$  can be replaced by  $w$  throughout. Also, the following expressions may be written:

$$S_o = S_g = S''_o = S''_g = 0 \quad (9.20)$$

$$S_w = S''_w = 1 \quad (9.21)$$

$$\frac{\partial p_o}{\partial t} = \frac{\partial p_g}{\partial t} = 0 \quad (9.22)$$

$$\bar{p} = p_w \quad (9.23)$$

Incorporating (9.20) to (9.23) into (9.3), the simplified form of the equilibrium equation is as follows:

$$\int_{\Omega_1} \delta \boldsymbol{\epsilon}_1^T \mathbf{D}_T \frac{\partial \boldsymbol{\epsilon}_1}{\partial t} d\Omega_1 - \int_{\Omega_1} \delta \boldsymbol{\epsilon}_1^T \mathbf{m} \frac{\partial p_w}{\partial t} d\Omega_1 + \frac{\partial \hat{\mathbf{f}}}{\partial t} = 0 \quad (9.24)$$

where

$$d\hat{\mathbf{f}} = - \int_{\Omega} \delta U^T d\mathbf{b} d\Omega - \int_{\Gamma} \delta \mathbf{u}^T d\hat{\mathbf{t}} d\Gamma + \int_{\Omega} \delta \boldsymbol{\varepsilon}^T d\sigma_o d\Omega \quad (9.25)$$

For this particular case the number of continuity equations reduces to one. The continuity equation may be derived by recalling equation (9.5). In the following analysis, the relative permeability  $k_{rw}$  is assumed equal to unity. Furthermore, the time derivative of the water density may be replaced by

$$\frac{1}{\rho_w} \frac{\partial \rho_w}{\partial t} = \frac{1}{K_w} \frac{\partial p_w}{\partial t} \quad (9.26)$$

where  $K_w$  is the bulk modulus of water. Similarly, the isothermal compressibility of water

$$C_w = - \frac{1}{\rho_w} \frac{\partial \rho_w}{\partial p_w} = \frac{-1}{K_w}$$

has been used in the literature. However, we use bulk modulus instead of compressibility in this case, as they can be used interchangeably. Returning to (9.26), dividing by  $\rho_w$  and incorporating it in (9.5), we obtain

$$\begin{aligned} \nabla^T \left( \frac{-\mathbf{K}_w}{\mu_w} \nabla p_{w\alpha} \right) + (-1)^{\alpha-1} \frac{\bar{\alpha} \mathbf{K}_1}{\mu_w} (p_{w1} - p_{w2}) + (\alpha - 1) Q_w \\ + \frac{\phi_\alpha}{K_w} \frac{\partial p_w}{\partial t} + (2 - \alpha) \mathbf{m}^T \frac{\partial \boldsymbol{\varepsilon}_\alpha}{\partial t} = 0 \end{aligned} \quad (9.27)$$

Boundary conditions, similar to the previous case, may also be applied to the two governing equations for single-phase flow.

## 9.5 DISCRETISATION IN SPACE

The discretisation process of the resulting governing equations may be implemented based on the method discussed in Chapters 3 and 8 and will include two distinct stages: discretisation in the space domain and discretisation in the time domain. As the first step, the unknown variables possessing continuous values should be approximated in terms of a limited number of nodal values. Hence the expressions for  $\mathbf{u}$ ,  $\boldsymbol{\varepsilon}$ ,  $p_1$  and  $p_2$  in equations (9.3), (9.7), (9.11) and (9.15) can now take the form

$$\begin{aligned} \mathbf{u}_1 &\approx \bar{\mathbf{u}}_1 = \mathbf{N} \bar{\mathbf{u}}_1 \rightarrow \boldsymbol{\varepsilon}_1 = \mathbf{L} \mathbf{u}_1 \approx \mathbf{L} \mathbf{N} \bar{\mathbf{u}}_1 = \mathbf{B} \bar{\mathbf{u}}_1 \\ p_1 &\approx \hat{p}_1 = \bar{\mathbf{N}} \bar{\mathbf{p}}_1 \\ p_2 &\approx \hat{p}_2 = \bar{\mathbf{N}} \bar{\mathbf{p}}_2 \end{aligned} \quad (9.28)$$

where  $\mathbf{L}$  is the matrix relating the displacements  $\mathbf{u}_1$  to the strains  $\boldsymbol{\varepsilon}_1$  and is defined as follows:

$$\mathbf{L} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \end{bmatrix} \quad (9.29)$$

Note that, in the finite element procedure, the shape functions used for the various variables existing in the formulation are not necessarily the same. Hence the two different types are distinguished by  $\mathbf{N}$  and  $\bar{\mathbf{N}}$  in (9.28).

Substituting the expression given in (9.28) into the weak form of the equilibrium equation will produce

$$\begin{aligned} \delta\bar{\mathbf{u}}_1^T \left\{ \left( \int_{\Omega_1} \mathbf{B}_1^T \mathbf{D}_T \mathbf{B} d\Omega_1 \right) \frac{d\bar{\mathbf{u}}_1}{dt} - \left( \int_{\Omega_1} \mathbf{B}^T S''_{w1} \mathbf{m} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{w1}}{dt} \right. \\ - \left( \int_{\Omega_1} \mathbf{B}^T S''_{o1} \mathbf{m} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{o1}}{dt} - \left( \int_{\Omega_1} \mathbf{B}^T S''_{g1} \mathbf{m} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{g1}}{dt} \\ \left. + \left[ - \left( \int_{\Omega_1} \mathbf{N}^T \frac{d\mathbf{b}}{dt} d\Omega_1 \right) - \left( \int_{\Gamma_1} \mathbf{N}^T \frac{d\hat{\mathbf{t}}}{dt} d\Gamma_1 \right) + \left( \int_{\Omega_1} \mathbf{B}^T \frac{d\sigma_o}{dt} d\Omega_1 \right) \right] \right\} = 0 \end{aligned} \quad (9.30)$$

It is immediately obvious that

$$\int_{\Omega} \mathbf{F} d\Omega = \sum \int_{\Omega_e} \mathbf{F} d\Omega \quad (9.31)$$

in which  $\Omega$  and  $\Omega_e$  represents the whole region and the element domain, respectively. We see that the secret of the approximation lies simply in the requirement of writing the approximation in an integral form. Thus, all the  $\Omega_1$  subscripts must be replaced by  $\Omega_{1e}$ . However, bearing in mind the previous point, we maintain the notation just to avoid confusion, as the formulation is identically applicable both for the whole domain and the subdivisions.

Equations (9.30) is valid for any value of the virtual displacements  $\delta\bar{\mathbf{u}}_1$  and can be further simplified as

$$\mathbf{K}_1 \frac{d\bar{\mathbf{u}}_1}{dt} + \mathbf{L}_{w1} \frac{d\bar{\mathbf{p}}_{w1}}{dt} + \mathbf{L}_{o1} \frac{d\bar{\mathbf{p}}_{o1}}{dt} + \mathbf{L}_{g1} \frac{d\bar{\mathbf{p}}_{g1}}{dt} - \frac{d\mathbf{f}_1}{dt} = 0 \quad (9.32)$$

where

$$\mathbf{K}_1 = - \int_{\Omega_1} \mathbf{B}_1^T \mathbf{D}_T \mathbf{B} d\Omega_1 \quad (9.33)$$

$$\mathbf{L}_{w1} = \int_{\Omega_1} \mathbf{B}^T S''_{w1} \mathbf{m} \bar{\mathbf{N}} d\Omega_1 \quad (9.34)$$

$$\mathbf{L}_{o1} = \int_{\Omega_1} \mathbf{B}^T S''_{o1} \mathbf{m} \bar{\mathbf{N}} d\Omega_1 \quad (9.35)$$

$$\mathbf{L}_{g1} = \int_{\Omega_1} \mathbf{B}^T S''_{g1} \mathbf{m} \bar{\mathbf{N}} d\Omega_1 \quad (9.36)$$

$$d\mathbf{f}_1 = - \int_{\Omega_1} \mathbf{N}^T d\mathbf{b} d\Omega_1 - \int_{\Gamma_1} \mathbf{N}^T d\bar{\mathbf{t}} d\Gamma_1 + \int_{\Omega_1} \mathbf{B}^T d\sigma_o d\Omega_1 \quad (9.37)$$

The same procedure used for the equilibrium equation is also applicable to the continuity equations. For the oil phase within the matrix continuum, i.e. where  $\alpha = 1$ , equation (9.7) will take the form

$$\begin{aligned} & \left( \int_{\Omega_1} (\nabla a)^T \frac{-\mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} \nabla \bar{\mathbf{N}} d\Omega_1 \right) \bar{\mathbf{p}}_{o1} + \left( \int_{\Omega_1} \mathbf{a}^T \frac{\bar{\alpha} \mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{o1}}{dt} \\ & - \left( \int_{\Omega_1} \mathbf{a}^T \frac{\bar{\alpha} \mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{o2}}{dt} + \left( \int_{\Omega_1} \mathbf{a}^T \lambda_{oo1} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{o1}}{dt} \\ & + \left( \int_{\Omega_1} \mathbf{a}^T \lambda_{wo1} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{w1}}{dt} + \left( \int_{\Omega_1} \mathbf{a}^T \lambda_{go1} \bar{\mathbf{N}} d\Omega_1 \right) \frac{d\bar{\mathbf{p}}_{g1}}{dt} \\ & + \left( \int_{\Omega_1} \mathbf{a}^T \frac{S_{o1}}{B_{o1}} \mathbf{m}^T \mathbf{B} d\Omega_1 \right) \frac{d\bar{\mathbf{u}}_1}{dt} = 0 \end{aligned} \quad (9.38)$$

On replacing the arbitrary function  $\mathbf{a}$  by the shape function  $\mathbf{N}$ , equation (9.38) then becomes

$$\mathbf{H}_{p1} \bar{\mathbf{p}}_{o1} + \mathbf{H}_l \bar{\mathbf{p}}_{o2} + \mathbf{H}_{o1} \frac{d\bar{\mathbf{p}}_{o1}}{dt} + \mathbf{H}_{w1} \frac{d\bar{\mathbf{p}}_{w1}}{dt} + \mathbf{H}_{g1} \frac{d\bar{\mathbf{p}}_{g1}}{dt} + \mathbf{H}_{u1} \frac{d\bar{\mathbf{u}}_1}{dt} = 0 \quad (9.39)$$

where

$$\mathbf{H}_{p1} = \int_{\Omega_1} (\nabla \bar{\mathbf{N}})^T \frac{-\mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} (\nabla \bar{\mathbf{N}}) d\Omega_1 \quad (9.40)$$

$$\mathbf{H}_l = \int_{\Omega_1} \bar{\mathbf{N}} \frac{\bar{\alpha} \mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} \bar{\mathbf{N}} d\Omega_1 \quad (9.41)$$

$$\mathbf{H}_{o1} = \int_{\Omega_1} \bar{\mathbf{N}} \lambda_{oo1} \bar{\mathbf{N}} d\Omega_1 \quad (9.42)$$

$$\mathbf{H}_{w1} = \int_{\Omega_1} \bar{\mathbf{N}} \lambda_{wo1} \bar{\mathbf{N}} d\Omega_1 \quad (9.43)$$

$$\mathbf{H}_{g1} = \int_{\Omega_1} \bar{\mathbf{N}} \lambda_{go1} \bar{\mathbf{N}} d\Omega_1 \quad (9.44)$$

$$\mathbf{H}_{u1} = \int_{\Omega_1} \bar{\mathbf{N}} \frac{S_{o1}}{B_{o1}} \mathbf{m}^T \mathbf{B} d\Omega_1 \quad (9.45)$$

Also, on application of the finite element discretisation for the oil phase within the fractured continuum, i.e.  $\alpha = 2$ , equation (9.7) leads to

$$\mathbf{H}_{p2} \bar{\mathbf{p}}_{o1} - \mathbf{H}_l \bar{\mathbf{p}}_{o1} + \mathbf{H}_l \bar{\mathbf{p}}_{o2} + \mathbf{H}_{o2} \frac{d\bar{\mathbf{p}}_{o2}}{dt} + \mathbf{H}_{w2} \frac{d\bar{\mathbf{p}}_{w2}}{dt} + \mathbf{H}_{g2} \frac{d\bar{\mathbf{p}}_{g2}}{dt} = \bar{\mathbf{F}}_{o2} \quad (9.46)$$

where

$$\mathbf{H}_{p2} = \int_{\Omega_2} (\nabla \bar{\mathbf{N}})^T \frac{-\mathbf{k}_2 k_{ro2}}{\mu_o B_{o2}} (\nabla \bar{\mathbf{N}}) d\Omega_2 \quad (9.47)$$

$$\mathbf{H}_l = \int_{\Omega_2} \bar{\mathbf{N}}^T \frac{\bar{\alpha} \mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} \bar{\mathbf{N}} d\Omega_2 \quad (9.48)$$

$$\mathbf{H}_{o2} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{oo2} \bar{\mathbf{N}} d\Omega_2 \quad (9.49)$$

$$\mathbf{H}_{w2} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{wo2} \bar{\mathbf{N}} d\Omega_2 \quad (9.50)$$

$$\mathbf{H}_{g2} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{go2} \bar{\mathbf{N}} d\Omega_2 \quad (9.51)$$

$$\bar{\mathbf{F}}_{o2} = - \int_{\Omega_2} \bar{\mathbf{N}}^T Q_o d\Omega_2 - \int_{\Gamma_2} \bar{\mathbf{N}} q_o d\Gamma_2 \quad (9.52)$$

A similar result is obtained for the water-phase equation within the matrix continuum by using (9.11), and this can be written in a discretised form as

$$\mathbf{W}_{p1} \bar{\mathbf{p}}_{w1} + \mathbf{W}_l \bar{\mathbf{p}}_{w1} - \mathbf{W}_l \bar{\mathbf{p}}_{w2} + \mathbf{W}_{w1} \frac{d\bar{\mathbf{p}}_{w1}}{dt} + \mathbf{W}_{o1} \frac{d\bar{\mathbf{p}}_{o1}}{dt} + \mathbf{W}_{g1} \frac{d\bar{\mathbf{p}}_{g1}}{dt} + \mathbf{W}_{u1} \frac{d\bar{\mathbf{u}}_1}{dt} = 0 \quad (9.53)$$

where

$$\mathbf{W}_{p1} = \int_{\Omega_1} (\nabla \bar{\mathbf{N}})^T \frac{-\mathbf{k}_1 k_{rw1}}{\mu_w B_{w1}} (\nabla \bar{\mathbf{N}}) d\Omega_1 \quad (9.54)$$

$$\mathbf{W}_l = \int_{\Omega_1} \bar{\mathbf{N}} \frac{\bar{\alpha} \mathbf{k}_1 k_{rw1}}{\mu_w B_{w1}} \bar{\mathbf{N}} d\Omega_1 \quad (9.55)$$

$$\mathbf{W}_{w1} = \int_{\Omega_1} \bar{\mathbf{N}} \lambda_{ww1} \bar{\mathbf{N}} d\Omega_1 \quad (9.56)$$

$$\mathbf{W}_{\text{o1}} = \int_{\Omega_1} \bar{\mathbf{N}} \lambda_{\text{ow1}} \bar{\mathbf{N}} d\Omega_1 \quad (9.57)$$

$$\mathbf{W}_{\text{g1}} = \int_{\Omega_1} \bar{\mathbf{N}} \lambda_{\text{gw1}} \bar{\mathbf{N}} d\Omega_1 \quad (9.58)$$

$$\mathbf{W}_{\text{u1}} = \int_{\Omega_1} \bar{\mathbf{N}} \frac{S_{\text{w1}}}{B_{\text{w1}}} \mathbf{m}^T \mathbf{B} d\Omega_1 \quad (9.59)$$

Also, for the water phase in the fractured continuum, we obtain

$$\mathbf{W}_{p2}\bar{\mathbf{p}}_{\text{w2}} + \mathbf{W}_l\bar{\mathbf{p}}_{\text{w1}} - \mathbf{W}_l\bar{\mathbf{p}}_{\text{w2}} + \mathbf{W}_{\text{w2}} \frac{d\bar{\mathbf{p}}_{\text{w2}}}{dt} + \mathbf{W}_{\text{o2}} \frac{d\bar{\mathbf{p}}_{\text{o2}}}{dt} + \mathbf{W}_{\text{g2}} \frac{d\bar{\mathbf{p}}_{\text{g2}}}{dt} = \bar{\mathbf{F}}_{\text{w2}} \quad (9.60)$$

where

$$\mathbf{W}_{p2} = \int_{\Omega_2} (\nabla \bar{\mathbf{N}})^T \frac{-\mathbf{k}_2 k_{\text{rw2}}}{\mu_w B_{\text{w2}}} (\nabla \bar{\mathbf{N}}) d\Omega_2 \quad (9.61)$$

$$\mathbf{W}_l = \int_{\Omega_2} \bar{\mathbf{N}}^T \frac{\bar{\mathbf{x}}\mathbf{k}_1 k_{\text{rw1}}}{\mu_w B_{\text{w1}}} \bar{\mathbf{N}} d\Omega_2 \quad (9.62)$$

$$\mathbf{W}_{\text{w2}} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{\text{ww2}} \bar{\mathbf{N}} d\Omega_2 \quad (9.63)$$

$$\mathbf{W}_{\text{o2}} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{\text{ow2}} \bar{\mathbf{N}} d\Omega_2 \quad (9.64)$$

$$\mathbf{W}_{\text{g2}} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{\text{gw2}} \bar{\mathbf{N}} d\Omega_2 \quad (9.65)$$

$$\bar{\mathbf{F}}_{\text{w2}} = - \int_{\Omega_2} \bar{\mathbf{N}}^T Q_{\text{w}} d\Omega_2 - \int_{\Gamma_2} \bar{\mathbf{N}} q_{\text{w}} d\Gamma_2 \quad (9.66)$$

In a similar manner, the discretised form of the continuity equation for the gas phase within the matrix continuum may be derived from (9.15) as follows:

$$\begin{aligned} & \mathbf{G}_{p1}\bar{\mathbf{p}}_{\text{g1}} + \mathbf{G}_{k1}\bar{\mathbf{p}}_{\text{o1}} + \mathbf{G}_l(\bar{\mathbf{p}}_{\text{g1}} - \bar{\mathbf{p}}_{\text{o2}}) + \mathbf{G}_m(\bar{\mathbf{p}}_{\text{o1}} - \bar{\mathbf{p}}_{\text{o2}}) \\ & + \mathbf{G}_{\text{g1}} \frac{d\bar{\mathbf{p}}_{\text{g1}}}{dt} + \mathbf{G}_{\text{o1}} \frac{d\bar{\mathbf{p}}_{\text{o1}}}{dt} + \mathbf{G}_{\text{w1}} \frac{d\bar{\mathbf{p}}_{\text{w1}}}{dt} + \mathbf{G}_{\text{u1}} \frac{d\bar{\mathbf{u}}_1}{dt} = 0 \end{aligned} \quad (9.67)$$

where

$$\mathbf{G}_{p1} = \int_{\Omega_1} (\nabla \bar{\mathbf{N}})^T \frac{-\mathbf{k}_1 k_{\text{rg1}}}{\mu_g B_{\text{g1}}} (\nabla \bar{\mathbf{N}}) d\Omega_1 \quad (9.68)$$

$$\mathbf{G}_{k1} = \int_{\Omega_1} (\nabla \bar{\mathbf{N}})^T R_{\text{sol}} \frac{-\mathbf{k}_1 k_{\text{ro1}}}{\mu_o B_{\text{o1}}} (\nabla \bar{\mathbf{N}}) d\Omega_1 \quad (9.69)$$

$$\mathbf{G}_l = \int_{\Omega_1} \bar{\mathbf{N}}^T R_{\text{sol}} \frac{\bar{\mathbf{x}}\mathbf{k}_1 k_{\text{rg1}}}{\mu_g B_{\text{g1}}} \bar{\mathbf{N}} d\Omega_1 \quad (9.70)$$

$$\mathbf{G}_m = \int_{\Omega_1} \bar{\mathbf{N}}^T R_{sol} \frac{\bar{\alpha} \mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} \bar{\mathbf{N}} d\Omega_1 \quad (9.71)$$

$$\mathbf{G}_{g1} = \int_{\Omega_1} \bar{\mathbf{N}}^T \lambda_{gg1} \bar{\mathbf{N}} d\Omega_1 \quad (9.72)$$

$$\mathbf{G}_{o1} = \int_{\Omega_1} \bar{\mathbf{N}}^T \lambda_{og1} \bar{\mathbf{N}} d\Omega_1 \quad (9.73)$$

$$\mathbf{G}_{w1} = \int_{\Omega_1} \bar{\mathbf{N}}^T \lambda_{wg1} \bar{\mathbf{N}} d\Omega_1 \quad (9.74)$$

$$\mathbf{G}_{u1} = \int_{\Omega_1} \bar{\mathbf{N}}^T C_{go} \mathbf{m}^T \mathbf{B} d\Omega_1 \quad (9.75)$$

And finally, the discretised continuity equation for gas phase within the fractured continuum will lead to

$$\begin{aligned} & \mathbf{G}_{p2} \bar{\mathbf{p}}_{g2} + \mathbf{G}_{k2} \bar{\mathbf{p}}_{o2} - \mathbf{G}_t (\bar{\mathbf{p}}_{g1} - \bar{\mathbf{p}}_{g2}) - \mathbf{G}_m (\bar{\mathbf{p}}_{o1} - \bar{\mathbf{p}}_{o2}) \\ & + \mathbf{G}_{g2} \frac{d\bar{\mathbf{p}}_{g1}}{dt} + \mathbf{G}_{o2} \frac{d\bar{\mathbf{p}}_{o1}}{dt} + \mathbf{G}_{w2} \frac{d\bar{\mathbf{p}}_{w1}}{dt} = \bar{\mathbf{F}}_{g2} \end{aligned} \quad (9.76)$$

where

$$\mathbf{G}_{p2} = \int_{\Omega_2} (\nabla \bar{\mathbf{N}})^T \frac{-\mathbf{k}_2 k_{rg2}}{\mu_g B_{g2}} (\nabla \bar{\mathbf{N}}) d\Omega_2 \quad (9.77)$$

$$\mathbf{G}_{k2} = \int_{\Omega_2} (\nabla \bar{\mathbf{N}})^T R_{so2} \frac{-\mathbf{k}_2 k_{ro2}}{\mu_o B_{o2}} (\nabla \bar{\mathbf{N}}) d\Omega_2 \quad (9.78)$$

$$\mathbf{G}_t = \int_{\Omega_2} \bar{\mathbf{N}}^T R_{sol} \frac{\bar{\alpha} \mathbf{k}_1 k_{rg1}}{\mu_g B_{g1}} \bar{\mathbf{N}} d\Omega_1 \quad (9.79)$$

$$\mathbf{G}_m = \int_{\Omega_2} \bar{\mathbf{N}}^T R_{sol} \frac{\bar{\alpha} \mathbf{k}_1 k_{ro1}}{\mu_o B_{o1}} \bar{\mathbf{N}} d\Omega_1 \quad (9.80)$$

$$\mathbf{G}_{g2} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{gg2} \bar{\mathbf{N}} d\Omega_2 \quad (9.81)$$

$$\mathbf{G}_{o2} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{og2} \bar{\mathbf{N}} d\Omega_2 \quad (9.82)$$

$$\mathbf{G}_{w2} = \int_{\Omega_2} \bar{\mathbf{N}}^T \lambda_{wg2} \bar{\mathbf{N}} d\Omega_2 \quad (9.83)$$

$$\bar{\mathbf{F}}_{g2} = \int_{\Omega_2} \bar{\mathbf{N}}^T (Q_g + R_{so2} Q_o) d\Omega_2 + \int_{\Gamma_2} \bar{\mathbf{N}}^T (q_g + R_{so2} q_o) d\Gamma_2 \quad (9.84)$$

Equations (9.32), (9.39), (9.46), (9.53), (9.60), (9.67) and (9.76) represent a set of ordinary differential equations in time. For convenience, the equations are written in the following concise form:

$$\mathbf{C}\mathbf{X} + \mathbf{B} \frac{d}{dt} \mathbf{X} = \mathbf{F} \quad (9.85)$$

where

$$C = \left[ \begin{array}{cccc|ccc} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{p1} + \mathbf{W}_l & \mathbf{0} & \mathbf{0} & -\mathbf{W}_l & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_{p1} + \mathbf{H}_l & \mathbf{0} & \mathbf{0} & -\mathbf{H}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_{k1} + \mathbf{G}_m & \mathbf{G}_{p1} + \mathbf{G}_l & \mathbf{0} & -\mathbf{G}_m & -\mathbf{G}_l \\ \hline \mathbf{0} & -\mathbf{W}_l & \mathbf{0} & \mathbf{0} & \mathbf{W}_{p2} + \mathbf{W}_l & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{H}_l & \mathbf{0} & \mathbf{0} & \mathbf{H}_{p2} + \mathbf{H}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{G}_m & -\mathbf{G}_l & \mathbf{0} & \mathbf{G}_{k2} + \mathbf{G}_m & \mathbf{G}_{p2} + \mathbf{G}_l \end{array} \right] \quad (9.86)$$

$$B = \left[ \begin{array}{c|cccc|ccc} \mathbf{K}_1 & \mathbf{L}_{w1} & \mathbf{L}_{o1} & \mathbf{L}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{W}_{u1} & \mathbf{W}_{w1} & \mathbf{W}_{o1} & \mathbf{W}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{u1} & \mathbf{H}_{w1} & \mathbf{H}_{o1} & \mathbf{H}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{G}_{u1} & \mathbf{G}_{w1} & \mathbf{G}_{o1} & \mathbf{G}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W}_{w2} & \mathbf{W}_{o2} & \mathbf{W}_{g2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{H}_{w2} & \mathbf{H}_{o2} & \mathbf{H}_{g2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_{w2} & \mathbf{G}_{o2} & \mathbf{G}_{g2} \end{array} \right] \quad (9.87)$$

$$\mathbf{X} = \begin{bmatrix} \bar{\mathbf{u}}_1 \\ \hline \bar{\mathbf{p}}_{w1} \\ \bar{\mathbf{p}}_{o1} \\ \hline \bar{\mathbf{p}}_{g1} \\ \hline \bar{\mathbf{p}}_{w2} \\ \bar{\mathbf{p}}_{o2} \\ \hline \bar{\mathbf{p}}_{g2} \end{bmatrix} \quad (9.88)$$

$$F = \begin{bmatrix} \frac{d\mathbf{f}_1}{dt} \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \bar{\mathbf{F}}_{w2} \\ \bar{\mathbf{F}}_{o2} \\ \hline \bar{\mathbf{F}}_{g2} \end{bmatrix} \quad (9.89)$$

The values  $\bar{\mathbf{u}} = \bar{\mathbf{u}}_t$  and  $\bar{\mathbf{p}} = \bar{\mathbf{p}}(t)$  at different time values may now be obtained by means of suitable time-stepping algorithms.

Having discretised the governing equation in the space domain which finally results in (9.85), we may now apply the same numerical technique used in Section 8.4 for time discretisation of the derived equation. Equation (9.85) therefore takes the form

$$[\mathbf{B} + \mathbf{C}\Delta t_n \theta] \bar{\mathbf{X}}_{t_n + \Delta t_n} = [\mathbf{B} - \mathbf{C}\Delta t_n(1 - \theta)] \bar{\mathbf{X}}_{t_n} + \mathbf{F}\Delta t_n \quad (9.90)$$

where

$$\theta = \frac{\int_{t_n}^{t_n + \Delta t_n} \bar{g}\alpha(t)dt}{\int_{t_n}^{t_n + \Delta t_n} \bar{g}dt} \quad (9.91)$$

and

$$\alpha = \frac{t - t_n}{\Delta t_n} \quad (9.92)$$

Now we may replace matrices in (9.90) by their original forms given in equations (9.86) to (9.89). This leads to the following expanded form:

$$\begin{bmatrix} \mathbf{K}_1 & \mathbf{L}_{w1} & \mathbf{L}_{o1} & \mathbf{L}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W}_{u1} & \mathbf{W}'_1 & \mathbf{W}_{o1} & \mathbf{W}_{g1} & \mathbf{W}' & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{u1} & \mathbf{H}_{w1} & \mathbf{H}'_1 & \mathbf{H}_{g1} & \mathbf{0} & \mathbf{H}' & \mathbf{0} \\ \mathbf{G}_{u1} & \mathbf{G}_{w1} & \hat{\mathbf{G}}'_1 & \mathbf{G}'_1 & \mathbf{0} & \hat{\mathbf{G}}' & \hat{\mathbf{G}}' \\ \mathbf{0} & \mathbf{W}' & \mathbf{0} & \mathbf{0} & \mathbf{W}'_2 & \mathbf{W}_{o2} & \mathbf{W}_{g2} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}' & \mathbf{0} & \mathbf{H}_{w2} & \mathbf{H}'_2 & \mathbf{H}_{g2} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{G}}' & \mathbf{G}' & \mathbf{G}_{w2} & \hat{\mathbf{G}}'_2 & \hat{\mathbf{G}}'_2 \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}}_1 \\ \bar{\mathbf{p}}_{w1} \\ \bar{\mathbf{p}}_{o1} \\ \bar{\mathbf{p}}_{g1} \\ \hline \bar{\mathbf{p}}_{w2} \\ \dot{\mathbf{p}}_{o2} \\ \bar{\mathbf{p}}_{g2} \end{bmatrix}_{t_n + \Delta t_n} =$$

$$\begin{bmatrix} \mathbf{K}_1 & \mathbf{L}_{w1} & \mathbf{L}_{o1} & \mathbf{L}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W}_{u1} & \mathbf{W}''_1 & \mathbf{W}_{o1} & \mathbf{W}_{g1} & \mathbf{W}'' & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{u1} & \mathbf{H}_{w1} & \mathbf{H}''_1 & \mathbf{H}_{g1} & \mathbf{0} & \mathbf{H}'' & \mathbf{0} \\ \mathbf{G}_{u1} & \mathbf{G}_{w1} & \hat{\mathbf{G}}''_1 & \mathbf{G}''_1 & \mathbf{0} & \hat{\mathbf{G}}'' & \hat{\mathbf{G}}'' \\ \mathbf{0} & \mathbf{W}'' & \mathbf{0} & \mathbf{0} & \mathbf{W}''_2 & \mathbf{W}_{o2} & \mathbf{W}_{g2} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}'' & \mathbf{0} & \mathbf{H}_{w2} & \mathbf{H}''_2 & \mathbf{H}_{g2} \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{G}}'' & \mathbf{G}'' & \mathbf{G}_{w2} & \hat{\mathbf{G}}''_2 & \hat{\mathbf{G}}''_2 \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}}_1 \\ \bar{\mathbf{p}}_{w1} \\ \bar{\mathbf{p}}_{o1} \\ \bar{\mathbf{p}}_{g1} \\ \hline \bar{\mathbf{p}}_{w2} \\ \dot{\mathbf{p}}_{o2} \\ \bar{\mathbf{p}}_{g2} \end{bmatrix}_{t_n} + \begin{bmatrix} \frac{d\mathbf{f}_1}{dt} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \hline \bar{\mathbf{F}}_{w2} \\ \bar{\mathbf{F}}_{o2} \\ \bar{\mathbf{F}}_{g2} \end{bmatrix} \Delta t_n \quad (9.93)$$

where

$$\mathbf{W}'_1 = \mathbf{W}_{w1} + \theta \Delta t_n (\mathbf{W}_{p1} + \mathbf{W}_l) \quad (9.94)$$

$$\mathbf{H}'_1 = \mathbf{H}_{o1} + \theta \Delta t_n (\mathbf{H}_{p1} + \mathbf{H}_l) \quad (9.95)$$

$$\hat{\mathbf{G}}'_1 = \mathbf{G}_{o1} + \theta \Delta t_n (\mathbf{G}_{k1} + \mathbf{G}_m) \quad (9.96)$$

$$\mathbf{G}'_1 = \mathbf{G}_{g1} + \theta \Delta t_n (\mathbf{G}_{p1} + \mathbf{G}_l) \quad (9.97)$$

$$\mathbf{W}'_2 = \mathbf{W}_{w2} + \theta \Delta t_n (\mathbf{W}_{p2} + \mathbf{W}_l) \quad (9.98)$$

$$\mathbf{H}'_2 = \mathbf{H}_{o2} + \theta \Delta t_n (\mathbf{H}_{p2} + \mathbf{H}_l) \quad (9.99)$$

$$\hat{\mathbf{G}}'_2 = \mathbf{G}_{o2} + \theta \Delta t_n (\mathbf{G}_{k2} + \mathbf{G}_m) \quad (9.100)$$

$$\mathbf{G}'_2 = \mathbf{G}_{g2} + \theta \Delta t_n (\mathbf{G}_{p2} + \mathbf{G}_l) \quad (9.101)$$

$$\mathbf{W}' = -\theta \Delta t_n \mathbf{W}_l \quad (9.102)$$

$$\mathbf{H}' = -\theta \Delta t_n \mathbf{H}_l \quad (9.103)$$

$$\hat{\mathbf{G}}' = -\theta \Delta t_n \mathbf{G}_m \quad (9.104)$$

$$\mathbf{G}' = -\theta \Delta t_n \mathbf{G}_l \quad (9.105)$$

$$\mathbf{W}''_1 = \mathbf{W}_{w1} - (1 - \theta) \Delta t_n (\mathbf{W}_{p1} + \mathbf{W}_l) \quad (9.106)$$

$$\mathbf{H}''_1 = \mathbf{H}_{o1} - (1 - \theta) \Delta t_n (\mathbf{H}_{p1} + \mathbf{H}_l) \quad (9.107)$$

$$\hat{\mathbf{G}}''_1 = \mathbf{G}_{o1} - (1 - \theta) \Delta t_n (\mathbf{G}_{k1} + \mathbf{G}_m) \quad (9.108)$$

$$\mathbf{G}''_1 = \mathbf{G}_{g1} - (1 - \theta) \Delta t_n (\mathbf{G}_{p1} + \mathbf{G}_l) \quad (9.109)$$

$$\mathbf{W}''_2 = \mathbf{W}_{w2} - (1 - \theta) \Delta t_n (\mathbf{W}_{p2} + \mathbf{W}_l) \quad (9.110)$$

$$\mathbf{H}''_2 = \mathbf{H}_{o2} - (1 - \theta) \Delta t_n (\mathbf{H}_{p2} + \mathbf{H}_l) \quad (9.111)$$

$$\hat{\mathbf{G}}''_2 = \mathbf{G}_{o2} - (1 - \theta) \Delta t_n (\mathbf{G}_{k2} + \mathbf{G}_m) \quad (9.112)$$

$$\mathbf{G}''_2 = \mathbf{G}_{g2} - (1 - \theta) \Delta t_n (\mathbf{G}_{p2} + \mathbf{G}_l) \quad (9.113)$$

$$\mathbf{W}'' = (1 - \theta) \Delta t_n \mathbf{W}_l \quad (9.114)$$

$$\mathbf{H}'' = (1 - \theta) \Delta t_n \mathbf{H}_l \quad (9.115)$$

$$\hat{\mathbf{G}}''_1 = (1 - \theta) \Delta t_n \mathbf{G}_m \quad (9.116)$$

$$\mathbf{G}''_1 = (1 - \theta) \Delta t_n \mathbf{G}_l \quad (9.117)$$

To facilitate the numerical calculation, the equilibrium equations can be solved for  $\Delta \bar{\mathbf{u}}_1^{t_n + \Delta t_n} = (\bar{\mathbf{u}}_1^{t_n + \Delta t_n} - \bar{\mathbf{u}}_1^{t_n})$ . Hence (9.93) can be further simplified to

$$\begin{bmatrix}
\mathbf{K}_1 & \mathbf{L}_{w1} & \mathbf{L}_{o1} & \mathbf{L}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{W}_{u1} & \mathbf{W}'_1 & \mathbf{W}_{o1} & \mathbf{W}_{g1} & \mathbf{W}' & \mathbf{0} & \mathbf{0} \\
\mathbf{H}_{u1} & \mathbf{H}_{w1} & \mathbf{H}'_1 & \mathbf{H}_{g1} & \mathbf{0} & \mathbf{H}' & \mathbf{0} \\
\mathbf{G}_{u1} & \mathbf{G}_{w1} & \hat{\mathbf{G}}'_1 & \mathbf{G}'_1 & \mathbf{0} & \hat{\mathbf{G}}' & \mathbf{G}' \\
\hline
\mathbf{0} & \mathbf{W}' & \mathbf{0} & \mathbf{0} & \mathbf{W}'_2 & \mathbf{W}_{o2} & \mathbf{W}_{g2} \\
\mathbf{0} & \mathbf{0} & \mathbf{H}' & \mathbf{0} & \mathbf{H}_{w2} & \mathbf{H}'_2 & \mathbf{H}_{g2} \\
\mathbf{0} & \mathbf{0} & \hat{\mathbf{G}}' & \mathbf{G}' & \mathbf{G}_{w2} & \hat{\mathbf{G}}'_2 & \mathbf{G}'_2
\end{bmatrix}
\begin{bmatrix}
\Delta \mathbf{u}_1 \\
\bar{\mathbf{p}}_{w1} \\
\bar{\mathbf{p}}_{o1} \\
\bar{\mathbf{p}}_{g1} \\
\hline
\bar{\mathbf{p}}_{w2} \\
\bar{\mathbf{p}}_{o2} \\
\bar{\mathbf{p}}_{g2}
\end{bmatrix}_{t_n + \Delta t_n} = \frac{d\mathbf{f}_1}{dt} \quad (9.118)$$
  

$$\begin{bmatrix}
\mathbf{L}_{w1} & \mathbf{L}_{o1} & \mathbf{L}_{g1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{W}''_1 & \mathbf{W}_{o1} & \mathbf{W}_{g1} & \mathbf{W}'' & \mathbf{0} & \mathbf{0} \\
\mathbf{H}_{w1} & \mathbf{H}''_1 & \mathbf{H}_{g1} & \mathbf{0} & \mathbf{H}'' & \mathbf{0} \\
\mathbf{G}_{w1} & \hat{\mathbf{G}}''_1 & \mathbf{G}''_1 & \mathbf{0} & \hat{\mathbf{G}}'' & \mathbf{G}'' \\
\hline
\mathbf{W}'' & \mathbf{0} & \mathbf{0} & \mathbf{W}''_2 & \mathbf{W}_{o2} & \mathbf{W}_{g2} \\
\mathbf{0} & \mathbf{H}'' & \mathbf{0} & \mathbf{H}_{w2} & \mathbf{H}''_2 & \mathbf{H}_{g2} \\
\mathbf{0} & \hat{\mathbf{G}}'' & \mathbf{G}'' & \mathbf{G}_{w2} & \hat{\mathbf{G}}''_2 & \mathbf{G}''_2
\end{bmatrix}
\begin{bmatrix}
\bar{\mathbf{p}}_{w1} \\
\bar{\mathbf{p}}_{o1} \\
\bar{\mathbf{p}}_{g1} \\
\hline
\bar{\mathbf{p}}_{w2} \\
\bar{\mathbf{p}}_{o2} \\
\bar{\mathbf{p}}_{g2}
\end{bmatrix}_{t_n} + \frac{\bar{\mathbf{F}}_{w2}}{\bar{\mathbf{F}}_{o2}} \Delta t_n = \frac{\bar{\mathbf{F}}_{g2}}{\bar{\mathbf{F}}_{g2}}$$

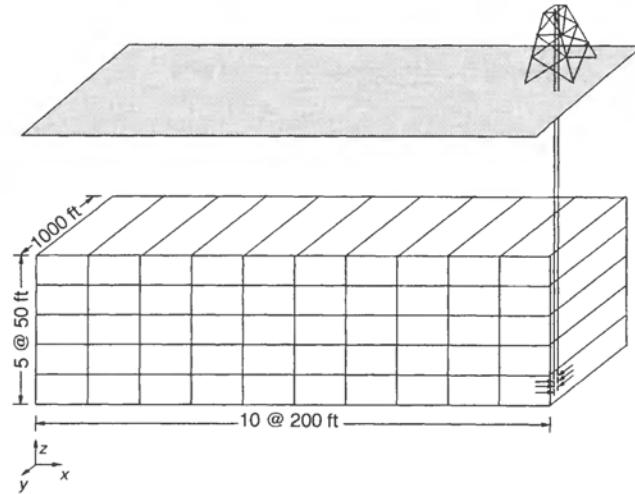
This incremental displacement is added after each time step to the total displacement known at the start of the time step.

Equations (9.118) are the final discretised form of the governing equations which are applied at all nodes within the domain and on the boundary where the unknowns are not prescribed. Therefore, the number of equations is equal to the number of unknown parameters. These equations represent a fully coupled and highly non-linear system for three-phase flow in a deforming fractured porous medium. All the non-linear coefficient matrices depend on the values of the sought unknowns, so iterative procedures are performed within each time step to obtain the final solution.

## 9.6 VALIDATION OF THE MODEL

In the absence of any analytical solution for the proposed model, program validation was carried out by making a comparison between the results obtained by the present simulator and those of similar models found in the open literature. For this purpose, a three-dimensional field-scale example is presented here to demonstrate the utility of the model and to illustrate the nature of fluid flow in fractured reservoirs. The problem is also used to compare fractured reservoir models as a part of the sixth Society of Petroleum Engineers (SPE) comparative solution project for double-porosity simulators [16].

A linear section of reservoir was modelled (Figure 9.2). Vertically, five layers each having a height of 50 ft were used. Horizontally, the reservoir was divided into ten 200 ft



**Figure 9.2** Finite element mesh for the SPE example (not to scale)

**Table 9.1** Basic data for the sixth SPE comparative solution projects<sup>a</sup>

Layer	Matrix block size (ft)	Fracture permeability $k_2$ (mD)	Shape factor (ft <sup>-2</sup> )	$J$ (RB-cP/D-psi)
1	25	10	0.040	1
2	25	10	0.040	1
3	5	90	1.000	9
4	10	20	0.025	2
5	10	20	0.025	2

<sup>a</sup> Other relevant data:

$k_1$ (mD)	1
$\phi_1$	0.29
$\Phi_2$	0.01
Initial oil pressure, psi (MPa)	6000 (41.37)
Initial water-oil capillary pressure, psi (Pa)	0.87 (6000)
Initial gas-oil capillary pressure, psi (Pa)	0.02 (500)
Matrix compressibility (vol/vol-psi)	$3.5 \times 10^{-6}$
Fracture compressibility (vol/vol-psi)	$3.5 \times 10^{-6}$
Z-direction transmissibilities	multiply calculated values by 0.1

grid blocks. A uniform thickness of 1000 ft was used in the  $Y$ -direction. The layer description for the cross-section is given in Table 9.1. Whereas both the shape factors and fracture permeabilities may be determined in terms of the matrix block size [47], direct use was made of the values given in the literature [16] and presented in Table 9.1. The oil production rate was calculated as follows,

$$Q_o = I_p \Delta p \quad (9.119)$$

where  $\Delta p$  is the well drawdown in psi and  $I_p$  is the productivity index, defined as

$$I_p = \frac{k_r J}{B\mu} \quad (9.120)$$

in which  $k_r$  is the oil relative permeability  $\mu$  (cP) and  $B$ (RB/STB) are respectively the viscosity and formation volume factor of produced oil at the well bottom hole pressure (BHP). Coefficient  $J$  has dimensions of RB-cP/D-psi, or darcy-ft, and its corresponding values for each layer are given in Table 9.1. The water and gas rates were obtained in terms of the mobility ratios as follows:

$$Q_w = M_{wo} Q_o \quad (9.121)$$

$$Q_g = M_{go} Q_o + R_{so} Q_o \quad (9.122)$$

where  $M_{wo}$  and  $M_{go}$  are the mobility ratios water/oil and gas/oil, respectively, and are calculated as follows:

$$M_{wo} = \frac{k_{rw}/\mu_w B_w}{k_{ro}/\mu_o B_o} \quad (9.123)$$

$$M_{go} = \frac{k_{rg}/\mu_g B_g}{k_{ro}/\mu_o B_o} \quad (9.124)$$

The production rates are calculated using data evaluated at the beginning of the time step, i.e. explicit loading. Depletion runs were carried out to a maximum of 10 years, or whenever the production declined to less than 1 STB per day. The production well has a maximum rate of 500 STB/D per day and was limited by a maximum drawdown of 100 psi. This well was located at the far right column and perforated only in the bottom layer.

Basic PVT data for the example are given in Table 9.2 [45]. Also Figures 9.3 to 9.6 illustrate capillary pressure–saturation and saturation–relative permeability curves for water–oil and gas–oil systems, respectively. These curves correspond to the matrix continuum only. For the fracture continuum, as in most double-porosity models, the capillary pressure was assumed to be zero. However, because of the limitation associated with the formulation, a pseudo-capillary pressure curve with a negligible effect on the depletion behaviour was assumed. The variation of gas–oil interfacial tension  $\sigma$  with pressure, as given in Table 9.2, has also been incorporated in the problem.

For this purpose, the gas–oil capillary pressure is directly related to the interfacial tension (IFT), so the gas–oil capillary pressure should be adjusted according to the ratio of the IFT at the reservoir pressure divided by the IFT specified for the capillary pressure–saturation curve, i.e. bubble point pressure, or in a mathematical form as

$$p_{cg} = \frac{\sigma}{\sigma_f} p_{cgI} \quad (9.125)$$

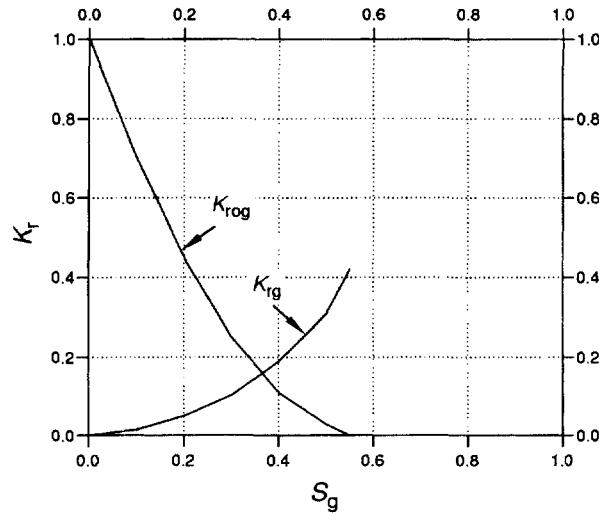
where the term  $p_{cgI}$  corresponds to input capillary pressure values calculated using the surface tension  $\sigma_f$ . Initially one hundred equally spaced time steps each equal to 0.1

**Table 9.2** PVT data for the sixth SPE comparative solution project<sup>a</sup>

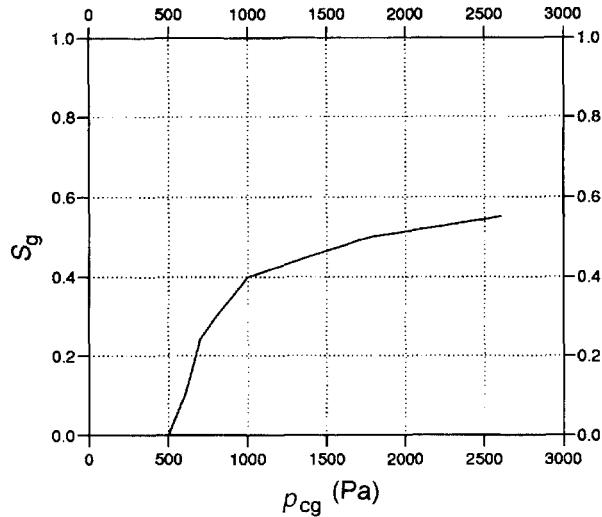
Pressure (psig)	$B_o$ (RB/STB)	$B_g$ (RB/SCF)	$R_s$ (SCF/STB)	$\mu_o$ (cP)	$\mu_g$ (cP)	$\sigma$ (dyne cm <sup>-1</sup> )
1674.0	1.3001	0.00198	367.0	0.529	0.0162	6.0
2031.0	1.3359	0.00162	447.0	0.487	0.0171	4.7
2530.0	1.3891	0.00130	564.0	0.436	0.0184	3.3
2991.0	1.4425	0.00111	679.0	0.397	0.0197	2.2
3553.0	1.5141	0.000959	832.0	0.351	0.0213	1.28
4110.0	1.5938	0.000855	1000.0	0.310	0.0230	0.72
4544.0	1.6630	0.000795	1143.0	0.278	0.0244	0.444
4935.0	1.7315	0.000751	1285.0	0.248	0.0255	0.255
5255.0	1.7953	0.000720	1413.0	0.229	0.0265	0.155
5545.0	1.8540	0.000696	1530.0	0.210	0.0274	0.090
7000.0	2.1978	0.000600	2259.0	0.109	0.0330	0.050

<sup>a</sup> Other relevant data:

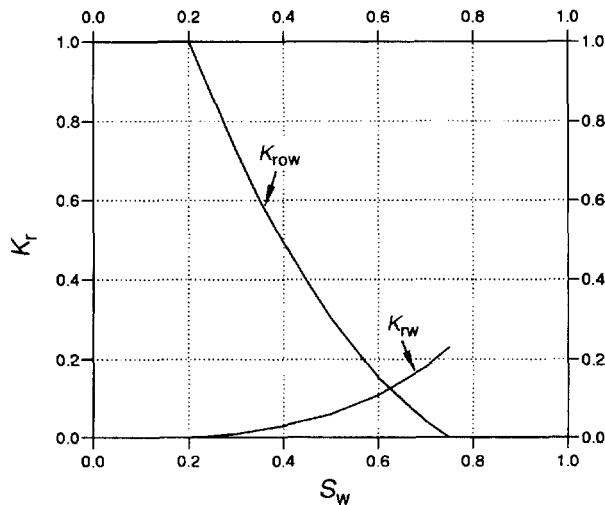
Original bubble point, psig (MPa)	5545 (38.23)
Density of stock-tank oil (lbm/cu ft)	51.14
Gas density at standard condition (lbm/cu ft)	0.058
Water formation volume factor	1.07
Water compressibility (vol/vol-psi)	$3.5 \times 10^{-6}$
Water viscosity (cP)	0.35

**Figure 9.3** Saturation versus relative permeability in the oil–gas system

years were used. Then, in a trial-and-error manner, the time steps were modified interactively if a diverging solution occurred. Finally, for the present example, the time steps shown in Table 9.3 were found to be satisfactory. The domain was assumed to be sealed and insulated at all boundaries; and except at the top surface, the displacements



**Figure 9.4** Capillary pressure versus saturation in the oil–gas system



**Figure 9.5** Saturation versus relative permeability in the oil–water system

were assumed to be zero in the directions perpendicular to the faces. Also, the base was assumed to have no movement in all directions.

The Young's modulus of elasticity and Poisson's ratio for the solid porous body were estimated in terms of the compressibility values of the matrix and fractures [45], and using the relationship

$$C = \frac{3(1 - 2\nu)}{E_s} \quad (9.126)$$

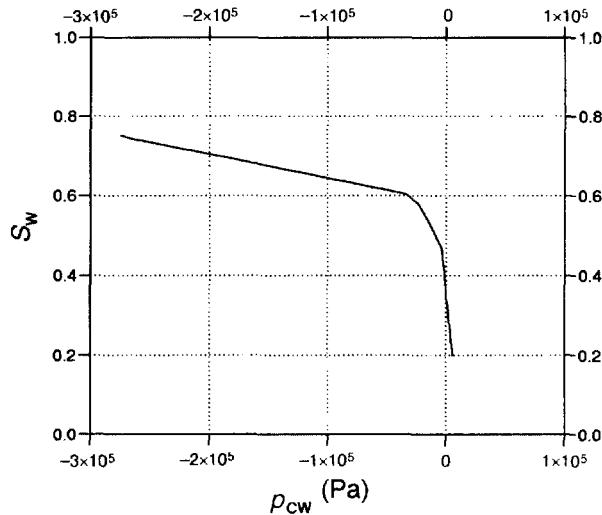


Figure 9.6 Capillary pressure versus saturation in the oil–water system

Table 9.3 Time steps for different time intervals

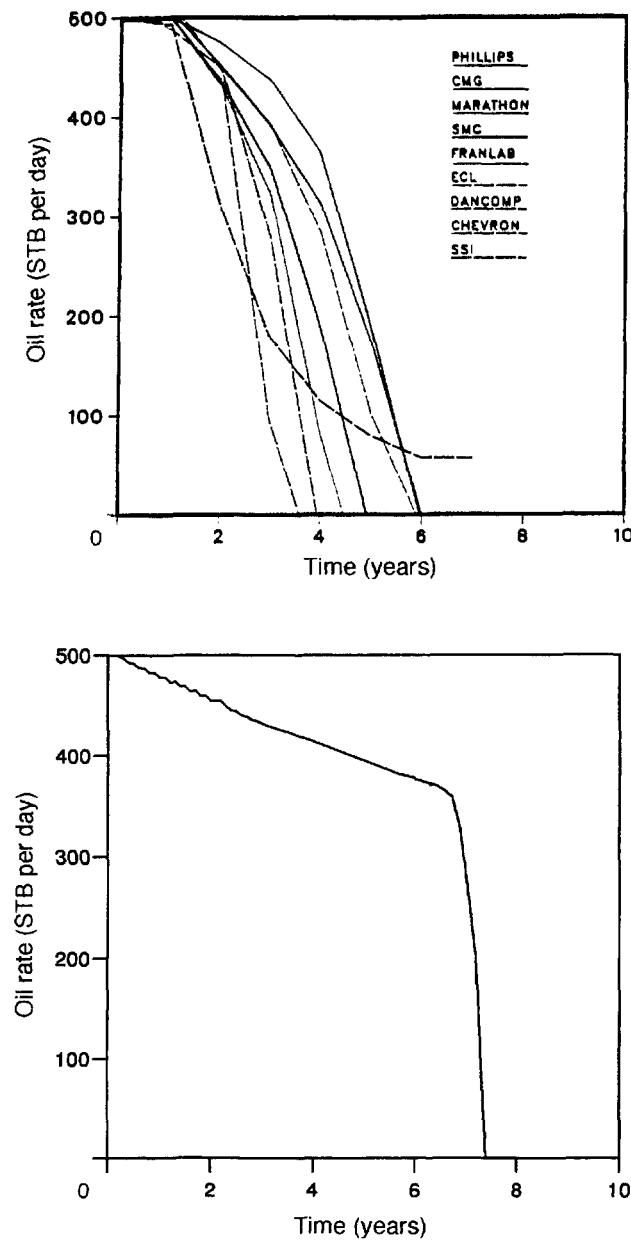
Time step (year)	Number of time steps
$1.0 \times 10^{-1}$	80
$1.0 \times 10^{-2}$	65
$1.0 \times 10^{-3}$	77
$1.0 \times 10^{-4}$	15

where  $C$  is the compressibility;  $E_s$  and  $v$  are Young's modulus and Poisson's ratio, respectively.

This reservoir test problem was designed as the sixth SPE comparative solution. The project was to compare various simulators developed using the double-porosity model. Ten organisations participated in the solution project. A full description of the models can be found in the literature [16]. Generally, all the models were developed using the finite difference method, and none account for the coupling effects between the rock deformation and the fluid flow. Hence the two salient features of the model described in this chapter were the use of a fully coupled finite element model as the numerical scheme for the system. The following results were reported by the ten participants:

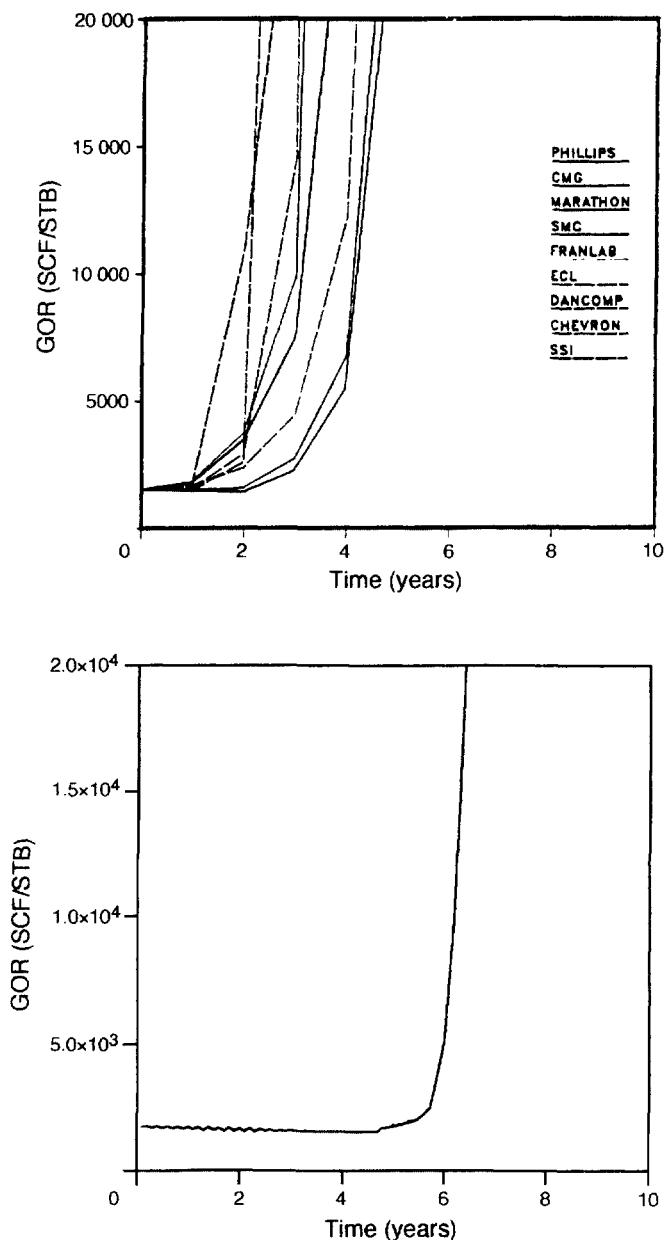
1. Oil production rate
2. Gas/oil ratio
3. Pressure at grid block (5,1,1) each year

Also, in order to make a comprehensive comparison, the total number of time steps and the global iterations for all runs were reported.



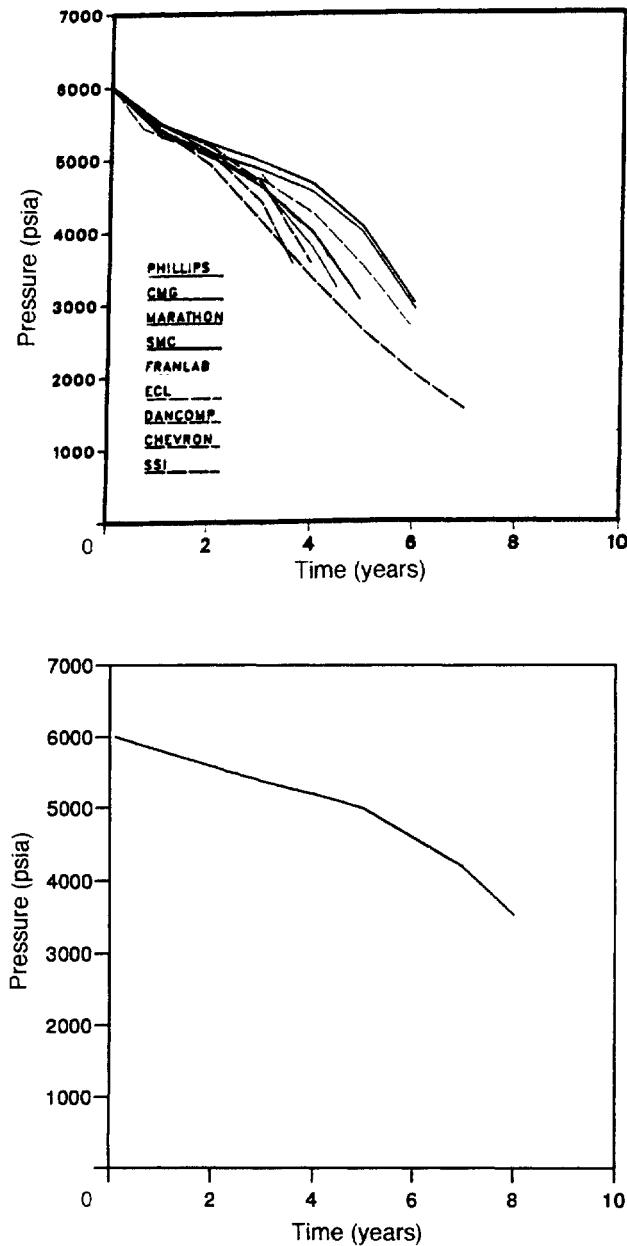
**Figure 9.7** Oil rate versus time for the SPE example

Figure 9.7(a) shows the plot of oil production rate versus the elapsed time. With few exceptions, most models show similar trends in the reported results. However, the results obtained by the coupled model, illustrated in Figure 9.8(b), show a significantly delayed reduction of the production rate. A negligibly small oscillation of the production rate



**Figure 9.8** Gas/oil ratio versus time for the SPE example

was observed during the early time steps; this could be attributed to the fairly large time intervals in the beginning, where the production rate was still high. The oscillation soon damped out after a few time steps. Whereas most of the previous models predicted a sharp decline of the production rate after 2–6 years, the coupled model predicted a



**Figure 9.9** Pressure at gridlock (5, 1, 1) versus time for the SPE example

much longer period with a mild slope at the early stages which finally ends after almost 7.5 years.

The same trend may be observed for the gas/oil ratio (GOR) in Figures 9.8(a) and 6.12(b). Figure 9.8(a) suggests a sharp increase in GOR after a few years, due to a

**Table 9.4** Total number of time steps and global iterations for different simulators

Company <sup>a</sup>	Number of steps	Simulation period (years)	Mean number of step per year	Mean number of iterations per step
Phillips	48	4.4	10.91	2.42
CMG	63	6.0	10.50	3.80
SMC	53	6.0	8.83	5.00
Marathon	66	5.0	13.20	2.12
Franlab	60	5.0	12.00	3.33
ECL	63	5.9	10.67	4.37
Chevron	36	4.0	9.00	5.40
SSI	75	3.6	20.83	4.11
Dancomp	50	7.0	7.14	5.20
JOE	—	—	—	—
Present model	237	7.6	31.18	2.41

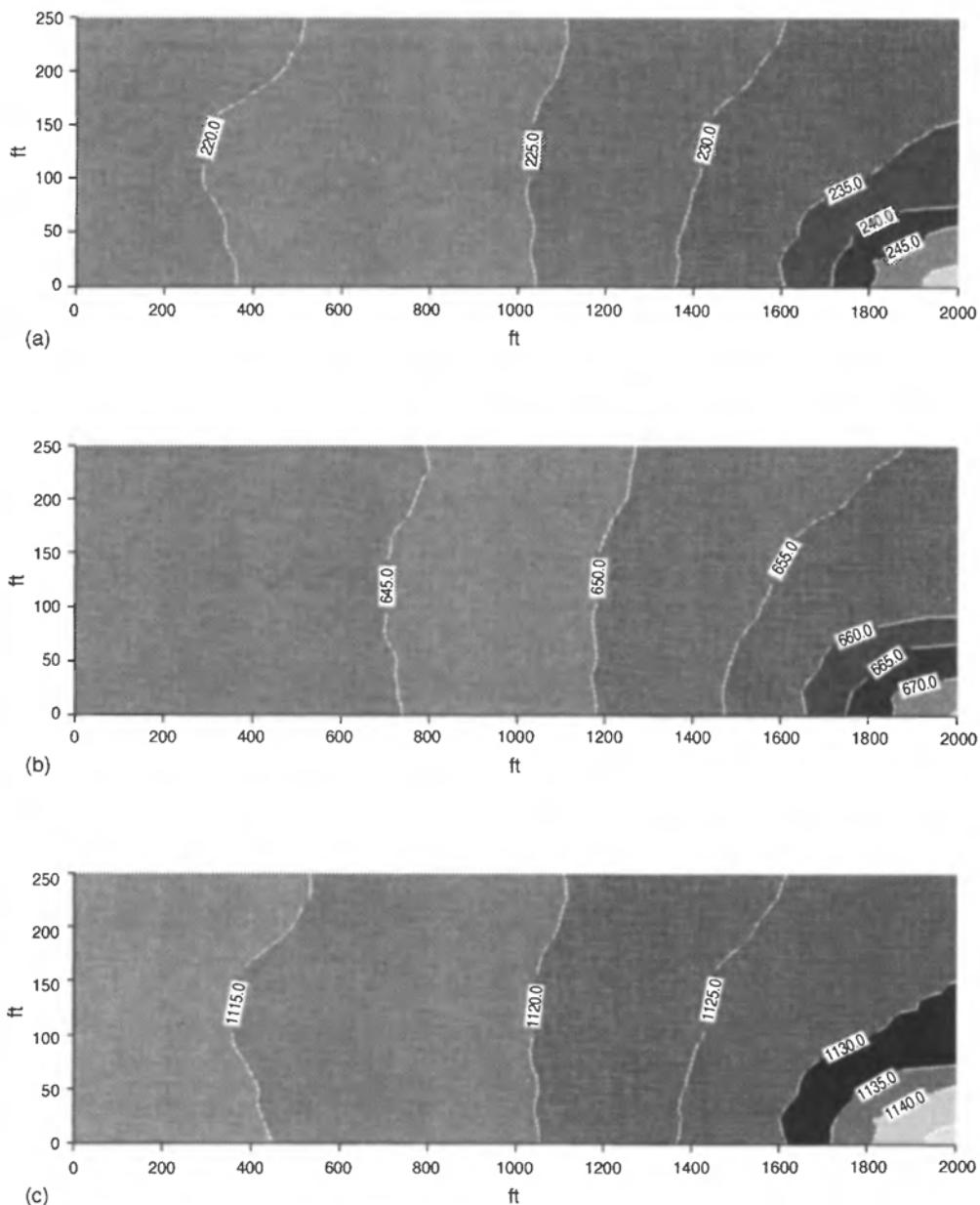
<sup>a</sup> The full names of the companies are as follows:

Phillips	Phillips Petroleum Co. (USA)
CMG	Computer Modelling Group (Canada)
SMC	Simulation and Modelling Consultancy Ltd (England)
Marathon	Marathon Oil Co. (USA)
Franlab	Franlab (France)
ECL	Exploration Consultants Ltd (England)
Chevron	Chevron Oil Field Research Co. (USA)
SSI	Scientific Software-Intercomp (USA)
Dancomp	Dancomp A/S (Denmark)
JOE	Japan Oil Engineering Co. (Japan)

decline in the reservoir pressure; but in the finite element model, where coupling effects are also considered, this was postponed until the sixth year, as shown in Figure 9.8(b).

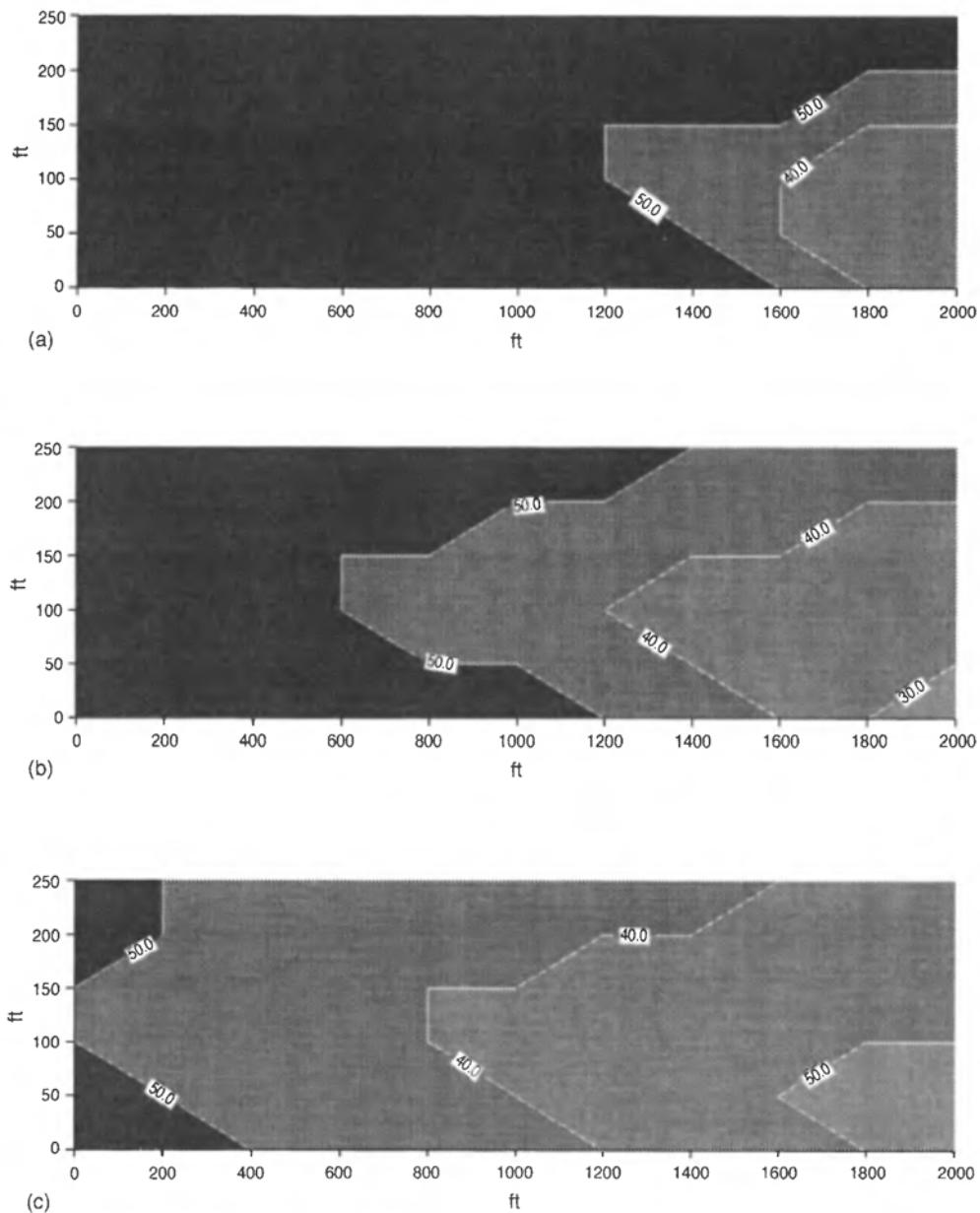
In order to investigate the reservoir pressure drawdown, the average pressure values in grid block (5,1,1), obtained by the various simulators, are compared in Figure 9.9. The rate of pressure decline was significantly slower where the coupling effect was also incorporated (Figure 9.9(b)), whereas all uncoupled models predicted almost a similar trend with a rapidly decreasing pressure (Figure 9.9(a)). This confirms the considerable impact of soil deformation on maintaining the initial reservoir pressure for a relatively longer time, which can improve long-term economical productivity of a reservoir. The effect, which is ignored in conventional uncoupled models, has particularly great significance where the oil-bearing strata are quite deformable.

The next observation compares the number of time steps used in various simulators for the whole time span. The number of time steps and their corresponding sizes are determined by the user, based on common engineering sense, so they are usually a matter of trial and error. As a general rule, in order to ensure convergence, smaller time steps must be used where a higher production rate or injection rate is expected at a single node. In return, a larger time step is allowed where the production rate or the injection rate is relatively small. However, bear in mind that the number of time steps must be kept as low as possible to avoid an uneconomic run time. On the other hand, the number



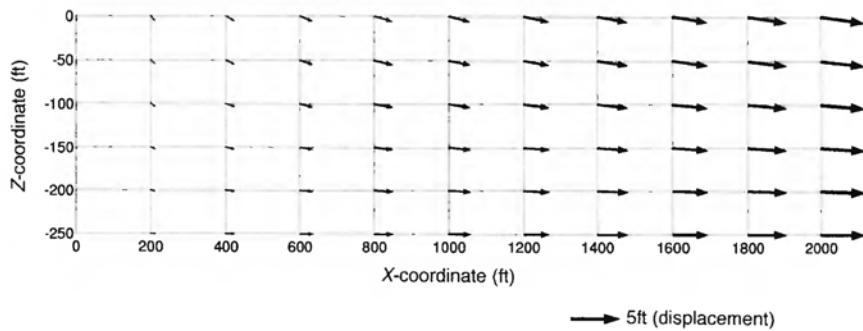
**Figure 9.10** Pressure drawdown contours (psi) in the reservoir cross-section at (a)  $T = 1$  year, (b)  $T = 3$  years and (c)  $T = 5$  years

of iterations is directly proportional to the size of the time step: the larger the time step, the more iterations are needed for convergence. Since each iterations needs almost the same amount of run time, a balance should be made between the number of iterations and the time step size, so an optimised solution can be obtained. Table 9.4 compares the



**Figure 9.11** Oil saturation contours (percent) in the reservoir cross-section at (a)  $T = 1$  year, (b)  $T = 3$  years and (c)  $T = 5$  years

number of time steps used in the present study with the number used in several uncoupled models. It implies that the average number of time steps per year is remarkably higher for the coupled finite element model.



**Figure 9.12** Nodal displacements of the finite element grid

Isopressure contours in the cross-section of the reservoir are shown in Figure 9.10 for 1, 3 and 5 years. The pressure declines towards the producing well with a relatively high pressure gradient near the well bore. The pressure gradient in the region adjacent to the well bore is usually very high compared to the far field. This is firstly due to a very high flow rate per unit area in this region. Moreover, in many producing wells the value of absolute permeability for the formation close to the well bore is greater than the value further back in the reservoir. This causes an even greater pressure drawdown around the well bore, the so-called *skin effect*. Due to the combined effect of these two factors, a very rapid drawdown usually occurs around the well. However, this effect was not predicted in the present simulation. This is because the chosen elements were much larger than the region affected by the phenomenon, which is usually of the order of a few metres. This implies that the pressure values obtained around the well may be significantly different from the real values. More accurate values may be obtained by refining the mesh in this region. However, the pressure values in the other regions are by no means dramatically affected.

Note that, although the pressure values vary with the elapsed time, the rate of variation for all points within the reservoir is almost the same, so the pressure difference between any two given points of the reservoir at any specified instant is not very great when compared to the initial pressure values. This can be attributed to the presence of the fractured network, which facilitates the fluid flow and the pressure communication within the reservoir and therefore causes a rather uniform pressure distribution profile for the entire domain.

Figure 9.11 indicates the oil saturation contour profiles for 1, 3 and 5 years. Saturation values increase towards the reservoir flank, where the effect of oil production is felt later when compared to the region close to the well bore. The contours elongate horizontally along the mid-depth of the reservoir, where the highly fractured layer 3 is located. The higher permeability of the layer, caused by the intensively fractured network rapidly drains the fluid from the oil-bearing matrix. The drained fluid is then conducted towards the producing well through the fractured network, thus creating a less saturated area within and around layer 3.

A unique feature of the coupled model, among the models in this study, is that the displacements are considered as primary unknowns. Thus, one may determine the soil deformation at any specified point of interest. Figure 9.12 is a graphical illustration of the displacements. The soil deformation is induced by the drag force caused only by the

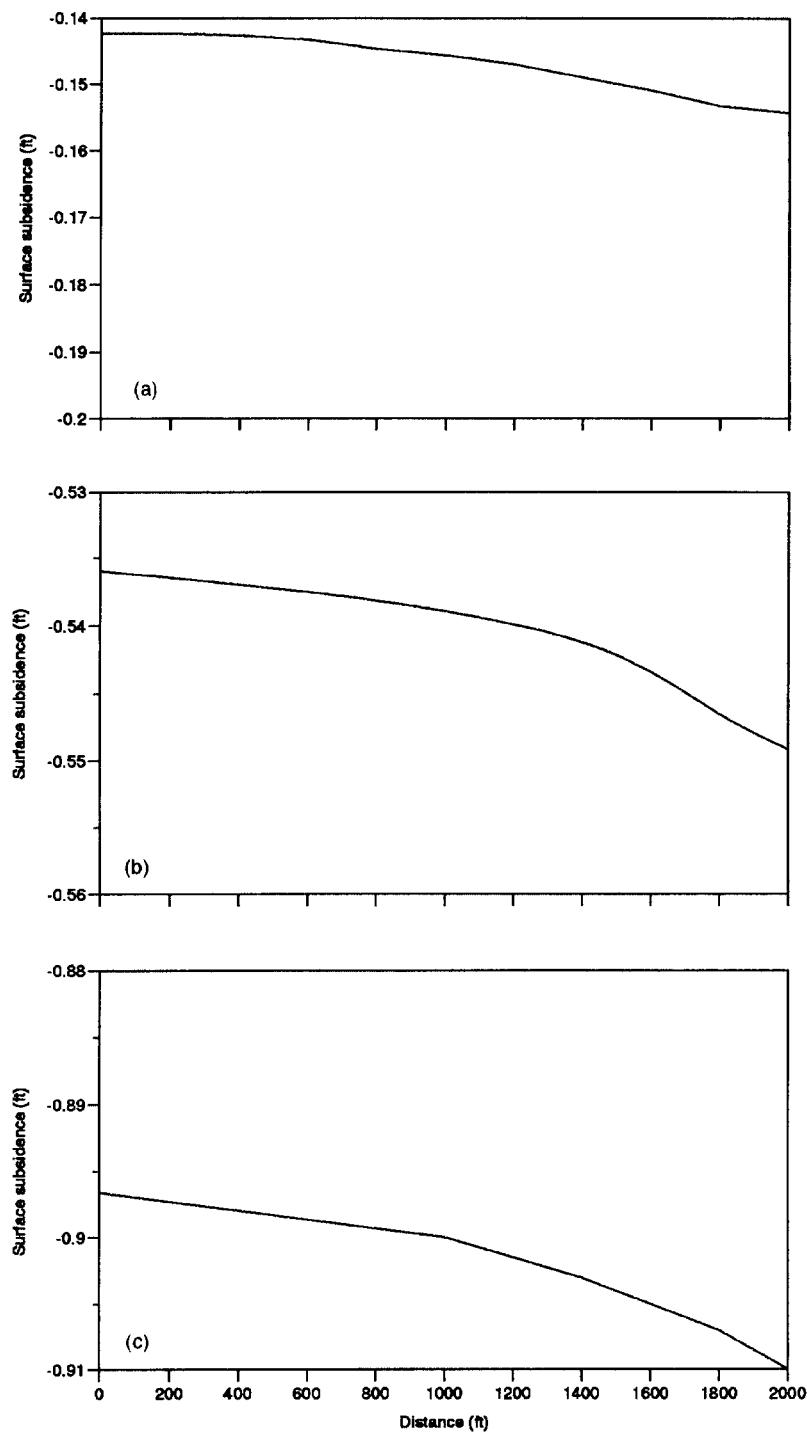
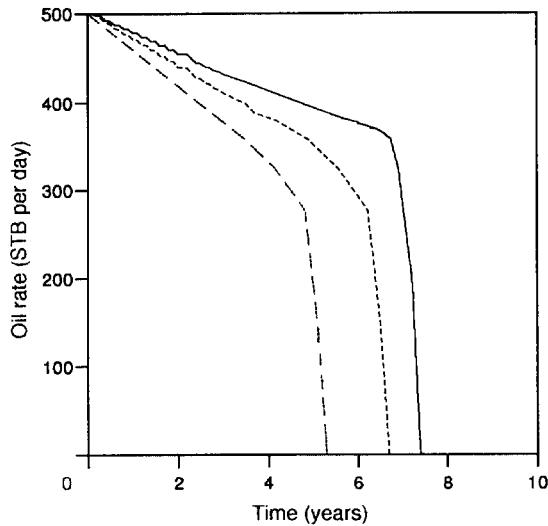
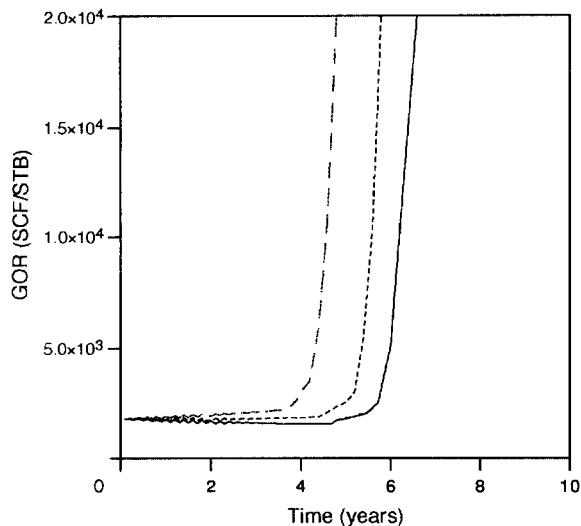


Figure 9.13 Subsidence values at (a)  $T = 1$  year, (b)  $T = 4$  years and (c)  $T = 8$  years

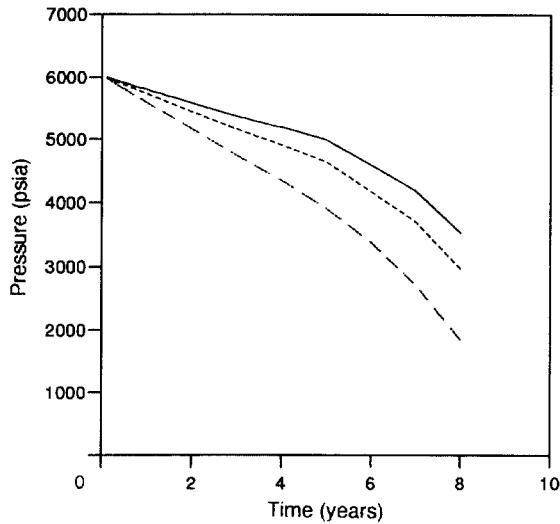


**Figure 9.14** Oil production rate versus time for different moduli: (—)  $E = 2.0 \times 10^9$  Pa; (---)  $E = 2.0 \times 10^{10}$  Pa; (· · · ·)  $E = 2.0 \times 10^{11}$  Pa

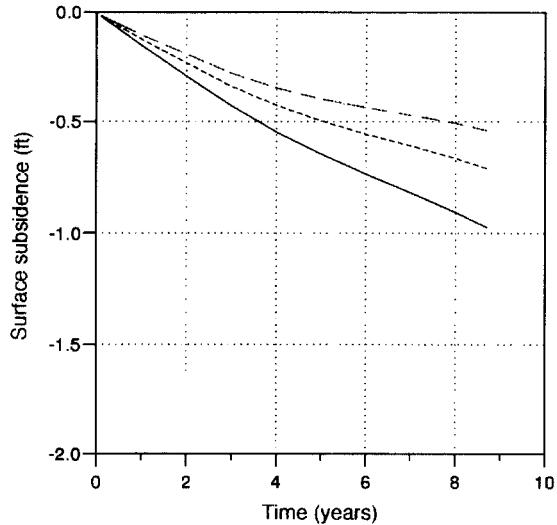


**Figure 9.15** Gas/oil ratio versus time for different moduli: (—)  $E = 2.0 \times 10^9$  Pa; (---)  $E = 2.0 \times 10^{10}$  Pa; (· · · ·)  $E = 2.0 \times 10^{11}$  Pa

fluid depletion; no other external load is applied. Notice that the maximum displacements take place close to the well bore, where the maximum drag force exists. Also, the displacement values increase towards the top surface. This implies that the influence of the displacements in the lower layers was transmitted to the upper layers, thus increasing the total settlement value at any given depth above the layer.



**Figure 9.16** Pressure at gridlock (5, 1, 1) versus time for different moduli: (—)  $E = 2.0 \times 10^9 \text{ Pa}$ ; (---)  $E = 2.0 \times 10^{10} \text{ Pa}$ ; (····)  $E = 2.0 \times 10^{11} \text{ Pa}$



**Figure 9.17** Subsidence profiles for different moduli: (—)  $E = 2.0 \times 10^9 \text{ Pa}$ ; (---)  $E = 2.0 \times 10^{10} \text{ Pa}$ ; (····)  $E = 2.0 \times 10^{11} \text{ Pa}$

The subsidence profile of the top surface is shown in Figure 9.13 at 1, 4 and 8 years. Notice that the reservoir pressure declined, causing a gradual settlement of the ground surface over the production time. This profile was not generated by the other models.

The impact of coupling between the deformation of the porous media and the fluid flow is particularly important where a highly deformable formation is to be modelled. In

this case the soil deformation plays a vital role in maintaining the initial reservoir pressure for a longer time, thus allowing a higher economic productivity for the reservoir. This mechanism, called a *compaction drive*, was not predicted by the conventional uncoupled models. In order to investigate the significance of the mechanism, three extra runs were also carried out with different moduli of elasticity. The results are shown in Figs 9.14 and 9.17. As indicated in Figure 9.14, a longer economic oil production was obtained for the formation having the lowest modulus. Figure 9.15 shows that the sharp increase in gas/oil ratio occurs later for the higher moduli, i.e. for the more deformable formations. The reason may be identified using Figure 9.16, where the reservoir pressure dissipation appears to be slower for the lower moduli. Finally, Figure 9.17 indicates the amount of subsidence versus time for three different moduli. As expected, the greater the modulus, the smaller the subsidence.

## REFERENCES

1. Aifantis, E. C. (1977) Introducing a multi-porous medium. *Developments in Mechanics*, **8**, 209–11.
2. Aifantis, E. C. (1979) On the response of fissured rocks. *Developments in Mechanics*, **10**, 249–53.
3. Auriault, J. L. and Boutin, C. (1992) Deformable porous media with double porosity. Quasi-statics. I: coupling effects., *Transport in Porous Media*, **7**, 63–82.
4. Bai, M., Ma, Q. and Roegiers, J. C. (1994) Dual porosity behavior of naturally fractured reservoirs. *Int. J. Num. Anal. Meth. Geomech.*, **18**, 359–76.
5. Bai, M., Elsworth, D. and Roegiers, J. C. (1993) Modeling of naturally fractured reservoirs using deformation dependent flow mechanism. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.*, **30**, 1185–91.
6. Barenblatt, G. I., Zheltov, I. P. and Kochina, I. N. (1960) Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks. *J. Appl. Math. Mech. USSR*, **24**(5), 1286–1303.
7. Bibby, R. (1981) Mass transport of solutes in dual-porosity media. *Water Resources Research*, **17**, 1075–81.
8. Coats, K. H. (1989) Implicit compositional simulation of single-porosity and dual-porosity reservoirs. Paper SPE 18427 presented at SPE Reservoir Simulation Symposium, Houston TX.
9. Daniel, E. J. (1954) Fractured reservoirs of Middle East. *Bull. Amer. Assoc. Petrol Geol.*, **38**(5), 774–815.
10. Davadant, D. (1971) Méthodes d'étude de la fissuration par diagraphies et essais de puits (application of logs and well tests to the study of fractured reservoirs). Institut Français du Pétrole (IFP), No. 19676 (in French).
11. Dean, R. H. and Lo, L. L. (1988) Simulation of naturally fractured reservoirs. *SPERE*, May, 638–48.
12. Duguid, J. O. and Abel, J. (1974) Finite element Galerkin method for flow in fractured porous media, in *Finite Element Methods in Flow Problems*, J. T. Oden, O. C. Zienkiewicz, R. H. Gallagher and C. Taylor (eds), UAH Press, Huntsville AL, pp. 599–615.
13. Duguid, J. O. and Lee, P. C. Y. (1977) Flow in fractured porous media., *Water Resources Research*, **13**, 558–66.
14. Dutra, T. V. Jr. and Aziz, K. (1992) A new double-porosity reservoir model for oil/water flow problems. *SPERE*, Nov., 419–25.
15. Elsworth, D. and Bai, M. (1992) Flow–deformation response of dual-porosity media. *J. Geotech. Engng ASCE*, **118**, 107–24.

16. Firoozabadi, A. and Kent Thomas, L. (1990) Sixth SPE comparative solution project: dual-porosity simulators. *J. Petro. Tech.*, **42**(6), 710–15.
17. Fung, L. S.-K. (1991) Simulation of block-to-block processes in naturally fractured reservoirs. *SPERE Nov.*, 477–84.
18. Fung, L. S.-K. (1993) Numerical simulation of naturally fractured reservoirs. Paper SPE 25616 presented at the SPE Middle East Oil Conference and Exhibition, Bahrain.
19. Fung, L. S.-K. and Collins, D. A. (1991) An evaluation of the improved dual porosity model for the simulation of gravity effects in naturally fractured reservoirs. *J. Can. Petro. Tech.*, **30**(3), 61–68.
20. Ghafouri, H. R. and Lewis, R. W. (1996) A finite element double porosity model for heterogeneous deformable porous media. *Int. J. Num. Anal. Meth. Geomech.*, **20**, 831–44.
21. Ghafouri, H. R. and Lewis, R. W. (1997) A Fully coupled implicit model for 3-phase fluid flow through deformable fractured porous media, in *Proceedings of the Fourth International Conference on Civil Engineering*, Tehran, May 1997.
22. Huyakorn, P. S., Lester, B. H. and Faust, C. R. (1983) Finite element techniques for modelling groundwater flow in fractured aquifers. *Water Resources Research*, **19**(4), 1019–35.
23. Kazemi, H., Merrill, L., Porterfield, K. and Zaman, P. (1976) Numerical simulation of water–oil flow in naturally fractured reservoirs. *Soc. Petro. Eng. J., Dec.*, 317–26; *Trans. AIME*, **261**.
24. Kazemi, H., Gilman, J. R. and Elsharkawy, A. M. (1992) Analytical and numerical solution of oil recovery from fractured reservoirs with empirical transfer functions. *SPERE, May*, 219–27.
25. Khalili-Naghadeh, N. (1991) Numerical modelling of flow through fractured media. PhD thesis, University of NSW, Australia.
26. Lewis, R. W. and Ghafouri, H. R. (1997) A novel finite element double porosity model for multi-phase flow through deformable fractured porous media. *Int. J. Num. Anal. Meth. Geomech.*, **21**, 789–816.
27. Litvak, B. L. (1985) Simulation and characterization of naturally fractured reservoirs, in *Proceeding of the Reservoir Characterization Conference*, Dallas TX, Academic Press, New York, pp. 561–83.
28. Litvak, B. L., Satter, A. and Etebar, S. (1988) An analysis of naturally fractured reservoir performance using a novel fractured reservoir simulator. Paper SPE 17615 presented at the SPE, International Meeting on Petroleum Engineering, Tianjin.
29. Long, J. C. S., Remer, J. S., Wilson, C. R. and Witherspoon, P. A. (1982) Porous-media equivalents for networks of discontinuous fractures. *Water Resources Research*, **18**(3), 645–58.
30. Narasimhan, T. N. (1982) Multidimensional numerical simulation of fluid flow in fractured porous media. *Water Resources Research*, **18**, 1235–47.
31. Odeh, A. S. (1965) On steady-state behavior of naturally fractured reservoirs. *Trans. AIME*, **234**, 60.
32. Ramstad, L. R. (1977) Geological modelling of fractured hydrocarbon reservoirs. University of Trondheim, Report 77–4.
33. Reiss, L. H. (1980) *The Reservoir Engineering Aspects of Fractured Formations* (translation from the French). Institut Français du Pétrole (IFP). Editions Technip, Paris.
34. Romm, E. S. (1966) *Fluid Flow in Fractured Rocks*, Nedra, Moscow.
35. Rossen, R. J. (1977) Simulation of naturally fractured reservoirs with semi-implicit source terms. *Soc. Pet. Eng. J.*, **17**, 201–10.
36. Rossen, R. H., and Shen, E. I. (1989) Simulation of gas/oil drainage and water/oil imbibition in Naturally fractured reservoirs. *SPERE, Nov.*, 464–70.
37. Saidi, A. M. (1975) Mathematical simulation Model describing Iranian fractured reservoirs and its application to Haft Kel Field, in *Proceedings of the World Petroleum Congress*, Tokyo, Vol. 4, PD13(3), pp. 209–19.

38. Saidi, A. M. (1983) Simulation of naturally fractured reservoirs. Paper SPE 12270 presented at the 1983 SPE Reservoir Simulation Symposium, San Francisco CA.
39. Saidi, A. M. (1987) *Reservoir Engineering of Fractured Reservoirs*, Total Press, p. 708.
40. Saidi, A. M., Tehrani, D. H. and Wit, K. (1979) Mathematical simulation of fractured reservoir performance, based on physical model, in *Proceedings of the Tenth World Petroleum Congress*, Bucharest, Vol. 3, PD10(3), pp. 225–33.
41. Sheldon, J. W., Zondek, B. and Cardwell, W. T. (1959) One-dimensional, incompressible, non-capillary, two-phase fluid flow in a porous medium. *Trans. SPE AMIE*, **216**, 290–96.
42. Reiss, L. H. (1980) The Reservoir Engineering Aspects of Functional Formations, translated from the French by Max Crenot. Paris, Editions Technip.
43. Sonier, F., Souillard, P. and Blaskovich, F. T. (1988) Numerical simulation of naturally fractured reservoirs. *SPERE*, 1114–22.
44. Sonier F. and Eymard, R. (1987) A new simulator for naturally fractured reservoirs. Paper SPE 16006 presented at the 1987 SPE Reservoir Simulation Symposium, San Antonio TX.
45. Thomas, L. K., Dixon, T. N. and Pierson, R. G. (1988) Fractured reservoir simulation. *SPERE*, May, 638–48.
46. Van Golf-Racht, T. D. (1982) *Fundamentals of Fractured Reservoir Engineering*, Elsevier, New York.
47. Warren, J. E. and Root, P. J. (1963) The behavior of naturally fractured reservoirs. *Soc. Petro. Eng. J.*, Sept., 245–55; *Trans. SPE of AIME*, Vol. 228.
48. Wilson, R. K. and Aifantis, E. C. (1982) On the theory of consolidation with double porosity. *Int. J. Engng Sci.*, **20**(9), 1009–35.

# 10

## Heat and Fluid Flow in Deforming Porous Media

### 10.1 INTRODUCTION

This chapter considers the effects of heat flow and fluid flow through porous media. It allows investigation of land subsidence in connection with geothermal energy production for a given geothermal system and it enables estimates to be made of the recoverable energy and the optimum extraction rate of mass and energy. Analyses of this type can also be applied to the design of thermal and/or hydraulic fracturing stimulation of oil reservoirs and for a more accurate interpretation of well tests when thermal effect are taken into account [1].

Another important area of application is the disposal of radioactive waste in geological formations. The case of metal canisters containing vitrified waste buried deep in clay layers is chosen as example. The decay of radioactive material produces heat, causing a temperature rise and an expansion in both the pore water and the soil skeleton. The volume increase of the pore water is generally greater than the volume increase of the voids in the soil skeleton; this increases the pore pressure and therefore reduces the effective stress, so the clay beds become consolidated. When the effective stress reduction is too great, a progressive failure may occur; this has to be avoided since the clay is the main long-term barrier against the diffusion of radionucleides [2]. Hence a tool for predictive analyses is important in this case.

The study of geothermal reservoirs has received much attention but the deformation characteristics have been considered only in a few papers, especially if coupled with the heat and mass transfer equations. The governing equations in earlier studies on this subject are reviewed and thoroughly discussed by Corapcioglu and Karahanoglu [3]. Some of the more recent studies relevant to this chapter are now briefly discussed. The effect of geothermal production, as well as reinjection, on the deformation of geothermal systems has been discussed by Lippman *et al.* [4], where the mathematical model for the mass and energy equations is combined with the numerical solution of Terzaghi's consolidation equation. The reservoir temperature and pore pressure are solved by interlacing equations for mass and energy, and the pore pressure changes are then used to determine the volumetric and vertical deformation. Brownell *et al.* [5] discussed the full interaction of a porous solid matrix and the fluid in geothermal systems, including momentum and energy transfer and the dependence of porosity and permeability upon fluid and solid stresses. The ground surface subsidence of Wairakei has been examined

in terms of the calculated two-phase fluid flow and the local geology by Pritchett *et al.* [6]. A similar study for a geopressured reservoir was presented by Garg *et al.* [7], considering dissolved methane in the fluid. Aktan and Farouq Ali [8] studied thermal stresses induced by hot-water injection using a thermoelastic stress-strain constitutive relationship. Similarly, Ertekin [9] has presented a two-dimensional two-phase fluid flow, a three-dimensional heat flow and a two-dimensional displacement model for a hot-water-flooded oil reservoir. He used the method of finite differences to solve for the fluid and energy flow equations, then applied a finite element model to determine the displacements.

Bear and Corapcioglu [10] developed a mathematical model to simulate the areal distribution of fluid pressure, temperature, land subsidence and horizontal displacements due to hot-water injection into a thermoelastic confined and leaky aquifer. Their mathematical model was derived by averaging a three-dimensional model over the vertical thickness of the aquifer under the assumption that the aquifers were thin in relation to the horizontal distance. The numerical solution of such an averaged model is shown in Section 6.2 for the isothermal case.

Borsetto *et al.* [1] presented a numerical model for the heat and mass transfer and stress field in transient coupled conditions using an elastic stress-strain relationship. The governing equations were partitioned before the discretisation in the time domain, and for their solution they require a predictor of the displacements, the pressure and the temperature. A different partitioning procedure for the governing equations is presented in Section 10.4.2.

In connection with the analysis of nuclear waste disposal in a geological clay formation, Borsetto *et al.* [11] discuss the constitutive relationship for clay under the combined action of heating, elastoplastic deformation and groundwater flow. The dependence of the coefficient of permeability  $k$  on the temperature is also included. The resulting governing equations can be solved in principle to obtain displacements, pressure, temperature and porosity. Since the solution was expected to require a considerable numerical effort, Borsetto *et al.* proposed simplifications such as the uncoupling of the heat flow equation, thereby reducing the number of independent unknowns to be solved simultaneously. Booker and Savvidou [2] studied the consolidation around a heat source in thermoelastic soil. The temperature field was uncoupled from the determination of the displacements and pressure by neglecting the mechanical contributions to the energy balance as well as the convective terms. Analytical solutions are given for a spherical heat source and a point heat source buried deep in clay. The point source solution was then integrated over the volume of a cylindrical canister; Section 10.7 compares it with the numerical model presented in this chapter.

Thermoelastic consolidation was also investigated by Aboustit, *et al.* [12] using a general variational principle. The fluid was assumed to be incompressible and convection was neglected. The coupling terms between the pressure and temperature values do not appear in their formulation. Only the coupling between displacements and temperature and displacements and pressure are taken into account, which gives a symmetric coefficient matrix. Numerical results are presented for a one-dimensional thermoelastic consolidation and for the Centralia Coal gasification site.

Two-phase flow in a non-isothermal deforming porous medium will now be discussed in more detail. Besides the linear momentum balance equation for the whole mixture

(equilibrium equation), a comprehensive model for heat and mass transfer in a partially saturated porous medium should consider an energy balance equation taking into account the latent heat transfer and at least two mass conservation equations. These last two equations result from the sum of the mass conservation equations of the separate phases considered in the simulation. Earlier methods used only one mass conservation equation and more complex transfer coefficients [13].

Very few published solutions exist for the combined influence of heat and mass transfer in unsaturated deforming porous media. This appears clearly from an extensive survey carried out by Alonso *et al.* [14] on the general topic of partially saturated porous media. The following publications typify the different approaches present in the literature.

A one-dimensional heat and moisture transfer model in a rigid unsaturated soil was presented by Dakshanamurthy and Fredlund [15]. The model considered liquid flow (no vapour flow), gas flow and heat flow equations; the heat flow equation is uncoupled from the other two conservation equations. The liquid flow equation ignores the influence of a temperature gradient and the heat flow equation considers only the flow caused by conduction. The permeability and the density of the mixture were assumed to be constant. Thomas [16] used one moisture continuity equation which described the liquid and the vapour flows. The de Vries theory, with modifications along the lines of Luikov's work [17], was used in the heat conservation equation and no solid deformation was considered. The chosen variables were the temperature and volumetric moisture content.

Capillary potential and temperature were taken to be the main variables in a more recent paper by Thomas and King [18] for the heat and mass transfer analysis of a rigid unsaturated soil. Baggio *et al.* [19] presented a model for a rigid matrix with three conservation equations: one for energy, one for dry air and one for water phase (liquid and vapour). The latent heat transfer and convection terms were accounted for in the energy balance equation. Geraminegad and Saxena [20] put forward a thermoelastic model for heat and mass transfer in a partially saturated soil. A modified version of the Philip and de Vries formulation was applied in this case to the heat transfer equation. Two mass transfer equations were used and the model variables were temperature, capillary pressure and gas pressure. The model incorporated only volumetric soil deformation and neglected the soil deformation due to any external loading. A model for heat and mass transfer in deforming geothermal reservoirs was also presented by Lewis *et al.* [21]. In this instance the capillary pressure effects could be neglected so that a unique continuity equation for the steam and water phases was used.

A fully coupled numerical model to simulate soil deformation, water flow, airflow and heat flow in a porous medium was developed by Schrefler *et al.* [22]. In this model the heat transfer via conduction and convection was taken into account as well as the flow, viscosity and density variation of the fluid due to temperature gradients. However, phase change and latent heat transfer (evaporation and condensation) influences were neglected. These effects were later introduced into a subsequent model by Gawin *et al.* [23], and form the basis of the model given in Section 10.5. This model is aimed at handling situations which span from fully saturated to almost dry conditions. Finally, Olivella *et al.* [24] presented a model for the non-isothermal multiphase flow of brine and gas in a deformable saline medium.

The next section gives the equations for a fully saturated non-isothermal consolidation problem. It is similar to the model developed by Lewis and Karahanoglu [25] for the analysis of subsidence in fluid-dominated hot-water geothermal systems. Small strain and displacement assumptions are valid throughout this chapter and inertia terms are neglected.

## 10.2 NON-ISOTHERMAL FULLY SATURATED CONSOLIDATION

This model is suitable for many practical situations and deserves a separate treatment from the more comprehensive but complicated model where phase change is also taken into account. A partitioned solution procedure, which is obviously also applicable to the full model, is discussed in Section 10.4.2 in relation to this simpler model.

### 10.2.1 Governing equations

The linear momentum balance equation in terms of total stresses is

$$\mathbf{L}^T \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \quad (10.1)$$

where, using (2.310), the density of the two-phase medium may be defined as

$$\rho = (1 - n)\rho_s + n\rho_w \quad (10.2)$$

while the effective stress is given by (2.278) with  $S_w = 1$  and  $S_g = 0$ , i.e.

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' - \alpha \mathbf{m} p^w \quad (10.3)$$

The mass balance equation of the water phase is obtained from (2.318) by neglecting the gradient of water density and taking into account that  $S_w = 1$ ,  $k^{rw} = 1$ , and  $\dot{m} = 0$  for fully saturated medium. Thus we have

$$\left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} + \alpha \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} - \beta_{sw} \frac{\partial T}{\partial t} + \nabla^T \left[ \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] = 0 \quad (10.4)$$

where  $\beta_{sw} = (\alpha - n)\beta_s + n\beta_w$  is the thermal expansion coefficient of the medium. The enthalpy balance equation in this case is of the form

$$(\rho C_p)_{\text{eff}} \frac{\partial T}{\partial t} + \left( \rho_w C_p^w \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right) \cdot \nabla T - \nabla^T (\chi_{\text{eff}} \nabla T) = 0 \quad (10.5)$$

where the effective heat capacity is described by

$$(\rho C_p)_{\text{eff}} = \rho_s C_p^s + \rho_w C_p^w \quad (10.6)$$

The constitutive relationship for the solid, assuming linear isotropic elastic behaviour, can be written as

$$\boldsymbol{\sigma}'' = \mathbf{D}_e(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_0 - \boldsymbol{\varepsilon}^T) \quad (10.7)$$

where the total strain  $\boldsymbol{\varepsilon} = \mathbf{L}\mathbf{u}$ , thermal strain  $\boldsymbol{\varepsilon}^T = \mathbf{m}(\beta_s/3)T$ , and the term  $\boldsymbol{\varepsilon}_0$  represents all other strains which are not directly associated with the stress changes. For non-linear behaviour of the solid phase, equation (10.7) becomes

$$d\boldsymbol{\sigma}'' = \mathbf{D}_T(d\boldsymbol{\varepsilon} - d\boldsymbol{\varepsilon}_0 - d\boldsymbol{\varepsilon}^T) \quad (10.8)$$

where the total strain increment  $d\boldsymbol{\varepsilon} = \mathbf{L}d\mathbf{u}$ , and the thermal strain increment  $d\boldsymbol{\varepsilon}^T = \mathbf{m}(\beta_s/3)dT$ .

The capillary pressure relationship for this case is given by

$$S_w = S_w(p^w, T) \quad (10.9)$$

and the relative permeability term is

$$k^{rw} = k^{rw}(S_w, T) \quad \text{or} \quad k^{rw} = k^{rw}(p^w, T) \quad (10.10)$$

For all other constitutive relationships the reader is advised to read Section 2.4.

### 10.2.2 Initial and boundary conditions

The initial and boundary conditions for both the displacement and water pressure values are similar to those described in Section 3.2.2, but for the non-isothermal problem they must be complemented by an appropriate condition for heat flow. Thus, the initial conditions, which specify the full fields of displacement, water pressure and temperature values at time  $t = 0$ , are

$$\mathbf{u} = \mathbf{u}_0 \quad p^w = p_0^w \quad T = T_0 \quad \text{in } \Omega \text{ and on } \Gamma \quad (10.11)$$

The Dirichlet boundary conditions are given by

$$\begin{aligned} \mathbf{u} &= \hat{\mathbf{u}} && \text{on } \Gamma_u \\ p^w &= \hat{p}^w && \text{on } \Gamma_w \\ T &= \hat{T} && \text{on } \Gamma_t \end{aligned} \quad (20.12)$$

and the Neumann boundary conditions can be written as follows.

For the stresses:

$$\mathbf{l}^T \boldsymbol{\sigma} = \bar{\mathbf{t}} \quad \text{on } \Gamma_u^q \quad (10.13)$$

For the water phase:

$$\rho^w \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g})^T \cdot \mathbf{n} = q^w \quad \text{on } \Gamma_w^q \quad (10.14)$$

And for the heat transfer:

$$(-\chi_{\text{eff}} \nabla T)^T \cdot \mathbf{n} = q^T + \alpha_c(T - T_\infty) \quad \text{on } \Gamma_u^q \quad (10.15)$$

where  $q^T$  is the imposed heat flux  $\alpha_c$  the convective heat transfer coefficient and  $T_\infty$  the temperature in the far field of the indisturbed gas phase. The last term of (10.15), which corresponds to Newton's law of cooling, describes the conditions occurring in most practical situations at the interface between a porous medium and the surrounding fluid (air in this case).

### 10.3 DISCRETISATION FOR NON-ISOTHERMAL CONSOLIDATION OF SATURATED POROUS MEDIA

After application of the weighted residual method to the linear momentum balance equation for the multiphase system, we obtain

$$\int_{\Omega} (\mathbf{Lw})^T \boldsymbol{\sigma} d\Omega = \int_{\Omega} \mathbf{w}^T \rho g d\Omega + \int_{\Gamma_u^q} \mathbf{w}^T \bar{\mathbf{t}} d\Gamma \quad (10.16)$$

With a proper choice of weighting functions for the mass balance equation and the boundary conditions, and then application of Green's theorem (Chapter 3), the weighted residual method gives

$$\begin{aligned} & \int_{\Omega} \left[ \mathbf{w}^{*T} \left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} + \mathbf{w}^{*T} \alpha \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} - \mathbf{w}^{*T} \beta_{sw} \frac{\partial T}{\partial t} \right] d\Omega \\ & - \int_{\Omega} \left[ (\nabla \mathbf{w}^*)^T \left( -\frac{\mathbf{k}}{\mu^w} \nabla p^w + \frac{\mathbf{k}}{\mu^w} \rho^w \mathbf{g} \right) \right] d\Omega + \int_{\Gamma_w^q} \mathbf{w}^{*T} \frac{q^w}{\rho^w} d\Gamma = 0 \end{aligned} \quad (10.17)$$

were the only new term, in comparison with (3.33), is related to the thermal expansion of the medium. The same procedure when applied to the energy balance equation, results in

$$\begin{aligned} & \int_{\Omega} \left\{ \mathbf{w}^{oT} \left[ (\rho C_p)_{\text{eff}} \frac{\partial T}{\partial t} \right] + \mathbf{w}^{oT} \left[ \left( \rho^w C_p^w \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right) \cdot \nabla T \right] \right\} d\Omega \\ & - \int_{\Omega} \left\{ \nabla \mathbf{w}^{oT} (-\chi_{\text{eff}} \nabla T) \right\} d\Omega + \int_{\Gamma_u^q} \mathbf{w}^{oT} [q^T + \alpha_c(T - T_\infty)] d\Gamma = 0 \end{aligned} \quad (10.18)$$

Now, the finite element method may be applied to spatially discretise the governing equations in their weak form. The state variables are expressed in terms of the nodal values and shape functions as

$$\mathbf{u} = \mathbf{N}_u \bar{\mathbf{u}} \quad p^w = \mathbf{N}_p \bar{p}^w \quad T = \mathbf{N}_t \bar{T} \quad (10.19)$$

The introduction of (10.19) into (10.16) to (10.18) and application of the Galerkin method, where the weighting functions  $\mathbf{w}$ ,  $\mathbf{w}^*$  and  $\mathbf{w}^o$  are replaced by the corresponding

shape functions  $\mathbf{N}_u$ ,  $\mathbf{N}_p$ ,  $\mathbf{N}_t$ , gives the following set of equations:

$$\int_{\Omega} (\mathbf{L}\mathbf{N}_u)^T (\boldsymbol{\sigma}'' - \alpha \mathbf{m} \mathbf{p}^w \mathbf{N}_p) d\Omega = \int_{\Omega} \mathbf{N}_u^T \rho \mathbf{g} d\Omega + \int_{\Gamma_u^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \quad (10.20)$$

$$\begin{aligned} & \int_{\Omega} \left[ \mathbf{N}_p^T \left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \mathbf{N}_p^T \alpha \mathbf{m}^T \mathbf{L} \mathbf{N}_u \frac{\partial \bar{\mathbf{u}}}{\partial t} - \mathbf{N}_p^T \beta_{sw} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right] d\Omega \\ & + \int_{\Omega} \left[ -(\nabla \mathbf{N}_p)^T \frac{\mathbf{k}}{\mu^w} \rho^w \mathbf{g} + (\nabla \mathbf{N}_p)^T \frac{\mathbf{k}}{\mu^w} \nabla \mathbf{N}_p \bar{\mathbf{p}}^w \right] d\Omega - \int_{\Gamma_w^q} \mathbf{N}_p^T \frac{q^w}{\rho^w} d\Gamma = 0 \end{aligned} \quad (10.21)$$

$$\begin{aligned} & \int_{\Omega} \left\{ \mathbf{N}_t^T [(\rho C_p)_{eff} \mathbf{N}_t] \frac{\partial \bar{\mathbf{T}}}{\partial t} + \mathbf{N}_t^T \left[ \left( \rho^w C_p^w \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right) \cdot \nabla \mathbf{N}_t \right] \bar{\mathbf{T}} \right\} d\Omega \\ & - \int_{\Omega} \nabla \mathbf{N}_t^T (-\chi_{eff} \nabla \mathbf{N}_t) \bar{\mathbf{T}} d\Omega + \int_{\Gamma_t^q} \mathbf{N}_t^T [q^T + \alpha_c (T - T_\infty)] d\Gamma = 0 \end{aligned} \quad (10.22)$$

These equations can be rewritten as

$$\int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega - \mathbf{Q} \bar{\mathbf{p}}^w = \mathbf{f}^u \quad (10.23)$$

$$\mathbf{H} \bar{\mathbf{p}}^w + \mathbf{Q}^T \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{S} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \mathbf{R} \frac{\partial \bar{\mathbf{T}}}{\partial t} = \mathbf{f}^p \quad (10.24)$$

$$\mathbf{K}_t \bar{\mathbf{T}} + \mathbf{C}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} = \mathbf{f}^t \quad (10.25)$$

where the matrices  $\mathbf{B}$ ,  $\mathbf{Q}$ ,  $\mathbf{H}$  and vector  $\mathbf{f}^p$  are given by (3.40), and other matrices are defined as

$$\mathbf{f}^u = \int_{\Omega} \mathbf{N}_u^T \rho \mathbf{g} d\Omega + \int_{\Gamma_u^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma + \int_{\Omega} \mathbf{B}^T \mathbf{D}_e \boldsymbol{\varepsilon}_0 d\Omega \quad (10.26)$$

$$\mathbf{R} = \int_{\Omega} (-\mathbf{N}_p^T \beta_{sw} \mathbf{N}_t) d\Omega \quad (10.27)$$

$$\mathbf{S} = \int_{\Omega} \left[ \mathbf{N}_p^T \left( \frac{\alpha - n}{K_s} + \frac{n}{K_w} \right) \mathbf{N}_p \right] d\Omega \quad (10.28)$$

$$\mathbf{K}_t = \int_{\Omega} \left\{ \mathbf{N}_t^T \left[ \rho^w C_p^w \frac{\mathbf{k}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] \cdot \nabla \mathbf{N}_t - \nabla \mathbf{N}_t^T (-\chi_{eff} \nabla \mathbf{N}_t) \right\} d\Omega \quad (10.29)$$

$$\mathbf{C}_t = - \int_{\Gamma_t^q} \mathbf{N}_t^T [q^T + \alpha_c (T - T_\infty)] d\Gamma \quad (10.30)$$

Using (3.42) and (3.43), together with the constitutive relations for the isotropic linear elastic case (10.7), the internal force matrix is

$$\begin{aligned} \mathbf{P}(\bar{\mathbf{u}}) &= \int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega = \int_{\Omega} \mathbf{B}^T \mathbf{D}_e (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_0 - \boldsymbol{\varepsilon}^T) d\Omega \\ &= \mathbf{K}_e \bar{\mathbf{u}} + \mathbf{K}_{te} \bar{\mathbf{T}} - \int_{\Omega} \mathbf{B}^T \mathbf{D}_e \boldsymbol{\varepsilon}_0 d\Omega \end{aligned} \quad (10.31)$$

where

$$\mathbf{K}_e = \int_{\Omega} \mathbf{B}^T \mathbf{D}_e \mathbf{B} d\Omega \quad (10.32)$$

and

$$\mathbf{K}_{te} = - \int_{\Omega} \mathbf{B}^T \mathbf{D}_e \mathbf{m} \frac{\beta_s}{3} \mathbf{N}_t d\Omega \quad (10.33)$$

For non-linear behaviour of the solid phase, only the tangential matrices can be established, i.e.

$$\frac{\partial \mathbf{P}(\bar{\mathbf{u}})}{\partial \bar{\mathbf{u}}} = \mathbf{K}_T = \int_{\Omega} \mathbf{B}^T \mathbf{D}_T \mathbf{B} d\Omega \quad (10.34)$$

$$\frac{\partial \mathbf{P}(\bar{\mathbf{u}})}{\partial \bar{T}} = \mathbf{K}_{tT} = - \int_{\Omega} \mathbf{B}^T \mathbf{D}_T \mathbf{m} \frac{\beta_s}{3} \mathbf{N}_t d\Omega \quad (10.35)$$

So, for a unified treatment, it is useful to time differentiate the linear momentum balance equation. The governing equations can then be written for the general case as

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_t \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \\ \bar{T} \end{Bmatrix} + \begin{bmatrix} \mathbf{K}_T & -\mathbf{Q} & \mathbf{K}_{tT} \\ \mathbf{Q}^T & \mathbf{S} & \mathbf{R} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_t \end{bmatrix} \frac{\partial}{\partial t} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \\ \bar{T} \end{Bmatrix} = \begin{Bmatrix} \frac{\partial \mathbf{f}^u}{\partial t} \\ \mathbf{f}^p \\ \mathbf{f}' \end{Bmatrix} \quad (10.36)$$

The next step is the discretisation in time and the solution of the resulting system of equations.

## 10.4 SOLUTION PROCEDURES

The analysis of heat and fluid flow through a fully saturated, deforming porous medium represents a field problem with three degrees of freedom. The most obvious solution procedure of the three coupled, semidiscrete equations (10.36) developed in the previous section, consists of adding the temperature  $T$  as an additional variable to the existing nodal variables, i.e. the displacements  $\mathbf{u}$  and the pressure  $\mathbf{p}$  of the isothermal case and to solve simultaneously the system of equations. This has been done, as shown by (10.36), and is known as the monolithic augmentation approach [26], first proposed by Lewis and Karahanoglu [25]. This results in an unsymmetric coefficient matrix for the combined equations after discretisation in time. Another solution procedure consists of partitioning the combined system of equations after the integration in the time domain. This procedure is known as algebraic partitioning, and was already used in Section 6.2.3. It has been shown that this method provides a superior implementation flexibility when compared with differential partitioning. It is well known that a great number of partitions are possible in three-field problems. One particular partition, proposed by Schrefler [27], produces a symmetric coefficient matrix and an effective gain in

computing time, as shown in Section 10.7.3. This partitioning procedure is described in Section 10.4.2.

Both the monolithic augmentation approach and the partitioned solution procedure refer to the same discretisation process for the time domain. This discretisation is dealt with in the next section, along with the monolithic augmentation approach (direct solution).

### 10.4.1 Monolithic Augmentation Approach

For the discretisation in time of (10.36) the single-step finite difference operator (3.49) is applied, i.e.

$$[\mathbf{B} + \theta \Delta t \mathbf{C}]_{n+0} \mathbf{X}_{n+1} - [\mathbf{B} - (1-\theta) \Delta t \mathbf{C}]_{n+0} \mathbf{X}_n - \Delta t \mathbf{F}_{n+0} = 0 \quad (10.37a)$$

or

$$\mathbf{A} \mathbf{X}_{n+1} - \mathbf{B} \mathbf{X}_n - \mathbf{f} = 0 \quad (10.37b)$$

The matrices  $\mathbf{B}$ ,  $\mathbf{C}$  and vectors,  $\mathbf{X}$ ,  $\mathbf{F}$  of this operator are as follows:

$$\mathbf{B} = \begin{bmatrix} \mathbf{K}_T & -\mathbf{Q} & \mathbf{K}_{tT} \\ \mathbf{Q}^T & \mathbf{S} & \mathbf{R} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \end{bmatrix} \quad (10.38)$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_t \end{bmatrix} \quad (10.39)$$

$$\mathbf{X} = \left\{ \begin{array}{c} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \\ \bar{\mathbf{T}} \end{array} \right\} \quad \mathbf{F} = \left\{ \begin{array}{c} \frac{\partial \mathbf{f}^u}{\partial t} \\ \mathbf{f}^p \\ \mathbf{f}' \end{array} \right\} \quad (10.40)$$

The definitions of  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{f}$  are obvious from inspection of (10.37a). But equation (10.36) is non-linear because of the convective term occurring in matrix  $\mathbf{K}$ , (10.29), so an appropriate solution procedure is required even for linear solid-phase behaviour. One method which has been extensively used is the Newton–Raphson method (Chapter 4 and Section 10.6). Furthermore, the system is coupled because of the form of  $\mathbf{K}_t$ , and if the convective term is of importance, then the use of a fully implicit scheme ( $\theta = 1$ ) is advisable [23]. The stability, consistency and an error analysis of the time discretisation scheme are extensively discussed in Section 3.7.

### 10.4.2 Partitioned solution procedures

Matrix partitioning, together with staggered or simultaneous solution schemes, plays an important role in the design of numerical models for coupled problems. This section

looks at a matrix partitioning technique that is particularly advantageous here and it discusses the standard staggered solution procedure. The standard staggered procedure has been described and analysed in the literature [26–30]; it is essentially an iteration method of the Gauss–Seidel type which can be used to solve large, coupled sets of algebraic equations resulting from a discretisation of the appropriate field equations.

The advantage of the staggered procedure is that it permits a sequential solution of the equations, allowing the use of available numerical codes for simpler problems. The main concept of the staggered strategy is to begin by solving a block of equations for the first set of field variables while the other variables remain fixed. The remaining block of equations are then solved for the second set of field variables while the updated first variables are kept fixed. This procedure is obtained by performing an appropriate partitioning of the matrices (operators) on the left-hand side and transferring components to the right hand side of the equation.

The idea of partitioning has also been associated with fractional step methods and in this context was analysed by Armero and Simo [31]. The fractional step method was first introduced into the Russian literature by Yanenko [32] and Marchuk [33]. This is a scheme which first reduces a complicated problem into a sequence of simpler ones by splitting up (partitioning) the operator of the equations requiring solution. In this way we obtain differential equations of a smaller order or simpler form. To preserve the convergence and stability properties of the full operator, the component operators must exhibit appropriate group characteristics [33–35].

A modification of the standard staggered procedure is the staggered Newton procedure, which has been formulated and implemented for weakly coupled thermomechanical problems [36–39]. This staggered Newton procedure is briefly discussed in Appendix 10A. Extensive numerical experiments on consolidation problems [40] have shown that staggered solutions are in general only advantageous if the coupling is weak. This is precisely the case if the strongly coupled displacement and pressure fields are solved implicitly while partitioning is used between these two fields and the temperature field.

This partitioning technique will be considered next after briefly recalling the necessary steps required for the partitioning process. The matrix  $[\mathbf{B} + \theta\Delta t\mathbf{C}]$  appearing on the left-hand side of (10.37) is represented as a sum, i.e.

$$[\mathbf{B} + \theta\Delta t\mathbf{C}] = [\mathbf{B}^L + \theta\Delta t\mathbf{C}^L] + [\mathbf{B}^R + \theta\Delta t\mathbf{C}^R] \quad (10.41)$$

Rearranging (10.37) and on the right-hand side denoting  $\mathbf{X}_{n+1}$  as  $\mathbf{X}_{n+1}^{(P)}$ , we obtain the partitioned equation

$$[\mathbf{B}^L + \theta\Delta t\mathbf{C}^L]\mathbf{X}_{n+1} = [\mathbf{B} - (1-\theta)\Delta t\mathbf{C}]\mathbf{X}_n - [\mathbf{B}^R + \theta\Delta t\mathbf{C}^R]\mathbf{X}_{n+1}^{(P)} - \Delta t\mathbf{F}_{n+0} \quad (10.42)$$

where  $\mathbf{X}_{n+1}^{(P)}$  is a suitable predictor, which usually depends in a linear manner on the already calculated values  $\mathbf{X}_{n-i}$ , i.e.

$$\mathbf{X}_{n+1}^{(P)} = \sum_{i=0}^n a_i \mathbf{X}_{n-i} \quad (10.43)$$

Within each time step, a predictor–corrector (PC) method may be used to solve (10.42). First we set the predictor  $\mathbf{X}_{n+1}^{(p)} = \mathbf{X}_{n+1}^0$ , perhaps using (10.43), then we carry out the correction. That is,  $\mathbf{X}_{n+1}^0$  is substituted into the right-hand side of (10.42) and the term  $\mathbf{X}_{n+1}^1$  is evaluated. This in turn is substituted into the right-hand side of (10.42), which is re-evaluated, and a new value  $\mathbf{X}_{n+1}^2$  is produced. The cycle is repeated until successive values differ by less than a required accuracy  $\epsilon$ :

$$|\mathbf{X}_{n+1}^s - \mathbf{X}_{n+1}^{s-1}| < \epsilon$$

where  $s$  is the iteration number and  $\mathbf{X}_{n+1}$  is taken as  $\mathbf{X}_{n+1} = \mathbf{X}_{n+1}^s$ . Some predictor formulae allow the most recent iterate to be updated before performing the subsequent corrector step [33].

There are many ways of partitioning the matrix  $[\mathbf{B} + \theta\Delta t \mathbf{C}]$  in (10.41), so there are many different forms of (10.42) [34]. We have chosen a partitioning scheme that enables us to use the staggered procedure, i.e. to solve the equations in sequence. In order to exploit the symmetry of the partitioned matrices for the case of a symmetric stiffness matrix, the first equation is multiplied by  $(-1)$  and the partitioning used is given by

$$\mathbf{A}^L = [\mathbf{B}^L + \theta\Delta t \mathbf{C}^L] = \begin{bmatrix} -\mathbf{K}_T & \mathbf{Q} & \mathbf{0} \\ \mathbf{Q}^T & \mathbf{S} + \theta\Delta t \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_t + \theta\Delta t \mathbf{K}_t \end{bmatrix} \quad (10.44)$$

$$\mathbf{A}^R = [\mathbf{B}^R + \theta\Delta t \mathbf{C}^R] = \begin{bmatrix} \mathbf{0} & \mathbf{0} & -\mathbf{K}_{tT} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (10.45)$$

Substitution into (10.42) yields

$$\begin{aligned} & \begin{bmatrix} -\mathbf{K}_T & \mathbf{Q} & \mathbf{0} \\ \mathbf{Q}^T & \mathbf{S} + \theta\Delta t \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_t + \theta\Delta t \mathbf{K}_t \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}}_{n+1} \\ \bar{\mathbf{p}}_{n+1}^w \\ \bar{\mathbf{T}}_{n+1} \end{Bmatrix} \\ &= \begin{bmatrix} -\mathbf{K}_T & \mathbf{Q} & \mathbf{K}_{tT} \\ \mathbf{Q}^T & \mathbf{S} - (1-\theta)\Delta t \mathbf{H} & \mathbf{R} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_t - (1-\theta)\Delta t \mathbf{K}_t \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}}_n \\ \bar{\mathbf{p}}_n^w \\ \bar{\mathbf{T}}_n \end{Bmatrix} \\ & - \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{K}_{tT} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{0} \\ \bar{\mathbf{T}}_{n+1}^{(p)} \end{Bmatrix} + \Delta t \begin{Bmatrix} \frac{\partial \mathbf{f}^u}{\partial t} \\ \mathbf{f}^p \\ \mathbf{f}^t \end{Bmatrix} \end{aligned} \quad (10.46)$$

Notice that the systems split into the first two sets of equations associated with  $\{\bar{\mathbf{u}}, \bar{\mathbf{p}}^w\}^T$  and the third set of equations associated with  $\bar{\mathbf{T}}$ .

The staggered procedure is initiated by setting the predictor  $\bar{\mathbf{T}}_{n+1}^{(p)}$ , then evaluating  $\bar{\mathbf{u}}_{n+1}$  and  $\bar{\mathbf{p}}_{n+1}^w$  from the first two sets of equations (10.46). Subsequently, the last set of equations, where the evaluated  $\bar{\mathbf{p}}_{n+1}^w$  is substituted in  $\mathbf{K}_T$ , is finally solved for  $\bar{\mathbf{T}}_{n+1}$ . This is known as the corrector step and the cycle must be repeated until the PC procedure

converges. The next time integration step proceeds in the same way. Note that iterations within one time step generally cannot be avoided, as discussed in the next section.

It is also important that a suitable predictor is chosen for  $\bar{\mathbf{T}}_{n+1}$ . A predictor of the form  $\bar{\mathbf{T}}_{n+1}^{(p)} = 2 \cdot \bar{\mathbf{T}}_n - \bar{\mathbf{T}}_{n-1}$  results in a very rapidly convergent scheme, whereas the last term predictor  $\bar{\mathbf{T}}_{n+1}^{(p)} = \bar{\mathbf{T}}_n$  causes an uncoupling of the system.

If the convective terms in matrix  $\mathbf{K}_t$  (10.29) are neglected, then the temperature field will uncouple from the other two fields and no iteration will be necessary between the two groups of fields.

#### 10.4.2.1 Numerical properties of partitioned procedures

It is assumed that the consistency and convergence characteristics of the finite element discretisation in space is maintained by a proper choice of elements (Section 3.5), so it is only necessary to establish the convergence of the temporal discretisation. The numerical properties of partitioned difference equations depend not only on the time-stepping formula but also on the partition and on the predictor formula [26].

The error analysis is carried out along the lines of Section 3.10, following Turska *et al.* [30] and Turska and Schrefler [41], where the number of iterations are taken into account. This is not the case in conventional error analysis, as in Section 6.2.3 and in Park and Felippa [26].

The partitioning of matrix  $\mathbf{A}$  in (10.37b) is chosen as follows:

$$\mathbf{A} = \mathbf{A}^L + \mathbf{A}^R \quad (10.47)$$

$$\mathbf{A}^L = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{a}_{22} \end{bmatrix} \quad \mathbf{A}^R = \begin{bmatrix} \mathbf{0} & \mathbf{a}_{12} \\ \mathbf{a}_{21} & \mathbf{0} \end{bmatrix} \quad (10.48)$$

Then (10.37b) can be written as

$$\mathbf{A}^L \mathbf{X}_{n+1} = \mathbf{B} \mathbf{X}_n + \hat{\mathbf{F}} - \mathbf{A}^R \mathbf{X}_{n-1} \quad (10.49)$$

Some rearrangement leads to the partitioned form of (10.37):

$$\mathbf{x}_{n+1} = -\mathbf{G}_1 \mathbf{y}_{n+1} + \mathbf{H}_1 \mathbf{X}_n + \mathbf{F}_1 \quad \mathbf{y}_{n+1} = -\mathbf{G}_2 \mathbf{x}_{n+1} + \mathbf{H}_2 \mathbf{X}_n + \mathbf{F}_2 \quad (10.50)$$

where

$$\begin{aligned} \mathbf{X}_{n+1} &= [\mathbf{x}_{n+1}, \mathbf{y}_{n+1}]^T \\ \hat{\mathbf{F}} &= [\hat{\mathbf{F}}_1, \hat{\mathbf{F}}_2]^T \quad \mathbf{F}_1 = \mathbf{a}_{11}^{-1} \hat{\mathbf{F}}_1 \quad \mathbf{F}_2 = \mathbf{a}_{22}^{-1} \hat{\mathbf{F}}_2 \\ \mathbf{G}_1 &= \mathbf{a}_{11}^{-1} \mathbf{a}_{12} \quad \mathbf{G}_2 = \mathbf{a}_{22}^{-1} \mathbf{a}_{21} \\ \mathbf{H}_1 &= \mathbf{a}_{11}^{-1} [\mathbf{b}_{11}, \mathbf{b}_{12}] \quad \mathbf{H}_2 = \mathbf{a}_{22}^{-1} [\mathbf{b}_{21}, \mathbf{b}_{22}] \end{aligned} \quad (10.51)$$

$$\text{and } \mathbf{G} = \begin{bmatrix} \mathbf{0} & \mathbf{G}_1 \\ \mathbf{G}_2 & \mathbf{0} \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix}$$

Matrices  $\mathbf{G}_1, \mathbf{G}_2$  depend on the time increment  $\Delta t$ .

We define  $K$  as the number of iterations performed,  $K = K(n)$ . The predictor  $\mathbf{x}_{n+1}^0$  has the form

$$\mathbf{x}_{n+1}^0 = \sum_{i=0}^m \beta_i \mathbf{x}_{n-i}^K \quad (10.52)$$

The staggered scheme is given by

$$\mathbf{x}_{n+1}^k = -\mathbf{G}_1 \mathbf{y}_{n+1}^{k-1} + \mathbf{H}_1 \mathbf{X}_n^K + \mathbf{F}_1 \quad (10.53)$$

$$\mathbf{y}_{n+1}^{k-1} = -\mathbf{G}_2 \mathbf{x}_{n+1}^{k-1} + \mathbf{H}_2 \mathbf{X}_n^K + \mathbf{F}_2 \quad (10.54)$$

The global error,  $\mathbf{e}_{n+1}^K = \mathbf{X}_{n+1} - \tilde{\mathbf{X}}_{n+1}^K$ , where  $\tilde{\mathbf{X}}_{n+1}^K$  is the numerical solution of (10.53) and (10.54), satisfies

$$\mathbf{e}_{n+1}^K = (\mathbf{G})^{2K} \mathbf{e}_{n+1}^0 + (\mathbf{I} - (\mathbf{G})^{2K})(\mathbf{I} + \mathbf{G})^{-1} \mathbf{H} \mathbf{e}_n^K + (\mathbf{I} - (\mathbf{G})^{2K})(\mathbf{I} + \mathbf{G})^{-1} \mathbf{r}_{n+1} \quad (10.55)$$

where  $\mathbf{r}_{n+1}$  is the local truncation error.

The sufficient conditions for global convergence are iteration convergence  $\|\mathbf{G}\| < 1$ , stability  $\|\mathbf{Q}_n\| = \|(\mathbf{I} - (\mathbf{G})^{2K})(\mathbf{I} + \mathbf{G})^{-1} \mathbf{H}\| < 1$  and consistency  $\mathbf{r}_{n+1} = \mathbf{O}(\Delta t^2)$ . In a concise form, the error for the staggered scheme (10.55) can be written as

$$\mathbf{e}_{n+1}^K = \mathbf{P}_{n+1} \mathbf{e}_{n+1}^0 + \mathbf{Q}_{n+1} \mathbf{e}_n^K + \mathbf{R}_{n+1} \mathbf{r}_{n+1} \quad (10.56)$$

Recursively for  $n$ , we have from (10.55) that

$$\begin{aligned} \mathbf{e}_{n+1}^K &= \mathbf{P}_{n+1} \mathbf{e}_{n+1}^0 + \sum_{l=0}^n \mathbf{Q}_{n+1} \dots \mathbf{Q}_{n+1-l} \mathbf{P}_{n-l} \mathbf{e}_{n-l}^0 + \mathbf{Q}_{n+1} \dots \mathbf{Q}_1 \mathbf{e}_0^0 \\ &\quad + \sum_{l=0}^n \mathbf{Q}_{n+1} \dots \mathbf{Q}_{n+1-l} \mathbf{R}_{n-l} \mathbf{r}_{n-l} + \mathbf{R}_{n+1} \mathbf{r}_{n+1} \end{aligned} \quad (10.57)$$

Notice that  $\mathbf{e}_0^0$  is the ‘initial’ round-off error.

For the staggered scheme, even for a matrix  $\mathbf{G}$  which has real and positive eigenvalues and  $\|\mathbf{G}\| < 1$ , the stability condition for one iteration, i.e.  $\|(\mathbf{I} - \mathbf{G})\mathbf{H}\| < 1$ , does not result in a stable solution for an arbitrary value of  $K$ . This is caused by the fact that  $\|(\mathbf{I} - (\mathbf{G})^{2K})(\mathbf{I} - (\mathbf{G})^2)^{-1}\|$  may be greater than 1, and therefore

$$\|(\mathbf{I} - (\mathbf{G})^{2K})(\mathbf{I} + \mathbf{G})^{-1} \mathbf{H}\| = \|(\mathbf{I} - (\mathbf{G})^{2K})(\mathbf{I} - (\mathbf{G})^2)^{-1}(\mathbf{I} - \mathbf{G})\mathbf{H}\|$$

may be also greater than 1.

Finally, note that in partitioned solution procedures it is not possible to obtain a required accuracy simply by reducing the size of the time step and performing a constant number of iterations. For partitioned solution procedures which use a constant number of iterations within one time step, Turska *et al.* [30] has shown that the global error is

bounded but does not tend to zero. This suggests that, for a fixed number of iterations, the partitioned solution procedure converges to the solution of another equation. Turska and Schrefler [41] have reported these equations for the case of a last solution predictor, and an example is given in Schrefler *et al.* [42].

Furthermore, in some cases there exists lower bound for  $\Delta t/h^2$ , which implies it is not always possible to decrease  $\Delta t$  without also decreasing the mesh size  $h$ . Such a bound was independently found for the solution of consolidation problems using an iteration convergence analysis by Turska *et al.* [28] and Saetta *et al.* [43]. Also, Murti *et al.* [44] obtained such a bound by using a conditioning number, and Rank *et al.* [45] achieved a similar result for other problems by means of the discrete maximum principle.

## 10.5 NON-ISOTHERMAL AIRFLOW AND WATER FLOW IN A DEFORMING POROUS MEDIUM

This Section details the full model for heat and mass transfer in partially saturated media, as developed in Section 10.2, and solves some examples numerically. There are many applications of this numerical model, as will be shown in Section 10.7. The model has also been successfully applied to the drying of porous building materials [46].

There are various possible sets of primary variables which can be chosen for this type of problem. Firstly, a variable describing the gas flow is required, which can either be the total gas pressure  $p^g$  or the dry air partial pressure  $p^{ga}$ . Then, for the water species flow, the water pressure  $p^w$ , the capillary pressure  $p^c$ , the vapour partial pressure  $p^{gw}$ , or the vapour density  $\rho^{gw}$  may be used. Note that all combinations of these variables are possible, because they uniquely describe the state of liquid water, the gas phase and the water vapour (applying the perfect gas law and/or the Kelvin equation). For the heat flow equation, the preferred choice is usually temperature  $T$ ; and for the mechanical behaviour, the preferred choice is usually the displacement vector  $\mathbf{u}$ . Other possible variables for the fluid phases are the respective degrees of saturation, but they are discontinuous at the interface between different materials. In this section the gas pressure  $p^g$ , the water pressure  $p^w$ , the temperature  $T$  and the displacement vector  $\mathbf{u}$  are chosen as the primary variables.

### 10.5.1 Governing equations

The linear momentum balance equation, expressed in terms of total stresses, has the same form as for the fully saturated case (10.1), i.e.

$$\mathbf{L}^T \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \quad (10.58)$$

However, when invoking the effective stress principle, the gas phase must also be considered:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}'' - \alpha \mathbf{m}^T (S_w p^w + S_g p^g) \quad (10.59)$$

and the density of the multiphase medium is

$$\rho = (1 - n)\rho^s + nS_w\rho^w + nS_g\rho^g \quad (10.60)$$

The mass balance equation of the water species (vapour and liquid water) in its full form is given by (2.259), i.e.

$$\begin{aligned} & \left( \frac{\alpha - n}{K_s} S_w (\rho^{gw} S_g + \rho^w S_w) + \rho^w S_w \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_g (\rho^{gw} S_g + \rho^w S_w) \frac{\partial p^g}{\partial t} \\ & + \alpha (\rho^{gw} S_g + \rho^w S_w) \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} - S_w \beta_{swg} \frac{\partial T}{\partial t} + S_g n \frac{\partial}{\partial t} \left( \frac{M_w}{\theta R} p^{gw} \right) \\ & + \left[ \frac{\alpha - n}{K_s} (\rho^{gw} S_g p^c + \rho^w S_w p^w - \rho^w S_w p^c) + n(\rho^w - \rho^{gw}) \right] \frac{\partial S_w}{\partial t} \\ & - \nabla^T \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \nabla \left( \frac{p^{gw}}{p^g} \right) \right] \\ & + \nabla^T \left[ \rho^{gw} \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] + \nabla^T \left[ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] = 0 \end{aligned} \quad (10.61)$$

where

$$\beta_{swg} = \beta_s(\alpha - n)(\rho^{gw} S_g + \rho^w S_w) + n\beta_w \rho^w S_w \quad (10.62)$$

Note how this equation takes into account any phase change, i.e. evaporation-condensation phenomena, even if the mass source term is not explicitly present. The mass source term was eliminated by summing both the mass conservation equations of the liquid water equation (2.200) and the vapour equation (2.213).

The mass balance equation of gas is in fact the dry air conservation equation when the source term is expressed through the vapour mass conservation equation; it reads

$$\begin{aligned} & \frac{\alpha - n}{K_s} S_w S_g \rho^g \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_g^2 \rho^g \frac{\partial p^g}{\partial t} - \left( \frac{\alpha - n}{K_s} S_g p^c + n \right) \rho^g \frac{\partial S_w}{\partial t} - \rho^g \beta_{sg} \frac{\partial T}{\partial t} \\ & + \alpha S_g \rho^g \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + S_g n \frac{\partial}{\partial t} \left[ \frac{1}{\theta R} (p^{ga} M_a + p^{gw} M_w) \right] \\ & + \nabla^T \left[ \rho^g \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] = \dot{m} \end{aligned} \quad (10.63)$$

where

$$\begin{aligned} \dot{m} = & \frac{\alpha - n}{K_s} S_w S_g \rho^{gw} \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_g^2 \rho^{gw} \frac{\partial p^g}{\partial t} - \left( \frac{\alpha - n}{K_s} S_g p^c + n \right) \rho^{gw} \frac{\partial S_w}{\partial t} \\ & - \rho^{gw} \beta_{sg} \frac{\partial T}{\partial t} + \alpha S_g \rho^{gw} \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + S_g n \frac{\partial}{\partial t} \left( \frac{1}{\theta R} p^{gw} M_w \right) \\ & + \nabla^T \left[ \rho^{gw} \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] - \nabla^T \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \nabla \left( \frac{p^{gw}}{p^g} \right) \right] \end{aligned} \quad (10.64)$$

$$\beta_{sg} = \beta_s(\alpha - n)S_g \quad (10.65)$$

$$\beta_{sw} = \rho^w S_w [\beta_s(\alpha - n) + n\beta_g] \quad (10.66)$$

The enthalpy balance equation of the multiphase medium takes the following form:

$$(\rho C_p)_{\text{eff}} \frac{\partial T}{\partial t} + (nS_w \rho^w C_p^w \mathbf{v}^w + nS_g \rho^g C_p^g \mathbf{v}^g) \cdot \nabla T - \nabla^T (\chi_{\text{eff}} \nabla T) = -\dot{m} \Delta H_{\text{vap}} \quad (10.67)$$

where the effective heat capacity in this case is given by

$$(\rho C_p)_{\text{eff}} = (1 - n)\rho^s C_p^s + nS_w \rho^w C_p^w + nS_g \rho^g C_p^g \quad (10.68)$$

and the mass source term emanates from the water mass conservation equation

$$\begin{aligned} \dot{m} = & -\rho^w \left( \frac{\alpha - n}{K_s} S_w^2 + S_w \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} - \rho^w \frac{\alpha - n}{K_s} S_g S_w \frac{\partial p^g}{\partial t} - \alpha \rho^w S_w \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} \\ & + \beta_{sw} \frac{\partial T}{\partial t} - \rho^w \left[ \frac{\alpha - n}{K_s} S_w p^w - \frac{\alpha - n}{K_s} S_w p^g + n \right] \frac{\partial S_w}{\partial t} \quad (10.69) \\ & - \nabla^T \left[ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] \end{aligned}$$

The constitutive relationship for a linear isotropic material is given by (10.7), and for non-linear behaviour by (10.8). As in Section 10.2.1, the capillary pressure relationship has the form of (10.9) with the additional condition that

$$S_w + S_g = 1 \quad (10.70)$$

Here we need both the relative permeability for liquid water (10.10) and for the gas phase, i.e.

$$k^{rg} = k^{rg}(S_g, T) \quad \text{or} \quad k^{rg} = k^{rg}(p^g, T) \quad (10.71)$$

### 10.5.2 Initial and boundary conditions

In comparison with the non-isothermal saturated medium (Section 10.2.2), the initial and boundary conditions are required with appropriate conditions for the gas phase, modified to take into account any phase change which may take place. As usual, the initial conditions describe the full field of variables, i.e. displacements, water pressures, gas pressures and temperatures at time zero:

$$\mathbf{u} = \mathbf{u}_0 \quad p^w = p_0^w \quad p^g = p_0^g \quad T = T_0 \quad \text{in } \Omega \text{ and on } \Gamma \quad (10.72)$$

The boundary conditions can be of the Dirichlet type, i.e.

$$\begin{aligned} \mathbf{u} &= \hat{\mathbf{u}} && \text{on } \Gamma_u \\ p^w &= \hat{p}^w && \text{on } \Gamma_w \\ p^g &= \hat{p}^g && \text{on } \Gamma_g \\ T &= \hat{T} && \text{on } \Gamma_t \end{aligned} \quad (10.73)$$

Or the Neumann type, i.e.

$$\mathbf{l}^T \boldsymbol{\sigma} = \hat{\mathbf{t}} \quad \text{on } \Gamma_n^q \quad (10.74)$$

Or the respective fluxes of each phase may be required, i.e. for the water flux:

$$\begin{aligned} &\left[ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g})^T + \rho^g w \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g})^T + \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \nabla \left( \frac{p^{gw}}{p^g} \right) \right] \cdot \mathbf{n} \\ &= q^w + q^{gw} + \beta_c (\rho^{gw} - \rho_\infty^{gw}) \quad \text{on } \Gamma_w^q \end{aligned} \quad (10.75)$$

for the gas flux:

$$\left[ \rho^{ga} \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g})^T - \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \nabla \left( \frac{p^{gw}}{p^g} \right) \right] \cdot \mathbf{n} = q^{ga} \quad \text{on } \Gamma_g^q \quad (10.76)$$

and for the heat flux:

$$(\rho_w v^w \Delta H_{vap} - \chi_{eff} \nabla T) \cdot \mathbf{n} = q^T + \alpha_c (T - T_\infty) \quad \text{on } \Gamma_T^q \quad (10.77)$$

where  $q^{ga}$  is the imposed mass flux of dry air,  $\alpha_c$  is the convective mass transfer coefficient and  $\rho_\infty^{gw}$  is the partial density of vapour in the surrounding air at a great distance from the interface. The last term of (10.75) is analogous to the last term in (10.15); it causes cooling as the porous bodies dry, which is well known in practice. Evaporation of water from the surface of the body requires energy, the so-called latent heat of evaporation, which comes from the internal energy of the medium.

Convective heat and mass transfer are analogous processes, i.e. they are governed by dimensionless equations of the same type [47]. This implies that the boundary layer profiles for the temperature and vapour gradients must be of the same functional form. This has a very important practical consequence, because the convectional heat and mass transfer coefficients are strictly related to each other. For moist air in a temperature range which is typical for most practical situations, i.e. about 20 °C, then

$$\frac{\beta_c}{\alpha_c} = \frac{D_g}{\lambda_g} \cong \frac{2.6 \times 10^{-5} [\text{m}^2 \text{s}^{-1}]}{2.63 \times 10^{-2} [\text{W m}^{-1} \text{K}^{-1}]} \cong 10^{-3} [\text{K m J}^{-1}] \quad (10.78)$$

Hence, knowing the value of either one of the two convection transfer coefficients, the value of the other may easily be calculated from (10.78).

## 10.6 DISCRETISATION FOR AIRFLOW AND WATER FLOW IN A DEFORMING POROUS MEDIUM

An application of the weighted residual method to the linear momentum balance equation for the multiphase system results in

$$\int_{\Omega} (\mathbf{Lw})^T \boldsymbol{\sigma} d\Omega = \int_{\Omega} \mathbf{w}^T \rho \mathbf{g} d\Omega + \int_{\Gamma_w^q} \mathbf{w}^T \bar{\mathbf{t}} d\Gamma \quad (10.79)$$

With appropriate choice of weighting functions and subsequent application of Green's theorem to the water mass conservation equation and the corresponding boundary equation, the weighted residual method yields

$$\begin{aligned} & \int_{\Omega} \mathbf{w}^{*T} \left[ \left( \frac{\alpha - n}{K_s} S_w (\rho^{gw} S_g + \rho^w S_w) + \rho^w S_w \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} \right. \\ & \quad \left. + \frac{\alpha - n}{K_s} S_g (\rho^{gw} S_g + \rho^w S_w) \frac{\partial p^g}{\partial t} \right] d\Omega \\ & + \int_{\Omega} \mathbf{w}^{*T} \left[ \alpha (\rho^{gw} S_g + \rho^w S_w) \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} - S_w \beta_{swg} \frac{\partial T}{\partial t} + S_g n \frac{\partial}{\partial t} \left( \frac{M_w}{\theta R} p^{gw} \right) \right] d\Omega \\ & + \int_{\Omega} \mathbf{w}^{*T} \left\{ \left[ \frac{\alpha - n}{K_s} (\rho^{gw} S_g p^c + \rho^w S_w p^w - \rho^w S_w p^c) + n(\rho^w - \rho^{gw}) \right] \frac{\partial S_w}{\partial t} \right\} d\Omega \\ & - \int_{\Omega} (\nabla \mathbf{w}^*)^T \left[ -\rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \nabla \left( \frac{p^{gw}}{p^g} \right) + \rho^{gw} \frac{\mathbf{k} k^{rw}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] d\Omega \\ & - \int_{\Omega} (\nabla \mathbf{w}^*)^T \left[ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] d\Omega \\ & + \int_{\Gamma_w^q} \mathbf{w}^{*T} [q^w + q^{gw} + \beta_c (\rho^{gw} - \rho_\infty^{gw})] d\Gamma = 0 \end{aligned} \quad (10.80)$$

The same procedure for the gas mass conservation equation, taking into account that  $\rho^{ga} = \rho^g - \rho^{gw}$ , gives

$$\begin{aligned} & \int_{\Omega} \mathbf{w}^{*T} \left[ \frac{\alpha - n}{K_s} S_w S_g \rho^{ga} \frac{\partial p^w}{\partial t} + \frac{\alpha - n}{K_s} S_g^2 \rho^{ga} \frac{\partial p^g}{\partial t} - \left( \frac{\alpha - n}{K_s} S_g p^c + n \right) \rho^{ga} \frac{\partial S_w}{\partial t} \right. \\ & \quad \left. - \rho^{ga} \beta_{sg} \frac{\partial T}{\partial t} \right] d\Omega + \int_{\Omega} \mathbf{w}^{*T} \left[ \alpha S_g \rho^{ga} \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + S_g n \frac{\partial}{\partial t} \left( \frac{p^{ga} M_a}{\theta R} \right) \right] d\Omega \\ & - \int_{\Omega} (\nabla \mathbf{w}^*)^T \left\{ \left[ \rho^{ga} \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] + \left[ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \nabla \left( \frac{p^{gw}}{p^g} \right) \right] \right\} d\Omega \\ & + \int_{\Gamma_g^q} \mathbf{w}^{*T} q^{ga} d\Gamma = 0 \end{aligned} \quad (10.81)$$

And the energy conservation equation can be expressed as

$$\begin{aligned}
 & \int_{\Omega} \mathbf{w}^{\circ T} (\rho C_p)_{\text{eff}} \frac{\partial T}{\partial t} d\Omega + \int_{\Omega} (\nabla \mathbf{w}^{\circ})^T \chi_{\text{eff}} \nabla T d\Omega \\
 & + \int_{\Omega} \mathbf{w}^{\circ T} \left\{ \left[ n S_w \rho^w C_p^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g})^T \right. \right. \\
 & \left. \left. + n S_g \rho^g C_p^g \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] \cdot \nabla T \right\} d\Omega \\
 & + \int_{\Omega} \mathbf{w}^{\circ T} \Delta H_{\text{vap}} \left\{ - \rho^w \left( \frac{\alpha - n}{K_s} S_w^2 + S_w \frac{n}{K_w} \right) \frac{\partial p^w}{\partial t} \right. \\
 & \left. - \rho^w \frac{\alpha - n}{K_s} S_g S_w \frac{\partial p^g}{\partial t} - \alpha \rho^w S_w \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} \right\} d\Omega \\
 & + \int_{\Omega} \mathbf{w}^{\circ T} \Delta H_{\text{vap}} \left\{ \beta_{sw} \frac{\partial T}{\partial t} - \rho^w \left[ \frac{\alpha - n}{K_s} S_w p^w - \frac{\alpha - n}{K_s} S_w p^g + n \right] \frac{\partial S_w}{\partial t} \right\} d\Omega \\
 & + \int_{\Omega} (\nabla \mathbf{w}^{\circ})^T \Delta H_{\text{vap}} \left[ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g}) \right] d\Omega \\
 & + \int_{\Gamma_q} \mathbf{w}^{\circ T} [q^T + \alpha_c (T - T_{\infty})] d\Gamma = 0
 \end{aligned} \tag{10.82}$$

Discretisation in space of governing equations (10.79) to (10.82) is again performed by means of the finite element method. As usual, the primary variables are expressed by their nodal values and the shape functions

$$\mathbf{u} = \mathbf{N}_u \bar{\mathbf{u}} \quad p^w = \mathbf{N}_p \bar{\mathbf{p}}^w \quad p^g = \mathbf{N}_p \bar{\mathbf{p}}^g \quad T = \mathbf{N}_t \bar{\mathbf{T}} \tag{10.83}$$

Then we introduce (10.83) into (10.79) to (10.82) and apply the standard Galerkin procedure, replacing the weighting functions,  $\mathbf{w}$ ,  $\mathbf{w}^*$ ,  $\mathbf{w}^*$  and  $\mathbf{w}^{\circ}$  by the corresponding shape functions  $\mathbf{N}_u$ ,  $\mathbf{N}_p$ ,  $\mathbf{N}_p$  and  $\mathbf{N}_t$ . Taking into account (10.58) and (10.59), as well as the following relationships:

$$\frac{\partial S_w}{\partial t} = \frac{\partial S_w}{\partial p^c} \frac{\partial p^c}{\partial t} + \frac{\partial S_w}{\partial T} \frac{\partial T}{\partial t} = \frac{\partial S_w}{\partial p^c} \frac{\partial p^g}{\partial t} - \frac{\partial S_w}{\partial p^c} \frac{\partial p^w}{\partial t} + \frac{\partial S_w}{\partial T} \frac{\partial T}{\partial t} \tag{10.84}$$

$$\frac{\partial}{\partial t} \left( \frac{M_w}{\theta R} p^{gw} \right) = \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \left( \frac{\partial p^g}{\partial t} - \frac{\partial p^w}{\partial t} \right) + \frac{M_w}{\theta R} \left( \frac{\partial p^{gw}}{\partial T} - \frac{p^{gw}}{\theta} \right) \frac{\partial T}{\partial t} \tag{10.85}$$

$$\nabla \left( \frac{p^{gw}}{p^g} \right) = \frac{\nabla p^{gw}}{p^g} - \frac{p^{gw}}{(p^g)^2} \nabla p^g = \frac{1}{p^g} \frac{\partial p^{gw}}{\partial p^c} (\nabla p^g - \nabla p^w) - \frac{p^{gw}}{(p^g)^2} \nabla p^g \tag{10.86}$$

$$\begin{aligned}
\frac{\partial}{\partial t} \left( \frac{p^g M_a}{\theta R} \right) &= \frac{\partial}{\partial t} \left[ \frac{(p^g - p^{gw}) M_a}{\theta R} \right] \\
&= \frac{M_a}{\theta R} \frac{\partial p^g}{\partial t} - \frac{M_a}{\theta^2 R} \frac{\partial T}{\partial t} - \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \left( \frac{\partial p^g}{\partial t} - \frac{\partial p^w}{\partial t} \right) \\
&\quad - \frac{M_w}{\theta R} \left( \frac{\partial p^{gw}}{\partial T} - \frac{p^{gw}}{\theta} \right) \frac{\partial T}{\partial t}
\end{aligned} \tag{10.87}$$

we finally obtain the following set of equations:

$$\int_{\Omega} (\mathbf{L}\mathbf{N}_u)^T \boldsymbol{\sigma}'' d\Omega - \int_{\Omega} \alpha \mathbf{m}^T (S_w p^w \mathbf{N}_p + S_g p^{gw} \mathbf{N}_p) d\Omega = \int_{\Omega} \mathbf{N}_u^T \rho \mathbf{g} d\Omega + \int_{\Gamma_u^q} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \tag{10.88}$$

$$\begin{aligned}
&\int_{\Omega} \mathbf{N}_p^T \left[ \left( \frac{\alpha - n}{K_s} S_w (\rho^{gw} S_g + \rho^w S_w) + \rho^w S_w \frac{n}{K_w} \right) \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} \right. \\
&\quad \left. + \frac{\alpha - n}{K_s} S_g (\rho^{gw} S_g + \rho^w S_w) \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} \right] d\Omega \\
&+ \int_{\Omega} \mathbf{N}_p^T \left[ \alpha (\rho^{gw} S_g + \rho^w S_w) \mathbf{m}^T \mathbf{L} \mathbf{N}_u \frac{\partial \bar{\mathbf{u}}}{\partial t} - S_w \beta_{sgw} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right. \\
&\quad \left. + S_g n \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \mathbf{N}_p \left( \frac{\partial \bar{\mathbf{p}}^g}{\partial t} - \frac{\partial \bar{\mathbf{p}}^w}{\partial t} \right) \right] d\Omega + \int_{\Omega} \mathbf{N}_p^T \left[ S_g n \frac{M_w}{\theta R} \left( \frac{\partial p^{gw}}{\partial T} - \frac{p^{gw}}{\theta} \right) \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right] d\Omega \\
&+ \int_{\Omega} \mathbf{N}_p^T \left\{ \begin{array}{l} \left[ \frac{\alpha - n}{K_s} (\rho^{gw} S_g p^c + \rho^w S_w p^w - \rho^w S_w p^c) + n(\rho^w - \rho^{gw}) \right] \\ \left( \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} - \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \frac{\partial S_w}{\partial T} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right) \end{array} \right\} d\Omega \\
&- \int_{\Omega} (\nabla \mathbf{N}_p)^T \left\{ -\rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \left[ \frac{1}{p^g} \frac{\partial p^{gw}}{\partial p^c} (\nabla \mathbf{N}_p \bar{\mathbf{p}}^g - \nabla \mathbf{N}_p \bar{\mathbf{p}}^w) - \frac{p^{gw}}{(p^g)^2} \nabla \mathbf{N}_p \bar{\mathbf{p}}^g \right] \right\} d\Omega \\
&- \int_{\Omega} (\nabla \mathbf{N}_p)^T \left\{ \rho^{gw} \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla \mathbf{N}_p \bar{\mathbf{p}}^g + \rho^g \mathbf{g}) \right\} d\Omega \\
&- \int_{\Omega} (\nabla \mathbf{N}_p)^T \left[ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla \mathbf{N}_p \bar{\mathbf{p}}^w + \rho^w \mathbf{g}) \right] d\Omega \\
&+ \int_{\Gamma_w^q} \mathbf{N}_p^T [q^w + q^{gw} + \beta_c (\rho^{gw} - \rho_\infty^{gw})] d\Gamma = 0
\end{aligned} \tag{10.89}$$

$$\begin{aligned}
& \int_{\Omega} \mathbf{N}_p^T \left[ \frac{\alpha - n}{K_s} S_w S_g \rho^{ga} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \frac{\alpha - n}{K_s} S_g^2 \rho^{ga} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} - \rho^{ga} \beta_{sg} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right] d\Omega \\
& - \int_{\Omega} \mathbf{N}_p^T \left[ \left( \frac{\alpha - n}{K_s} S_g p^c + n \right) \rho^{ga} \left( \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} - \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \frac{\partial S_w}{\partial T} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right) \right] d\Omega \\
& + \int_{\Omega} \mathbf{N}_p^T \left[ \alpha S_g \rho^{ga} \mathbf{m}^T \mathbf{L} \mathbf{N}_u \frac{\partial \bar{\mathbf{u}}}{\partial t} + S_g n \frac{M_a}{\theta R} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} - S_g n \frac{M_a}{\theta^2 R} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right] d\Omega \\
& - \int_{\Omega} \mathbf{N}_p^T \left[ S_g n \frac{M_w}{\theta R} \left( \frac{\partial p^{gw}}{\partial T} - \frac{p^{gw}}{\theta} \right) \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right] d\Omega \\
& + \int_{\Omega} \mathbf{N}_p^T \left[ -S_g n \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} + S_g n \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} \right] d\Omega \\
& - \int_{\Omega} (\nabla \mathbf{N}_p)^T \left[ \rho^{ga} \frac{\mathbf{k} k^{vg}}{\mu^g} (-\nabla \mathbf{N}_p \bar{\mathbf{p}}^g + \rho^g \mathbf{g}) \right] d\Omega + \int_{\Gamma_g^q} \mathbf{N}_p^T q^{ga} d\Gamma \\
& - \int_{\Omega} (\nabla \mathbf{N}_p)^T \left\{ \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \left[ \frac{1}{p^g} \frac{\partial p^{gw}}{\partial p^c} (\nabla \mathbf{N}_p \bar{\mathbf{p}}^g - \nabla \mathbf{N}_p \bar{\mathbf{p}}^w) - \frac{p^{gw}}{(p^g)^2} \nabla \mathbf{N}_p \bar{\mathbf{p}}^g \right] \right\} d\Omega = 0
\end{aligned} \tag{10.90}$$

$$\begin{aligned}
& \int_{\Omega} \mathbf{N}_t^T (\rho C_p)_{eff} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} d\Omega + \int_{\Omega} (\nabla \mathbf{N}_t)^T \chi_{eff} \nabla \mathbf{N}_t \bar{\mathbf{T}} d\Omega \\
& + \int_{\Omega} \mathbf{N}_t^T \left\{ \left[ n S_w \rho^w C_p^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g})^T + n S_g \rho^g C_p^g \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g}) \right] \cdot \nabla \mathbf{N}_t \bar{\mathbf{T}} \right\} d\Omega \\
& + \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \left\{ -\rho^w \left( \frac{\alpha - n}{K_s} S_w^2 + S_w \frac{n}{K_w} \right) \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} - \rho^w \frac{\alpha - n}{K_s} S_g S_w \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} \right\} d\Omega \\
& - \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \alpha \rho^w S_w \mathbf{m}^T \mathbf{L} \mathbf{N}_u \frac{\partial \bar{\mathbf{u}}}{\partial t} d\Omega \\
& + \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \left\{ \beta_{sw} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} - \rho^w \left[ \frac{\alpha - n}{K_s} S_w p^w - \frac{\alpha - n}{K_s} S_w p^g + n \right] \right\} d\Omega \\
& \quad \left( \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^g}{\partial t} - \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \frac{\partial S_w}{\partial T} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right) \\
& + \int_{\Omega} (\nabla \mathbf{N}_t)^T \Delta H_{vap} \left[ \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla \mathbf{N}_p \bar{\mathbf{p}}^w + \rho^w \mathbf{g}) \right] d\Omega \\
& + \int_{\Gamma_t^q} \mathbf{N}_t^T [q^T + \alpha_c (T - T_{\infty})] d\Gamma = 0
\end{aligned} \tag{10.91}$$

For a linear elastic solid,

$$\mathbf{P}(\bar{\mathbf{u}}) = \int_{\Omega} \mathbf{B}^T \boldsymbol{\sigma}'' d\Omega = \mathbf{K}_e \bar{\mathbf{u}} + \mathbf{K}_{te} \bar{\mathbf{T}} \tag{10.92}$$

After time differentiation we obtain the discretised mechanical equilibrium equation in the following form:

$$\mathbf{K}_e \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{K}_{te} \frac{\partial \bar{\mathbf{T}}}{\partial t} - C_{sw} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} - C_{sg} \frac{\partial \bar{\mathbf{p}}^g}{\partial t} = \frac{\partial \mathbf{f}''}{\partial t} \quad (10.93)$$

where the matrices  $\mathbf{K}_e$ ,  $\mathbf{K}_{te}$ ,  $\mathbf{C}_{sw}$ ,  $\mathbf{C}_{sg}$ ,  $\mathbf{f}''$  are defined by (10.32), (10.33) and (3.119), (3.120).

Again, for the non-linear solid-phase behaviour, only the appropriate tangent matrices can be established (10.34 and 10.35), where  $\mathbf{K}_e$  and  $\mathbf{K}_{te}$  may be replaced by  $\mathbf{K}_T$  and  $\mathbf{K}_{iT}$  respectively. Similar to (3.124) and (3.125), the discretised mass and energy conservation equations can be written as

$$\mathbf{C}_{ws} \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{P}_{ww} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \mathbf{C}_{wg} \frac{\partial \bar{\mathbf{p}}^g}{\partial t} + \mathbf{C}_{wt} \frac{\partial \bar{\mathbf{T}}}{\partial t} + \mathbf{H}_{ww} \bar{\mathbf{p}}^w + \mathbf{K}_{wg} \bar{\mathbf{p}}^g = \mathbf{f}^w \quad (10.94)$$

$$\mathbf{C}_{gs} \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{C}_{gw} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \mathbf{P}_{gg} \frac{\partial \bar{\mathbf{p}}^g}{\partial t} + \mathbf{C}_{gt} \frac{\partial \bar{\mathbf{T}}}{\partial t} + \mathbf{K}_{gw} \bar{\mathbf{p}}^w + \mathbf{H}_{gg} \bar{\mathbf{p}}^g = \mathbf{f}^g \quad (10.95)$$

$$\mathbf{C}_{ts} \frac{\partial \bar{\mathbf{u}}}{\partial t} + \mathbf{C}_{tw} \frac{\partial \bar{\mathbf{p}}^w}{\partial t} + \mathbf{C}_{tg} \frac{\partial \bar{\mathbf{p}}^g}{\partial t} + \mathbf{P}_{tt} \frac{\partial \bar{\mathbf{T}}}{\partial t} + \mathbf{K}_{tw} \bar{\mathbf{p}}^w + \mathbf{H}_{tt} \bar{\mathbf{T}} = \mathbf{f}' \quad (10.96)$$

where the matrices are defined in Appendix 10B.

Equations (10.93) to (10.96) can finally be written in the following matrix form:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{ww} & \mathbf{K}_{wg} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{gw} & \mathbf{H}_{gg} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{tw} & \mathbf{0} & \mathbf{H}_{tt} \end{bmatrix} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \\ \bar{\mathbf{p}}^g \\ \bar{\mathbf{T}} \end{Bmatrix} + \begin{bmatrix} \mathbf{K}_T & -\mathbf{C}_{sw} & -\mathbf{C}_{sg} & \mathbf{K}_{iT} \\ \mathbf{C}_{ws} & \mathbf{P}_{ww} & \mathbf{C}_{wg} & \mathbf{C}_{wt} \\ \mathbf{C}_{gs} & \mathbf{C}_{gw} & \mathbf{P}_{gg} & \mathbf{C}_{gt} \\ \mathbf{C}_{ts} & \mathbf{C}_{tw} & \mathbf{C}_{tg} & \mathbf{P}_{tt} \end{bmatrix} \frac{\partial}{\partial t} \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \\ \bar{\mathbf{p}}^g \\ \bar{\mathbf{T}} \end{Bmatrix} = \begin{Bmatrix} \frac{\partial \mathbf{f}''}{\partial t} \\ \mathbf{f}^w \\ \mathbf{f}^g \\ \mathbf{f}' \end{Bmatrix} \quad (10.97)$$

As shown in Section 10.4.1, the single-step finite difference operator (3.49) is applied to achieve the discretisation in time of this equation set. Thus (10.97) is transformed as follows:

$$\Psi(\mathbf{X}_{n+1}) = [\mathbf{B} + \theta \Delta t \mathbf{C}]_{n+0} \mathbf{X}_{n+1} - [\mathbf{B} - (1-\theta) \Delta t \mathbf{C}]_{n+\theta} \mathbf{X}_n - \Delta t \mathbf{F}_{n+\theta} = 0 \quad (10.98)$$

where

$$\mathbf{X}_{n+\theta} = (1-\theta) \mathbf{X}_n + \theta \mathbf{X}_{n+1} \quad (10.99)$$

and

$$\Psi(\mathbf{X}_{n+1}) = [\Psi^u(\mathbf{X}_{n+1}), \Psi^w(\mathbf{X}_{n+1}), \Psi^g(\mathbf{X}_{n+1}), \Psi'( \mathbf{X}_{n+1})]^T \quad (10.100)$$

The matrices  $\mathbf{B}$ ,  $\mathbf{C}$  and vectors  $\mathbf{X}$ ,  $\mathbf{F}$  of this operator are

$$\mathbf{B} = \begin{bmatrix} \mathbf{K}_T & -\mathbf{C}_{sw} & -\mathbf{C}_{sg} & \mathbf{K}_{iT} \\ \mathbf{C}_{ws} & \mathbf{P}_{ww} & \mathbf{C}_{wg} & \mathbf{C}_{wt} \\ \mathbf{C}_{gs} & \mathbf{C}_{gw} & \mathbf{P}_{gg} & \mathbf{C}_{gt} \\ \mathbf{C}_{ts} & \mathbf{C}_{tw} & \mathbf{C}_{tg} & \mathbf{P}_u \end{bmatrix} \quad (10.101)$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{ww} & \mathbf{K}_{wg} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{gw} & \mathbf{H}_{gg} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{tw} & \mathbf{0} & \mathbf{H}_u \end{bmatrix} \quad (10.102)$$

$$\mathbf{X} = \begin{Bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}}^w \\ \bar{\mathbf{p}}^g \\ \bar{\mathbf{T}} \end{Bmatrix} \quad \mathbf{F} = \begin{Bmatrix} \frac{\partial \mathbf{f}^u}{\partial t} \\ \mathbf{f}^w \\ \mathbf{f}^g \\ \mathbf{f}^t \end{Bmatrix} \quad (10.103)$$

An examination of equations (10.79) to (10.96) reveals that the coupled equation system (10.98) is non-linear; it requires linearising by an iterative Newton–Raphson procedure of the form

$$\Psi^\kappa(\mathbf{X}_{n+1}^i) = -\frac{\partial \Psi^\kappa}{\partial \mathbf{X}} \Big|_{\mathbf{X}_{n+1}^i} \Delta \mathbf{X}_{n+1}^i \quad (\kappa = u, w, g, t). \quad (10.104)$$

where the Jacobian matrix is defined by

$$\frac{\partial \Psi}{\partial \mathbf{X}} \Big|_{\mathbf{X}_{n+1}^i} = \begin{bmatrix} \frac{\partial \Psi^u}{\partial \bar{\mathbf{u}}} & \frac{\partial \Psi^u}{\partial \bar{\mathbf{p}}^w} & \frac{\partial \Psi^u}{\partial \bar{\mathbf{p}}^g} & \frac{\partial \Psi^u}{\partial \bar{\mathbf{T}}} \\ \frac{\partial \Psi^w}{\partial \bar{\mathbf{u}}} & \frac{\partial \Psi^w}{\partial \bar{\mathbf{p}}^w} & \frac{\partial \Psi^w}{\partial \bar{\mathbf{p}}^g} & \frac{\partial \Psi^w}{\partial \bar{\mathbf{T}}} \\ \frac{\partial \Psi^g}{\partial \bar{\mathbf{u}}} & \frac{\partial \Psi^g}{\partial \bar{\mathbf{p}}^w} & \frac{\partial \Psi^g}{\partial \bar{\mathbf{p}}^g} & \frac{\partial \Psi^g}{\partial \bar{\mathbf{T}}} \\ \frac{\partial \Psi^t}{\partial \bar{\mathbf{u}}} & \frac{\partial \Psi^t}{\partial \bar{\mathbf{p}}^w} & \frac{\partial \Psi^t}{\partial \bar{\mathbf{p}}^g} & \frac{\partial \Psi^t}{\partial \bar{\mathbf{T}}} \end{bmatrix} \Big|_{\mathbf{X}=\mathbf{X}_{n+1}^i} \quad (10.105)$$

and the increment vector of the primary variables is

$$\Delta \mathbf{X}_{n+1}^i = [\Delta \bar{\mathbf{u}}_{n+1}^i, (\Delta \bar{\mathbf{p}}^w)_{n+1}^i, (\Delta \bar{\mathbf{p}}^g)_{n+1}^i, \Delta \bar{\mathbf{T}}_{n+1}^i]^T \quad (10.106)$$

where  $i$  is the iteration count.

During the computation for each time step, the primary variable vector  $\mathbf{X}_{n+1}$  is updated after each iteration, i.e.

$$\mathbf{X}_{n+1}^{i+1} = \mathbf{X}_{n+1}^i + \Delta \mathbf{X}_{n+1}^i \quad (10.107)$$

In the Newton–Raphson method, the elements of the Jacobian matrix are updated after each iteration, which is computationally expensive. Thus, in practice the modified Newton–Raphson method is often preferred, where the Jacobian matrix is calculated once only at the beginning of each time step or after a fixed number of iterations.

## 10.7 NUMERICAL EXAMPLES

This section presents a number of examples dealing with heat and mass transfer in a deforming porous medium. The first three deal with the fully saturated case, and the last two investigate partially saturated conditions including phase change. The solution of the first example is compared with the thermoelastic solution of Aboustit *et al.* [12]. A second example deals with thermoelastoplastic behaviour, and the third example is compared against an analytical solution given by Booker and Savvidou [2]. For this case the partitioned solution is also given in order to compare computing time for both methods, and a solution with infinite elements is shown. In the fourth example the first problem is again solved, but this time for a partially saturated condition and the influence of phase change is also shown. Finally, in the last example the solutions given by two different computer programs are compared [22,23], and the need for appropriate boundary conditions is shown when dealing with realistic situations.

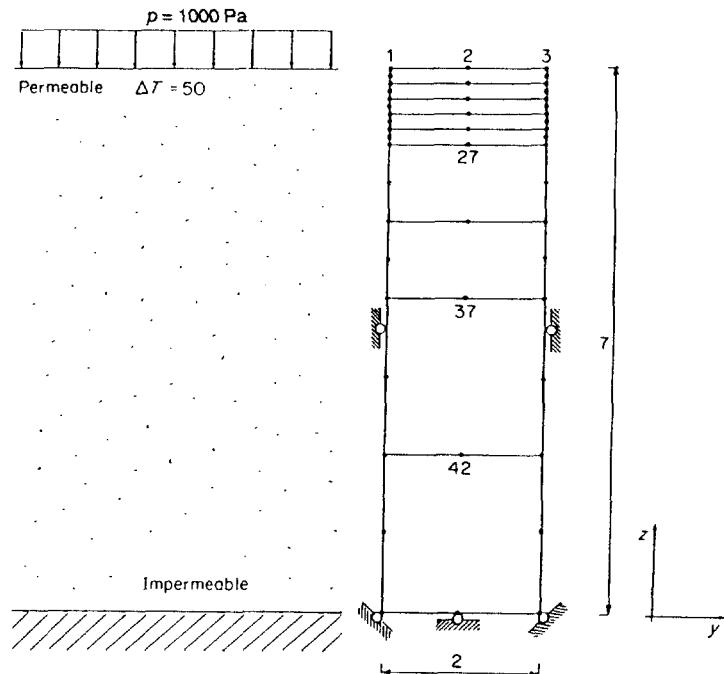
### 10.7.1 Thermo-elastic consolidation

The first example is a thermoelastic one-dimensional consolidation problem solved previously by Aboustit *et al.* [12]. A column of linear elastic material is subjected to a unit surface pressure and a constant surface temperature of  $T = 50^\circ\text{C}$ . The following values were assumed:  $L = 7 \text{ m}$ ,  $E = 6000 \text{ kPa}$ ,  $\nu = 0.4$ ,  $K = 4 \times 10^{-6} \text{ m h}^{-1}$ ,  $\lambda = 0.2 \text{ kcal m}^{-1} \text{ K}^{-1} \text{ s}^{-1}$ ,  $(pc)_a = 40 \text{ kcal m}^{-3} \text{ K}^{-1}$  where  $(pc)_a$  indicates the average heat capacity,  $\beta_s = 0.9 \times 10^{-6}$ . The coupling matrix  $\mathbf{R}$  of (10.38) is equal to zero;  $\mathbf{R}$  is not taken into account by Aboustit *et al.* We do not include the matrix related to energy generation due to deformation; this matrix was retained in the energy transport equation by Aboustit *et al.* Its contribution is negligible, as shown by comparing our values with those of Aboustit *et al.* (Figure 10.6).

The finite element discretisation of the problem is shown in Figure 10.1. The pore pressure is equal to zero at the top surface; everywhere else the surface of the body is sealed and insulated. The time steps for scheme 1 are shown in Table 10.1.

Scheme 2 is a more refined scheme; it assumes  $10 \times 1.0$  steps after the  $10 \times 0.1$  steps of scheme 1. Scheme 2 produced only slight differences during the few time steps after the change from  $\Delta t = 0.1$  to  $\Delta t = 1.0$ , hence it was not used again.

The values of  $\theta$  used in schemes 1 and 2 are 0.5 and 0.875. Spatial oscillation of the temperature distribution was observed with  $\theta = 0.5$  and a  $2 \times 2$  spatial integration, after suddenly applying the prescribed temperature. These oscillations disappeared after a few time steps and they were not observed when a  $3 \times 3$  Gaussian integration was used. The temporal pore pressure oscillations vanished with  $\theta = 0.875$ . The errors associated with a sudden change in the size of the time step also diminished within a few time steps.



**Figure 10.1** One-dimensional model and finite element mesh

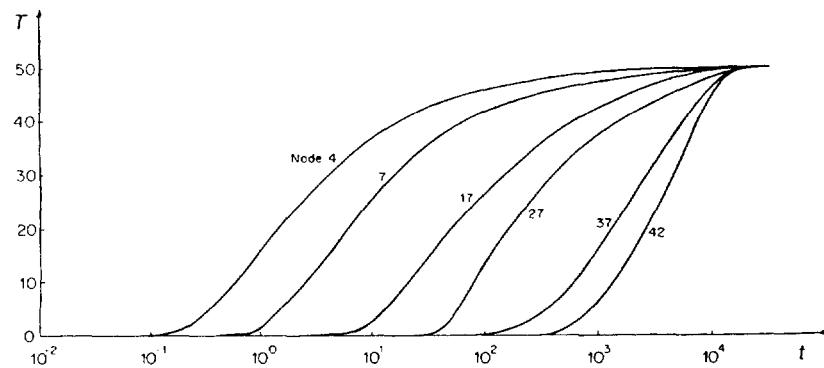
**Table 10.1** Times steps for different time intervals  
(scheme 1)

Time interval	Number of time steps
0.01	10
0.1	10
10	10
100	10
1000	20

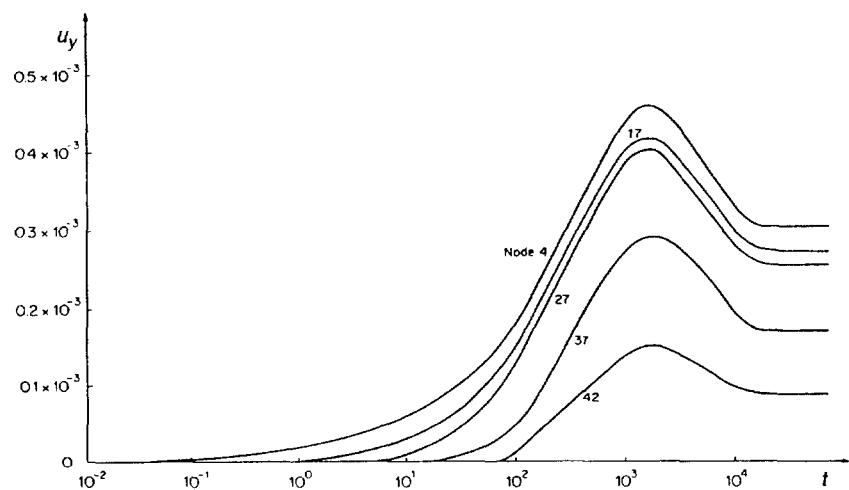
Temperature versus time for some nodal points is plotted in Figure 10.2; notice the very slow response. After 1000 days the temperature in the bottom half of the sample is still below 20 °C. This explains the settlement histories of Figures 10.3, 10.6 and 10.7, where a reversal of the displacement due to the thermal expansion is observed only after that time. Figure 10.4 shows the pore pressure values versus time at several nodes.

The damping effect of the  $3 \times 3$  Gaussian integration on the temperature distribution along the column is shown in Figure 10.5, where a comparison with a  $2 \times 2$  integration is made at two different times values. After  $t = 0.1$  the temperature distributions are the same; this justifies the use of the  $2 \times 2$  rule throughout the solution.

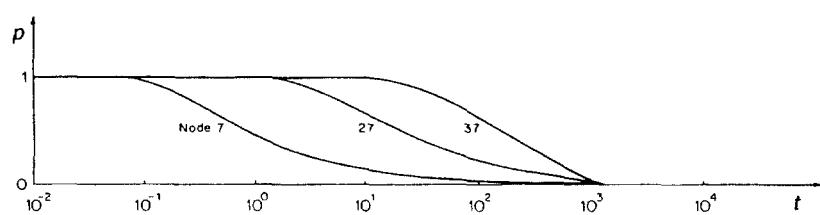
The settlement of the top surface versus time is shown in Figure 10.6 for isothermal consolidation, applied surface temperature, temperature change and surface load applied



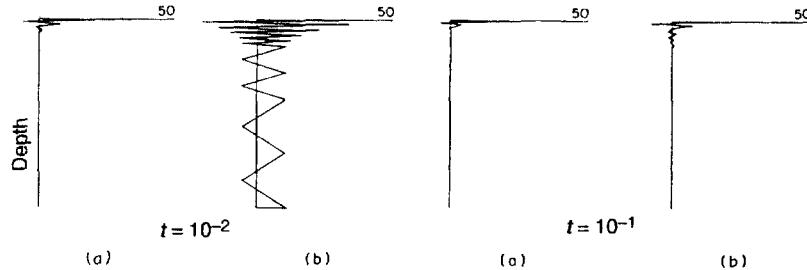
**Figure 10.2** Thermoelastic consolidation: temperature versus time at different nodes



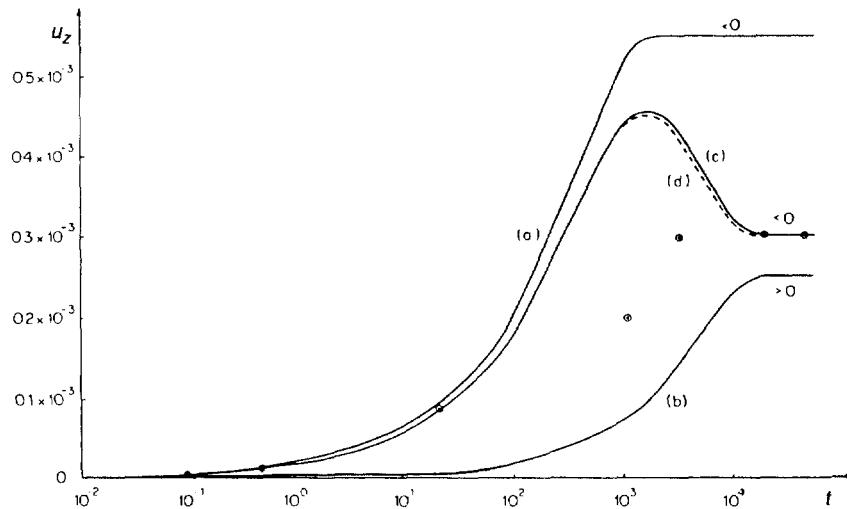
**Figure 10.3** Displacement versus time at different nodes



**Figure 10.4** Pore pressure versus time at different nodes



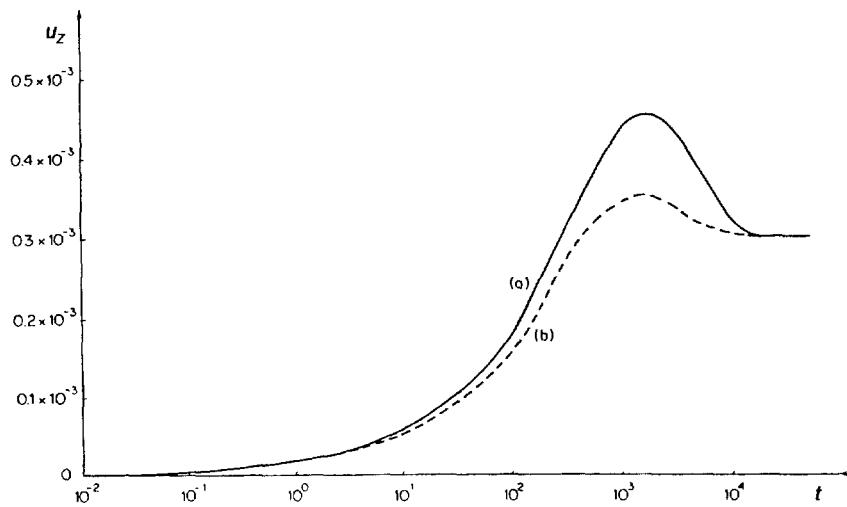
**Figure 10.5** Pore pressures at two different times using (a)  $3 \times 3$  Gauss points and (b)  $2 \times 2$  Gauss points in the spatial integration



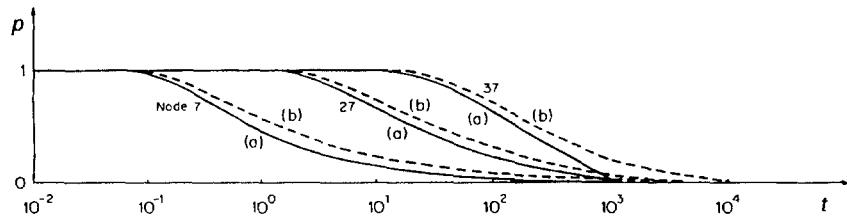
**Figure 10.6** Surface settlement versus time for the one-dimensional example of Figure 10.1: (a) isothermal consolidation only; (b) applied surface temperature only; (c) non-isothermal consolidation with incompressible fluid and solid grains; (d) non-isothermal consolidation with the actual values of the bulk modulus obtained from Combarnous and Bories [48]; the results of Aboustit *et al.* [12] are indicated by dots.

simultaneously. These values were obtained with the assumption of an incompressible fluid and solid grains. The effects of the compressibility of the fluid and solid grains bulk moduli  $K_w = 0.43 \times 10^{10}$  and  $K_s = 0.14 \times 10^{10}$  are also shown in Figure 10.6. And for comparison it shows the results obtained by Aboustit *et al.*; these results from Table I in the original work [12]. The comparison is excellent except for two values. This difference is probably due to an error in the original table [12]; in fact, for a time value 10 times smaller there is exact agreement between our values and those of Aboustit *et al.*

The influence of the thermal expansion coefficient of the fluid, hence of the first derivative of  $T$  in the flow equation (10.24), was investigated by assuming  $\beta_w = 0.63 \times 10^{-5}$ . This results in a non-zero coupling matrix  $\mathbf{R}$ . Figure 10.7 shows



**Figure 10.7** Surface settlement versus time: (a) thermal expansion coefficient of the fluid  $\beta_w = 0$  (curve c in Fig. 10.6); (b)  $\beta_w = 0.63 \times 10^{-5}$



**Figure 10.8** Pore pressure versus time obtained at three different nodes: (a)  $\beta_w = 0$ , (b)  $\beta_w = 0.63 \times 10^{-5}$

the effects on the displacements and Figure 10.8 shows the effects on the pore pressure values.

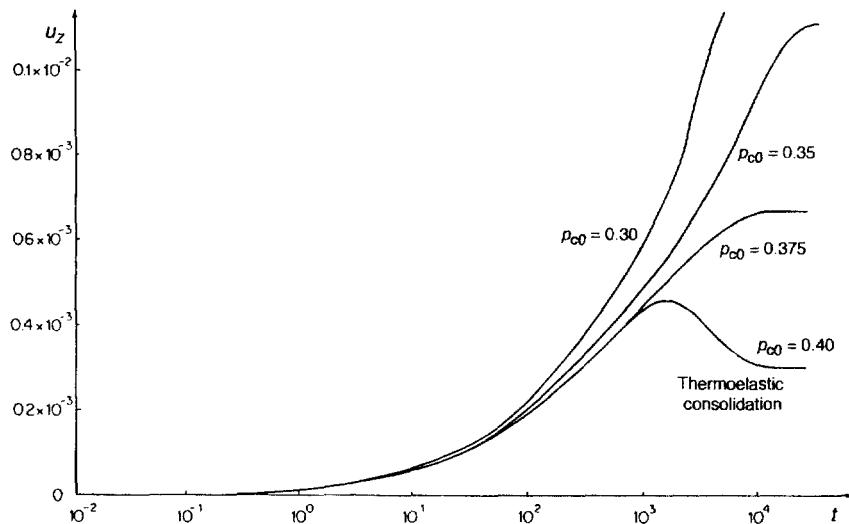
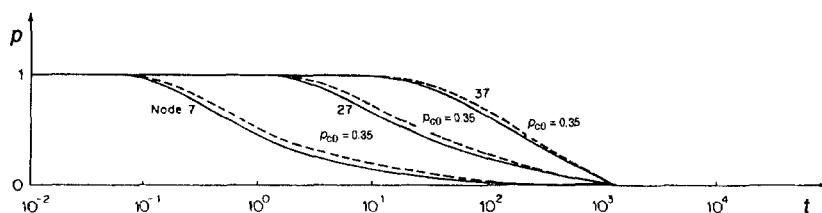
The effects of convection have also been investigated and found to be negligible in this case. The Petrov–Galerkin method along with upwinding techniques were used for this purpose. The effect of the heat generated by elastic deformation is negligible. This was also verified by introducing the matrix  $K_{fT}^T$  into the third equation of our model.

### 10.7.2 Thermo-elastoplastic Consolidation

The same problem as in the previous example was also investigated assuming elastoplastic constitutive behaviour with a temperature-independent yield function. A critical state yield surface of the form defined by (4.46) is used, together with the following material properties:  $M_{cs} = 10.5$ ,  $\lambda = 0.14$ ,  $k = 0.05$ ,  $e_0 = 0.90$  [49]. The initial elastic modulus, Poisson's ratio and the permeability tensor were the same as for the previous example. The time steps are shown in Table 10.2.

**Table 10.2** Times steps for different time intervals

Time interval	Number of time steps
0.01	10
0.1	10
1	10
2	10
5	10
10	15
20	35

**Figure 10.9** Thermo-elastoplastic consolidation of the model in Fig. 10.1: settlement versus time for different values of the preconsolidation pressure**Figure 10.10** Thermo-elastoplastic consolidation: pore pressure versus time for three different nodes,  $p_{c0} = 0.35$ 

The settlement history for different values of the preconsolidation pressure is shown in Figure 10.9. For  $P_{c0} = 0.40$  some yielding appeared at the beginning of the transient behaviour, then the material behaviour was simply elastic; with  $p_{c0} = 0.30, 0.35, 0.375$  a reversal of the settlement was not observed. This implies that the plastic flow due to the applied unit pressure is higher than the thermal expansion throughout the transient.

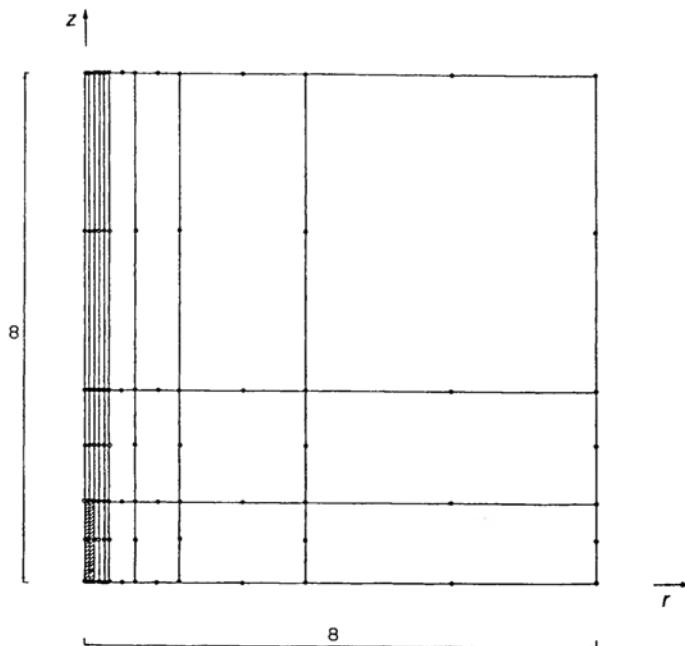
The pore pressure versus time is given in Figure 10.10 for  $p_{c0} = 0.35$ . The other values of  $p_{c0}$  led to similar behaviour. Note that non-isothermal condition would also require a temperature-dependent yield function, which is omitted here.

### 10.7.3 Thermoelastic Consolidation Around a Cylindrical Heat Source

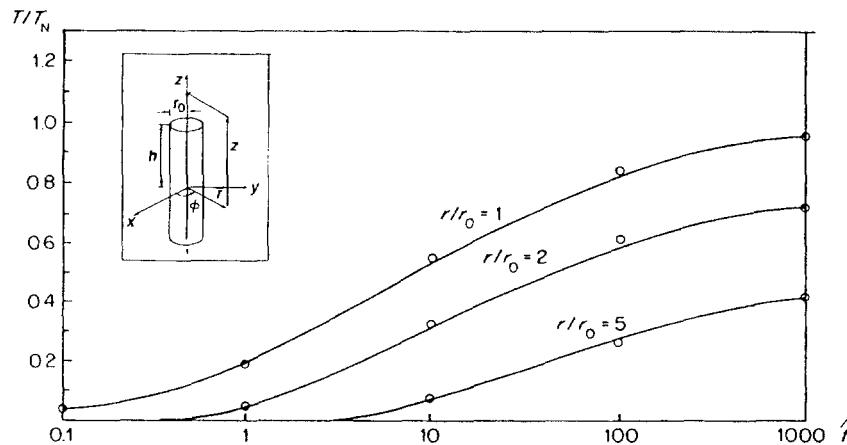
In the previous example the flow of pore water and temperature with time was in the vertical direction only. This meant that nodes at the same depth had identical pore pressure and temperature histories. This example investigates an axisymmetric problem.

The effects of a cylindrical radiating source, buried in a thermoelastic soil, were studied by Booker *et al.* [2] An analytic solution for a point heat source was numerically integrated over the surface of a cylindrical canister.

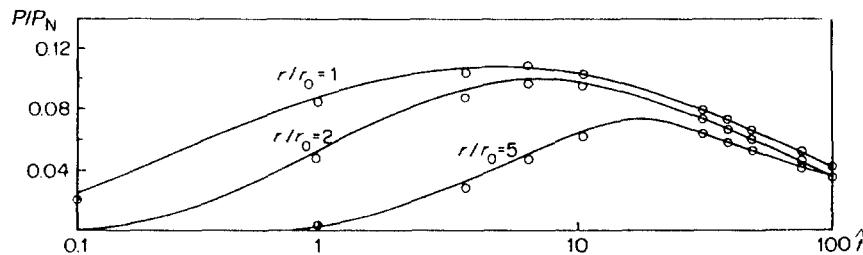
The solution was shown for the particular case of cylindrical heat source in which  $a'/a_u = \frac{1}{4}$ ,  $v = 0.4$ ,  $c/k = 2$ , where  $a'$  is the coefficient of volume expansion of the soil,  $a_u = \beta_s(1-n) + \beta_w n$ ,  $c$  is the coefficient of consolidation and  $k$  is the coefficient of thermal diffusivity. The solution was used as a basis for comparison with our numerical model. The finite element mesh is shown in Figure 10.11, where  $r_0$  is the radius of the heat source. A set of possible field data which satisfies the above ratios is  $E = 6000$ ,  $v = 0.4$ ,  $K = 0.4 \times 10^{-5}$ ,  $\lambda = 1.02857$ ,  $(\rho c)_a = 40$ ,  $\beta_s = 0.9 \times 10^6$ ,  $\beta_w = 0.63 \times 10^{-5}$ . The heat source was simulated by a constant heat input of 1000 W for each of the two elements of the source.



**Figure 10.11** Finite element discretisation of the radial cross-section containing a cylindrical heat source



**Figure 10.12** Temperature versus time for a cylindrical heat source: (—) analytical solution of Booker and Savvidou [2]; (○) finite element solution



**Figure 10.13** Pore pressure versus time for a cylindrical heat source: (—) analytical solution of Booker and Savvidou [2]; (○) finite element solution

The temperature and pressure histories for three values of  $r/r_0$  are shown in Figures 10.12 and 10.13. The temperature values are normalised with respect to the maximum value  $T_N$  at the surface of the heat source. The pressure values are quoted as ratios of  $p_N$ , the maximum pore water pressure at the surface of the heat source if the soil were permeable [2];  $i$  is given by  $i = r_0^2 t/k$ . The agreement between the analytical solution and the finite element solution is remarkable.

The finite element modelling was sensitive to the ratio between the characteristic lengths of the domain (finite in this model) and the size of the heat source. So three different cases were investigated. The axisymmetric domain (Figure 10.11) was chosen as  $8 \times 8$  in section and the radius of the heat source was set to  $r_0 = \sqrt{k}$  so that  $i = t$ . A solution was then obtained for cylinder lengths of 0.5, 1.25 and 2.0.

For a cylinder length of 2.0 the temperature variation was sufficiently close to that predicted by the analytical model; the maximum temperature at  $r/r_0 = 1$  and  $z = 0$  was  $T = 35.53$ . The pore pressure values were underestimated until  $i = 1$ , overestimated after  $i = 1$ , and at  $i = 100$  they were approximately twice as large as the analytical results. For a cylinder of length 0.5 there was an overall underestimation of the

**Table 10.3** Displacements at  $r/r_0 = 1, 2, 5$  for  $z = 0$ 

Time	$r/r_0 = 1$	$r/r_0 = 2$	$r/r_0 = 5$
0.1	0.1961	0.1516	0.0580
1	0.6689	0.9103	0.6036
10	1.0755	1.7819	2.7405
100	1.1352	1.9160	3.2868
1 000	1.0142	1.6758	2.7263
10 000	1.0000	1.6462	2.6928

**Table 10.4** Material data used in the computations

Radius of the source	$r_0 = 0.1604 \text{ m}$
Height of the source	$2h = 0.3208 \text{ m}$
Young's modulus of the soil	$E = 60 \text{ MPa}$
Poisson's ratio	$\nu = 0.4$
Permeability	$k = 0.4 \times 10^{-4} \text{ m per day}$
Soil porosity	$n = 0.5$
Fluid density	$\rho_w = 1000 \text{ kg m}^{-3}$
Thermal expansion coefficient	$\beta_g = 0.9 \times 10^{-6} \text{ }^\circ\text{K}^{-1}$
Thermal expansion coefficient of fluid	$\beta_w = 0.63 \times 10^{-5} \text{ }^\circ\text{K}^{-1}$
Specific heat of fluid	$C_w = 40 \text{ cal kg K}^{-1}$
Specific heat of solid	$C_g = 40 \text{ cal kg K}^{-1}$
Heat conductivity	$\lambda = 1.02857 \text{ kcal m}^{-1} \text{ }^\circ\text{K}^{-1} \text{ per day}$
Coefficient of consolidation	$c = 2.057 \text{ } 14 \text{ m}^2 \text{ per day}$

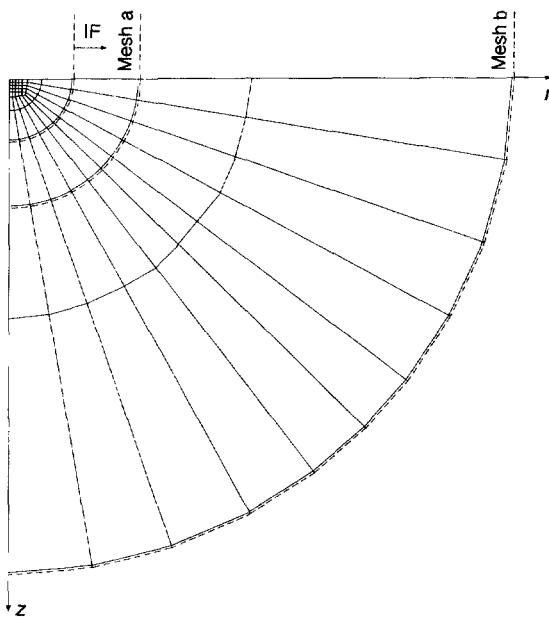
temperature, which attained a maximum of  $T = 21.68$  at  $r/r_0 = 1$  and  $z = 0$ ; the pore pressures were very high at  $\hat{t} = 1$  then dissipated rapidly, showing a substantial coincidence with the analytical curves between  $\hat{t} = 10$  and  $\hat{t} = 100$ .

A cylinder length of 1.25 was the most satisfactory in terms of both temperature and pore pressure. Their variations are shown in Figures 10.12 and 10.13. The displacements of three nodes along the  $r$ -axis are given in Table 10.3 (normalized with respect to  $u_\infty$ ,  $r/r_0 = 1$ ).

Because of the strong influence of the size of the domain, i.e. the position of the far-field boundary, the problem was investigated again with different meshes, also using the infinite element technology described in Section 6.3.1. The analytical solution is in fact valid for an unbounded domain. This analytical solution by Brooker and Savvidou [2] has been re-evaluated for the data set of Simoni and Schrefler [50] given in Table 10.4.

The discretised heat source is simulated in the numerical solution by a constant heat input of  $1000 \text{ kcal m}^{-3} \text{ per day}$  for each of the four elements. The following meshes are used for the spatial discretisation in the numerical tests (Figure 10.14):

*Mesh a:* finite element mesh with the boundary introduced at a distance of  $15r_0$  from the heat source (reduced mesh, 640 d.o.f.).

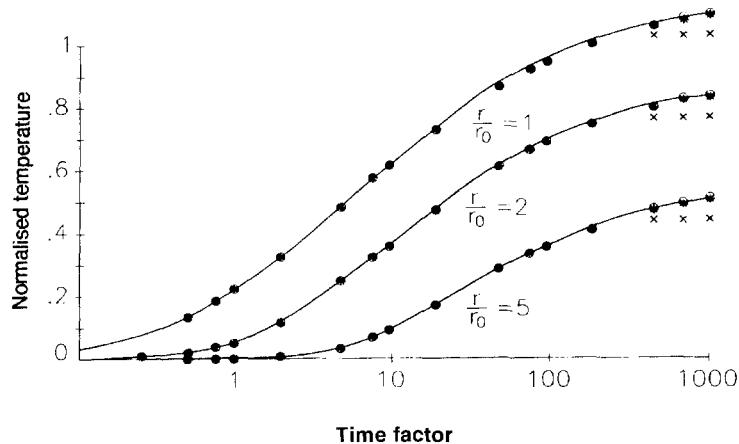


**Figure 10.14** Spatial discretisation of the radial cross-section containing a cylindrical heat source

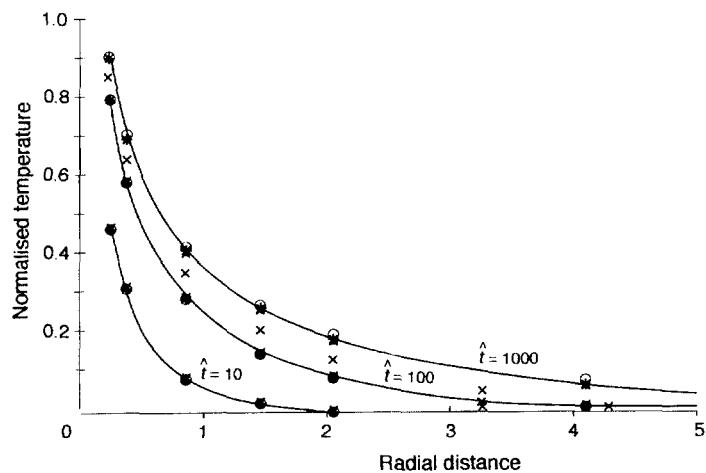
*Mesh b:* finite element mesh with the boundary introduced at a distance of  $115r_0$  from the heat source (extended mesh, 1020 d.o.f.). In this case the far-field approximation is performed in an approximate way using large finite element sizes.

*Mesh c:* finite element mesh in conjunction with infinite elements (728 d.o.f.). Pore pressure and temperature values are approximated using both linear and quadratic expansions in the parent element, whereas displacements are approximated using second-order polynomials (the same is true for the analytical solution at very large time values). Both the linear and quadratic interpolations of pressure and temperature gave similar results. The spatial discretisation is performed in such a way that the poles of all the infinite elements are part of the contour of the heat source.

Figure 10.15 represents the temperature histories for three values of the ratio  $r/r_0$ . The differences between the three solutions appear at high time values, depending on the characteristics of the temperature field. For instance, the error in the temperature value at  $r = r_0$  and  $\hat{t} = 1000$  is approximately +1.5 percent using infinite elements (in this way the steady state is slightly advanced); in mesh a this error is approximately -6 percent and decreases in the mesh b to about -0.5 percent. The situation is the same in the radial direction ; the respective errors are +5, -4 and -60 percent for meshes c, b and a at  $\hat{t} = 1000$  and  $r/r_0 = 10$  (Figure 10.16). However, for certain small time values, the results of the three numerical solutions are practically the same and coincide with the analytical solution.

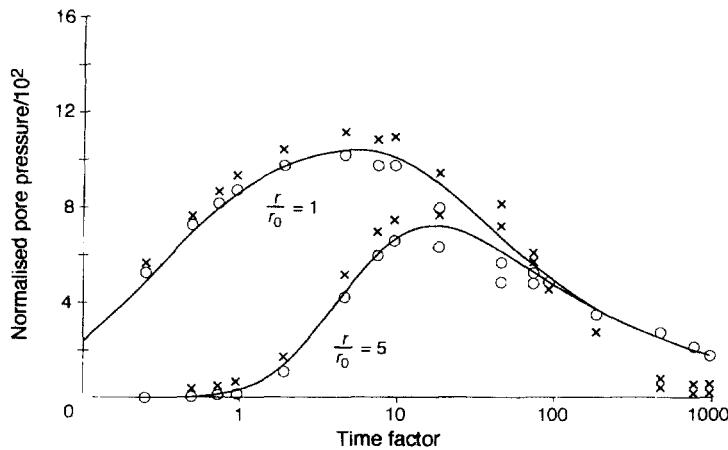


**Figure 10.15** Cylindrical heat source: normalised temperature versus time at  $z = 0$  for three values of  $r/r_0$ ; (—) analytical solution; ( $\times \times \times \times$ ) truncation approximation, mesh a; ( $\ast \ast \ast \ast$ ) truncation approximation, mesh b; (.....) FE + IE, mesh c

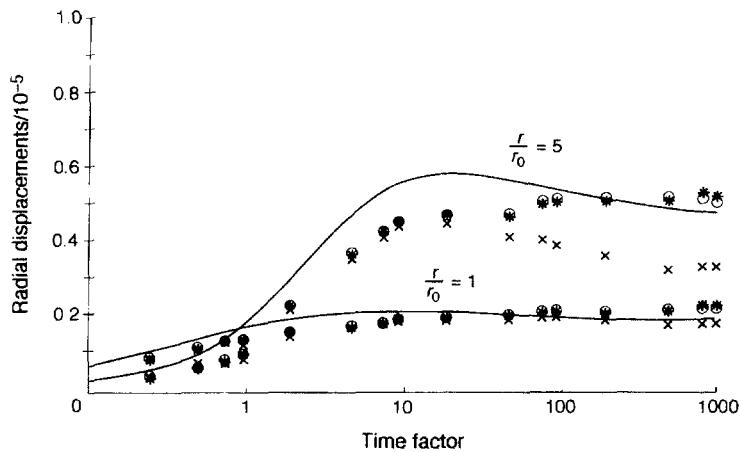


**Figure 10.16** Cylindrical heat source: normalised temperature versus time for two values of the dimensionless time parameter; (—) analytical solution; ( $\times \times \times \times$ ) truncation approximation, mesh a; ( $\ast \ast \ast \ast$ ) truncation approximation, mesh b; (.....) FE + IE, mesh c

Figure 10.17 compares the normalised pore pressure histories for two values of  $r/r_0$ . The maximum pore pressure using infinite elements at  $r = r_0$  is equal to the maximum pore pressure for the analytical solution, which is also true for the extended mesh with the truncation approach. Using truncation in the reduced mesh, we obtain pore pressure values larger than the analytical values for  $\hat{t} < 1000$  with an error of approximately 5 percent for the maximum pressure value at  $r = r_0$ . The discrepancies at  $\hat{t} = 10$  are due to sharp increase in the time step necessary to reduce the computer time for the simulation.



**Figure 10.17** Cylindrical heat source: normalised pore pressure versus time for two values of  $r/r_0$ : (—) analytical solution; ( $\times \times \times \times$ ) truncation approximation, mesh a; (.....) FE + IE, mesh c



**Figure 10.18** Cylindrical heat source: radial displacement versus time at  $z = 0$  for two values of  $r/r_0$ : (—) analytical solution; ( $\times \times \times \times$ ) truncation approximation, mesh a; ( $\ast \ast \ast$ ) truncation approximation, mesh b; (.....) FE + IE, mesh c

The radial displacements (not normalised) in the plane of symmetry are presented in Figure 10.18 for two values of the ratio  $r/r_0$ . These numerical results are less satisfactory than the solutions obtained for the pore pressure and the temperature. On comparing the maximum values from the numerical simulation at  $r/r_0 = 5$ , we obtain respective errors of -20, -12 and -3 percent for meshes a,b and c. The displacement values may be improved by using different discretisations in space and in time for the three fields.

This problem was also solved for the first set of data by using the staggered procedure of Section 10.4.2. A speed-up of 20 percent is obtainable by keeping the strongly coupled displacement and pressure fields together and introducing the partitioning between these two fields and the weakly coupled temperature field (Table 10.5) [51]. It

**Table 10.5** CPU times for one time step using the monolithic approach and the partitioned approach

solution	CPU time(s)	CPU times(s) for solving the equation system
Two-level monolithic	154.35	142.75
Two-level partitioned	123.22	91.92

is generally unwise to introduce operator splitting for two strongly coupled fields. This is demonstrated by Mateazzi *et al.* [40] for a linear consolidation problem; their partitioning between the displacement and pore pressure fields required over 40 iterations per time step.

#### 10.7.4 Non-isothermal consolidation

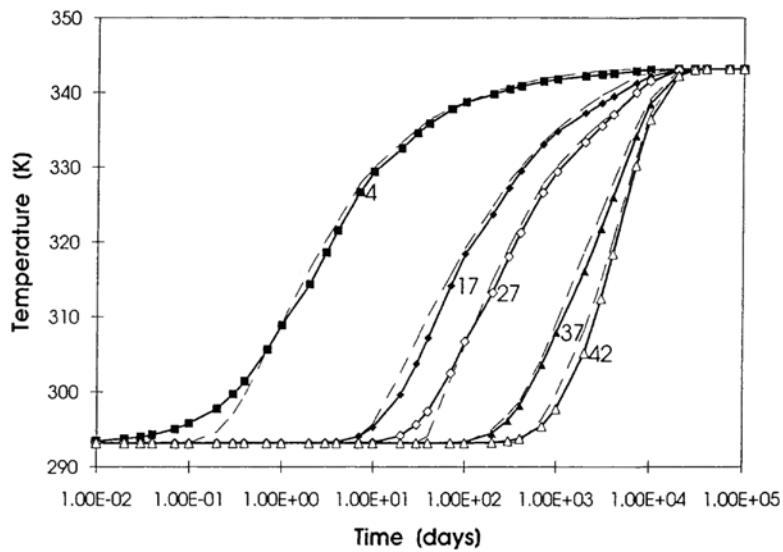
The same problem as given in Section 10.7.1 will now be discussed but for partially saturated conditions. Schrefler *et al.* [22] solved this problem but neglected the influences of latent heat and phase change. These two factors are taken into account in the present analysis; also important are ensuing changes in later stages of the simulation in terms of displacements and especially saturations and capillary pressures [23].

The relationships proposed by Brooks and Corey [52] were used between capillary pressure, saturation of water and relative permeabilities of water and gas. The water and solid phases were assumed to be incompressible and the remaining data for the porous medium were assumed to be the same as in Section 10.7.1.

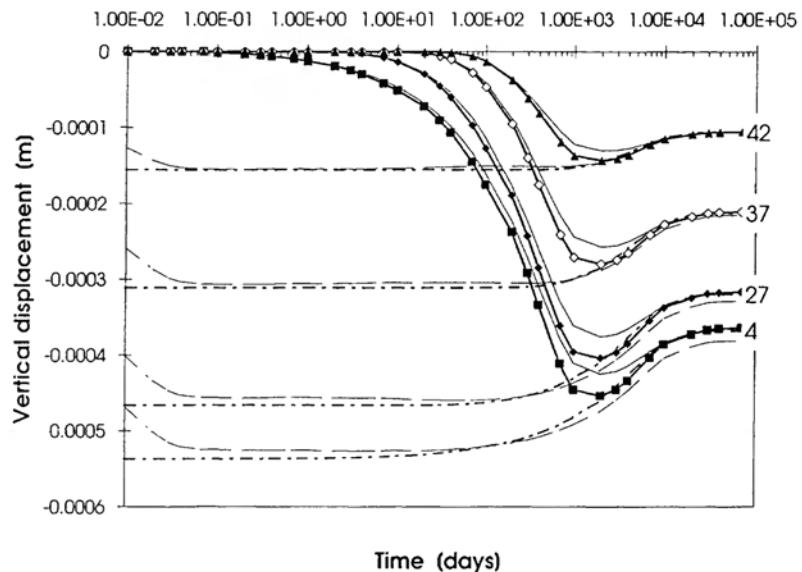
The boundary conditions were assumed to be for the lateral surface  $q^T = 0$ ,  $u_h = 0$ ; for the top surface  $T = 343.15$  K,  $p^g = p_{\text{atm}}$ ,  $p^c = p^{c*}$  where  $p^{c*}$  means capillary pressure corresponding to saturation  $S = 0.92$ ; and for the bottom surface  $q^T = 0$ ,  $u_v = 0$ . For numerical purposes the column was simulated by 9 and then by 18 eight-noded isoparametric elements giving practically the same results. A  $3 \times 3$  Gaussian integration scheme was used in both cases and the temporal discretisation was performed with an initial step of 0.01 days for the first 100 steps then multiplied by 10 thereafter for each set of 100 time-step solutions until an elapsed time of  $10^7$  days was achieved. The problem was solved for two different cases:

1. An initial homogeneous saturation of water  $S_w = 0.92$  and a capillary pressure–saturation curve which was independent of temperature.
2. An initial homogeneous saturation of water  $S_w = 0.92$  and a modified capillary pressure–saturation relationship [52]. It was assumed that these results were measured at temperature  $T = 293.15$  K, with regard to any changes of capillary pressure due to the dependence of surface tension on temperature [53].

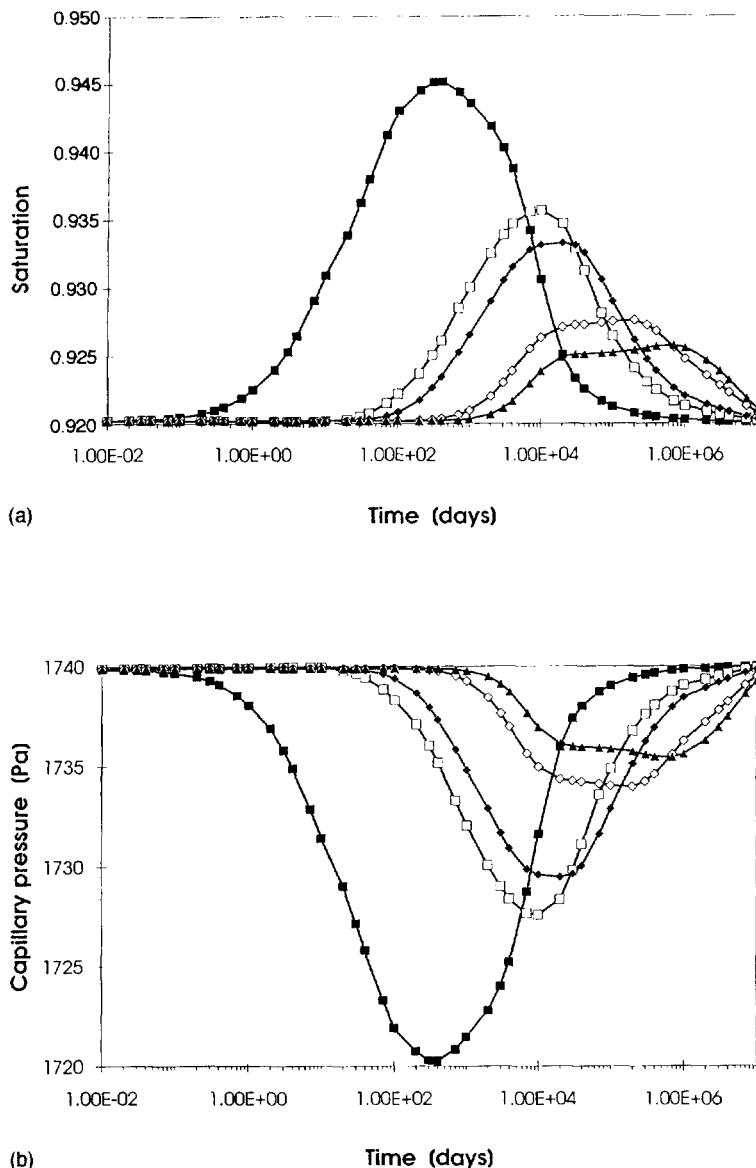
The resulting temperature profiles are practically the same for the two cases and also for the fully saturated case given in Section 10.7.1 (Figures 10.2 and 10.19). This is because of the identical averaged thermal conductivity and relatively high thermal capacity values assumed in Section 10.7.1. In all the figures, the numbers 4, 7, 17, 27, 37 and 42



**Figure 10.19** The resulting profiles of temperature (solid lines) compared with the solution of Aboustit *et al.* [12] (dashed lines). The numbers on the curves are the nodal points selected within the mesh



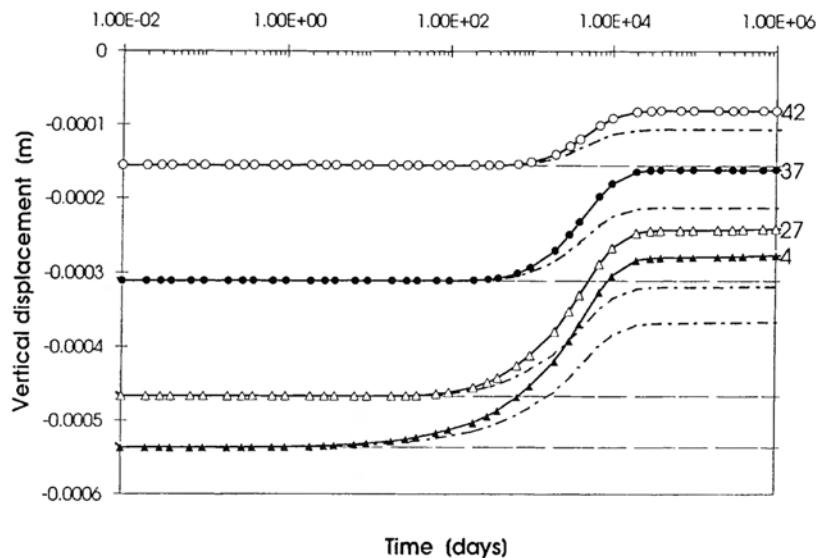
**Figure 10.20** Comparison between the saturated case (solid lines) and the second partially saturated case (chain dots): heavy lines are the present solution; light lines are the solution of Schrefler *et al.* [22]. The numbers on the curves are the nodal points selected within the mesh



**Figure 10.21** The resulting profiles of (a) saturation and (b) capillary pressure for the first partially saturated case. The nodal points selected within the mesh are as follows: (■) 4, (○) 17, (◆) 27, (◇) 37, (▲) 42

refer to selected nodal points within the mesh of Figure 10.1, respectively placed at 0.1, 0.2, 0.6, 1.0, 3.0 and 5.0 m from the upper surface of the column.

Figure 10.20 shows the settlement history of the semisaturated medium with a temperature-independent capillary pressure–saturation curve and compares it with the fully saturated case. The resulting profiles of both the capillary pressure and water

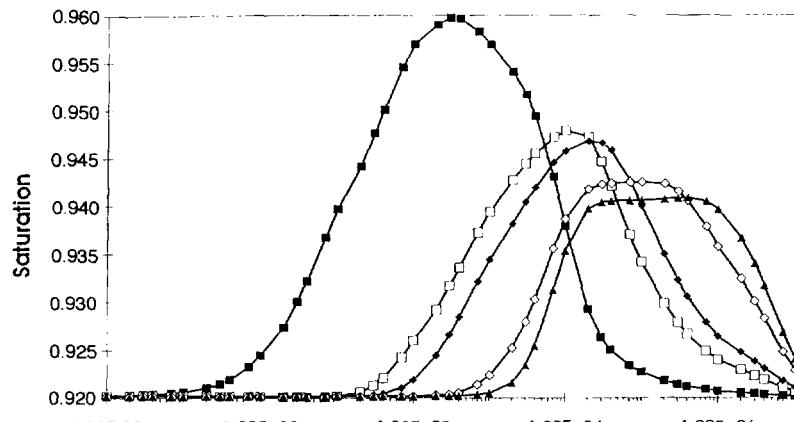


**Figure 10.22** Comparison between the second partially saturated case (solid line) and the first partially saturated case (chain dots); the light dashes indicate the isothermal solutions

saturation values given in Figure 10.21 show the influence of phase change on the phenomenon. The characteristic shape of the curves is influenced by the prior condensation and succeeding evaporation caused by temperature changes. This phenomenon was not observed in the solution given by Schrefler *et al.* [22] where phase change was not taken into account. The settlement histories of the two analysed semisaturated cases are compared in Figure 10.22 and indicate the significant differences caused by the temperature dependence of the capillary pressure–saturation curve; this aspect was omitted by Schrefler *et al.* [22]. The resulting profiles of capillary pressure and saturation of water for the second case (Figure 10.23) show that the phase change phenomenon is more pronounced, whereas the capillary pressure history is strongly influenced by the temperature history, causing additional vertical displacement (Figure 10.22).

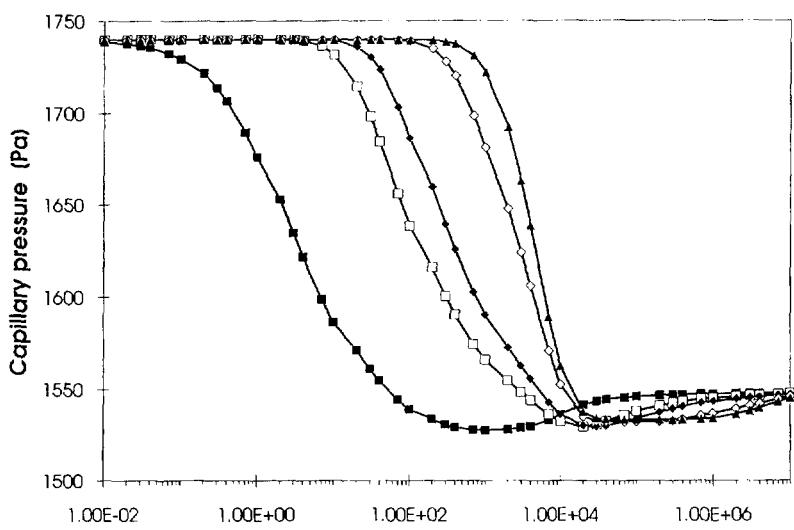
### 10.7.5 Thermoelastic consolidation of partially saturated clay

This problem deals with the thermoelastic consolidation of partially saturated clay for which previous solutions are known. Dakshanamurthy and Fredlund [15] studied in detail the case of constant permeability and thermal property values, whereas Schrefler *et al.* [22] solved the problem for the relationships between the relative permeabilities of water and air, the saturation of water and the capillary pressure, as proposed by Brooks and Corey [52]. Dakshanamurthy and Fredlund analysed four different cases: isothermal consolidation, isothermal swelling, non-isothermal consolidation and non-isothermal swelling whereas here only the non-isothermal consolidation problem has been investigated. The effects of appropriate boundary conditions were also studied in some detail.



(a)

Time (days)



(b)

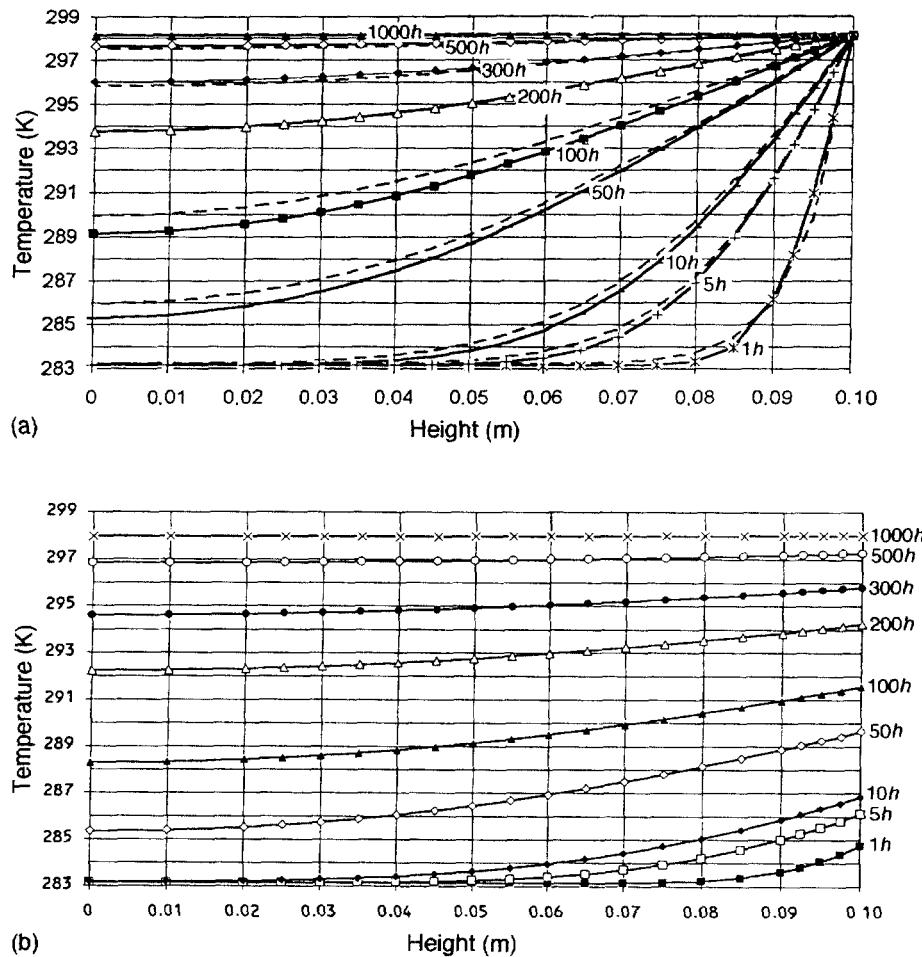
Time (days)

**Figure 10.23** The resulting profiles of (a) saturation and (b) capillary pressure for the second partially saturated case. The nodal points selected within the mesh are as follows: (■) 4, (□) 17, (◆) 27, (◇) 37, (▲) 42

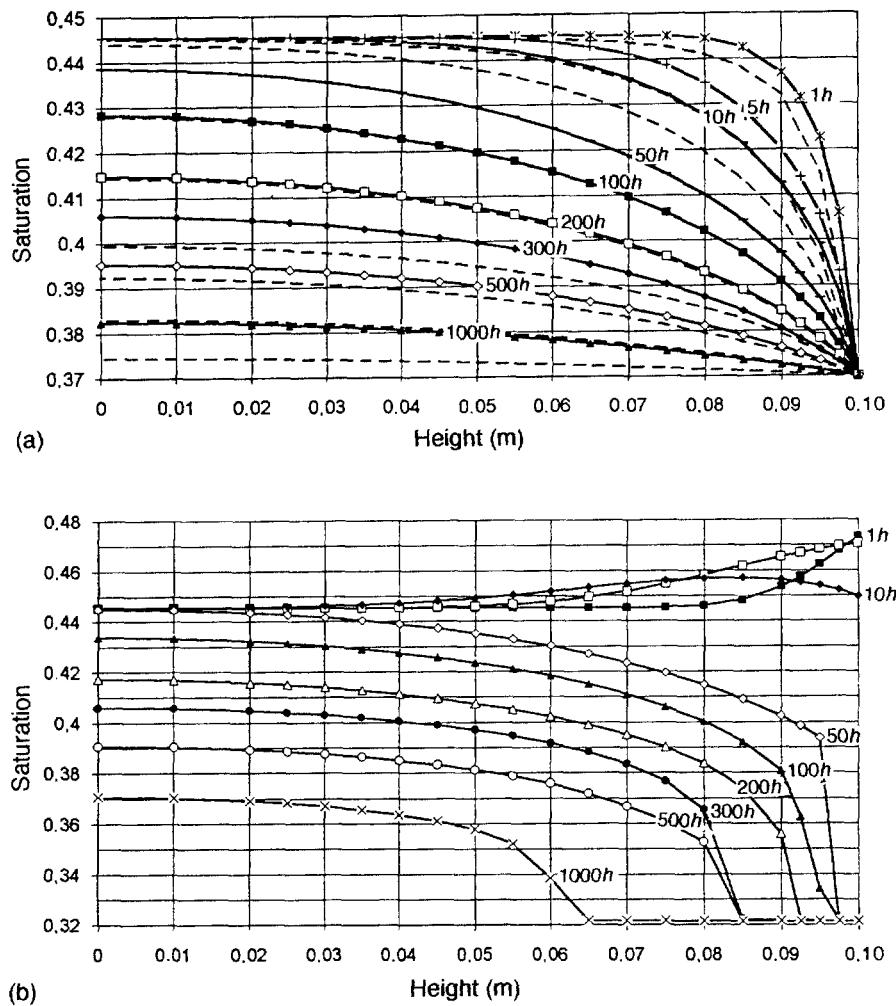
A layer of partially saturated subgrade soil ( $S_w = 0.445$ ), of thickness 10 cm, initially in a state of equilibrium, was subjected to a sudden environmental change. A surface temperature change of 15 K and a capillary pressure change of 140 kPa cause simultaneous heat and mass transfer, which corresponds physically to a drying process caused by a warm flux of dry air. Porosity  $n = 0.5$ , density of solid  $\rho^s = 2000 \text{ kg m}^{-3}$ ,

absolute permeability  $k = 4 \times 10^{-3} \text{ m}^2$ ; thermal conductivity  $\lambda = 0.2 \text{ kcal m}^{-1} \text{ K}^{-1} \text{ s}^{-1}$  and thermal capacity  $C_p = 40 \text{ kcal K}^{-1} \text{ m}^{-3}$  of the soil were assumed as in the literature [15]. A Young's modulus  $E = 60 \text{ MPa}$ , a Poisson's ratio  $\nu = 0.2857$  and a cubic thermal expansion coefficient  $\beta_s = 0.9 \times 10^{-6}$ , lying in the range of possible values for this kind of clay [54], were also assumed. The relationships given by Brooks and Corey [52] for the relative permeabilities of water and gas, as well as for the saturation-capillary pressure curves are as follows:

$$\begin{aligned} k^{rw} &= S_e^{(2+3\lambda)/\lambda} & k^{rg} &= (1 - S_e)^2 (1 - S_e^{(2-\lambda)/\lambda}), \\ S_e &= (S_w - S_{irr}) / (1 - S_{irr}) & p^c &= p^b S_e^{1/\lambda} \end{aligned}$$



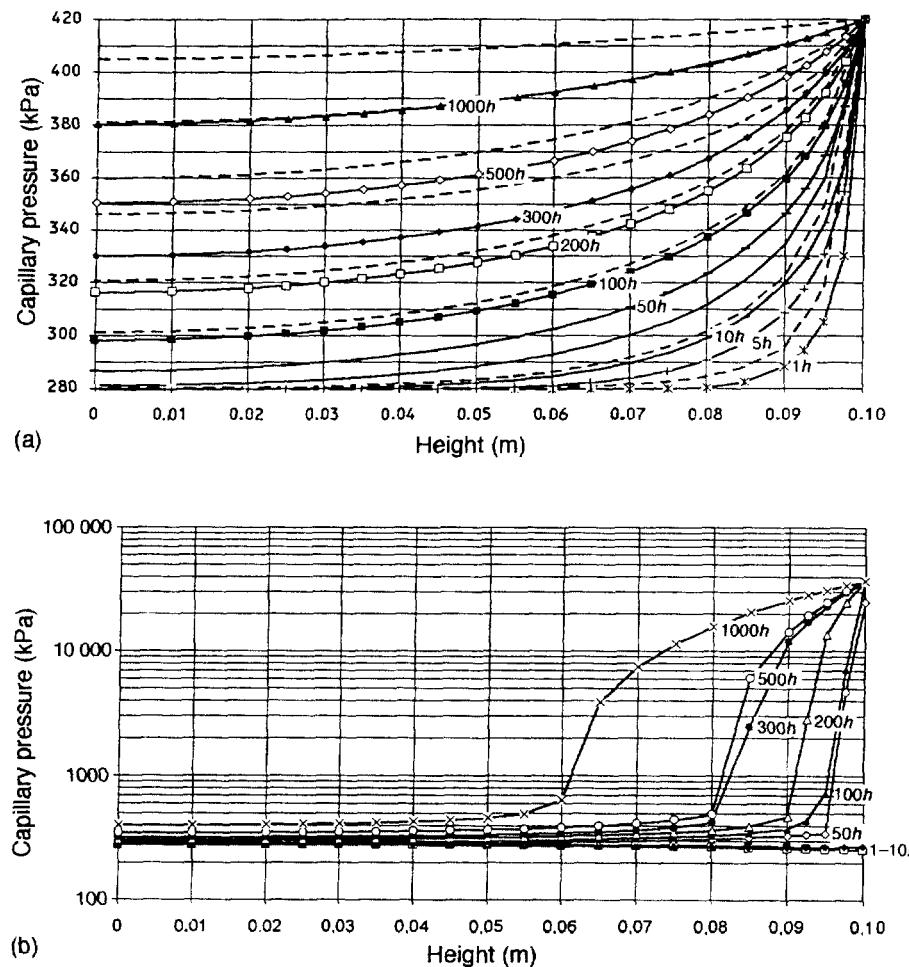
**Figure 10.24** The resulting temperature profiles during thermoelastic consolidation of clay: (a) solved for boundary conditions of the first kind (solid lines) compared with the solution of Schrefler *et al.* [22] (dashed lines); (b) solved for boundary conditions of the third kind



**Figure 10.25** The resulting saturation profiles during thermoelastic consolidation of clay: (a) solved for boundary conditions of the first kind (solid lines) compared with the solution of Schrefler *et al.* [22] (dashed lines); (b) solved for boundary conditions of the third kind

where  $S_e$  is the effective saturation; the bubbling pressure  $p^b = 133\,813$  Pa, the irreducible saturation  $S_{irr} = 0.3216$  and  $\lambda = 2.308$ , corresponding to the given saturation–capillary pressure values in the literature [15].

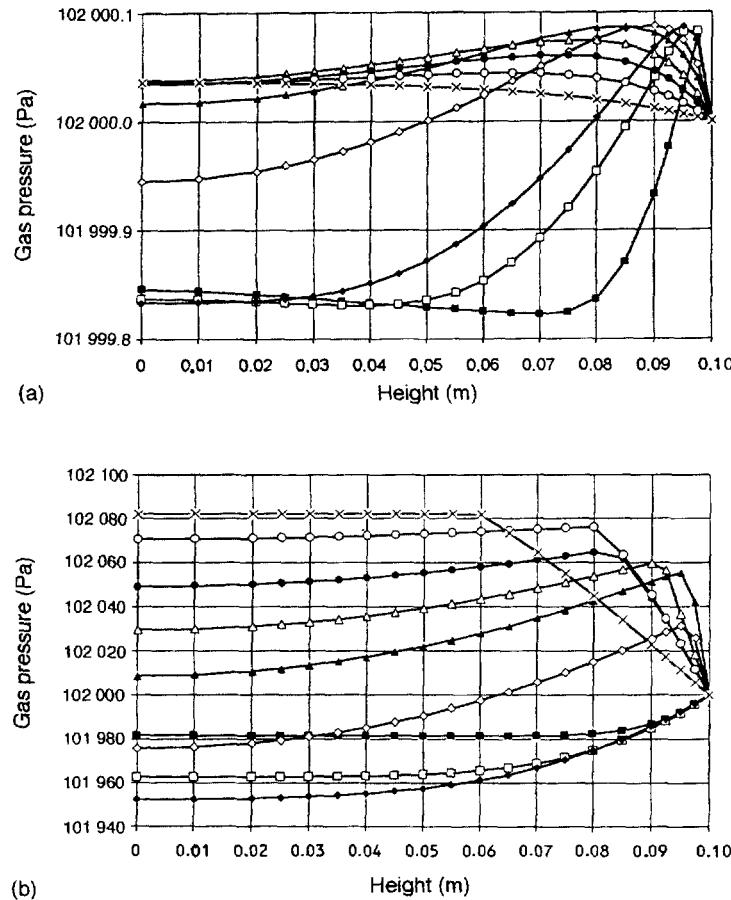
The initial conditions are  $T = 283.15$  K,  $p^g = 102$  kPa,  $p^c = 280$  kPa and the deformations are assumed to be consistent with a state of mechanical equilibrium. The boundary conditions for the lateral surface  $q^T = 0$ ,  $q^w = 0$ ,  $q^g = 0$ , horizontal displacements  $u_h = 0$ ; for the top surface were  $T = 298.15$  K,  $p^g = 102$  kPa,  $p^c = 420$  kPa; and for the bottom surface,  $q^T = 0$ ,  $q^w = 0$ ,  $q^g = 0$ , vertical displacements  $u_v = 0$ . The column of soil was simulated by 10 eight-noded isoparametric elements of different height ( $2 \times 0.5$  cm,  $7 \times 1.0$  cm,  $1 \times 2$  cm). The



**Figure 10.26** The resulting capillary pressure profiles during thermoelastic consolidation of clay: (a) solved for boundary conditions of the first kind (solid lines) compared with the solution of Schrefler *et al.* [22] (dashed lines); (b) solved for boundary conditions of the third kind

temporal discretisation was performed with an initial step of 0.001 h for the first 100 steps and multiplied thereafter by 10 for every subsequent 100 steps until a total elapsed time of 1000 h.

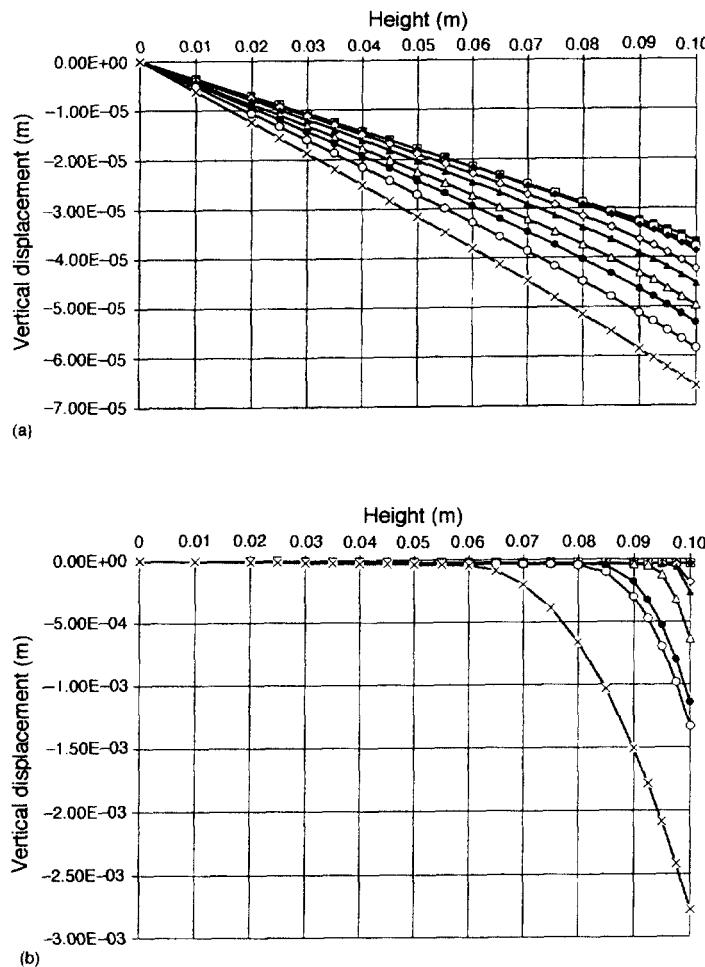
The resulting profiles of temperature, saturation and water pressure are compared in Figures 10.24(a), 10.25(a) and 10.26(a) with the results obtained by Schrefler *et al.* [22]. The slight differences between the temperature profiles (Figure 10.24(a)) could be caused by a different temporal discretisation (much finer for the present solution) and the inclusion of latent heat transfer. The visible differences between the saturation profiles (Figure 10.25(a)) and between the capillary pressure profiles (Figure 10.26(a)) are caused by the different curve shapes assumed for the relative permeability–capillary pressure and saturation–capillary pressure relationships used by Schrefler *et al.* [22].



**Figure 10.27** The resulting gas pressure profiles during thermoelastic consolidation of clay: (a) solved for boundary conditions of the first kind; (b) solved for boundary conditions of the third kind. Key: (■) 1 h, (□) 5 h, (●) 10 h, (◇) 50 h, (▲) 100 h, (○) 200 h, (●) 300 h, (○) 500 h, (×) 1000 h

(different parameters  $p^b$ ,  $S_{irr}$  and  $\lambda$  for the relations of Brooks and Corey [52]). The gas pressure profiles (Figure 10.27(a)) show that the gas pressure is practically constant in this case. The resulting profiles of vertical displacement are presented in Figure 10.28a.

The boundary conditions assumed at the top surface of the soil are very artificial from a physical point of view [53, 55]. Therefore, the same example was subsequently solved for more appropriate, natural boundary conditions: convective heat transfer coefficient  $\alpha_c = 23.1 \text{ W m}^{-2} K^{-1}$ , ambient temperature  $T = 298.15 \text{ K}$ ,  $p^g = 102 \text{ kPa}$ , convective mass transfer coefficient  $\beta_c = 0.0209 \text{ kg m}^{-3}$  and the relative humidity of surrounding air  $\text{RH} = 50$  percent. Also, the temperature dependence was taken into account for the capillary pressure–saturation and the thermal conductivity and capacity–water saturation relationships. The same space discretisation was used, but the temporal discretisation was performed with an initial step of 0.01 h during the first 1000 steps, then 0.1 h during the next 900 steps for the remaining period.



**Figure 10.28** The resulting vertical displacement profiles during thermoelastic consolidation of clay: (a) solved for boundary conditions of the first kind; (b) solved for boundary conditions of the third kind. Key: (■) 1 h, (□) 5 h, (◆) 10 h, (◇) 50 h, (▲) 100 h, (○) 200 h, (●) 300 h, (△) 500 h, (×) 1000 h

The resulting profiles of temperature, saturation, capillary pressure, gas pressure and vertical displacement are respectively shown in part (b) of Figures 10.24 to 10.28; notice the differences between part (a) and part (b) of each figure. The temperature gradients are much smaller in Figure 10.24(b) and the final value of the surface temperature was reached after approximately 1000 h (not immediately as previously assumed). Due to the temperature increase, the saturation values in the zone close to the upper surface increase slightly at the beginning (Figure 10.25(b)) then they decrease, reaching a final surface value after about 50 h (not immediately as previously assumed). Compared with the previous solution, the saturation gradients in the zone close to the surface are significantly higher (Figure 10.25). The capillary pressure profiles behave accordingly

(Figure 10.26(b), causing considerable displacements firstly near the surface (Figure 10.28(b)), though these values should not be compared with the previous solution because of the different final value of the capillary pressure at the surface. The changes in the values of the gas pressure are greater than for the previous solution, although they still do not exceed 85 Pa (Figure 10.27(b)). Notice that the whole process is strongly influenced by the boundary conditions. The behaviour of the soil for these more physically realistic values is completely different.

## 10.8 CONCLUSIONS

The full model described in Chapter 2 has been solved numerically for fully and partially saturated conditions. The model fully takes into account the interaction between the displacement, pressure and temperature fields. Appropriate solution techniques for such interaction problems has been given and their numerical properties discussed. The model has been applied to the cases of plane-strain thermoelastic and thermoplastic consolidation, consolidation around a cylindrical heat surface buried in saturated clay, and partially saturated non-isothermal consolidation including phase change. The need for appropriate boundary conditions has been demonstrated when simulating realistic problems.

## REFERENCES

1. Borsetto, M., Carradori, G. and Ribacchi R. (1981) Coupled seepage, heat transfer and stress analysis with applications to geothermal problems, in *Numerical Heat Transfer*, R. W. Lewis, K. Morgan and O. C. Zienkiewicz (eds), Wiley, Chichester.
2. Booker, J. R. and Savvidou, C. (1985) Consolidation around a heat source. *Int. J. Num. Anal. Meth. Geomech.*, **9**, 173–84
3. Corapcioglu, M. Y. and Karahanoglu, N. (1980) Simulation of geothermal production. in *Alternative Energy Sources II*, Vol. 5, T. Nejat Veziroglu (ed.), Hemisphere, New York, pp. 1895–1918.
4. Lippmann, M. J., Narasimhan, T. N. and Witherspoon, P. A. (1976) Numerical simulation of reservoir compaction in liquid-dominated geothermal systems, in *Proceedings of the Second International Symposium of Land Subsidence*, International Association of Hydrological Sciences, Anaheim CA, pp. 179–89.
5. Brownell, D. H., Garg, S. K. and Pritchett, J. W. (1977) Governing equations for geothermal reservoirs. *Water resources Research*, **13**, 929–34.
6. Pritchett, J. W., Garg, S. K. Brownell, D. H., Rice, L. F., Rice, M. H., Riney, T. D. and Hendrickson, R. R. (1976) Geohydrological environmental effects of geothermal power production. Phase IIa, Systems, Science and Software, SSS-R-77-2998, La Jolla CA.
7. Garg, S. K., Pritchett, J. W., Rice, M. H. and Riney, T. D. (1977) US Gulf Coast geopressured geothermal reservoir simulation. Final Report, Systems, Science and Software, SSS-R-77-3147, La Jolla CA.
8. Aktan, T. and Farouq Ali, S. M. (1978) Finite element analysis of temperature and thermal stresses induced by hot water injection. *Soc. Pet. Eng. J.*, **18**, 457–69
9. Ertekin, T. (1978) Numerical simulation of the compaction–subsidence phenomena in a reservoir for two-phase non-isothermal Flow., PhD thesis, Pennsylvania State University.

10. Bear, J. and Corapcioglu, M. Y. (1981) A mathematical model for consolidation in a thermoelastic aquifer due to hot water injection or pumping. *Water Resources Research*, **17**, 723–36.
11. Borsetto, M., Cricchi, D., Hueckel, T. and Peano, A. (1984) On numerical models for the analysis of nuclear waste disposal in geological clay formations, in *Numerical Methods for Transient and Coupled Problems*, R. W. Lewis, E. Hinton, P. Bettes and B. A. Schrefler (eds), Pineridge Press, Swansea, pp. 608–18.
12. Aboustit, B. L., Advani, S. H. and Lee, J. K. (1985) Variational principles and finite element simulations for thermo-elastic consolidation. *Int. J. Num. Anal. Meth. Geomech.*, **9**, 49–69.
13. Philip, J. R. and de Vries, D. A. (1957) Moisture movement in porous materials under temperature gradients. *Trans. Amer. Geophys. Union*, **38**, 222–32.
14. Alonso, E. E., Gens, A. and Hight, D. W. (1987) Special problem soils: general report, in *Proceedings of the 9th European Conference on Soil Mechanics and Foundation Engineering*, Vol. 3, Balkema, Rotterdam, pp. 1087–1146.
15. Dakshanamurthy, V. and Fredlund, D. G. (1981) A mathematical model for predicting moisture flow in an unsaturated soil under hydraulic and temperature gradients. *Water Resources Research*, **17**, 714–722.
16. Thomas, H. (1985) Modelling two-dimensional heat and moisture transfer in unsaturated soils, including gravity effects. *Int. J. Num. Anal. Meth. Geomech.*, **9**, 573–88.
17. Luikov, A. V. (1966) *Heat and Mass Transfer in Capillary porous Bodies*. Pergamon, Oxford.
18. Thomas, H. R. and King, S. D. (1991) Coupled temperature/capillary potential variation in unsaturated soils. *J. Engng Mech.*, **117**(11), 2475–90.
19. Baggio, P., Majorana, C. E. and Schrefler, B. A. (1993) Hydrothermomechanical analysis of concrete by a finite element method, in *Proceedings of the 8th International Conference on Numerical Methods for Thermal Problems*, R. W. Lewis (eds.), Pineridge Press, Swansea, pp. 847–59.
20. Geraminegad, M. and Saxena, S. K. (1986) A coupled thermoelastic model for saturated–unsaturated porous media. *Géotechnique*, **36**(4), 539–50.
21. Lewis, R. W., Roberts, P. J. and Schrefler, B. A. (1989) Finite element modelling of two-phase heat and fluid flow in deforming porous media. *Transport in Porous Media*, **4**, 319–34.
22. Schrefler, B. A., Zhan, X. Y. and Simoni, L. (1995) A coupled model for water flow, airflow and heat flow in deformable porous media. *Int. J. Num. Meth. Heat and Fluid Flow*, **5**, 531–47.
23. Gawin, D., Baggio, P. and Schrefler, B. A. (1995) Coupled heat, water and gas flow in deformable porous media. *Int. J. Num. Meth. Fluids*, **20**, 969–87.
24. Olivella, S., Carrera, J., Gens, A. and Alonso, E. E., (1994) Non-isothermal multiphase flow of brine and gas through saline media. *Transport in Porous Media*, **15**, 271–93.
25. Lewis, R. W. and Karahanoglu, N. (1981) Simulation of subsidence in geothermal reservoirs, in *Numerical Methods in Thermal Problems*, Vol. II, R. W. Lewis, K. Morgan and B. A. Schrefler (eds.), Pineridge Press, Swansea, pp. 326–35.
26. Park, K. C. and Felippa C. A. (1983) Partitioned analysis of coupled systems, in *Computational Methods for Transient Analysis*, T. Belytschko and T. R. J. Hughes (eds), Elsevier, Amsterdam, pp. 157–219.
27. Schrefler, B. A. (1985) A partitioned solution procedure for geothermal reservoir analysis. *CANM (Communications in Applied Numerical Methods)*, **1**, 53–56.
28. Turska, E. and Schrefler, B. A. (1993) On convergence conditions of partitioned solution procedures for consolidation problems. *Comp. Meth. Appl. Mech. Engng.*, **106**(1/2), 51–64.
29. Park, K. C. (1980) Partitioned transient analysis procedures for coupled-field problems: stability analysis. *ASME J. App. Mech.*, **47**, 370–76.
30. Turska, E., Wisniewski, K. and Schrefler, B. A. (1993) Error propagation of staggered solution procedures for transient problems. *Comp. Meth. Appl. Mech. Engng.*, **114**, 51–63.

31. Armero, F. and Simo, J. C. (1992) A new unconditionally stable fractional step method for nonlinear coupled thermo-mechanical problems. *Int. J. Num. Meth. Engng.*, **35**, 737–66.
32. Yanenko N. N. (1967) *Method of Fractional Steps for Solving Multi-Dimensional Problems of Mathematical Physics*, Nauka, Novosibirsk (In Russian).
33. Marchuk, G. I. (1975) *Methods of Numerical Analysis*, Springer-Verlag, New York.
34. Chorin, A. J., Hughes, T. J. R., McCracken, M. F. and Marsden, J. E. (1978) Product formulas and numerical algorithms. *Commun. Pure Appl. Math.*, **XXXI**, 205–56.
35. Strang, G. (1969) Approximating semigroups and the consistency of difference methods. *Proc. Amer. Math. Soc.*, **20**, 1–7.
36. Miehe, C. (1988) Zur numerischen Behandlung thermomechanischer Prozesse. University of Hanover Report Nr. F 88/6.
37. Miehe, C., Stein, E. and Wriggers, P. (1988) A general concept for thermomechanical coupling, in *Computational Mechanics 88* (Proceedings of an ICES Conference), S. N. Atluri and G. Yagawa (eds), Springer-Verlang, Berlin, Section 59, iv, pp. 1–4.
38. Simo, J. C. and Miehe, C. (1992) Associative coupled thermoplasticity at finite strains: formulation, numerical analysis and implementation. *Comp. Meth. Appl. Mech. Engng.*, **98**, 41–104.
39. Wriggers, P., Miehe, C., Kleiber, M. and Simo, J. C. (1993) On the coupled thermomechanical treatment of necking problems via finite element methods. *Int. J. Num. Meth. Engng.*, **33**, 869–83.
40. Matteazzi, R., Schrefler, B. A. and Vitaliani, R. (1996) Comparison of partitioned solution procedures for transient coupled problems in sequential and parallel processing, in *Advances in Finite Element Technology*, B. H. V. Topping (ed), Civil-Comp. Press, Edinburgh, pp. 351–57.
41. Turska, E. and Schrefler B. A. (1994) On consistency, stability and convergence of staggered solution procedures. *Rend. Mat. Acc. Lincei*, s. 9, **5**, 265–71.
42. Schrefler, B. A., Simoni, L. and Turska, E. (1997) Standard staggered and staggered Newton schemes in thermo-hydro-mechanical problems. *Comp. Meth. Appl. Mech. Engng.*, **144**, 93–109.
43. Saetta, A., Schrefler, B. A. and Vitaliani, R. (1991) Solution strategies for coupled consolidation analysis in porous media. *Int. J. Methods Numericos para Calculo y Diseño en Ingenieria*, **7**, 55–66.
44. Murti, V., Vallipapan, S. and Khalili-Naghadeh, N. (1989) Times step constraints in finite element analysis of the Poisson type equation. *Computers and Structures*, **31**, 269–73.
45. Rank, E., Katz, C., and Werner, H (1983) On the importance of the discrete maximum principle in transient analysis using finite element methods. *Int. J. Num. Meth. Engng.*, **19**, 1771–82.
46. Gawin, D., Baggio, P. and Schrefler, B. A. (1996) Modelling heat and moisture transfer in deformable porous building materials. *Arch. Civ. Engng.*, **XLII**(3), 325–49.
47. Incropera, F. P. and De Witt, D. P. (1990) *Fundamentals of Heat and Mass Transfer*, 3rd edn, Wiley, New York.
48. Combarous, N. A. and Bories, S. A. (1975) Hydrothermal convection in saturated porous media, in *Advances in Hydroscience*, Vol. 10, V. T. Chow (ed.), Academic Press, New York, pp. 231–307.
49. Siriwardane, H. J. and Desai, C. (1981) Two numerical schemes for nonlinear consolidation. *Int. J. Num. Meth. Engng.*, **17**, 405–26.
50. Simoni, L. and Schrefler, B. A. (1989) Numerical modelling of infinite domains in coupled field problems. *Meccanica*, **24**, 98–106.
51. Schrefler, B. A., Lewis, R. W. and Majorana, C. E. (1986) A comparison between different simulators for non-isothermal consolidation, in *Proceedings of the European Conference on Numerical Methods in Geomechanics*, University of Stuttgart, pp. 1–16.
52. Brooks, R. N. and Corey, A. T. (1966) Properties of porous media affecting fluid flow. *J. Irrig. Drain. Div. Amer. Soc. Civ. Engng.*, **92** (IR2), 61–68.

53. Krischer, O. and Kroell, K. (1978) *Trocknungstechnik*, Vol. 1, 3rd edn, *Die wissenschaftlichen Grundlagen der Trocknungstechnik*, Springer-Verlag, Berlin.  
 54. Lambe T. W. and Whitman, R. V. (1969) *Soil Mechanics*, Wiley, Chichester.  
 55. Gawin, D. and Klemm, P. (1994) A model of coupled heat and moisture transfer with phase changes in porous building material. *Arch. Civ. Engng.*, **40**(1), 89–104.

## APPENDIX 10A

### 10A.1 Standard Staggered Scheme

This appendix uses an iteration–convergence analysis [1] to compare the standard staggered procedure and the staggered Newton procedure. This means that the matrices  $\mathbf{G}$  of Section 10.4.2.1 are obtained and compared for the two procedures. Examples are given in Schrefler *et al.* [1].

For simplicity we shall consider a coupled problem consisting of two scalar fields  $x, y$ . The scalar fields can be easily generalised to vector arrays characterising two vector fields (e.g. displacements and pressures forming one field and temperatures the other (section 10.4)).

Let

$$\dot{\mathbf{X}} + \mathbf{C}\mathbf{X} = \mathbf{F} \quad (10A.1)$$

be the semidiscrete form of the governing equations (obtained from equation 10.37) where, in general, matrix  $\mathbf{C}$  can be non-symmetric and can depend on  $\mathbf{X}$ ,  $\mathbf{X} = [x, y]^T$ , and  $\mathbf{F} = [F^x, F^y]^T$ . The time-stepping algorithm is chosen to be the implicit backward Euler scheme:

$$\dot{x}_{n+1} = \frac{x_{n+1} - x_n}{\Delta t} \quad (10A.2)$$

$$\dot{y}_{n+1} = \frac{y_{n+1} - y_n}{\Delta t} \quad (10A.3)$$

This is a particular case of the implicit midpoint rule for  $\theta = 1$ . The use of this temporal integration is particularly useful in convective problems, but a generalisation is straightforward.

Thus from (10A.1), in conjunction with (10A.2) and (10A.3), we find

$$(\mathbf{I} + \Delta t \mathbf{C}_{n+1}) \mathbf{X}_{n+1} = \Delta t \mathbf{F}_{n+1} + \mathbf{X}_n \quad (10A.4)$$

Now we shall focus our attention on solving (10A.4) for a fixed time step  $n + 1$ . Let  $\mathbf{C}_{n+1} = [c_{ij}]$ , then (10A.4) in matrix form is written as

$$\begin{bmatrix} 1 + \Delta t c_{11} & \Delta t c_{12} \\ \Delta t c_{21} & 1 + \Delta t c_{22} \end{bmatrix} \begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \Delta t \begin{bmatrix} F_{n+1}^x \\ F_{n+1}^y \end{bmatrix} + \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (10A.5)$$

To perform the staggered iteration scheme, we first split the matrix on the left-hand side

of equation (10A.5):

$$\begin{bmatrix} 1 + \Delta t c_{11} & \Delta t c_{12} \\ \Delta t c_{21} & 1 + \Delta t c_{22} \end{bmatrix} = \begin{bmatrix} 1 + \Delta t c_{11} & 0 \\ 0 & 1 + \Delta t c_{22} \end{bmatrix} + \Delta t \begin{bmatrix} 0 & c_{12} \\ c_{21} & 0 \end{bmatrix} \quad (10A.6)$$

then we transfer part of it to the right-hand side. The iteration procedure starts with a predicted value of  $y_{n+1,0}$ ,  $x_{n+1,0}$  (only with  $y_{n+1,0}$  in the linear case) and provides new estimates  $x_{n+1,i}$  and  $y_{n+1,i}$  by the following equations:

$$x_{n+1,i} = (1 + \Delta t c_{11})^{-1} [\Delta t F_{n+1}^x + x_{n,K} - c_{12} y_{n+1,i-1}] \quad (10A.7)$$

where

$$c_{11} = c_{11}(x_{n+1,i-1}, y_{n+1,i-1}), c_{12} = c_{12}(x_{n+1,i-1}, y_{n+1,i-1})$$

$$y_{n+1,i} = (1 + \Delta t c_{22})^{-1} [\Delta t F_{n+1}^y + y_{n,K} - c_{21} x_{n+1,i}] \quad (10A.8)$$

where

$$c_{22} = c_{22}(x_{n+1,i}, y_{n+1,i-1}), c_{21} = c_{21}(x_{n+1,i}, y_{n+1,i-1})$$

where  $K$  denotes the number of performed iterations and depends on  $n$ ,  $K = K(n)$ .

The predictor is usually chosen as a linear extrapolation of the last obtained values, although to obtain better convergence properties an explicit one-step or multi-step formula is recommended [2].

In concise form (10A.7) and (10A.8) can written as

$$x_{n+1,i} P(x_{n+1,i-1}, y_{n+1,i-1}) \quad (10A.9)$$

$$y_{n+1,i} = Q(x_{n+1,i}, y_{n+1,i-1}) \quad (10A.10)$$

where  $P(x, y)$  denotes the right-hand side of (10A.7) and  $Q(x, y)$  denotes the right-hand side of (10A.8). As we see, it is a sequential successive substitution method. Let us assess the convergence properties of its estimates  $x_{n+1,i}$ ;  $y_{n+1,i}$  to the exact solution  $x_{n+1}$ ,  $y_{n+1}$  of equations (10A.9) and (10A.10).

Since the following considerations regard a fixed time instant  $t_{n+1} = (n+1)\Delta t$ , all the subscripts  $n+1$  shall be omitted, i.e.  $x_{n+1,i} = x_i$ ,  $y_{n+1,i} = y_i$ .

The iteration error terms are defined as

$$E_i^x = x_i - x \quad (10A.11)$$

$$E_i^y = y_i - y \quad (10A.12)$$

If  $P$ ,  $Q$  satisfy the necessary conditions, then by Taylor's theorem

$$P(x_{i-1}, y_{i-1}) = P(x, y) + P_{,x}(x, y)E_{i-1}^x + P_{,y}(x, y)E_{i-1}^y + O((E_{i-1}^x, E_{i-1}^y)^2) \quad (10A.13)$$

$$Q(x_i, y_{i-1}) = Q(x, y) + Q_{,x}(x, y)E_{i-1}^x + Q_{,y}(x, y)E_{i-1}^y + O((E_{i-1}^x, E_{i-1}^y)^2) \quad (10A.14)$$

From the fact that  $x, y$  are exact solutions of (10A.9) and (10A.10) i.e.  $x = P(x, y)$ ,  $y = Q(x, y)$ , and using (10A.9) and (10A.10), we derive expressions for the error terms:

$$E_i^x = P_{,x}E_{i-1}^x + P_{,y}E_{i-1}^y + O((E_{i-1}^x, E_{i-1}^y)^2) \quad (10A.15)$$

$$E_i^y = Q_{,x}E_i^x + Q_{,y}E_{i-1}^y + O((E_i^x, E_{i-1}^y)^2) \quad (10A.16)$$

or in matrix form

$$\begin{bmatrix} 1 & 0 \\ -Q_{,x} & 1 \end{bmatrix} \mathbf{E}_i \cong \begin{bmatrix} P_{,x} & P_{,y} \\ 0 & Q_{,y} \end{bmatrix} \mathbf{E}_{i-1} \quad (10A.17)$$

where  $\mathbf{E}_i = [E_i^x, E_i^y]^T$ .

Thus the staggered procedure is of first order (linear convergence) and its amplification matrix  $\mathbf{A}^S$  is

$$\mathbf{A}^S = \begin{bmatrix} P_{,x} & P_{,y} \\ Q_{,x}P_{,x} & Q_{,x}P_{,y} + Q_{,y} \end{bmatrix} \quad (10A.18)$$

For the errors  $E_i^x, E_i^y$  to decrease, the spectral radius of  $\mathbf{A}^S$  must be less than unity:

$$\rho(\mathbf{A}^S) < 1 \quad (10A.19)$$

This condition limits the applicability of the method and suggests the Newton procedure as a modification. In its usual form the Newton method is of second order and is always convergent close to the root.

## 10A.2 Staggered Newton Scheme

Let us determine the order of iteration convergence for the Newton staggered Scheme. Following Miehe [3] and Simo and Miehe [4], we define the Newton staggered procedure for the monolithic problem (10A.4) written as

$$(\mathbf{I} + \Delta t \mathbf{C}_{n+1}) \mathbf{X}_{n+1} - \mathbf{X}_n - \Delta t \mathbf{F}_{n+1} = \mathbf{0} \quad (10A.20)$$

or

$$p(x_{n+1}, y_{n+1}) = 0 \quad (10A.21)$$

$$q(x_{n+1}, y_{n+1}) = 0 \quad (10A.22)$$

where

$$p(x_{n+1}, y_{n+1}) = (1 + \Delta t c_{11})x_{n+1} + \Delta t c_{12}y_{n+1} - x_n - \Delta t F_{n+1}^x \quad (10A.23)$$

$$q(x_{n+1}, y_{n+1}) = \Delta t c_{21}x_{n+1} + (1 + \Delta t c_{22})y_{n+1} - y_n - \Delta t F_{n+1}^y \quad (10A.24)$$

As before, to simplify the equations, we shall omit subscript  $n+1$ :  $x_{n+1,i} = x_i, y_{n+1,i} = y_i$ . The staggered Newton Scheme is given by the iterations

$$x_i = x_{i-1} - p(x_{i-1}, y_{i-1})/p_{,x}(x_{i-1}, y_{i-1}) \quad (10A.25)$$

$$y_i = y_{i-1} - q(x_i, y_{i-1})/q_{,y}(x_i, y_{i-1}) \quad (10A.26)$$

where  $x_0, y_0$  are the starting values.

Expanding the quotients in (10A.25) and (10A.26) as a Taylor series about the point  $(x,y)$  and taking into account that  $x,y$  are exact solutions of (10A.25) and (10A.26), we obtain

$$E_i^x = -\frac{p_{,y}}{p_{,x}} E_{i-1}^x - O((E_{i-1}^x, E_{i-1}^y)^2) \quad (10A.27)$$

$$E_i^y = -\frac{q_{,x}}{q_{,y}} E_i^y - O((E_i^x, E_i^y)^2) \quad (10A.28)$$

where all the derivatives are taken at the point  $(x,y)$ .

In matrix form

$$\mathbf{E}_i \cong \mathbf{A}^N \mathbf{E}_{i-1} \quad (10A.29)$$

where

$$\mathbf{A}^N = \begin{bmatrix} 0 & -\frac{p_{,y}}{p_{,x}} \\ 0 & \frac{p_{,y}q_{,x}}{p_{,x}q_{,y}} \end{bmatrix}$$

is the amplification matrix. Condition

$$\rho(\mathbf{A}^N) < 1 \quad (10A.30)$$

must be satisfied to ensure convergence.

As we see, the staggered Newton scheme is only of first order; it has the same order as the standard staggered scheme but requires an additional evaluation of tangent operators.

For weakly coupled problems, i.e.  $p_{,y} = 0$ , the method is of second order. This was the case considered by Miehe [3], Simo and Miehe [4] and Wriggers *et al.* [5], where coupling terms were neglected and good numerical properties achieved.

The comparison of the maximum eigenvalues of the amplification matrices  $\mathbf{A}^S$  and  $\mathbf{A}^N$  tells us which iteration should proceed faster, i.e. the smaller the spectral radius, the faster the process. This is concluded from the fact that for any matrix norm  $\|\cdot\|$  compatible with a vector norm

$$\lim_{n \rightarrow \infty} \|\mathbf{A}\|^n = \lim_{n \rightarrow \infty} \rho(\mathbf{A})^n \quad (10A.31)$$

The eigenvalues of matrix  $\mathbf{A}^S$  are

$$\lambda_{1,2}^S = \frac{1}{2}(Q_{,x}P_{,y} + Q_{,y} + P_{,x} \pm \Delta) \quad (10A.32)$$

where

$$\Delta = [(Q_{,x}P_{,y} + Q_{,y} + P_{,x})^2 - 4P_{,x}Q_{,y}]^{0.5}$$

Matrix  $\mathbf{A}^N$  possesses eigenvalues

$$\lambda_{1,2}^N = 0 \quad \text{and} \quad \frac{P_{,y}Q_{,x}}{P_{,x}Q_{,y}} \quad (10A.33)$$

To be able to compare them we must find a relation between  $P, Q$  and  $p, q$ . From the definitions of  $P, Q$ , found in (10A.9) and (10A.10) in conjunction with (10A.25) and (10A.26), we obtain

$$p(x, y) = (1 + \Delta t c_{11})x - (1 + \Delta t c_{11})P(x, y) \quad (10A.34)$$

$$q(x, y) = (1 + \Delta t c_{22})y - (1 + \Delta t c_{22})Q(x, y) \quad (10A.35)$$

Taking into account that  $x, y$  is the exact solution of (10A.9), (10A.10) and (10A.25), (10A.26), we find that

$$p_{,x}(x, y) = (1 + \Delta t c_{11})(1 - P_{,x}(x, y)) \quad (10A.36)$$

$$p_{,y}(x, y) = -(1 + \Delta t c_{11})P_{,y}(x, y) \quad (10A.37)$$

$$q_{,x}(x, y) = -(1 + \Delta t c_{22})Q_{,x}(x, y) \quad (10A.38)$$

$$q_{,y}(x, y) = (1 + \Delta t c_{22})(1 - Q_{,y}(x, y)) \quad (10A.39)$$

Let us consider some particular cases:

1. In weakly coupled problems, i.e.  $p_{,y} = 0, q_{,x} \neq 0$ , we have  $\lambda_{1,2}^S = P_{,x}$  and  $Q_{,y}$ ,  $\mathbf{A}^N = 0$ ; the staggered Newton method is of second order.
2. In alternatively coupled problems, i.e.  $p_{,x} = 0, q_{,y} = 0$ , the staggered Newton method cannot be used.
3. In strongly coupled problems, where  $P_{,x} = 0, Q_{,y} = 0$ , then

$$\begin{aligned} \lambda_1^S &= 0 & \lambda_1^N &= P_{,y}Q_{,x} \\ \lambda_1^N &= 0 & \lambda_1^N &= p_{,y}q_{,x}/(1 + \Delta t c_{11})(1 + \Delta t c_{22}) = P_{,y}Q_{,x} \end{aligned}$$

The spectral radii are equal, so the standard method and the Newton method converge at the same rate.

4. When  $P_{,x} = 0$  the standard method may be faster than the Newton method. Then  $\lambda_{1,2}^S = 0, P_{,y}Q_{,x} + Q_{,y}$  and  $\lambda_{1,2}^N = 0, P_{,y}Q_{,x}/(1 - Q_{,y})$ . A typical example is the system of equations in Appendix C of Turska and Schrefler [6].
5. The staggered Newton scheme can be divergent, even for starting values very close to the roots.

## References

1. Schrefler, B. A., Simoni, L. and Turska, E. (1997) Standard staggered and staggered Newton schemes in thermo-hydro-mechanical problems. *Comp. Meth. in Appl. Mech. Engng.*, **144**, 93–109.

2. Turska, E., Wisniewski, K. and Schrefler, B. A. (1993) Error propagation of staggered solution procedures for transient problems. *Comp. Meth. Appl. Mech. Engng.*, **114**, 51–63.
3. Miehe, C. (1988) Zur numerischen Behandlung thermomechanischer Prozesse. University of Hanover Report Nr. F 88/6.
4. Simo, J. C. and Miehe, C. (1992) Associative coupled thermoplasticity at finite strains: formulation, numerical analysis and implementation. *Comp. Meth. Appl. Mech. Engng.*, **98**, 41–104.
5. Wriggers, P., Miehe, C., Kleiber, M. and Simo, J. C. (1993) On the coupled thermomechanical treatment of necking problems via finite element methods. *Int. J. Num. Meth. Engng.*, **33**, 869–83.
6. Turska, E. and Schrefler, B. A. (1993) On convergence conditions of partitioned solution procedures for consolidation problems. *Comp. Meth. Appl. Mech. Engng.*, **106**(1/2), 51–64.

## APPENDIX 10B

The matrices occurring in the discretised mass and energy conservation equations (10.94) to (10.96) are defined in this appendix.

$$\mathbf{C}_{ws} = \int_{\Omega} \mathbf{N}_p^T [\alpha(\rho^{gw} S_g + \rho^w S_w) \mathbf{m}^T \mathbf{L} \mathbf{N}_p] d\Omega \quad (10B.1)$$

$$\begin{aligned} \mathbf{P}_{ww} = & \int_{\Omega} \mathbf{N}_p^T \left[ \left( \frac{\alpha - n}{K_s} S_w (\rho^{gw} S_g + \rho^w S_w) + \rho^w S_w \frac{n}{K_w} \right) \mathbf{N}_p \right] d\Omega \\ & - \int_{\Omega} \mathbf{N}_p^T \left( S_g n \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \mathbf{N}_p \right) d\Omega \\ & - \int_{\Omega} \mathbf{N}_p^T \left\{ \left[ \frac{\alpha - n}{K_s} (\rho^{gw} S_g p^c + \rho^w S_w p^w - \rho^w S_w p^c) + n(\rho^w - \rho^{gw}) \right] \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \right\} d\Omega \end{aligned} \quad (10B.2)$$

$$\begin{aligned} \mathbf{C}_{wg} = & \int_{\Omega} \mathbf{N}_p^T \left[ \frac{\alpha - n}{K_s} S_g (\rho^{gw} S_g + \rho^w S_w) \mathbf{N}_p \right] d\Omega + \int_{\Omega} \mathbf{N}_p^T \left[ S_g n \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \mathbf{N}_p \right] d\Omega \\ & + \int_{\Omega} \mathbf{N}_p^T \left\{ \left[ \frac{\alpha - n}{K_s} (\rho^{gw} S_g p^c + \rho^w S_w p^w - \rho^w S_w p^c) + n(\rho^w - \rho^{gw}) \right] \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \right\} d\Omega \end{aligned} \quad (10B.3)$$

$$\begin{aligned} \mathbf{C}_{wt} = & - \int_{\Omega} \mathbf{N}_p^T \left( S_w \beta_{sgw} \mathbf{N}_t \frac{\partial \bar{\mathbf{T}}}{\partial t} \right) d\Omega + \int_{\Omega} \mathbf{N}_p^T \left[ S_g n \frac{M_w}{\theta R} \left( \frac{\partial p^{gw}}{\partial T} - \frac{p^{gw}}{\theta} \right) \mathbf{N}_t \right] d\Omega \\ & + \int_{\Omega} \mathbf{N}_p^T \left\{ \left[ \frac{\alpha - n}{K_s} (\rho^{gw} S_g p^c + \rho^w S_w p^w - \rho^w S_w p^c) + n(\rho^w - \rho^{gw}) \right] \frac{\partial S_w}{\partial T} \mathbf{N}_t \right\} d\Omega \end{aligned} \quad (10B.4)$$

$$H_{ww} = \int_{\Omega} (\nabla \mathbf{N}_p)^T \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} \nabla \mathbf{N}_p d\Omega + \int_{\Omega} (\nabla \mathbf{N}_p)^T \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \frac{1}{p^g} \frac{\partial p^{gw}}{\partial p^c} \nabla \mathbf{N}_p d\Omega \quad (10B.5)$$

$$\mathbf{K}_{wg} = - \int_{\Omega} (\nabla \mathbf{N}_p)^T \left[ -\rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \left( \frac{1}{p^g} \frac{\partial p^{gw}}{\partial p^c} \nabla \mathbf{N}_p - \frac{p^{gw}}{(p^g)^2} \nabla \mathbf{N}_p \right) - \rho^{gw} \frac{\mathbf{k} k^{rg}}{\mu^g} \nabla \mathbf{N}_p \right] d\Omega \quad (10B.6)$$

$$\begin{aligned} \mathbf{f}^w = & - \int_{\Gamma_g^q} \mathbf{N}_p^w [q^w + q^{gw} + \beta_c (\rho^{gw} - \rho_\infty^{gw})] d\Gamma + \int_{\Omega} (\nabla \mathbf{N}_p)^T \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} \rho^w \mathbf{g} d\Omega \\ & + \int_{\Omega} (\nabla \mathbf{N}_p)^T \rho^{gw} \frac{\mathbf{k} k^{rg}}{\mu^g} \rho^g \mathbf{g} d\Omega \end{aligned} \quad (10B.7)$$

$$\mathbf{C}_{gs} = \int_{\Omega} \mathbf{N}_p^T \alpha S_g \rho^{ga} \mathbf{m}^T \mathbf{L} \mathbf{N}_u d\Omega \quad (10B.8)$$

$$\begin{aligned} \mathbf{C}_{gw} = & \int_{\Omega} \mathbf{N}_p^T \frac{\alpha - n}{K_s} S_w S_g \rho^{ga} \mathbf{N}_p d\Omega + \int_{\Omega} \mathbf{N}_p^T \left[ \left( \frac{\alpha - n}{K_s} S_g p^c + n \right) \rho^{ga} \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \right] d\Omega \\ & + \int_{\Omega} \mathbf{N}_p^T S_g n \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \mathbf{N}_p d\Omega \end{aligned} \quad (10B.9)$$

$$\begin{aligned} \mathbf{P}_{gg} = & \int_{\Omega} \mathbf{N}_p^T \frac{\alpha - n}{K_s} S_g^2 \rho^{ga} \mathbf{N}_p d\Omega - \int_{\Omega} \mathbf{N}_p^T \left[ \left( \frac{\alpha - n}{K_s} S_g p^c + n \right) \rho^{ga} \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \right] d\Omega \\ & + \int_{\Omega} \mathbf{N}_p^T S_g n \frac{M_a}{\theta R} \mathbf{N}_p d\Omega - \int_{\Omega} \mathbf{N}_p^T S_g n \frac{M_w}{\theta R} \frac{\partial p^{gw}}{\partial p^c} \mathbf{N}_p d\Omega \end{aligned} \quad (10B.10)$$

$$\begin{aligned} \mathbf{C}_{gt} = & - \int_{\Omega} \mathbf{N}_p^T \rho^{ga} \beta_{sg} \mathbf{N}_t d\Omega - \int_{\Omega} \mathbf{N}_p^T \left( \frac{\alpha - n}{K_s} S_g p^c + n \right) \rho^{ga} \frac{\partial S_w}{\partial T} \mathbf{N}_t d\Omega \\ & - \int_{\Omega} \mathbf{N}_p^T \left[ S_g n \frac{M_a}{\theta^2 R} \mathbf{N}_t + S_g n \frac{M_w}{\theta R} \left( \frac{\partial p^{gw}}{\partial T} - \frac{p^{gw}}{\theta} \right) \mathbf{N}_t \right] d\Omega \end{aligned} \quad (10B.11)$$

$$\mathbf{K}_{gw} = \int_{\Omega} (\nabla \mathbf{N}_p)^T \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \frac{1}{p^g} \frac{\partial p^{gw}}{\partial p^c} \nabla \mathbf{N}_p d\Omega \quad (10B.12)$$

$$\begin{aligned} \mathbf{H}_{gg} = & \int_{\Omega} (\nabla \mathbf{N}_p)^T \rho^{ga} \frac{\mathbf{k} k^{rg}}{\mu^g} \nabla \mathbf{N}_p d\Omega \\ & - \int_{\Omega} (\nabla \mathbf{N}_p)^T \rho^g \frac{M_a M_w}{M_g^2} \mathbf{D}_g \left( \frac{1}{p^g} \frac{\partial p^{gw}}{\partial p^c} \nabla \mathbf{N}_p - \frac{p^{gw}}{(p^g)^2} \nabla \mathbf{N}_p \right) d\Omega \end{aligned} \quad (10B.13)$$

$$\mathbf{f}^g = - \int_{\Gamma_g^q} \mathbf{N}_p^T q^{ga} d\Gamma \quad (10B.14)$$

$$\mathbf{C}_{ts} = - \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \alpha \rho^w S_w \mathbf{m}^T \mathbf{L} \mathbf{N}_u d\Omega \quad (10B.15)$$

$$\begin{aligned}\mathbf{C}_{rw} = & - \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \rho^w \left( \frac{\alpha - n}{K_s} S_w^2 + S_w \frac{n}{K_w} \right) \mathbf{N}_p d\Omega \\ & + \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \left\{ \rho^w \left( \frac{\alpha - n}{K_s} S_w p^w - \frac{\alpha - n}{K_s} S_w p^g + n \right) \frac{\partial S_w}{\partial p^c} \mathbf{N}_p \right\} d\Omega\end{aligned}\quad (10B.16)$$

$$\begin{aligned}\mathbf{C}_{rg} = & - \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \rho^w \frac{\alpha - n}{K_s} S_g S_w \mathbf{N}_p d\Omega \\ & - \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \rho^w \left( \frac{\alpha - n}{K_s} S_w p^w - \frac{\alpha - n}{K_s} S_w p^g + n \right) \frac{\partial S_w}{\partial p^c} \mathbf{N}_p d\Omega\end{aligned}\quad (10B.17)$$

$$\begin{aligned}\mathbf{P}_n = & \int_{\Omega} \mathbf{N}_t^T (\rho C_p)_{\text{eff}} \mathbf{N}_t d\Omega \\ & + \int_{\Omega} \mathbf{N}_t^T \Delta H_{vap} \left[ \beta_{sw} \mathbf{N}_t - \rho^w \left( \frac{\alpha - n}{K_s} S_w p^w - \frac{\alpha - n}{K_s} S_w p^g + n \right) \frac{\partial S_w}{\partial T} \mathbf{N}_t \right] d\Omega\end{aligned}\quad (10B.18)$$

$$\mathbf{K}_{rw} = - \int_{\Omega} (\nabla \mathbf{N}_t)^T \Delta H_{vap} \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} \nabla \mathbf{N}_p d\Omega \quad (10B.19)$$

$$\begin{aligned}\mathbf{H}_n = & \int_{\Omega} \mathbf{N}_t^T \left[ n S_w \rho^w C_p^w \frac{\mathbf{k} k^{rw}}{\mu^w} (-\nabla p^w + \rho^w \mathbf{g})^T \right] \cdot \nabla \mathbf{N}_t d\Omega \\ & + \int_{\Omega} \mathbf{N}_t^T \left[ n S_g \rho^g C_p^g \frac{\mathbf{k} k^{rg}}{\mu^g} (-\nabla p^g + \rho^g \mathbf{g})^T \right] \cdot \nabla \mathbf{N}_t d\Omega + \int_{\Omega} (\nabla \mathbf{N}_t)^T \chi_{\text{eff}} \nabla \mathbf{N}_t d\Omega\end{aligned}\quad (10B.20)$$

$$\begin{aligned}\mathbf{f}' = & - \int_{\Omega} (\nabla \mathbf{N}_t)^T \Delta H_{vap} \rho^w \frac{\mathbf{k} k^{rw}}{\mu^w} \rho^w \mathbf{g} d\Omega - \int_{\Gamma_t'} \mathbf{N}_t^T [q^T + \alpha_c (T - T_{\infty})] d\Gamma = 0\end{aligned}\quad (10B.21)$$

# 11

## Secondary Consolidation: Creep in Soils

### 11.1 INTRODUCTION

Secondary consolidation in soils can be defined as the slow, continual compression of the soil skeleton which occurs after the excess pore pressures have been substantially dissipated. The rate of this delayed compression depends on the soil skeleton.

Settlements due to secondary consolidation normally account for only a small fraction of the total settlements; they represent a small percentage of the primary consolidation settlements by compression. However, where a very soft organic clay is under compression, secondary consolidation may be much more important. In an attempt to stiffen these soils, geotechnical engineers often apply a surcharge load to the soil over a long period of time, before any construction is carried out. After construction, much of the compression which may occur during the structure's existence may be attributed to secondary consolidation, since a relatively longer period of time will then be required to dissipate excess pore pressures.

The phenomenon of secondary consolidation was observed as early as 1925 by Terzaghi [1]. Since then several authors have investigated its occurrence. In 1936 Buisman established a logarithmic relationship to deal with consolidation without seepage resistance [2]. His empirical relationship can be written

$$\frac{\rho}{h\Delta\sigma} = \alpha_0 + \alpha_s \log_{10} t \quad (11.1)$$

in which  $\rho$  denotes the settlement,  $h$  the original height of the specimen,  $\Delta\sigma$  the pressure increment and  $t$  time.  $\alpha_0$  and  $\alpha_s$  are constants given by the graph of observed settlements versus the logarithm of time.

Mitchell and Singh [3] developed a creep formulation for soils from a series of experiments; it gave a linear relationship between strain and the logarithm of time, which included the influence of the deviatoric stress on the creep rate. Their relationship is

$$\varepsilon = \varepsilon_1 + A e^{\gamma D} \ln t \quad (11.2)$$

in which  $D$  is a function of the deviatoric stress,  $t$  is time, and  $\varepsilon_1$ ,  $A$  and  $\gamma$  are constants determined from the experimental data.

Equations (11.1) and (11.2) are purely empirical, so their extrapolation for a real soil would prove extremely difficult. Furthermore they are inappropriate for incorporation into a finite element model or similar numerical work. However, representation of secondary consolidation or creep by a viscous model could provide a simple but effective model. The following sections discuss such a rheological model and derive a creep formulation which can be incorporated into any standard finite element program.

## 11.2 FORMULATION OF SECONDARY CONSOLIDATION

For this formulation we assume a Gibson–Lo soil skeleton that consists of a non-linear Hookean spring in series with a Kelvin body (Figure 11.1). The Hookean spring is representative of any of the variable-elastic models which were discussed in Section 4.4.

The stress,  $\sigma_{ij}$ , is the same in both bodies and the total strain is given by

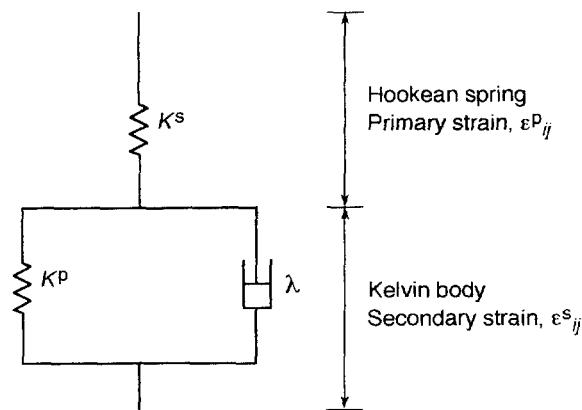
$$\varepsilon_{ij} = \varepsilon_{ij}^p + \varepsilon_{ij}^s \quad (11.3)$$

where p represents primary and s secondary.

When a load is applied suddenly, the elements of the Kelvin body exhibit no instantaneous deformation. If the elements are linear, the deformation gradually increases exponentially with time. Initially, all the load is carried by the dashpot, but is transferred at a decreasing rate to the spring, stiffness  $K^s$ , until after a finite time the spring eventually carries the entire load.

Two further assumptions are made in this formulation. Firstly, the Poisson's ratio for creep strains is assumed constant throughout the secondary consolidation process, and secondly the process is assumed to be independent of temperature.

For the completely general three-dimensional situation, it is convenient to separate the secondary strains  $\varepsilon_{ij}^s$  and the effective stress  $\sigma'_{ij}$  into isotropic and deviatoric components. If we now denote the isotropic component by 0 and the deviatoric



**Figure 11.1** Rheological model for creep

component by 1, and omit the prime for clarity, we have

$$\varepsilon_{ij}^s = \varepsilon_{ij}^{s0} + \varepsilon_{ij}^{s1} \quad (11.4a)$$

$$\sigma_{ij} = \sigma_{ij}^0 + \sigma_{ij}^1 \quad (11.4b)$$

where

$$\varepsilon_{ij}^{s0} = \sum_{k=1}^3 \frac{1}{3} \varepsilon_{kk}^s \delta_{ij} \quad (11.5a)$$

and

$$\sigma_{ij}^0 = \sum_{k=1}^3 \frac{1}{3} \sigma_{kk} \delta_{ij} \quad (11.5b)$$

The stress-strain relationship for the Kelvin body, in terms of the two components, is now given by

$$\sigma_{ij}^0 = 3K_b^s \varepsilon_{ij}^{s0} + 3\lambda_b \dot{\varepsilon}_{ij}^{s0} \quad (11.6a)$$

$$\sigma_{ij}^1 = 2G^s \varepsilon_{ij}^{s1} + 2\lambda_d \dot{\varepsilon}_{ij}^{s1} \quad (11.6b)$$

in which  $K_b^s$  is the bulk modulus for creep,  $G^s$  is the shear modulus for creep and  $\lambda_b$  and  $\lambda_d$  are damping coefficients.  $K_b^s$  and  $G^s$  both remain constant and · represents the rate of change in the strains with time.

For the two-dimensional situation, equation (11.6) can be rewritten in matrix form as

$$\boldsymbol{\sigma} = \mathbf{D}^s \boldsymbol{\varepsilon}^s + \mathbf{D}_\lambda \dot{\boldsymbol{\varepsilon}}^s \quad (11.7)$$

or as

$$\begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{Bmatrix} = \mathbf{D}_s \begin{Bmatrix} \varepsilon_{xx}^s \\ \varepsilon_{yy}^s \\ \gamma_{xy}^s \end{Bmatrix} + \mathbf{D}_\lambda \begin{Bmatrix} \dot{\varepsilon}_{xx}^s \\ \dot{\varepsilon}_{yy}^s \\ \dot{\varepsilon}_{xy}^s \end{Bmatrix} \quad (11.8)$$

in which

$$\mathbf{D}_s = \begin{bmatrix} \left(K_b^s + \frac{4}{3}G^s\right) & \left(K_b^s - \frac{2}{3}\right) & 0 \\ & \left(K_b^s + \frac{4}{3}G^s\right) & 0 \\ \text{Symmetry} & & G^s \end{bmatrix} \quad (11.9a)$$

$$\mathbf{D}_\lambda = \begin{bmatrix} \left(\lambda_b + \frac{4}{3}\lambda_s\right) & \left(\lambda_b - \frac{2}{3}\lambda_d\right) & 0 \\ & \left(\lambda_b + \frac{4}{3}\lambda_d\right) & 0 \\ \text{Symmetry} & & \lambda_d \end{bmatrix} \quad (11.9b)$$

The incremental stress-strain relationship for the whole system, i.e. Hookean spring and Kelvin body, is now written as

$$\Delta\boldsymbol{\sigma} = \mathbf{D}^p \Delta\boldsymbol{\varepsilon}^p = \mathbf{D}^p (\Delta\boldsymbol{\varepsilon} - \Delta\boldsymbol{\varepsilon}^s) \quad (11.10)$$

in which  $\mathbf{D}^p$  is the stress-strain matrix for the Hookean spring. By assuming an explicit formulation,  $\Delta\boldsymbol{\varepsilon}^s$  can be written as

$$\Delta\boldsymbol{\varepsilon}^s = \dot{\boldsymbol{\varepsilon}}^s \Delta t \quad (11.11)$$

and if we now substitute (11.11) into (11.7) we have

$$\Delta\boldsymbol{\varepsilon}^s = \mathbf{D}_\lambda^{-1} (\boldsymbol{\sigma} - \mathbf{D}^s \boldsymbol{\varepsilon}^s) \Delta t \quad (11.12)$$

If we now introduce the average secondary strain as

$$\bar{\boldsymbol{\varepsilon}}^s = (1 - \theta) \boldsymbol{\varepsilon}_t^s + \theta \boldsymbol{\varepsilon}_{t+\Delta t}^s \quad (11.13)$$

and use this value in (11.12), then the iteration on  $\boldsymbol{\varepsilon}_{t+\Delta t}^s$  at time  $t + \Delta t$  proves more convenient.

Therefore, if we assume that during a short time interval  $\Delta t$  both the stresses and the material properties remain unchanged, then the increment  $\Delta\boldsymbol{\varepsilon}^s$  in the creep strain can be calculated for the end of that time step from equation (11.12). The total creep strain vector  $\boldsymbol{\varepsilon}^s$  and the total (effective) strain vector  $\boldsymbol{\sigma}'$  can be accumulated after every step.

The incremental 'force' due to creep strains can be determined for use in the right-hand force vector from

$$\Delta\mathbf{C} = \int_{\Omega} \mathbf{B}^T \mathbf{D}_T d\boldsymbol{\varepsilon}^s d\Omega \quad (11.14)$$

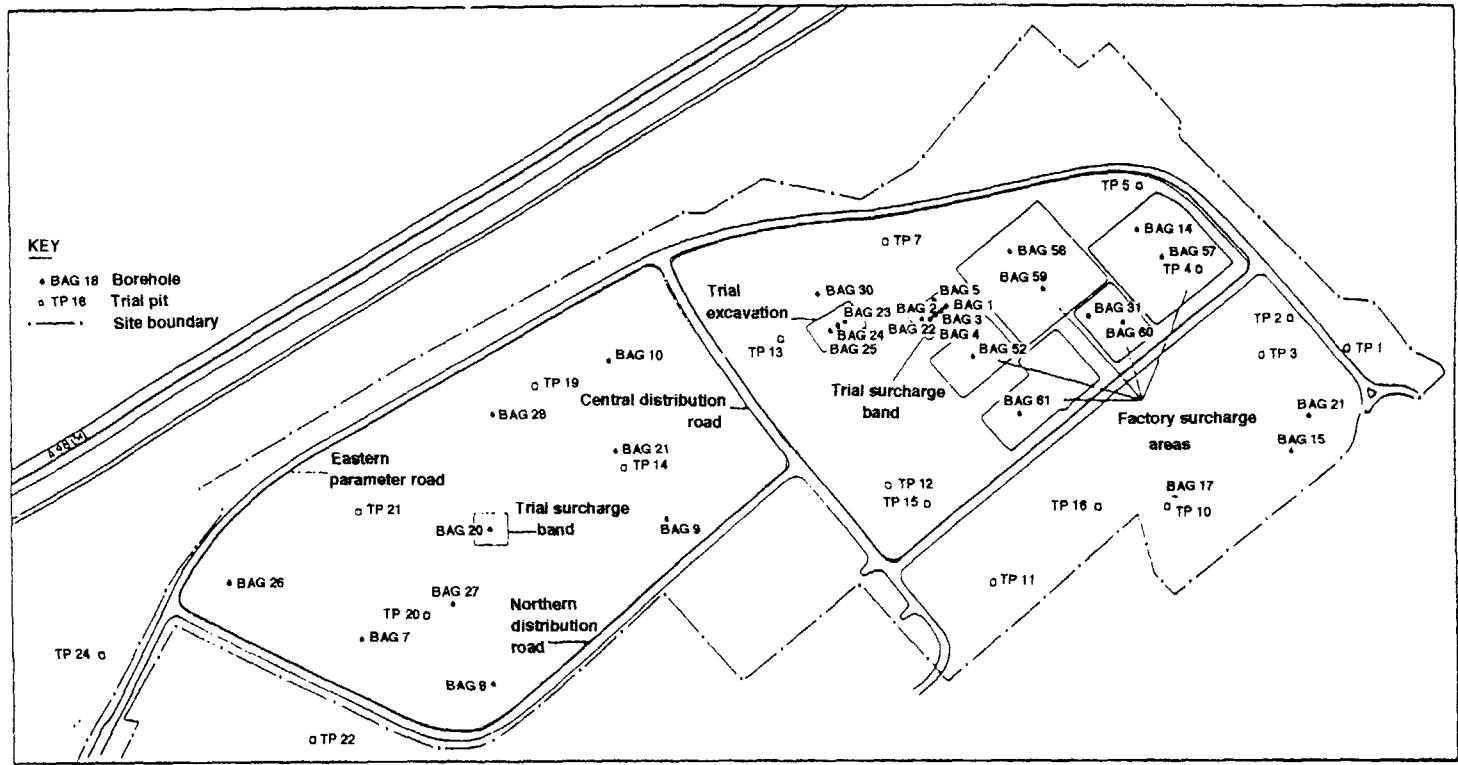
The two assumptions made earlier on the independence of secondary consolidation with temperature and material properties independent of time may clearly be a source of error since, in general, creep deformations are temperature dependent and creep parameters time dependent. Even so, temperature can easily be accounted for by expressing the parameters as functions of temperature, and in situations where temperature dependency is not important, a series of creep tests at constant stress and temperature will yield time-dependent parameters by curve-fitting procedures.

The creep model detailed here was incorporated into the hyperbolic stress-strain model outlined in Section 4.4.2.1. In the following section this program has been used to analyse a test embankment problem and the finite element results are compared with the available experimental results.

## 11.3 APPLICATION OF THE CREEP MODEL

### 11.3.1 Background

A practical problem which was strongly concerned with secondary consolidation arose at the Baglan Moors industrial estate, Swansea. A site investigation was designed and carried out in two stages. The first stage involved the determination of the nature and



**Figure 11.2** Schematic representation of the Baglan industrial park to show positions of boreholes, trial pits, trial excavation and two trial embankments: the embankment analysed by finite element analysis contains boreholes 1 to 5 and borehole 22

properties of the soils. A series of boreholes across the area (Figure 11.2) revealed that the soil profile was quite complex and tests showed that soft organic clays, silt and peat were highly compressible.

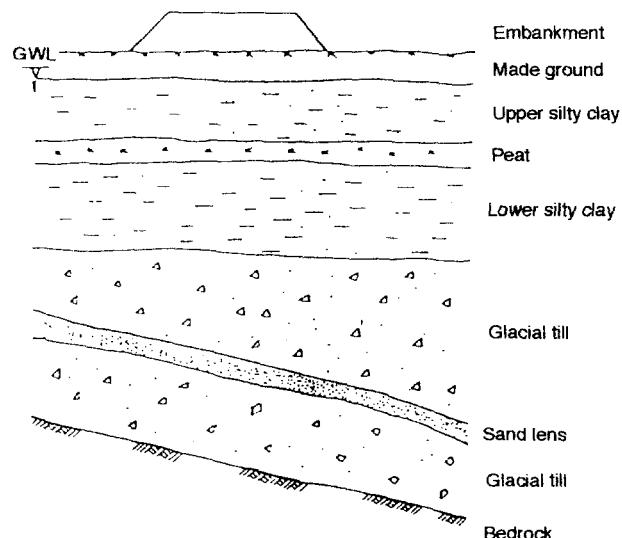
A second stage involved the construction of two trial embankments to a height of 3 m above a general fill level. This took place in November 1980 and field measurements of pore pressures and vertical displacements were recorded over a period of three years. Furthermore, long-term secondary consolidation of the peat was measured by laboratory creep tests.

It was proposed that a finite element analysis should be undertaken to simulate consolidation of the compressible soils under one of these embankments over a long period. The analysis would also involve an assessment of the laboratory test data and the pore pressure and settlement field measurements.

Boreholes around the test embankment revealed a layered system of different soil deposits. The stratigraphy is shown in Figure 11.3 and typical thickness of the strata below the made ground as follows:

Soft clay and silt	4.5 m
Fibrous peat	2.0 m
Soft to firm clay and silt	7.0 m
Dense glacial till, sands and gravel	13.0 m

Therefore, the ground consisted of 2 m of made ground, below which lay almost 14 m of highly compressible soil and 13 m of dense glacial till lying on the bedrock. Consolidation tests were carried out on samples from each layer and provided data on the physical properties of the soils (Table 11.1).



**Figure 11.3** Stratigraphy below the test embankment on the Baglan Moors industrial estate (not to scale)

**Table 11.1** Summary of laboratory test results

Materials	$W(\%)$	Bulk density ( $Mg\ m^{-3}$ )	$C_M$ ( $kNm^{-2}$ )	$C_v$ ( $m^2$ per year)	$M_v$ ( $m^3\ MN^{-1}$ )
Upper organic silty clay	49	1.72	19	1.08	0.86
Peat	371	0.99	32	$\geq 10$	1.98
Lower organic silty clay	49	1.72	14	2.03	0.50
			$C = 10$		
Sand lenses	34	1.87	$\phi = 30$	11	0.32
Glacial till	11.6	2.18	159(?)	17.5	0.09

### 11.3.2 Mathematical Modelling

The hyperbolic stress-strain model, incorporating the secondary consolidation and the elastoplastic model was used to analyse this particular problem, assuming plane strain and effective stress situations. Due to a lack of laboratory test data, some parameters still remained undetermined. Since an effective stress analysis is to be carried out, these parameters are equivalent to the parameters from drained tests.

The coefficient of consolidation and the coefficient of compressibility are often determined from the oedometer test. This gives the following relationship between the two parameters:

$$C_v = \frac{k}{\gamma_w M_v} \quad (11.15)$$

where  $C_v$  = Coefficient of consolidation ( $m^2$  per year)

$M_v$  = Coefficient of compressibility ( $m^2\ kN^{-1}$ )

$\gamma_w$  = Unit weight of water ( $kN\ m^{-3}$ )

$k$  = Coefficient of permeability (m per year)

As values of  $C_v$  and  $M_v$  are available for each soil (Table 11.1), then values for  $k$  can be computed from (11.15). Non-linear permeability has been proposed by several authors [4–6] and is adopted in Chapter 4, where the coefficient of permeability changes with the void ratio [7]. But this analysis assumes a constant coefficient of permeability, calculated from (11.5).

For a one-dimensional test, the constrained modulus is given by

$$D = \frac{1}{M_v} = \frac{E'(1 - v')}{(1 + v')(1 - 2v')} \quad (11.16)$$

where  $E$  = Young's Modulus

$v$  = Poisson's ratio

A prime denotes an effective stress parameter. Therefore, from (11.15) and (11.16), we can write

$$\frac{C_v \gamma_w}{k} = \frac{E'(1 - v')}{(1 + v')(1 - 2v')}$$

**Table 11.2** Summary of computed parameters

Materials	$E'$ (kN m <sup>-2</sup> )	$k$ (m per year)	$S(m^2 kN^{-1})$	Submerged unit weight (kN m <sup>-3</sup> )
Upper organic silty clay	863.8	$9.1 \times 10^{-3}$	$5 \times 10^{-4}$	7.06
Peat	375.4	$197 \times 10^{-3}$	$1 \times 10^{-2}$	1.77
Lower organic silty clay	1 385.7	$9.9 \times 10^{-3}$	$5 \times 10^{-4}$	7.06
Sand lenses	3 321.14	$34.5 \times 10^{-3}$	$1 \times 10^{-6}$	8.53
Glacial till	8 253.97	$15.5 \times 10^{-3}$	$1 \times 10^{-6}$	115.7

For normally consolidated soils, Lambe and Whitman [8] have found that the effective Poisson's ratio varies between 0.25 and 0.35. A typical value of 0.3 is adopted here for all soils and remains constant.

Due to lack of data, the bulk moduli of the soil skeleton and pore fluid (input in the program as compressibility factor  $S$ ) proved the most difficult parameters to determine. However, Naylor *et al.* [9] give a pore fluid modulus of 2200 M Pa for water. After several trial runs, a range of values of the compressibility factor between  $10^{-3}$  and  $10^{-4} m^2 kN^{-1}$  were found to be reasonable for undrained conditions. As the time step length increased, this became a less important problem, as  $(\theta H \Delta t)$  in Equation (3.50) is quite large in comparison with  $S$ . Table 11.2 summarises the values for these calculated parameters. The unit weight of water  $\gamma_w$  is assumed to be constant at  $10 \text{ kN m}^{-3}$ .

The dimensions of the test embankment were width 20 m, length 40 m and height 3 m. A plane-strain situation was considered most appropriate.

Although the laboratory report showed the soil's behaviour as highly non-linear, an initial analysis used the linear elastic model to provide useful first results. Later analyses used the modified hyperbolic stress-strain model to provide information on the secondary consolidation and an elastoplastic model to provide further comparison with the field measurements.

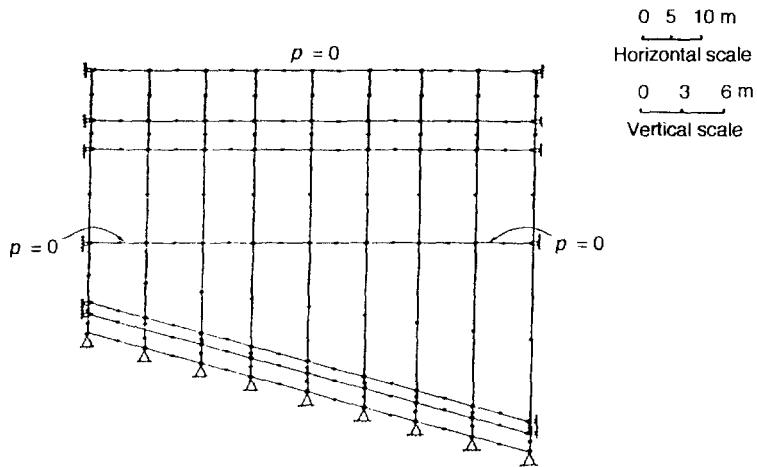
The finite element mesh which was adopted extended from below the made ground (which was considered as part of the loading, together with the embankment) down to the bedrock and up to 20 m beyond each edge of the embankment. Eight-noded isoparametric elements were used in the mesh, which comprised 48 elements and 173 nodes (Figure 11.4).

A free drainage surface is assumed to exist at the upper level of the mesh and along the upper surface of the glacial till. Hence excess pore water pressures are prescribed as zero along these two lines.

At the surface of the bedrock, we assume that no horizontal displacements take place and the rock is sufficiently rigid to prevent vertical displacements. At the vertical bounds of the mesh, we assume that the nodes are sufficiently distant from the embankment to prevent horizontal displacements.

It was intended to compare the finite element results with the field measurements for a period of up to one year. The first few time steps were taken as 0.0001 year and gradually increased to 0.2 year, giving a total of 18 time steps.

From the position of the water table shown in Figure 11.3, the soils below the embankment were assumed to be fully saturated with pore water. The body force vector



**Figure 11.4** Finite element mesh used in the test problem for the Baglan Moors industrial estate showing nodal positions and boundary conditions

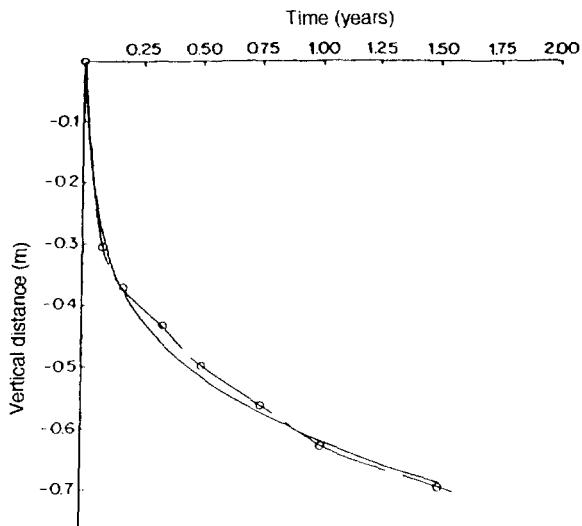
is therefore computed from the submerged unit weight of each soil and is only calculated in the first time step. The vertical effective stresses are considered to be due to an overburden and calculated from the submerged unit weight of the soils at any particular depth. The horizontal component can then be calculated by using the coefficient of earth pressure at rest  $K_0$ . A value for  $K_0$  between 0.8 and 1.0 was used, which might be considered a little high for normally consolidated soils [8].

From a load diagram provided by the field data, the time of construction is very small (less than one week) in comparison with the consolidation time. Furthermore, interest is concentrated on soils directly beneath the embankment, so a single lift analysis was assumed to be adequate.

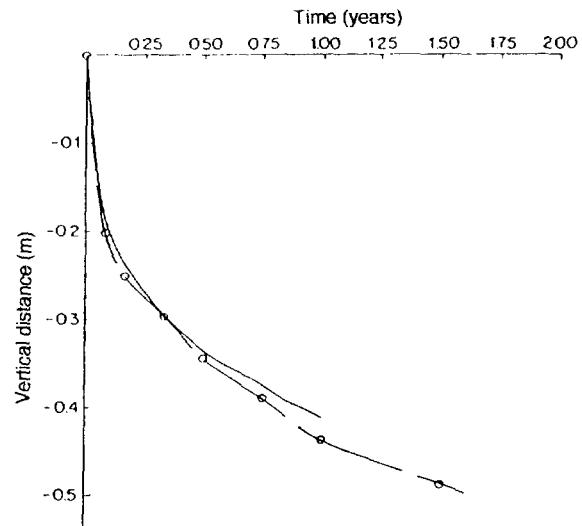
### 11.3.3 Results

As expected, the initial analysis, using the linear elastic models, proved to be the cheapest in terms of computer time and data preparation. But the results emphasised the high non-linearity of the soils. For example, in the worst case, the displacements were underestimated by as much as 40 percent. The model did provide some useful guides for the use of the hyperbolic model with the secondary consolidation model incorporated.

The results from the hyperbolic model were much better than the results from the initial linear elastic analysis, and Figures 11.5 to 11.9 show that very good agreement was achieved between these results and the field measurements for the period up to one year. In order to use this model in the analysis, the value of  $n$  in equation (4.17) was assumed to equal 0.2 for clay and 0.5 for the glacial till and sand. These values were recommended by Duncan and Chang [10] for use under the loaded area. A disadvantage of this model is that it required repeated trial runs to adjust the uncertain parameters



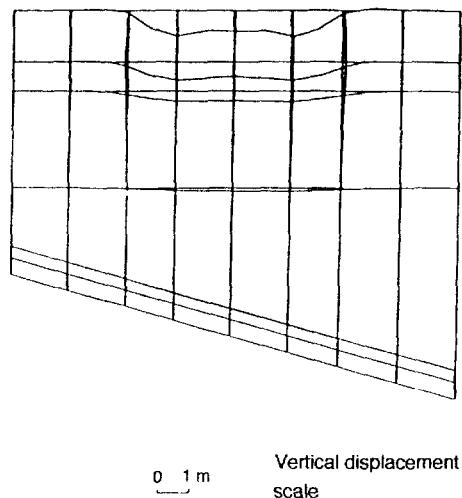
**Figure 11.5** Vertical displacement at the top of the upper silt clay for borehole BAG 2: (—) field results, (---) finite element results



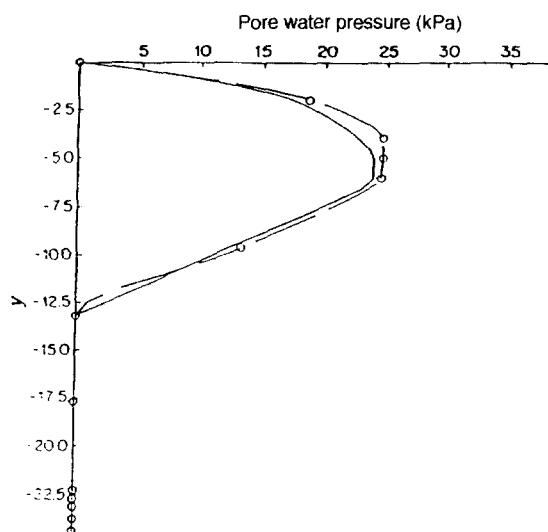
**Figure 11.6** Vertical displacement at the top of the peat for borehole BAG 2: (—) field results, (---) finite element results

which were not provided directly from field data. This aspect may give the inexperienced analyst some unwelcome difficulties. The model also proved more costly in computer time as it increased the number of iterations required for convergence in each time step.

Finally, the elastoplastic model effectively vindicated the secondary consolidation model. A Mohr–Coulomb yield criterion was used together with the isotropic hardening

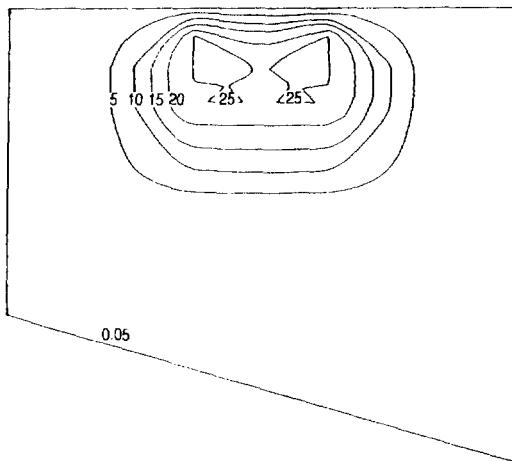


**Figure 11.7** Mesh deformation beneath the test embankment at time  $t = 1$  year (the geometric scale is as in Fig. 11.4)



**Figure 11.8** Pore pressure distribution along borehole BAG 2 at time  $t = 1$  year. (—) field results, (○) finite element results

rule. This required parameters  $C$  and  $\phi$ , and fortunately they were provided in the field data. The model appeared to be stable, exhibited rapid convergence and yielded accurate results. The fact it required fewer parameters, which were readily available to a high degree of certainty, may have been important in this problem.



**Figure 11.9** Pore pressure contours (kPa) within the finite element mesh boundaries at time  $t = 1$  year

## REFERENCES

1. Terzaghi, K. (1943) *Theoretical Soil Mechanics*, Wiley, New York.
2. Buisman, A. S. K. (1936) Results of long duration settlement tests, in *Proceedings of an International Conference on Soil Mechanics and Foundation Engineering*, Harvard University, Cambridge MA, vol. 1, pp. 103–6.
3. Mitchell, J. K. and Singh, A. (1968) General stress-strain-time function for soils. *J. Soil Mech. Found. Div. ASCE*, **94**, 21–46.
4. Gambolati, G. (1973) Equation for one-dimensional vertical flow of ground water, 2. Validity range of the diffusion equation. *Water Resources Research* **9**(5), 1385–95.
5. Finol, A. and Farouq-Ali, S. M. (1975) Numerical simulation of oil production with simultaneous ground subsidence. *SPE J.*, **15**, 411–24.
6. Monte, J. L. and Kritzen, R. J. (1976) One-dimensional mathematical model for large-strain consolidation. *Géotechnique*, **26** (3), 495–510.
7. Lewis, R. W., Roberts, G. K. and Zienkiewicz, O. C. (1976) A non-linear flow and deformation analysis of consolidation problems, in *Proceedings of the 2nd International Conference on Numerical Methods in Geomechanics*, ASCE, New York, pp. 1106–18.
8. Lambe, T. W. and Whitman, R. V. (1969) *Soil Mechanics*, Wiley, New York.
9. Naylor, D. J., Pande, G. N., Simpson, B. and Tabb, R. (1978) *Finite Elements in Geotechnical Engineering*, Pineridge Press, Swansea.
10. Duncan, J. M. and Chang, C. Y. (1970) Non-linear analyses of stress-strain in soils. *Proc. ASCE*, **96**, 1629–53.

# 12

## Soil–Structure Interaction

### 12.1 INTRODUCTION

The consolidation analysis of soils has long been recognised as an important problem in civil engineering design. However, most of the consolidation problems which have been analysed are limited by either assuming simple geometries or linear elastic constitutive behaviour, and often they are concerned only with the interests of geotechnical engineers. The influence of the interaction between the structure and the foundation are not effectively taken into account in these analyses.

There are two main reasons for omitting the interaction effects. First is the uncertainty of the load transfer and compatibility conditions at the interfaces between the soil mass and the superstructure. Second is the mathematical complexity which inevitably arises from a rigorous interaction analysis.

Some attempts have recently been made to solve these difficulties. Approximate solutions are available, albeit for very limited cases where simplifying assumptions are necessary. These interaction analyses are often carried out for important structures such as nuclear power plants subjected to earthquakes or offshore structures under severe wave loading [9,12]. The recent expansion of supercomputers and the well-developed finite element method have reduced many of the difficulties, even in the most complex situations. Interface elements can be used to study the behaviour of the contact boundary between the structure and the soil, and appropriate constitutive models are readily available to simulate the non-linearities of the materials involved.

Generally, there are two methods of analysing soil–structure interaction problems: the method of superposition and a true time-dependent interaction analysis. The superposition method first predicts the final shape of the soil profile then calculates the interaction of the superimposed structural members with the distorted profile. This method is simpler than the second technique and is cost-effective, but assumptions about rough or perfectly smooth interfacial contact can cause problems. Also, in propagating solutions in time, these interface boundaries with the originally assumed contact are no longer realistic. The second method marches forward in time and includes the interaction effects at every time step. This method is more costly than the superposition method in terms of computer memory storage. However, its use can be justified on the grounds of a better physical simulation and more confident interpretation of the interaction effects.

This chapter focuses on the method involving a true time interaction analysis during the consolidation of the soil mass. The inertia effects are omitted due to the slow

dissipation of excess pore water pressure. Other complexities, such as dynamic loading, will not be dealt with here and the interested reader should refer to the more general approaches as given in Chapters 2 and 14.

## 12.2 GOVERNING EQUATIONS

When a superstructure is supported by a compressible non-linear soil, the following factors are significant in deciding its behaviour, utility and reliability:

1. The relative settlements of the foundations.
2. The non-linear stress-strain behaviour of the soil.
3. The relative movements between the soil and structure at the interfaces.
4. The dissipation of pore pressure due to the pore fluid flowing out of the soil naturally or artificially.
5. The interaction between soil and structure leading to an eventual redistributions of stresses in the soil and the structural domain.

These factors contribute to making the structure and the soil behave in a totally different manner from the ideal state, where both are considered separately. A solution procedure needs to be developed in which the superstructure and the supporting medium are treated as parts of a single system.

A comprehensive formulation suitable for solution by the finite element method has been derived by several authors. Sandhu and Wilson [19] formulated the problem using the finite element idealisation in conjunction with the variational principle and this has been used by many subsequent researchers. A more general formulation was later proposed by Zienkiewicz *et al.* [17] and Lewis and Schrefler [20,21] to incorporate various non-linearities of the soil and the structural properties. It forms the basis for the finite element analysis in this chapter. The derivation of the formulation is well documented in the literature [5–7] and in Chapter 2. However, a brief discussion is presented here for the sake of completeness.

The formulation incorporates the total stress equilibrium and strain compatibility requirements during the soil consolidation process. The pore fluid effects are included in the equilibrium equation using the principle of effective stress:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' - \mathbf{m}\mathbf{p} \quad (12.1)$$

where  $\boldsymbol{\sigma}$  is the total stress vector,  $\boldsymbol{\sigma}'$  is the effective stress vector,  $\mathbf{m} = [1 \ 1 \ 1 \ 0 \ 0]^T$  and  $\mathbf{p}$  is the pore pressure.

The equilibrium of the total stresses, derived using the principle of virtual work in incremental form, is as follows:

$$\mathbf{K}_t \frac{d\ddot{\mathbf{u}}}{dt} - \mathbf{Q} \frac{d\mathbf{p}}{dt} = \frac{d\mathbf{f}^u}{dt} + \mathbf{C} \quad (12.2)$$

The continuity equation of pore fluid flow, derived using the compatibility condition of

volume change, is given by

$$\mathbf{H}\bar{\mathbf{p}} + \mathbf{S}\frac{d\mathbf{p}}{dt} + \mathbf{Q}^T\frac{d\bar{\mathbf{u}}}{dt} = \mathbf{f}^p \quad (12.3)$$

where  $\bar{\mathbf{u}}$  and  $\bar{\mathbf{p}}$  are the nodal displacement and pore fluid pressures, respectively,  $\mathbf{K}_t$  is the stiffness matrix of the solid phase,  $\mathbf{Q}$  is the coupling matrix between the solid and fluid phases,  $\mathbf{f}^u$  is the external force vector including the initial effects such as the initial stress in soil mass or the initial strain in the structure,  $\mathbf{C}$  is the force vector due to creep effects of the solid phase (Chapter 11),  $\mathbf{H}$  is the permeability matrix of the soil,  $\mathbf{S}$  is the compressibility matrix of the soil skeleton,  $\mathbf{f}^p$  is the outflow (or inflow) of pore fluid from the soil and  $t$  is the time.

Governing equations (12.2) and (12.3) can be solved simultaneously with a set of initial and boundary conditions:

*Initial conditions:* any prescribed  $\mathbf{u} = \mathbf{u}_p$  or  $\mathbf{p} = \mathbf{p}_p$ .

*Boundary conditions:* any imposed external forces on the boundary  $\Gamma_f$  and the continuity of pore fluid across the boundary  $\Gamma_q$ , which is

$$\rho^w \frac{\mathbf{k}}{\mu^w} (-\text{grad } p^w + \rho^w \mathbf{g})^T \cdot \mathbf{n} = q^w \quad (12.4)$$

where  $\mathbf{n}$  is the normal unit vector to the surface,  $\mathbf{k}$  is the permeability coefficient of the soil,  $\rho_w$  is the density of pore fluid,  $\mu^w$  is the dynamic viscosity and  $q^w$  is the outflow (or inflow) of pore fluid across the boundary.

Governing equations (12.2) and (12.3) can be rewritten and integrated in time using the well-known  $\theta$  method, which leads to the following equation:

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \left\{ (1 - \theta) \begin{bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}} \end{bmatrix}_{t_0} + \theta \begin{bmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{p}} \end{bmatrix}_{t_1} \right\} + \begin{bmatrix} \mathbf{K}_t & -\mathbf{Q} \\ \mathbf{Q}^T & \mathbf{S} \end{bmatrix} \begin{Bmatrix} \frac{\bar{\mathbf{u}}_{t_1} - \bar{\mathbf{u}}_{t_0}}{\Delta t} \\ \frac{\bar{\mathbf{p}}_{t_1} - \bar{\mathbf{p}}_{t_0}}{\Delta t} \end{Bmatrix} = \begin{Bmatrix} \frac{d\mathbf{f}^u}{dt} + \mathbf{C} \\ \mathbf{f}^p \end{Bmatrix} \quad (12.5)$$

where  $t_0$  and  $t_1$  are the time step levels and  $\Delta t$  is the time step.

It is easily seen that  $\theta = \frac{1}{2}$  is equivalent to the Crank–Nicolson scheme and  $\theta = 1$  leads to a fully implicit scheme. When  $\frac{1}{2} \leq \theta \leq 1$  the solution procedure is unconditionally stable [15].

Direct solvers, such as the profile or frontal solver, can be used to solve system (12.5) as is usually the case in many finite element applications. However, large computer memory storage and an excessive amount of computer time may result when dealing with a complicated three-dimensional problem. A staggered iteration scheme has been suggested as a means of circumventing this problem [4,16,3].

In the staggered iteration scheme, equations (12.2) and (12.3) are solved separately. A prediction of either the displacement field or the pore fluid pressure is required for each time step. The ensuing coefficient matrices remain symmetric and a consequent reduction in computer storage is achieved. The storage and efficiency may be improved further by partitioning each field using explicit–implicit elements [24].

The staggered scheme has been investigated in detail by Tran [7]. The stability condition for the scheme is  $\theta \geq \frac{1}{2}$ . However, the numerical stability of the partitioned equations also depends on the predictor formula at each time step, and the prediction of the pore fluid pressure is not always accurate, particularly for two- or three-dimensional problems under ‘undrained’ situations, i.e. the first time-step solution. Such inaccurate prediction of pore pressure could lead to fairly fast divergence of the solution scheme.

## 12.3 MATERIAL MODELS

Appropriate constitutive models for both the soil and structure must be chosen for any finite element interaction analysis. Of the two materials, the choice of the soil model is the more difficult task. Depending on the problem under consideration, care must also be taken in selecting suitable parameters for the structural material and the interface boundary. In most soil–structure interaction problems, the structure may be assumed to behave in a linear elastic manner. However, in the case of concrete this assumption must be carefully considered because of the possibility of long-term creep or cracking of the concrete. This is particularly important when considering the interaction problem during a slow process such as soil consolidation. A reduced modulus, in accordance with data obtained from creep tests, has been employed in the past to allow greater structural deflections. A Kelvin rheological model has been tried for simulating this long-term creep behaviour of the structural material [7].

### 12.3.1 Interface behaviour

In the conventional finite element method, the displacements along the boundary between adjacent elements should be compatible, i.e. there are no relative displacements. But in reality there are relative displacements between soil and structure, and they play an important role in the interaction analysis. Problems may arise at the interface for the following reasons:

1. The different stress–strain characteristics of the soil and structure.
2. The difficulty of interpreting the resultant stresses at the contact surface between the two media. This is because, at the beginning of the soil consolidation process, the stresses in the structure are much higher than the effective stresses in the surrounding soil.
3. The need to obtain the contact pressure explicitly. This is because it helps to explain the load transfer mechanism between the two media.
4. The need to simulate the relative movements between the soil and the structure. Slippage, or detachment, can occur where high tensions exist at the contact surface.
5. The problem of prescribing the interface boundary conditions.

All these problems can be overcome fairly easily by the use of interface elements. It has been confirmed by several authors [1,5] that, in the finite element analysis of a soil–structure interaction problem, proper estimates of the differential settlement and stress

distribution can be obtained using the integrated approach, which treats the superstructure and soil as a single continuous system with interface elements between both media. Any relative displacement, or slippage, between the foundation and soil affects the stress distribution in both. Relative displacements depend on the types of contact surface and the intensity of the normal and shear stresses at the interface.

A straight interface element with linear interpolation of slip was first developed by Goodman *et al.* [18]. This element has the disadvantage of producing an ill-conditioned stiffness matrix which leads to inaccurate solutions. Ghaboussi and Wilson's [13] interface element has the relative displacement as the unknown, hence it provides more accurate solutions. A six-noded interface element with relative displacement appears to be the best choice for most types of geotechnical problem [7].

The nature of the interface behaviour depends on the roughness of the materials involved in both the soil and the structure. Constitutive models for these interface elements can be determined from interface tests, such as a direct shear test, between the soil and the structure. The stress-strain relationship is expressed as follows:

$$\begin{Bmatrix} \sigma_n \\ \tau \end{Bmatrix} = \mathbf{D} \begin{Bmatrix} \varepsilon_n \\ \varepsilon_s \end{Bmatrix} \quad (12.6)$$

where for a non-dilatancy interface

$$\mathbf{D} = \begin{bmatrix} K_{nn} & 0 \\ 0 & K_{ss} \end{bmatrix} \quad (12.7)$$

$K_{nn}$  and  $K_{ss}$  are constants for a linear elastic analysis, or non-linear functions for a non-linear analysis.

1.  $K_{nn} = K_{ss} \simeq 0$  when separation occurs at the interface.
2.  $K_{nn} = E_f$  has a very high value when contact occurs; typically  $E_f = 10^6$ .

In the tangential direction, the stress-strain relationship can be assumed to be elastic-perfectly plastic using a Mohr-Coulomb yield criterion:

1.  $K_{ss} = G$  for stress  $\tau < c + \sigma_n \tan \phi$
2.  $K_{ss} \simeq 0$  for shear stress  $\tau \geq c + \sigma_n \tan \phi$

where  $c$  and  $\phi$  are the adhesion and friction angle respectively.

A hyperbolic relationship for the shearing behaviour of the interface has been successfully used by Tran [7] in the solution of a typical pile problem.

### 12.3.2 Soil behaviour

It is a difficult task to select a realistic constitutive model which fully represents the behaviour of the soil in the consolidation process. This is because the soil mass consists of soil grains and pore fluid, and they are coupled during consolidation. In the early applications of the finite element method to soil consolidation problems, the soil

skeleton was assumed to be homogeneous, isotropic and linear elastic. Sandhu and Wilson [19] were the first to study linear elastic soil consolidation by the finite element method. This model can give reasonable solutions for a single load path. However, in practical problems involving local failure of the soil, variable soil strength with depth and non-recoverable strains upon unloading, it is more appropriate to use either a variable elastic approach or an elastoplastic constitutive relationship. Many non-linear models have been developed in the soil mechanics context, e.g. variable elastic models such as the K–G model or the hyperbolic model [14] (See Chapter 4).

The simplicity of these models led to their application in many practical problems. However, they have certain drawbacks in that they do not predict the plastic deformation upon unloading, nor do they obey a flow rule which indicates the direction of the plastic strain. Hence, elastoplastic and/or elasto-viscoplastic models were used to overcome these shortcomings. Different yield criteria can be incorporated easily into these plastic models. Naylor *et al.* [6] studied the critical state model; Nayak ad Zienkiewicz [8] provided a detailed elastoplastic model; and a viscoplastic model was used by Lewis *et al.* [23] to study the stability of an embankment.

Another important characteristic behaviour of soft soil is secondary consolidation. As indicates in Chapter 11, this can be defined as the slowly continuing compression of the soil skeleton after the excess pore pressure has substantially dissipated. This behaviour can be modelled by a Kelvin rheological relationship [7] and has been studied in detail for soft soil [22].

In this section we have discussed only briefly the complexity of soil behaviour during consolidation; the interested reader should refer to Chapters 4 and 11 for further detail.

## 12.4 APPLICATIONS

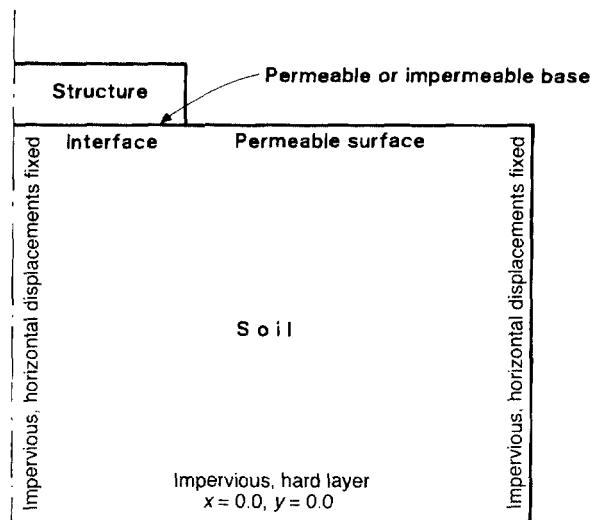
### 12.4.1 Test 1: Shallow foundation

A shallow foundation in the present context is defined as a foundation which rests on the ground surface or at a shallow depth below the ground level. A model of such a shallow foundation is analysed for the situation where a traction  $f$  is applied to the foundation at time  $t = 0$  and thereafter held constant with drainage occurring only at ground level. The experimental details are reproduced from Poulos and Davis [11]. A series of laboratory tests, where the footings rested on a soft consolidated clay, allowed control of the hydraulic boundary conditions and the stress state. Further details of the apparatus may be obtained in the original paper. The surface load of 71 kPa was applied via a steel plate with a radius of 63.5 mm. The diameter of the pressure vessel was 300 mm and the clay thickness was 71 mm. The soil–structure system is shown in Figure 12.1, and the typical finite element deformed mesh is shown in Figure 12.2. The material properties for the finite element analysis are as follows:

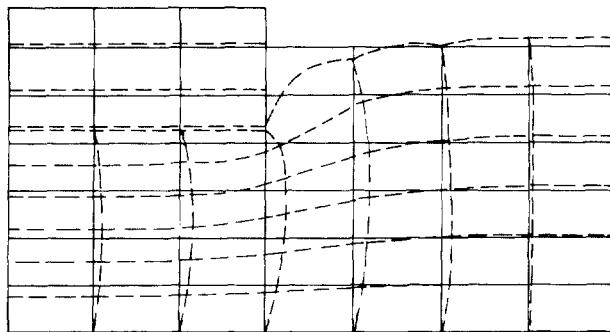
*Circular foundation*

$$E = 200 \text{ GPa}$$

$$\nu = 0.2$$



**Figure 12.1** Soil shallow foundation system: no flow across the interface boundary; no overlapping elements at the interface; no flow across the centreline; footing is free to move in both directions, except the centreline



**Figure 12.2** Deformed mesh at time 110 min (step 30); the maximum displacement is  $4.146 \times 10^{-4}$ , corresponding to 13.8 mm

*Clay*

$E' = 900$  kPa

$\nu' = 0.35$

$K/\gamma_w = 2.4 \times 10^{-8} \text{ m}^4 \text{ kN}^{-1}$  per minute

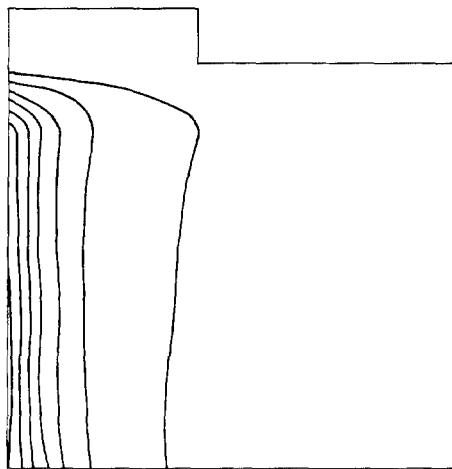
$C_u = 91.5$  kPa (undrained strength)

*Interface*

$K_{nn} = 10^{10}$  kPa

$K_{ss} = 10^{-6}$  for smooth base

$K_{ss} = 10^{10}$  for rough base



**Figure 12.3** Contour map of pore water pressure at time step 10: permeable base, interval = 5

With an undrained shear strength  $C_u = 91.5$  kPa, an estimated ultimate bearing capacity for this problem would be  $5.68C_u$ , i.e. 521 kPa. The tests used a footing stress of 71 kPa, giving a factor of safety of 7.6 against failure and 1.8 against local yield. It is reasonable to assume the foundation would still remain in the elastic stage. The finite element results are shown in Figure 12.2 to 12.5 and agree well with the measured values of footing settlements.

In order to carry out an investigation on the soil-structure interaction during the consolidation process, four cases were studied and their results compared with the measured results:

*Case 1:* a finite element analysis with the footing and the soil as a single system without any interface elements. This analysis represents the situation of a very rough (stick) base of the footing with no slippage at the interface. Results are given by curve 1 of Figure 12.5.

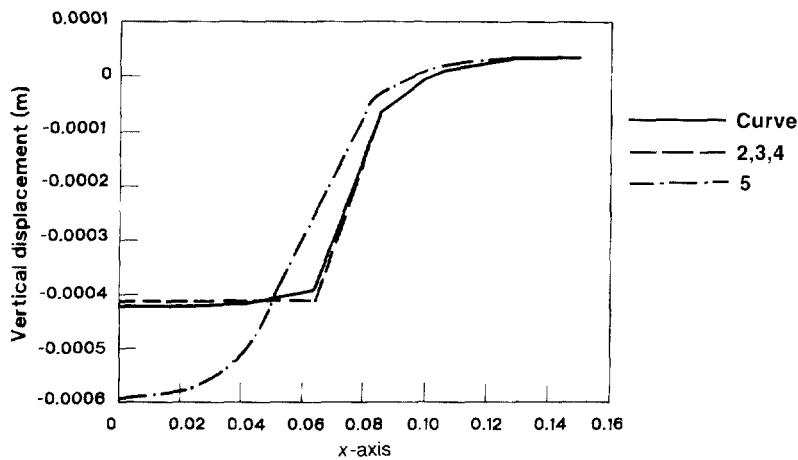
*Case 2:* a finite element analysis with the complete system being simulated. This consists of the foundation itself with interface and soil elements being present. A smooth interface between the footing and soil is assumed, i.e.  $K_w \approx 0$ . Results are given by curve 2.

*Case 3:* same as for case 2 but with non-linear materials, i.e. a hyperbolic interface behaviour and an elastoplastic soil. Results are given by curve 3.

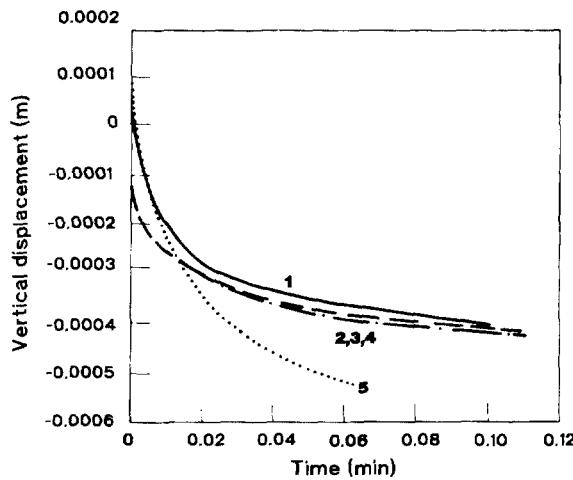
*Case 4:* a consolidation analysis was carried out on the soil domain only, i.e. the load is applied at the ground level but without the structure or interface elements. Results are given by curve 5.

*Test measurements:* results are given by curve 4 [11].

The results from these analyses are shown in Figures 12.4 and 12.5. Case 1 seems to underestimate the displacement, but on taking the foundation into account, the result is reasonably good. Case 4 seems to overestimate the settlement and the results are far too



**Figure 12.4** Surface settlement at time step 30: (1, 2, 3, 5) analytical results; (4) measured results from Poulos and Davis [11]



**Figure 12.5** Settlement at the bottom of the foundation: (1) footing and soil without interface; (2, 3) footing and soil with interface; (4) measured results from Poulos and Davis [11]; (5) soil consolidation without structure

much in error because the interaction between the soil and the foundation was not considered. Cases 2 and 3 agree very well with the experimental results. As the applied load was well below the value of the ultimate load, the results from the linear and non-linear analyses were almost identical.

A parametric study was subsequently made and conclusions drawn from the results. In general, an increase of the foundation stiffness leads to slightly higher pore water pressure values with a consequent reduction in the consolidation settlement values. This can be explained by the fact that the pore fluid pressure in the soil responded with a consequent transfer of the load intensity from the structure to the soil, whereas the

displacement response of the soil domain is attributed to the whole soil-structure stiffness system.

The contact pressure at the base of the footing can be obtained directly from the stress values at the interface elements. During the early stages of the consolidation process, the contact pressure is greatest at the centre of the footing with a gradual decrease towards the rim. This is due to the fact that, on first application of the load, high stresses are induced near the edge but the excess pore pressures dissipate faster in this region than at the centre. Thus, there is an initial tendency for the settlement at the edge of the footing to be greater than at the centre. In other words, there is an initial 'heave' at the centre of the foundation, inducing high contact pressures. This phenomenon is less pronounced for a flexible footing than for a rigid footing.

The high contact pressures at the centre of foundations during the early stages of consolidation have also been confirmed by Booker and Small [15]. Correspondingly, there is a development of 'negative' differential settlements, with the bending moment initially decreasing before it increases to its maximum value during the final stages of consolidation.

In general, a rough base tends to reduce the differential settlement values with a commensurate decrease in the bending stresses. And a rough-based footing tends to have a higher bearing capacity than a smooth-based footing. This phenomenon can be easily explained. As the load is applied, vertical settlements occur and horizontal compressive stresses are generated in the soil immediately below the footing. The soil consequently tends to move inwards towards the centre and is free to do so if there is no friction at the interface. For a rough or adhesive contact, this inward movement is restricted by the footing roughness, so interfacial shear stresses are developed. These shear stresses will reduce the displacement values along with the horizontal compressive stresses within the soil beneath the footing.

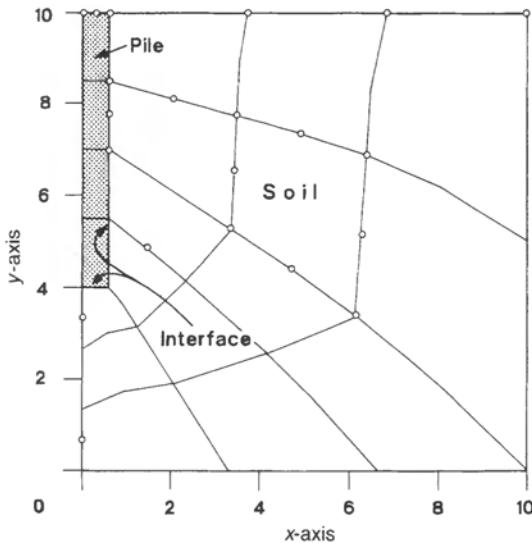
The effects of soil consolidation on the footing itself have also been studied. The shear force, bending moment and stresses developed with the footing are essentially due to the differential settlements of the soil during the consolidation process. A maximum value of the bending stress is often reached at the final stage of consolidation.

#### 12.4.2 Test 2: Pile-soil interaction

Conventional methods for computing the deformation and bearing capacity of piles cannot generally account for complexities such as insitu stresses, non-linear soil and interfacial behaviour, consolidation settlements, and variable skin friction on the pile surface as the soil mass consolidates. Our finite element model shows considerable promise in handling these complexities and a true time history of the deformation, skin friction and excess pore water pressures generated and dissipated beneath and around the pile can be easily analysed.

Although there appears to be no exact solution to the problem of a single pile loaded axially or horizontally, an approximate solution obtained by Poulos and Davis [10] is often used in practice. A simple problem is studied in this section and it would be interesting to compare the finite element result with that of Poulos and Davis.

A circular-section pile embedded in soft clay has an axial load of 200 kN. The problem is analysed assuming axisymmetric conditions with a mesh configuration



**Figure 12.6** Axisymmetric problem for a single pile axially loaded with 200 kN: pile diameter = 1.2 m, length = 6 m,  $L/D = 5$ ; the original mesh has 85 nodes and 24 elements (15 soil elements, 5 interface elements and 4 pile elements)

(Figure 12.6). The file itself was assumed to have linear elastic properties since the pile stiffness is very high compared to the stiffness of soft clay. The material properties are as follows:

*Pile*

$$E = 200 \text{ GPa}$$

$$\nu = 0.2 \text{ (typical for concrete pile)}$$

*Clay*

$$E' = 600 \text{ kPa}$$

$$\nu' = 0.35$$

$$K/\gamma_w = 2.5 \times 10^{-4} \text{ m}^4 \text{ kN}^{-1} \text{ per year (permeability)}$$

$$C_u = 91.5 \text{ kPa} \text{ (undrained strength)}$$

*Interface*

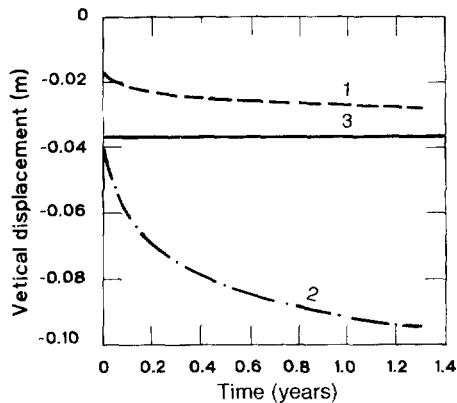
$$K_{nn} = 10^8 \text{ kPa}$$

$$K_{ss} = 10^{-5} \text{ for smooth base}$$

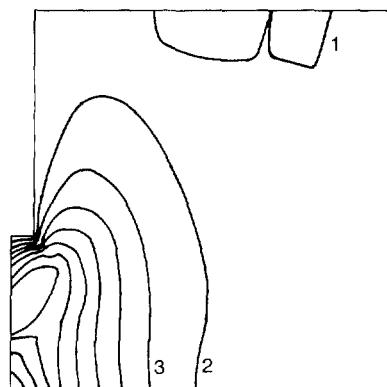
$$K_{ss} = 10^8 \text{ for rough base}$$

$$c = 15 \text{ kPa}, k = 10^4, n = 1, \phi = 30^\circ \text{ for hyperbolic interface model}$$

The resulting settlement of the pile tip is shown in Figure 12.7. The values of the excess pore water pressures generated in the soil mass at a typical time step are shown in Figure 12.8. The effective principal stresses and vertical displacement contours are given in Figures 12.9 and 12.10, respectively. The pile settlement was compared with the theoretical value reproduced from the paper by Poulos and Davis [10]; there appears to be a good correlation (Figure 12.7).



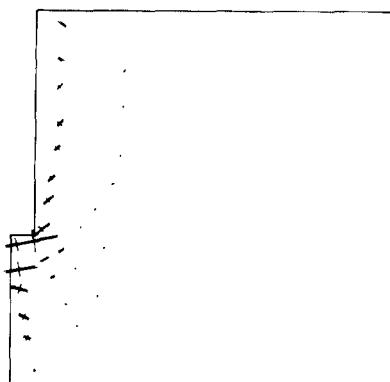
**Figure 12.7** Vertical displacement of the pile head: (1) rough pile, (2) smooth pile, (3) theoretical



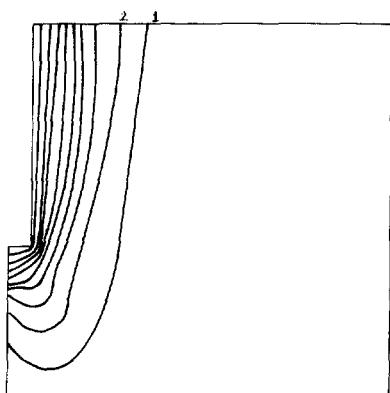
**Figure 12.8** Pore water pressure contours in the soil mass at time step 30: rough interface pile; 10 contours, interval = 0.93, first contour at 0.0

A parametric study we have carried out indicates that, in general, the pile stiffness has very little influence on the excess pore fluid pressure generated at the pile tip and the surrounding soil. The interface action between the pile shaft and the soil plays an important role in the behaviour of a pile in soft soil. Firstly, the bearing capacity of the pile is mainly governed by the skin friction; secondly, the consolidation settlement of the pile and the excess pore pressures generated in the surrounding soil depend upon the magnitude of the shear stresses on the skin surface as the load is applied.

A smooth pile, when loaded, induces large excess pore pressures in the soil below the pile tip, since the load transfer from the pile is mostly transmitted to the soil at the base (pile tip). At higher loads, tensile stresses are developed in the surrounding soil just above the pile tip, and negative excess pore pressures may be generated. This effect is less pronounced as the roughness of the pile surface increases. In this case the rate of the pile settlement is governed by the soil consolidation below the level of the pile tip.



**Figure 12.9** Effective principal stresses in soil at time step 30 (1.3 years): rough pile; scale = 50 kPa per millimetre

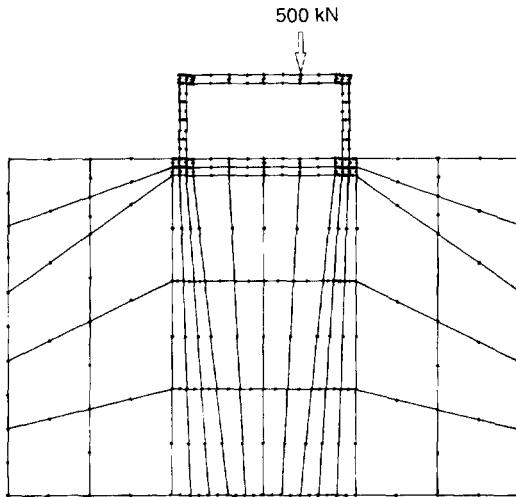


**Figure 12.10** Vertical displacement contours in the soil mass at time step 30 (1.3 years): rough interface pile; 10 contours, interval = 0.002 m, first contour at -0.0283 m

For a rougher pile surface, the carrying load is mainly transmitted to the soil via the skin surface, and the shear stresses at the interface drag down the surrounding soil. Positive excess pore pressures are generated along the pile shaft as well as beneath the pile tip. In this case the rate of the pile settlement is governed by the soil consolidation above the pile tip.

#### 12.4.3 Test 3: Frame on soft soil

The discussion of soil–structure interaction has so far concentrated on the analysis of shallow and pile foundations; little attention has been given to the effect of the superstructure. However, this effect can be significant and may even dominate the



**Figure 12.11** The original mesh: 90 elements, 348 nodes; eccentric loading = 500 kN; interface elements around foundation

foundation response. Consider a simple single-bay frame resting on soft soil. The frame is either loaded vertically or by a combination of vertical and horizontal loads.

Superstructure frames are conventionally designed by assuming their bases are completely rigid or hinged. However, foundations resting on deformable soil also undergo deformations that depend on the relative rigidity of the foundation, the superstructure and the soil. An interaction analysis is therefore necessary for correctly assessing the member forces in the superstructure itself.

A large frame, span of 20 m and height 10 m, is vertically loaded by a point load of 500 kN. The mesh configuration is shown in Figure 12.11. The analysis required 96 elements and 348 nodes. The structural material and the soil were assumed to respond elastically, since no data was available to perform a non-linear analysis. The interface elements were assumed to have ideally plastic characteristics.

The material properties for this problem are as follows:

*Frame and footings*

$$E = 200 \text{ GPa}$$

$$\nu = 0.2$$

*Clay*

$$E' = 6000 \text{ kPa}$$

$$\nu' = 0.3$$

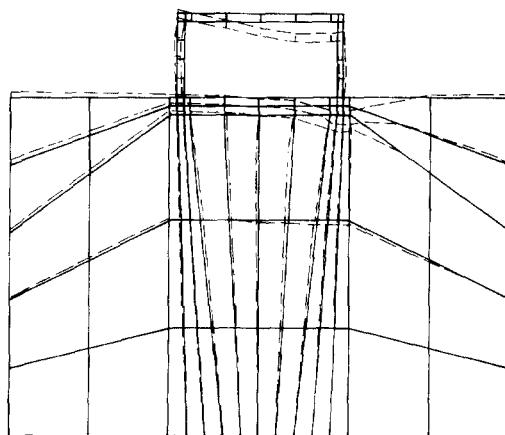
$$K/\gamma_w = 2 \times 10^{-6} \text{ m}^4 \text{ kN}^{-1} \text{ per year (permeability)}$$

*Interface*

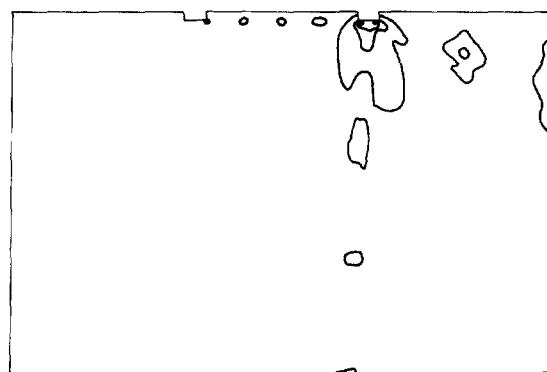
$$K_{nn} = 10^8 \text{ kPa}$$

$$K_{ss} = 2000 \text{ with ultimate shear strength } c_u = 20 \text{ kPa}$$

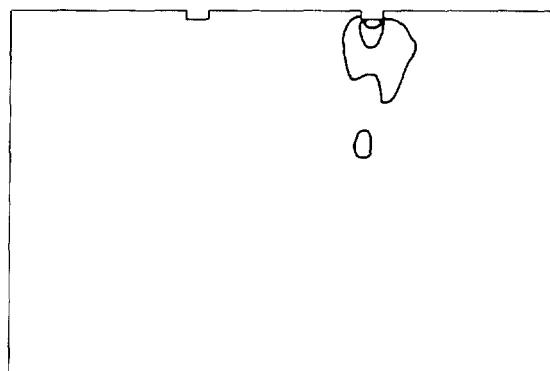
The results of the eccentrically loaded case are presented in Figures 12.12 to 12.16.



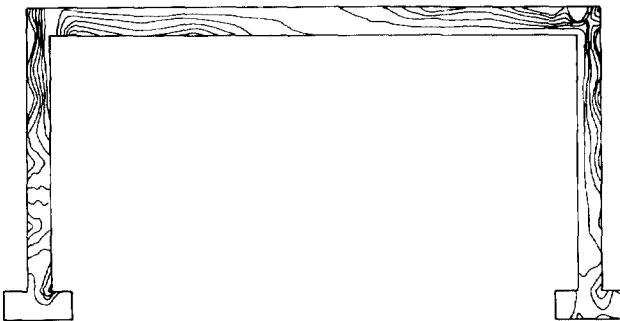
**Figure 12.12** The deformed mesh at time step 50 (4 years): displacement scale = 0.004 m per millimetre



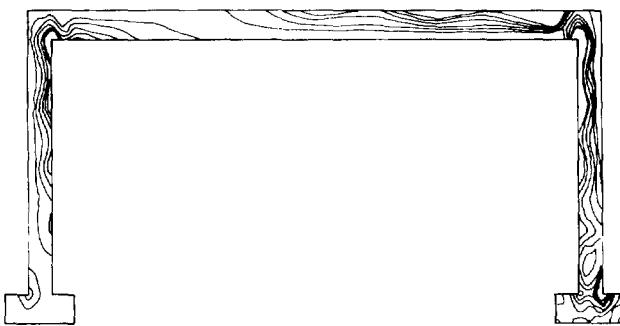
**Figure 12.13** Contours of pore water pressure in the soil mass at 1 year: interval = 20 kPa, first contour at 20 kPa



**Figure 12.14** Contours of pore water pressure in the soil mass at 1 year: interval = 20 kPa, first contour at 20 kPa



**Figure 12.15** Contours of major principal stress at 1 year: 6 contours, interval = 250 kPa, first contour at 1350 kPa compression



**Figure 12.16** Redistribution of major principal stresses after 4 years: 6 contours, interval = 250 kPa, first contour at 1350 kPa compression

A parametric study we have carried out indicates that, in general, a stiffer frame will exhibit lower total and differential settlements beneath the footings. However, note that the associated reduction in differential settlements will be accompanied by increased wall stresses and contact stresses at the corners of the footings, hence they should be accounted for at the design stage.

The frame stiffness also affects the moment at the foundation levels, which in turn affects the distribution and the magnitude of excess pore pressure generated in the soil surrounding the footings.

The consolidation settlements of the footings depend on many factors; here are the important ones.

#### 12.4.3.1 Interaction of two adjacent footings

By varying the frame span under the vertical loading, the results obtained by Tran [7] have shown that the settlements of the two footings, for large-span frames, are always less than those of smaller spans. This indicates that the footings behave as two separate entities in a large-span frame. The increase in settlements in the smaller-span frame can

only be due to the interaction between the footings, i.e. the settlement of one footing is affected by the load acting on the adjacent footing.

#### 12.4.3.2 Eccentric loading

Figure 12.12 shows that, as the right column settles due to the consolidation of the soil, it forces the left column, via the beam, to follow suit. The differential settlements between the footings influence the redistribution of the stresses in the frame, particularly at the corners of the frame. This indicates that a redistribution of the bending moment has taken place during the soil consolidation process.

#### 12.4.3.3 Horizontal loading

Differential settlements of footings also occur under the action of horizontal loading. The stresses transmitted from the frame to the soil are compressive at one footing, increasing the pore pressure values underneath the footing. In contrast, the stresses transmitted to the soil are tensile at the other footing, reducing the pore pressure values. The frame tends to rotate during the consolidation process and the differential settlement between the two footings increases with time.

## 12.5 CONCLUSIONS

A finite element model for soil–structure interaction during consolidation has been discussed along with tests for typical foundations resting on soft soil. The superstructure (or substructure) has a significant influence on the soil consolidation and vice versa; attempts to analyse either the superstructure or soil independently of each other could lead to serious errors. A true time interaction analysis using the model will therefore be very useful in design practice. Important factors for superstructure or substructure design, such as critical stress, contact pressure, displacement and pore fluid pressure, can be obtained explicitly from the developed finite element model. And it can easily incorporate any complexities caused by the initial stress effects, material non-linearities, and creep in the soil or structure. At present it does not incorporate the effects of cyclic loading on the consolidation process, but Chapter 14 gives more details about this aspect.

## REFERENCES

1. Cakmak, A. S. (ed). (1987) *Soil-structure Interaction, Development in Geotechnical Engineering*, **43**, Elsevier Computational Mechanics Publications, The Netherlands.
2. Murray, R. T. (1978) Developments in two and three dimensional consolidation theory. In *Developments in Soil Mechanics*, C. R. Scott (ed.). Applied Science Publishers, London.
3. Schrefler, B. A. (1985) A partitioned solution procedure for geothermal reservoir analysis. *Comm. Appl. Num. Meth.*, **1**, 1–4.

4. Felippa, C. A. and Park, K. C. (1974) Staggered transient analysis for a coupled mechanical system: formulation. *Comp. Meth. Appl. Mech. Engng.*, **24**, 61–111.
5. Desai, C. S. and Hargett, C. M. (1974) Analysis of pile by supported gravity dock. *J. Geotech. Engng Div. ASCE*, **100**, 1009–1029.
6. Naylor, D. J., Pande, G. N., Simpson, B. and Tabb, R. (1981) *Finite Elements in Geotechnical Engineering*, Pineridge Press, Swansea.
7. Tran, D. V. (1988) *A finite element analysis of soil–structure interaction during consolidation*. PhD thesis, University of Wales, Swansea.
8. Nayak, G. C. and Zienkiewicz, O. C. (1972) Elastoplastic stress analysis. *Int. J. Num. Meth. Engng.*, **5**, 113–35.
9. Seed, H. B., Lysmer, J. and Huang, R. (1975) Soil–structure interaction analysis for seismic response. *J. Geotech. Engng. Div. ASCE*, **101**, 439–457.
10. Poulos, H. G. and Davids, E. H. (1968) The settlement behaviour of single axially loaded incompressible piles and piers. *Géotechnique*, **18**, 351–71.
11. Poulos, H. G., de Ambousis, L. P. and Davis, E. H. (1976) Methods of calculating long-term creep settlements. *J. Geotech. Engng. Div., ASCE*, **102**, No. GT7, 787–804.
12. Smith, I. M. and Molenkamp, F. (1980) Dynamic displacements of offshore structures due to low frequency sinusoidal loading. *Géotechnique*, **30**, No. 2, 179–205.
13. Ghaboussi, J., Wilson, E. L. and Isenberg, J. (1973) Finite element for rock joints and interfaces. *Journal of the Geotech. Engng. Div., ASCE*, Vol. 99, No. SM10, 833–47.
14. Duncan, J. M. and Chang, C. Y. (1970) Nonlinear analysis of stress–strain in soils. *J. Soil Mech. Found. Div., ASCE*, Vol. 96, No. SM5, 1629–53.
15. Small, J. C. (1977) *Elastoplastic Consolidation of Soil*, Ph. D. Thesis, University of Sydney, N.S.W. Australia.
16. Park, K. C. (1980) A partitioned transient analysis procedure for coupled field problems: stability analysis. *J. Appl. Mech.*, **47**, 370–76.
17. Zienkiewicz, O. C., Lewis, R. W. and Humpheson, C. (1977) A unified approach to soil mechanics problems including plasticity and viscoplasticity. Invited Chapter in *Finite Elements in Geomechanics*. John Wiley & Sons, 151–77.
18. Goodman, R. E., Taylor, R. and Brekke, T. L. (1968) A model for the mechanics of jointed rocks. *J. Soil Mech. Found. Div., ASCE*, **94**, 637–59.
19. Sandhu, R. S. and Wilson, E. L. (1969) Finite element analysis of seepage in elastic media. *J. Engng Mech. Div., ASCE*, **95**, 641–52.
20. Lewis, R. W. and Schrefler, B. A. (1978) A fully coupled consolidation model of the subsidence of Venice. *Water Resources Research*, **14**, 223–29.
21. Lewis, R. W. and Schrefler, B. A. (1987) *The Finite Element Method in Deformation and Consolidation of Porous Media*, Wiley, New York.
22. Lewis, R. W. and Tran, D. V. (1989) Numerical simulation of Secondary consolidation of soils. *Int. J. Num. Anal. Meth. Geomech.*, **13**, 1–18.
23. Lewis, R. W., Zienkiewicz, O. C. and Roberts, G. K. (1976) A nonlinear flow and deformation analysis of consolidated problems, Vol. 2, 1106–18.
24. Hughes, T. J. R., Pister, K. S. and Taylor, R. L. (1979) Implicit–explicit finite element in nonlinear transient analysis. *Comp. Meth. Appl. Mech. Engng.*, **17/18**, 159–82.

# 13

## Back Analysis in Consolidation

### 13.1 INTRODUCTION

The design of most geotechnical engineering works is customarily carried out using 'stiff' and 'flexible' procedures [1]. The two approaches are based on different ideas and they are carried out in different ways.

In stiff procedures the experimental investigation needed to define the material parameters of the *in situ* geomaterial is organised during the design process. However, no *in situ* measurements are made during construction, so it will be impossible to modify the final design to accommodate unforeseen conditions met in the field. Stiff procedures assume the design process will provide meaningful values for the necessary parameters, and no significant deviations will be encountered in the field. But even the latest refinement to a design value is often unrealistic, and the result is called a stiff design. Because geotechnical problems are complex, stiff procedures can lead to serious drawbacks. In stiff procedures there is no interaction between the geotechnical consultant and the contractor, but some interaction is often necessary.

In order to reduce the many uncertainties embedded in practical problems and in order to design construction works more precisely, an alternative procedure was introduced by Terzaghi for the design of dams and large geotechnical works [3]. Once going by many titles, e.g. observational method, calibration method, inverse analysis and characterisation problem, Terzaghi's method has recently become known more widely as back analysis. It circumvents the difficulties with stiff procedures and it requires close cooperation between the contractor and the geotechnical consultant. Initially a measurement program is organised, so that the overall behavior of the geomaterial can be monitored during the construction. Subsequently, the experimental measurements (e.g. displacement and pore water pressure values) are checked against the calculated values on the basis of material parameters adopted during a previous stage of the design. If a significant discrepancy is observed between the two sets of data, the parameters are modified to minimise that discrepancy. This eventually leads to a new estimate of the material parameters, allowing the calculations to be further refined then perhaps the design will be modified.

Although it has certain limitations, back analysis does offer many possibilities for significant savings in time and money; can also provide the assurance of safe construction without the financial penalty attached to excessive safety considerations. Therefore, it has now become a preferred technique for validating theories, or more often for evaluating material parameters. The consolidation, or subsidence, problem is

but one of many and various geotechnical engineering problems, and it exhibits certain peculiarities due to the time dependency of the process. Consequently, back analysis of the consolidation problem has attracted a great deal of attention in geotechnical circles.

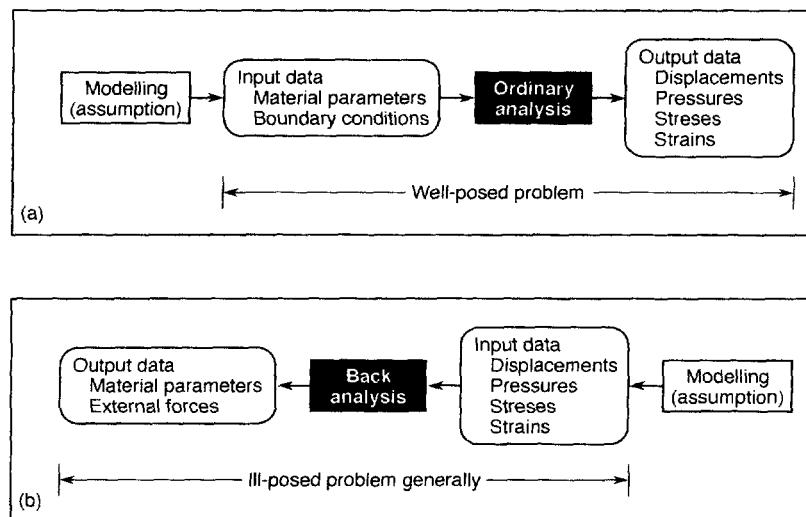
### 13.2 DEFINITION OF BACK ANALYSIS

A back analysis approach may be used to determine the values of the material parameters which, when introduced into an ordinary deformation analysis of the results (e.g. displacements and pore water pressures), will give the closest possible correspondence to the *in situ* measurements [1]. Figure 13.1 illustrates the difference in philosophy between ordinary analysis and back analysis.

In an ordinary analysis, assuming the physical model has been specified, the displacements, pore pressures, stresses and strains may be calculated once the values of the material parameters and boundary conditions are given. Normally, if a suitable algorithm has been developed, the solution to a particular problem has three distinct properties, such as existence, uniqueness and stability. This type of problem is called a well-posed problem.

In a back analysis, once a suitable physical model has been assumed, the values of the material parameters, relevant to the assumed model, are computed based on the measurement of displacements, pressures etc., at certain locations. However, in general, the solution of the problem will be ill-posed. If the location points and the number of required measurements are not chosen properly, the solution for the material parameters may not be stable.

The physical data of any practical problem is, at best, always an approximation to the true values. According to the classic concepts of mathematics, the solution obtained will be different from the true solution. Also, the ill-posedness of the problem is such that a



**Figure 13.1** (a) Ordinary analysis and (b) back analysis

discussion about the existence of the solution will not prove fruitful, and the solution itself will not be continuously dependent on the observed data. However, based on sound physical principles, a solution must exist which reflects the natural behaviour and possesses a certain form. Therefore, in order to avoid the question of solution existence, attempts should not be made to determine an exact solution which satisfies both the governing equations and the corresponding physical conditions (including the observed data). Instead a solution should be found which satisfies the governing equations and approximates to the corresponding condition, or a solution which approximately satisfies both the governing equations and the corresponding physical conditions. This type of solution will not be unique, but it is possible to obtain a solution which continuously depends on the observed data. This will require that some additional information is included to confine the solution domain so that the smoothness of the solution is improved.

It should be noted that parameters identified from a back analysis are indeed 'equivalent parameters' which synthesise the effects of various factors, such as the physical properties of geomaterials, geological conditions and other engineering factors. These are 'fuzzy' in the sense of their physical behaviour. The use of these 'equivalent parameters' will enable a calculation that will be relatively consistent with the practical measured values, and from which one can evaluate the engineering effectiveness of the design to predict dynamic processes during the construction. Traditionally, parameters identified from a back analysis are often compared with those obtained from experiments but these comparisons are often meaningless because the parameters are identified only with reference to the overall response produced by the observed data.

### 13.3 METHODOLOGY

Gioda and Sakurai [1] have presented a survey on back analysis techniques from a general geomechanics viewpoint, and some of these techniques are relevant to the particular case of consolidation problems. The techniques are the direct method, the indirect method and probabilistic approaches that allow certain practical aspects of the uncertainties to be taken into account. These three methods are briefly outlined in the following sections.

#### 13.3.1 Direct Method

For static problems in structural engineering, back analyses are normally designed on a contrary basis to ordinary stress analyses. The equations governing the ordinary problem are 'inverted' so that some of the quantities, which are unknown in the ordinary stress analysis but correspond to the available measurements, appear as sets of data. However, other quantities, having known values in an ordinary stress analysis (e.g. elastic modulus), appear in the group of unknowns. In most practical cases the number of *in situ* measurements exceeds the number of known parameters. This leads to an equation system having more equations than unknowns; the system can be solved using least squares techniques [1,4]. This approach was called an 'inverse' back analysis

by Gioda [4] and from a computational viewpoint it can be considered a direct method [5].

In order to use this technique, we should first obtain the ‘inversion’ of the equations which govern the corresponding stress analysis problem. Subsequently, different solution procedures have to be developed for the various back analyses, depending on the characteristics of the adopted models and on the type of parameters to be back calculated. However, we have not seen this technique applied to the back analysis of consolidation. This is probably due to the particular complexity of the problem, which has a strong dependence on the time history, unlike an ordinary static stress analysis.

### 13.3.2 Indirect Method

Instead of being inverted, the numerical formulation can remain the same as for an ordinary analysis. In this case, parameter search procedures are used to directly minimise the discrepancy between the field measurements and the corresponding numerical approximations. This was named ‘direct back analysis’ in 1985 [4] and ‘direct search procedure’ in 1987 [1]. However, Murakami and Hasegawa [5] called it the ‘indirect method’.

The general consolidation problem can be formulated as follows:

$$\begin{bmatrix} -\mathbf{K}_T & \mathbf{Q} \\ -\mathbf{Q}^T & -\mathbf{S} \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{p}} \end{Bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{H} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \mathbf{p} \end{Bmatrix} = \begin{Bmatrix} \dot{\mathbf{f}}^u + \mathbf{C} \\ \dot{\mathbf{f}}^p \end{Bmatrix} \quad (13.1)$$

where  $\mathbf{u}$  and  $\mathbf{p}$  are the nodal displacement vector and the nodal pore water pressure vector, respectively. The other variables are as defined in earlier chapters within the text.

The difference between the field measurements and the calculated values can be written as

$$E = \sum_{n=1}^{N_t} \left[ \sum_{i=1}^{N_d} (u_i^{*n} - u_i^n)^2 + \sum_{j=1}^{N_p} (p_j^{*n} - p_j^n)^2 \right] \quad (13.2)$$

where  $N_t$ ,  $N_d$  and  $N_p$  are the number of discretized time steps, measured displacements and measured pore pressures, respectively. The values  $u_i^{*n}$  and  $p_j^{*n}$  are the measured displacement at node  $i$  and the measured pore pressure at node  $j$ , at time step  $n$ ;  $u_i^n$  and  $p_j^n$  are the corresponding calculated values.

Equation (13.2) shows that the displacement and pressure terms have identical weight contributions to the error  $E$ . However, because of the difference in dimensions, they have different influences on the term  $E$ . In order to avoid this, the error  $E$  can be defined in the following non-dimensional form

$$E = \sum_{n=1}^{N_t} \left[ \sum_{i=1}^{N_d} \left( \frac{u_i^n}{u_i^{*n}} - 1 \right)^2 + \sum_{j=1}^{N_p} \left( \frac{p_j^n}{p_j^{*n}} - 1 \right)^2 \right] \quad (13.3)$$

More generally,  $E$  can also be defined using the weights  $w_{ui}^n$  and  $w_{pj}^n$ :

$$E = \sum_{n=1}^{N_t} \left[ \sum_{i=1}^{N_d} w_{ui}^n (u_i^{*n} - u_i^n)^2 + \sum_{j=1}^{N_p} w_{pj}^n (p_j^{*n} - p_j^n)^2 \right] \quad (13.4)$$

The literature [6–8] indicates some applications of the indirect method for the back analysis of consolidation problems. Both the direct method and the indirect method have advantages and shortcomings. The direct method generally converges towards the optimal values of the parameters faster and more ‘smoothly’ than the indirect method, and this makes it cheaper on computer time. However, it requires an ad hoc computer program, whereas the indirect method can be developed using a standard non-linear optimisation code and a subroutine module for ordinary consolidation analysis.

### 13.3.3 Probabilistic Approaches

The use of probabilistic approaches for back analyses is becoming increasingly popular in the field of geomechanics. In fact, many important practical aspects can be taken into account by means of these approaches, e.g. the uncertainties related to a definition of the soil layers and the influence of the experimental errors on the back-calculated parameters.

Among the various probabilistic approaches, the so-called Bayesian approach presents significant advantages. Firstly, the subjective judgement of an expert, or other valuable *a priori* information on the parameters to be back calculated, can be taken into account along with the experimental data. Secondly, the estimation of the unknown parameters can be ‘updated’ if additional experimental information becomes available. This factor is extremely useful when dealing with consolidation, which is time dependent. Subsequent field measurements can be used to obtain a continuous refinement of the estimated soil property values.

A variety of examples have been reported for the Bayesian approach [9–11]. The Kalman filter theory has recently been proposed as an alternative probabilistic approach to the back analysis of consolidation [5,14,15].

### 13.3.4 Alternative Methods

Due to the special characteristics involved in the back analysis of consolidation problems, the methodologies are distinct from those generally used in geomechanics. Particularly renowned among them is the Matsuo–Asaoka approach [9,12,13], an observational procedure for predicting consolidation behaviour.

For example, in the one-dimensional case, the consolidation settlement  $\rho(t)$ , at time  $t$ , is governed by the following ordinary differential equation:

$$\rho + c_1 \dot{\rho} + c_2 \ddot{\rho} + \dots + c_n \rho^{(n)} = f(t) \quad (13.5)$$

where  $c_i (i = 1, \dots, n)$  are constant parameters related to the consolidation coefficients and depths of the soil layers, and  $f(t)$  is a function that depends on the permeability condition and the loads.

Providing that  $f(t)$  can be expanded into a power series of finite order, the solution of (13.5) can be written as

$$\rho(t) = g(t) + \sum_{s=1}^n c_s \exp(\lambda_s t) \quad (13.6)$$

where  $g(t)$  is a polynomial of the same order as  $f(t)$ , and  $\lambda_s$  are the eigenvalues of (13.5). For the simplest case, i.e.  $f(t)$  is a constant and  $n = 1$ , then

$$\rho_j = \beta_0 + \beta_1 \rho_{j-1} \quad (13.7)$$

where  $\rho_j$  denotes  $\rho(t_j)$ ,  $t_j = \Delta t \cdot j (j = 1, 2, \dots)$ . Equation (13.7) has been used successfully to predict settlement characteristics by employing a chart [9,12,13].

### 13.4 PARAMETER IDENTIFICATION

We shall now concentrate on the parameter search approach in Section 13.3.2. All the terms in the error functions depend, via the calculated displacements  $u_i$  and pore water pressures  $p_j$ , on the physical parameters to be back calculated. If we define the vector of material parameters as  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_r]^T$ , then (13.4) can be rewritten as

$$E(\alpha) = \sum_{n=1}^{N_d} \left[ \sum_{i=1}^{N_d} w_{ui}^n (u_i^{*n} - u_i^n(\alpha))^2 + \sum_{j=1}^{N_p} w_{pj}^n (p_j^{*n} - p_j^n(\alpha))^2 \right] \quad (13.8)$$

For the sake of simplicity,  $\mathbf{q} = [q_1, q_2, \dots, q_s]^T$  denotes the vector consisting of physical measurements (including displacements and pressures), then there exists

$$E(\alpha) = \sum_{k=1}^N \sum_{l=1}^s W_l^k (q_l^{*k} - q_l^k(\alpha))^2 \quad (13.9)$$

where  $N$  is the number of observation points. Alternatively, the identification problem required to minimise  $E(\alpha)$  can be written as the following equivalent problem:

$$\sum_{k=1}^N \sum_{l=1}^s W_l^k \frac{\partial q_l^k}{\partial \alpha_m} (q_l^{*k} - q_l^k(\alpha)) = 0 \quad (m = 1, 2, \dots, r) \quad (13.10)$$

The error  $E(\alpha)$  is a complicated non-linear function of the unknown parameter  $\alpha$  and its analytical expression cannot be found. Therefore, the required minimisation algorithm

must be capable of dealing with general non-linear functions and it is not necessary to require the analytical evaluation of error function gradients with respect to the parameters  $\alpha$ .

The material parameters are normally identifiable within possible ranges, but also have other constraints. Therefore, the corresponding parameter identification process becomes a constraint optimisation problem. If methods involving penalty terms are used, then the constrained optimisation problem can become an unconstrained problem.

The back analysis technique is mathematically ill-posed in its initial formulation, in the sense that the existence, uniqueness, and/or stability of the ensuing solution are not guaranteed. A successful back analysis solution generally involves a reformulation of the problem to obtain an approximate solution that is well behaved.

Based on (13.9) the original problem of determining the material parameters is normally ill-posed with no unique solution. Therefore, a new problem should be constructed, essentially similar to the original problem. The purpose would then be to establish conditions such that the new problem is well posed. Specifically, conditions are required to ensure that the observed data  $q_l^{*k} (k = 1, 2, \dots, N; l = 1, 2, \dots, s)$  uniquely defines the estimated material parameters  $\alpha$ . One obvious unique requirement is that  $sN \geq n$ . However, some additional uniqueness requirements on the number of measurements can be expected when some measurements are not well located in the domain.

Overspecified data can smoothen the back analysis solution, i.e. they can improve the stability characteristics. Therefore, it has often been suggested that the conditions  $sN \gg n$  should be used. The specific criterion which defines the optimal number of measurements may be based on a statistical analysis and a knowledge of the standard deviation  $\sigma$  of the measurement errors. An estimate  $\hat{\sigma}$  of the true standard deviation  $\sigma$  may be given by the following expression:

$$\hat{\sigma} = \sqrt{\frac{1}{sN} \sum_{k=1}^N \sum_{l=1}^s W_l^k (q_l^{*k} - q_l^k(\alpha))^2} \quad (13.11)$$

A natural criterion for choosing the optimal value of  $sN$  is obtained by requiring that  $\hat{\sigma} \approx \sigma$ , which defines an optimal  $sN$  to sufficiently stabilise the solution of the back analysis. In a practical implementation of this technique, an iterative procedure can be initiated such that if  $\hat{\sigma}$  is greater (less) than  $\sigma$ , then  $sN$  is decreased (increased).

### 13.4.1 Optimisation Methods

Many algorithms can handle these optimisation problems; three of them are worth mentioning here [4,16].

#### 13.4.1.1 Simplex Method

In an  $n$ -dimensional space of the parameters to be evaluated, a simplex is a set of  $n + 1$  vertices such that  $n$  vectors from one vertex to all of the others are linearly independent.

Having chosen the first simplex (i.e.  $n + 1$  suitable trial sets of parameters), the error function is evaluated for each vertex. The new simplexes can be constructed using a projection of the vertex, according to the function values. Eventually, a vertex corresponding to a set of the estimated parameters can be obtained as a consequence of using the simplex method.

#### *13.4.1.2 Rosenbrock's Algorithm*

This algorithm minimises an error function with respect to each parameter in turn (linear search), keeping the remaining parameters fixed. It is known as a relaxation process.

#### *13.4.1.3 Levenberg–Marquardt Method*

This method employs information about the derivatives of the error function in order to direct the search in an optimal direction. A sequence of corrections to the parameters are determined until the convergence is achieved within a given tolerance.

Note that the derivative information can be obtained by numerical approximation derived from the error function values, e.g. via the finite difference method; there is no need to determine the derivatives from an explicit formulation of the problem concerned. For complicated material models and structures, the error function may have several minima. A local optimisation algorithm will probably lead to a minimum (a set of parameter values), but not necessarily the desired minimum.

### **13.4.2 Sensitivity Analysis**

In the back analysis of consolidation problems a very important aspect is the effectiveness of the algorithms. If one ignores the stability analysis (sensitivity analysis) and uses these algorithms to solve a practical engineering problem, then serious mistakes can ensue, primarily because the problem is ill-posed. The ill-posedness is generally characterised by the non-uniqueness and instability of the identified parameters. The instability of the back analysis procedure stems from the fact that small errors in measurements can sometimes cause serious errors in the evaluated parameters.

From a practical viewpoint, the analysis for the location of such optimal observation points is sure to attract further attention from analysts. For without suitable measured values, the whole back analysis procedure will make little sense, even with a practical engineering background.

Limited by the performance quality of the corresponding deformation analysis algorithm, and the errors included in the measurements, the back analysis will be subject to certain performance criteria. For example, there is a minimal step length requirement for calculating the derivative of the error function, if the chosen optimisation algorithm needs the derivative information.

The derivative of the vector  $\mathbf{q}(\alpha)$  with respect to  $\alpha$  is generally a matrix, usually called a sensitivity matrix:

$$\mathbf{S} = \left. \frac{\partial \mathbf{q}(\alpha)}{\partial \alpha} \right|_{\alpha_0} = \begin{bmatrix} \frac{\partial q_1}{\partial \alpha_1} & \dots & \frac{\partial q_1}{\partial \alpha_r} \\ \vdots & & \vdots \\ \frac{\partial q_s}{\partial \alpha_1} & \dots & \frac{\partial q_s}{\partial \alpha_r} \end{bmatrix} \quad (13.12)$$

where  $\alpha_0$  is a nominal parameter vector in which the sensitivity is relevant. If  $\mathbf{q}(\alpha)$  is a scale value for simple problems (e.g. a one-dimensional case only with displacements), then  $\mathbf{S}$  will become a sensitivity vector. As the consolidation process depends on time, we can further define a time-averaged sensitivity matrix as

$$\tilde{\mathbf{S}} = \frac{1}{T} \int_0^T \mathbf{S} dt = \frac{1}{T} \int_0^T \left. \frac{\partial \mathbf{q}(\alpha)}{\partial \alpha} \right|_{\alpha_0} dt \quad (13.13)$$

The sensitivity matrix  $\mathbf{S}$  can be formulated in terms of  $\alpha$ , but an explicit expression cannot be given in most cases, due to the physical nature of the problem, although a numerical evaluation is always possible, e.g. using a forward-difference approximation. Moreover, the sensitivity matrix depends on  $\alpha$  and the locations within the computational domain. If the components of this matrix are extremely small for a parameter at one position, then the corresponding back analysis is not efficient when evaluating the value of this particular parameter, given the measurements at that position. A back analysis is efficient only if the algorithm is sensitive to the parameters to be determined at locations where measurements can be made.

Furthermore, the concept of an ‘effectiveness index’ for the parameter identification can be introduced as

$$I = \sqrt{\frac{1}{rsN} \sum_{k=1}^N \sum_{l=1}^s W_l^k \sum_{m=1}^r (S_{lm}^k)^2} \quad (13.14)$$

and a ‘global effectiveness index’ over the time domain was

$$\tilde{I} = \sqrt{\frac{1}{rsN} \sum_{k=1}^N \sum_{l=1}^s W_l^k \sum_{m=1}^r (\tilde{S}_{lm}^k)^2} \quad (13.15)$$

These are the average magnitudes of the sensitivity matrices for a set of observation points within a period of time. They can also be used to gain general concepts regarding the practical problem, but they have no precise physical meaning.

## 13.5 CASE STUDY

### 13.5.1 Background

In order to investigate the effectiveness of the proposed method, a hypothetical case has been created (Figure 13.2). This is a typical shallow foundation problem in which a circular raft rests on saturated soil and a constant traction  $p$  is applied to the foundation with drainage occurring only at the free surface.

As described in the literature [17, 18], a series of laboratory model footing tests on normally consolidated clay were carried out in which the geometry, the hydraulic boundary condition and the stress state could be controlled. The pressure vessel was assumed to be rigid and impermeable. The load of 71 kPa ( $71 \text{ kN m}^{-2}$ ) was applied via a steel plate with a radius of 63.5 mm. The internal diameter of the pressure vessel was 300 mm and the clay thickness 71 mm. The geometry is illustrated in Figure 13.2 and the finite element mesh is shown in Figure 13.3.

An ordinary deformation analysis was executed, in which the constitutive law was assumed to be linear elastic and Biot's consolidation behaviour prevailed. The material parameters used in the analysis are given in Table 13.1.

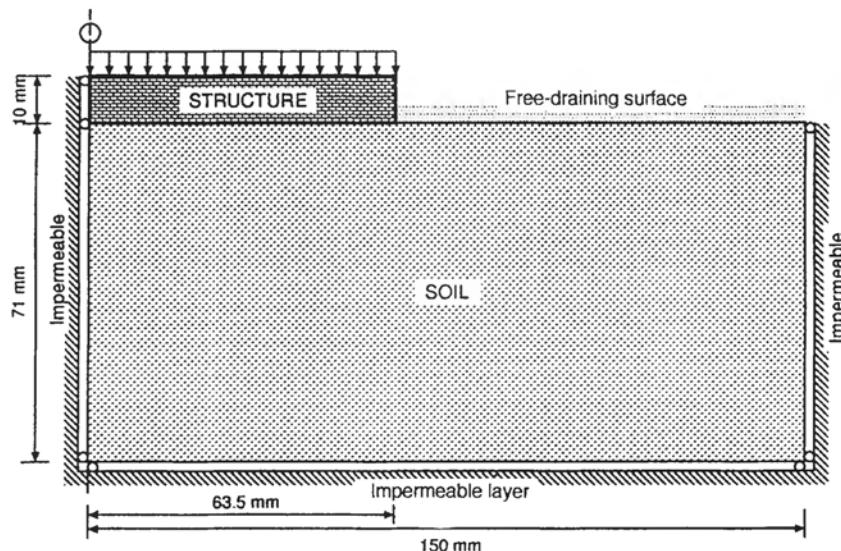


Figure 13.2 Geometry of the shallow foundation

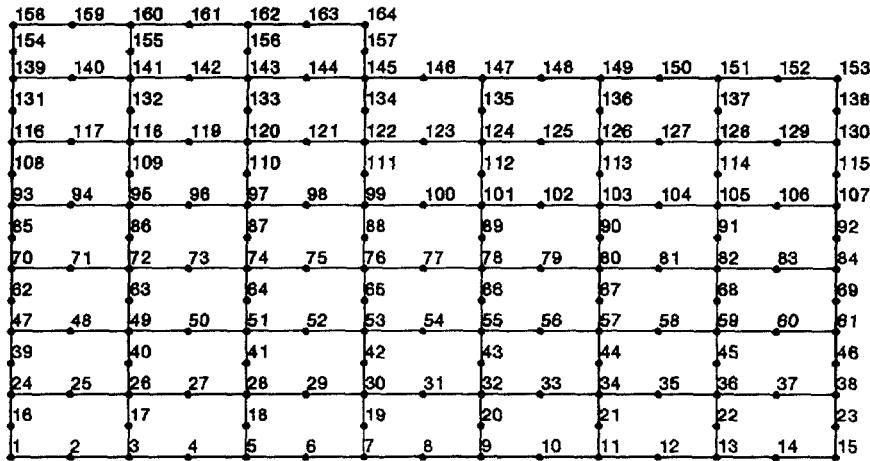


Figure 13.3 Finite element mesh

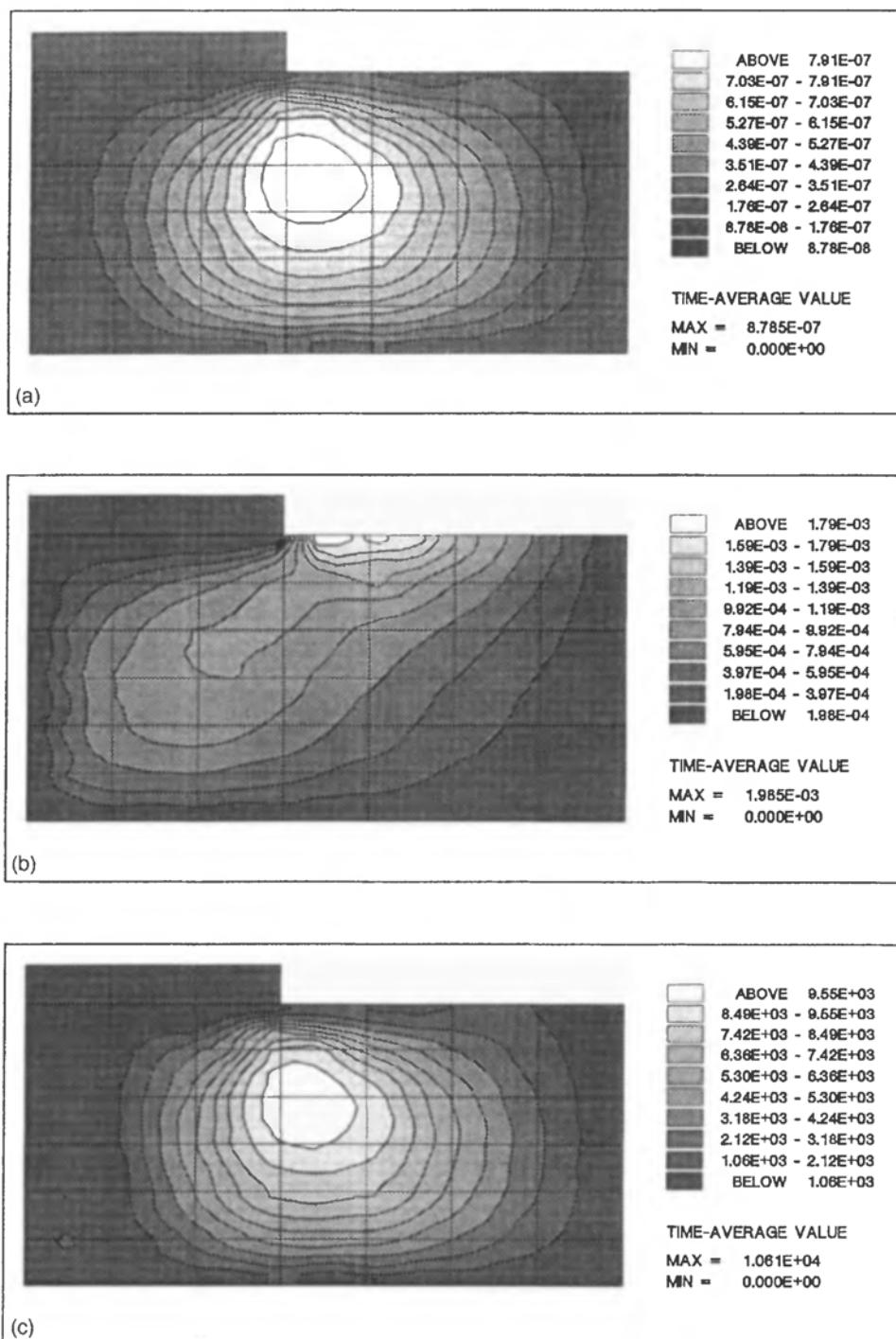
**Table 13.1** Material parameters used for the ordinary deformation analysis

Soil	$E = 900 \text{ kPa}$ $\nu = 0.35$ $k/\gamma_w = 2.4 \times 10^{-8} \text{ m}^4 \text{ kN}^{-1} \text{ min}^{-1}$
Structure	$E' = 200 \times 10^6 \text{ kPa}$ $\nu' = 0.2$

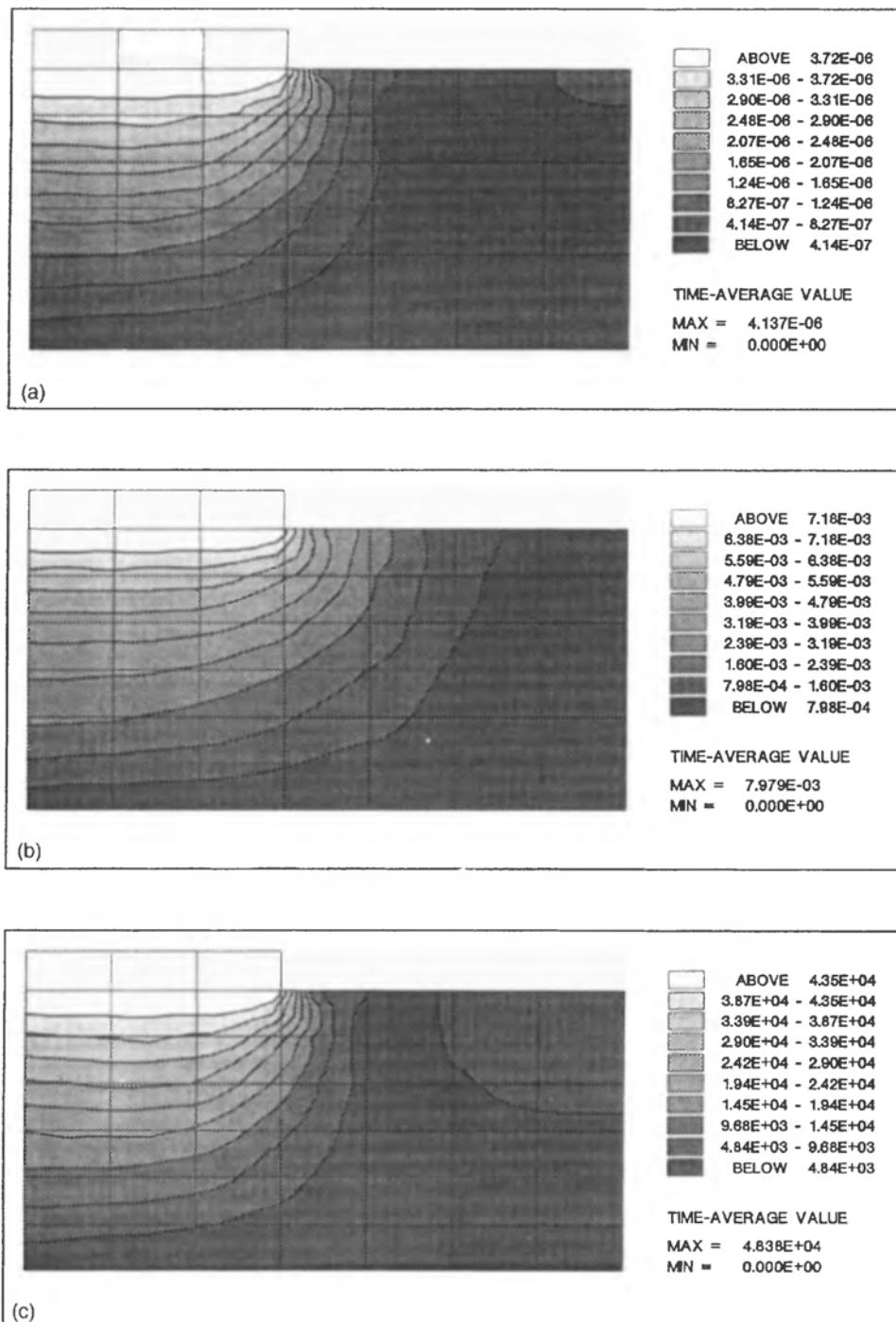
### 13.5.2 Hypothetical Case Study

We shall now consider a geotechnical problem in which the soil parameters are unknown, but the displacements and pore water pressures are measured at certain positions. In this hypothetical case the so-called measurements are all from the results of the ordinary deformation analysis.

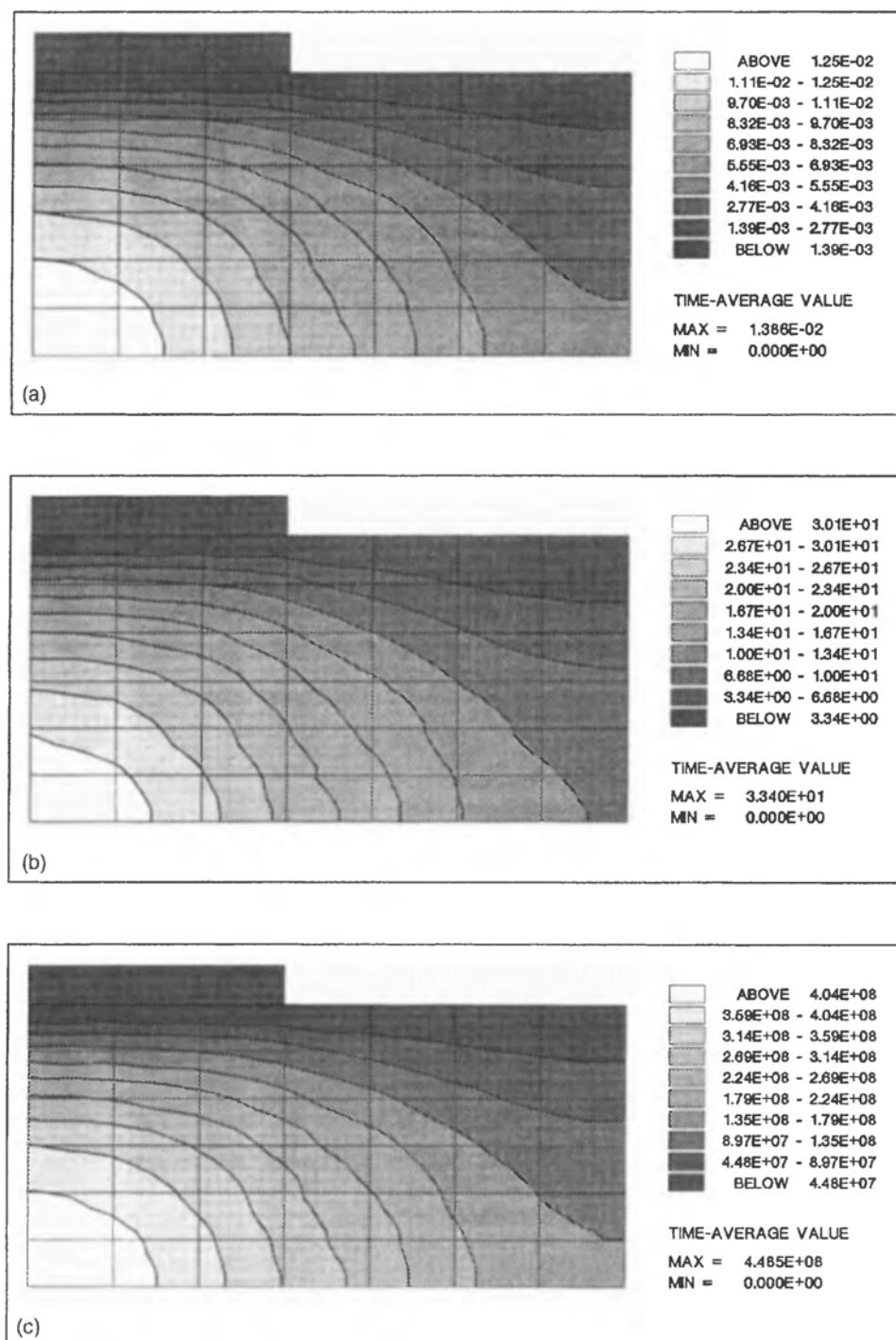
In order to understand the intricacies of the consolidation behaviour, a sensitivity analysis was first carried out. As already known, the sensitivity matrix varies according to the nominal parameters and the time history. Figures 13.4 to 13.6 indicate the time-averaged sensitivity distributions with the tabulated values of the material parameters. These distributions are calculated from a finite difference basis given that  $\Delta(E) = 90.0$ ,  $\Delta(\nu) = 0.035$ , and  $\Delta(k/\gamma_w) = 0.24 \times 10^{-8}$ . Notice that the most sensitive points are 1, 100, 139 (Figure 13.3) with respect to the corresponding pore water pressure, horizontal displacement and vertical displacement, respectively.



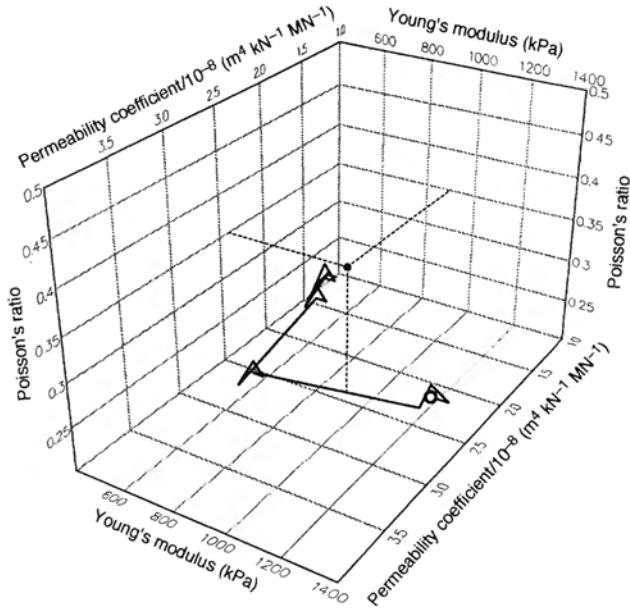
**Figure 13.4** Time-averaged sensitivity distribution for horizontal displacement: (a)  $x$ -displacement versus Young's modulus; (b)  $x$ -displacement versus Poisson's ratio; (c)  $x$ -displacement versus permeability coefficient



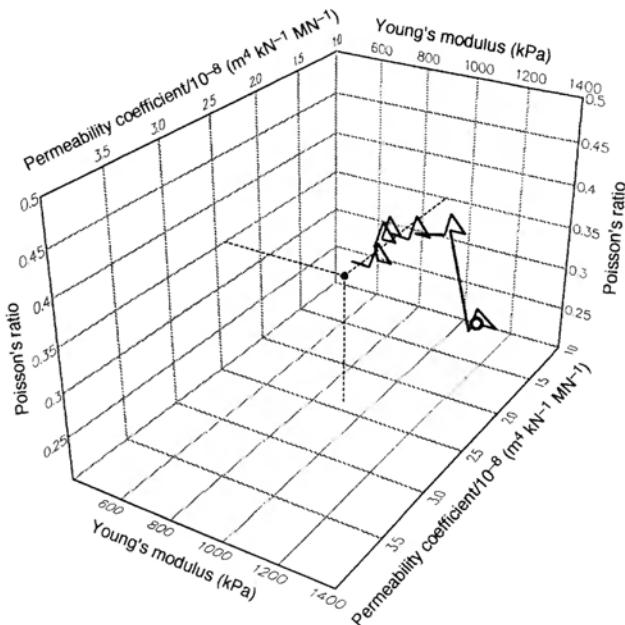
**Figure 13.5** Time-averaged sensitivity distribution for vertical displacement: (a) y-displacement versus Young's modulus; (b) y-displacement versus Poisson's ratio; (c) y-displacement versus permeability coefficient



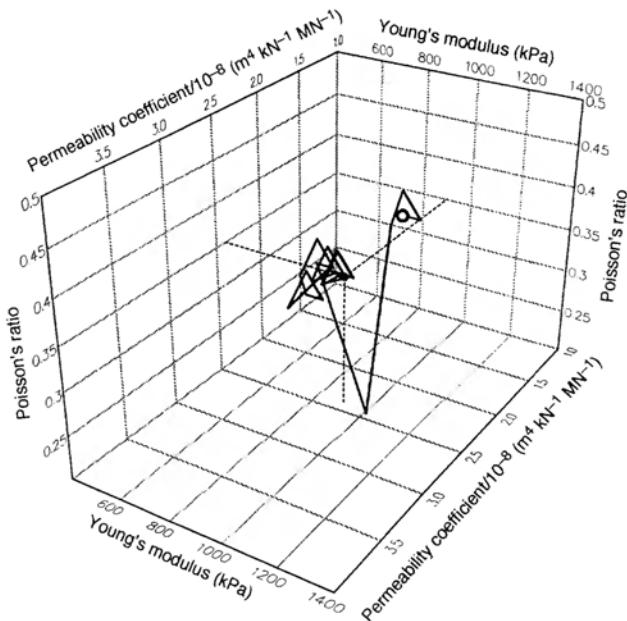
**Figure 13.6** Time-averaged sensitivity distribution for pore water pressure: (a) pore water pressure versus Young's modulus; (b) pore water pressure versus Poisson's ratio; (c) pore water pressure versus permeability



**Figure 13.7** Parameter identification of the first trial: the open circle denotes the starting-point of the search process; the filled circle denotes the parameters used in the ordinary deformation analysis



**Figure 13.8** Parameter identification of the second trial: the open circle denotes the starting-point of the search process; the filled circle denotes the parameters used in the ordinary deformation analysis



**Figure 13.9** Parameter identification of the third trial: the open circle denotes the starting-point of the search process; the filled circle denotes the parameters used in the ordinary deformation analysis

**Table 13.2** Starting points of the parameter identification processes

Trial no.	$E$ (kPa)	$\nu$	$(k/\gamma_w)/10^{-8} (\text{m}^4 \text{kN}^{-1} \text{min}^{-1})$
1	1400	0.3	3.0
2	1300	0.3	2.0
3	1000	0.4	2.0

The Levenberg–Marquardt algorithm has been employed as a search procedure in order to identify the necessary parameters. During the consolidation process, a series of pore water pressure values at point 1, horizontal displacements at point 100, and vertical displacements at point 139, have been used to determine the material parameters. Two weights of different values were assigned to the displacement contribution and to the pore water pressure contribution in (13.8), in order to balance these contributions. Figures 13.7 to 13.9 illustrate the results of the parameter identification from three trials; the starting-points of these identification processes are listed in Table 13.2.

## 13.6 SUMMARY

Back analysis in consolidation is of practical importance and has been receiving more and more attention from geotechnical engineers. Due to the ill-posed nature of the problem, special care should be taken to ensure the algorithm adequately. Sensitivity

analyses should be used to obtain a general idea on the optimal locations for measurements to be taken in order to improve the efficiency of the process.

## REFERENCES

1. Gioda, G. and Sakurai, S. (1987) Back analysis procedures for the interpretation of field measurements in geomechanics. *Int. J. Num. Anal. Meth. Geomech.*, **11**, 555–83.
2. Peck, R. B. (1969) Advantages and limitations of the observational method in applied soil mechanics. *Géotechnique*, **19**, 171–87.
3. Bjerrum, L. (1960) Some notes on Terzaghi's method of working, in *From Theory to Practice in Soil Mechanics*, Wiley, New York.
4. Gioda, G. (1985) Some remarks on back analysis and characterization problems in geomechanics, in *Proceedings of the 5th International Conference on Numerical Methods in Geomechanics*, Nagoya, pp. 47–61.
5. Murakami, A. and Hasegawa, T. (1987) Back analysis by Kalman filter–finite elements and a determination of optimal observed points location. *Proc. J. Soc. Civ. Engng.*, **388**, 227–35 (in Japanese).
6. Arai, K., Otha, H. and Kojima, K. (1984) Estimation of soil parameters based on monitored movement of subsoil under consolidation. *Soils and Foundations*, **24**, 95–108.
7. Arai, K., Otha, H., Kojima, K. and Wakasugi, M. (1986) Application of back-analysis to several test embankments on soft clay deposit. *Soils and Foundations*, **26**, 60–72.
8. Gioda, G. (1979) A numerical procedure for defining the values of soil parameters affecting consolidation, in *Proceedings of the 7th European Conference on SMFE*, Brighton, Vol. 1, pp. 169–72.
9. Asaoka, A. (1978) Observational procedure of settlement prediction. *Soils and Foundations*, **18**, 87–101.
10. Asaoka, A. and Matsuo, M. (1979) Bayesian approach to inverse problem in consolidation and its application to settlement prediction, in *Proceedings of the 3rd International Conference on Numerical Methods in Geomechanics*, Aachen, pp. 115–23.
11. Matsuo, M. and Asaoka, A. (1982) Bayesian calibration of embankment safety under earthquake loading. *Structural Safety*, **1**, 53–65.
12. Asaoka, A. and Matsuo, M. (1980) An inverse problem approach to settlement prediction. *Soils and Foundations*, **20**, 53–66.
13. Asaoka, A. and Matsuo, M. (1984) An inverse problem approach to the prediction of multi-dimensional consolidation behavior. *Soils and Foundations*, **24**, 49–62.
14. Murakami, A. and Hasegawa, T. (1985) Observational prediction of settlement using Kalman filter theory, in *Proceedings of the 5th International Conference on Numerical Methods in Geomechanics*, Nagoya, Vol. 3, pp. 1637–43.
15. Murakami, A. and Hasegawa, T. (1988) Back analysis using Kalman filter–finite elements and optimal location of observed points, in *Proceedings of the 6th International Conference on Numerical Methods in Geomechanics*, Innsbruck, Vol. 2, pp. 2051–58.
16. Schnur, D. S. and Zabaras, N. (1992) An inverse method for determining elastic material properties and a material interface. *Int. J. Num. Meth. Engng.*, **33**, 2039–57.
17. Davis, E. H. and Poulos, H. G. (1968) The use of elastic theory for settlement prediction under three-dimensional conditions. *Géotechnique*, **18**, 67–91.
18. Poulos, H. G., de Ambrosis, L. P. and Davis, E. H. (1976) Method of calculating long-term creep settlements. *J. Geotech. Engng Div. ASCE*, **102**, 787–805.

# 14

## Large-Strain Quasi-Static and Dynamic Soil Behaviour

### 14.1 INTRODUCTION

This chapter is devoted to the isothermal solution of dynamic and slow (quasi-static) phenomena in fully and partially saturated soil, under the assumption of finite strains. The model of the soil is described in Section 3.8, i.e. the air phase in the partially saturated zone always remains at atmospheric pressure, which is taken as the reference pressure. The assumption of finite strain is often important, especially in a partially saturated regime, as will be amplified by the examples.

An updated Lagrangian framework is used in this chapter, where the reference configuration is the last converged configuration of the solid phase. The addition of information needed to update the kinematic theory of Section 2.3.1 is shown in Section 14.2, and the necessary improvements to the constitutive relationships are dealt with in Section 14.3. The reader interested mainly in small-strain dynamics can move directly to Sections 14.4 and 14.6.

The approach adopted here is that presented by Zienkiewicz and coworkers for the dynamic analysis of fully saturated [1] and partially saturated [2] soil, and the extension to finite strain is attributable to Meroi and coworkers [3–5]. Several examples at the end of this chapter deal with dynamic strain localisation.

### 14.2 KINEMATIC EQUATIONS

The kinematic equations of Section 2.3.1 are now expanded for the purpose of this chapter. It is assumed that all functions which describe motion are continuous with their derivatives. Equation (2.35) can be rewritten as a function of the displacement vector  $\mathbf{u}^\pi(\mathbf{X}^\pi, t)$

$$\mathbf{x}^\pi = \mathbf{x}^\pi(\mathbf{X}^\pi, t) = \mathbf{X}^\pi + \mathbf{u}^\pi(\mathbf{X}^\pi, t) \quad (14.1)$$

Once the non-singularity of this Lagrangian relationship is stated, its inverse can be

written. The Jacobian  $J$  of the transformation is strictly positive and is given as

$$J = \det \begin{vmatrix} \frac{\partial \mathbf{x}_1}{\partial \mathbf{X}_j} & \frac{\partial \mathbf{x}_1}{\partial \mathbf{X}_2} & \frac{\partial \mathbf{x}_1}{\partial \mathbf{X}_3} \\ \frac{\partial \mathbf{x}_2}{\partial \mathbf{X}_1} & \frac{\partial \mathbf{x}_2}{\partial \mathbf{X}_2} & \frac{\partial \mathbf{x}_2}{\partial \mathbf{X}_3} \\ \frac{\partial \mathbf{x}_3}{\partial \mathbf{X}_1} & \frac{\partial \mathbf{x}_3}{\partial \mathbf{X}_2} & \frac{\partial \mathbf{x}_3}{\partial \mathbf{X}_3} \end{vmatrix} \quad (14.2)$$

This inverse results in the Eulerian description of motion

$$\mathbf{X}^\pi = \mathbf{X}^\pi(\mathbf{x}^\pi, t) = \mathbf{x}^\pi - \mathbf{u}^\pi(\mathbf{x}^\pi, t) \quad (14.3)$$

The deformation gradient tensor

$$\mathbf{F}^\pi = \text{Grad}^\pi \mathbf{x}^\pi \quad (14.4)$$

includes the rotation as well as the deformation components; in fact, by applying the polar decomposition theorem, we obtain

$$\mathbf{F}^\pi = \mathbf{R}^\pi \mathbf{U}^\pi = \mathbf{V}^\pi \mathbf{R}^\pi \quad (14.5)$$

where  $\mathbf{R}^\pi$  is the orthogonal rotation tensor. The symmetric and positive-definite tensors  $\mathbf{U}^\pi$  and  $\mathbf{V}^\pi$  are the right and left stretch tensors respectively, and they have the same eigenvalues; in fact

$$(\mathbf{U}^\pi)^2 = (\mathbf{F}^\pi)^T \mathbf{F}^\pi \quad (\mathbf{V}^\pi)^2 = \mathbf{F}^\pi (\mathbf{F}^\pi)^T \quad (14.6)$$

The adoption of stretch tensors as measures of deformation is not generally advantageous from a computational viewpoint. If the formulation is referred to the undeformed configuration, the symmetric Green–Lagrange strain tensor [6] is generally used:

$$\mathbf{E}^\pi = \frac{1}{2} [(\mathbf{F}^\pi)^T \mathbf{F}^\pi - \mathbf{I}] = \frac{1}{2} \left[ \left( \frac{d\mathbf{u}^\pi}{d\mathbf{X}^\pi} \right) + \left( \frac{d\mathbf{u}^\pi}{d\mathbf{X}^\pi} \right)^T + \left( \frac{d\mathbf{u}^\pi}{d\mathbf{X}^\pi} \right)^T \left( \frac{d\mathbf{u}^\pi}{d\mathbf{X}^\pi} \right) \right] \quad (14.7)$$

The corresponding tensor in the spatial approach is the Almansi–Eulerian strain tensor

$${}^*\mathbf{E}^\pi = \frac{1}{2} [\mathbf{I} - ((\mathbf{F}^\pi)^{-1})^T (\mathbf{F}^\pi)^{-1}] = \frac{1}{2} \left[ \left( \frac{d\mathbf{u}^\pi}{d\mathbf{x}^\pi} \right) + \left( \frac{d\mathbf{u}^\pi}{d\mathbf{x}^\pi} \right)^T - \left( \frac{d\mathbf{u}^\pi}{d\mathbf{x}^\pi} \right)^T \left( \frac{d\mathbf{u}^\pi}{d\mathbf{x}^\pi} \right) \right] \quad (14.8)$$

If we assume small dilatations  $\epsilon^\pi$ , the physical meaning of the Green–Lagrange strain tensor can easily be given as

$$\epsilon_i^\pi \cong E_{ii}^\pi = \frac{\partial u_i^\pi}{\partial X_i^\pi} + \frac{1}{2} \left[ \left( \frac{\partial u_1^\pi}{\partial X_i^\pi} \right)^2 + \left( \frac{\partial u_2^\pi}{\partial X_i^\pi} \right)^2 + \left( \frac{\partial u_3^\pi}{\partial X_i^\pi} \right)^2 \right] \quad (14.9)$$

Also, if the same hypothesis is assumed, then the angle change, if small, is given by

$$\gamma_{ij}^\pi \cong 2E_{ij}^\pi = \frac{\partial u_i^\pi}{\partial X_j^\pi} + \frac{\partial u_j^\pi}{\partial X_i^\pi} + \frac{\partial u_1^\pi}{\partial X_i^\pi} \frac{\partial u_1^\pi}{\partial X_j^\pi} + \frac{\partial u_2^\pi}{\partial X_i^\pi} \frac{\partial u_2^\pi}{\partial X_j^\pi} + \frac{\partial u_3^\pi}{\partial X_i^\pi} \frac{\partial u_3^\pi}{\partial X_j^\pi} \quad (14.10)$$

If the partial derivatives of the displacements with respect to the material coordinates  $\mathbf{X}$  are all small compared to unity, their squares and products may be neglected in comparison to the linear terms. The remaining terms give the small-strain tensor  $\boldsymbol{\epsilon}^\pi$ , which is used in the classical linearised theory of small strains and rotations  $\boldsymbol{\Omega}^\pi$ , where

$$\boldsymbol{\epsilon}^\pi = \frac{1}{2} \left[ \frac{\partial \mathbf{u}^\pi}{\partial \mathbf{X}^\pi} + \left( \frac{\partial \mathbf{u}^\pi}{\partial \mathbf{X}^\pi} \right)^T \right] \quad \boldsymbol{\Omega}^\pi = \frac{1}{2} \left[ \frac{\partial \mathbf{u}^\pi}{\partial \mathbf{X}^\pi} - \left( \frac{\partial \mathbf{u}^\pi}{\partial \mathbf{X}^\pi} \right)^T \right] \quad (14.11)$$

In these conditions, the Eulerian strain tensor  ${}^*\mathbf{E}^\pi$  also reduces to the form of the small-strain tensor, except that the derivatives are with respect to the spatial coordinates instead of the material coordinates. When the displacements and displacement gradients are sufficiently small, the distinction between the two small-strain definitions is usually ignored and the deformed configuration is virtually indistinguishable from the undeformed configuration, i.e.

$$\begin{aligned} \frac{\partial \mathbf{u}^\pi}{\partial \mathbf{x}^\pi} &= \mathbf{I} - (\mathbf{F}^\pi)^{-1} = \mathbf{I} - [\mathbf{I} + (\mathbf{F}^\pi - \mathbf{I})]^{-1} = \mathbf{I} - [\mathbf{I} - (\mathbf{F}^\pi - \mathbf{I}) + (\mathbf{F}^\pi - \mathbf{I})^2 + \dots] \\ &\cong (\mathbf{F}^\pi - \mathbf{I}) = \frac{\partial \mathbf{u}^\pi}{\partial \mathbf{X}^\pi} \end{aligned} \quad (14.12)$$

As far as deformations are concerned, attention will now be focused on the solid skeleton itself. If interest is devoted not only to the initial and final configurations, but also to the history of the deformation, then the so-called tangent motion must be evaluated in terms of the velocity components:

$$\mathbf{v}^s = \mathbf{v}^s(\mathbf{x}^s, t) = \dot{\mathbf{x}}^s(\mathbf{x}^s, t) = \dot{\mathbf{u}}^s(\mathbf{x}^s, t) \quad (14.13)$$

where, for the sake of brevity, the material time derivative is indicated by a dot. The deformation can be described by the velocity gradient tensor  $\mathbf{L}^s$ , as given in (2.50):

$$\mathbf{L}^s \equiv \text{grad } \mathbf{v}^s = \mathbf{D}^s + \mathbf{W}^s = (\dot{\mathbf{R}}^s \mathbf{U}^s + \mathbf{R}^s \dot{\mathbf{U}}^s)(\mathbf{R}^s \mathbf{U}^s)^{-1} \quad (14.14)$$

The symmetric part  $\mathbf{D}^s$  is the Eulerian strain rate tensor, which is related to pure straining as follows:

$$\mathbf{D}^s = \frac{1}{2} (\mathbf{L}^s + (\mathbf{L}^s)^T) = \frac{1}{2} \mathbf{R}^s [\dot{\mathbf{U}}^s (\mathbf{U}^s)^{-1} + (\mathbf{U}^s)^{-1} \dot{\mathbf{U}}^s] (\mathbf{R}^s)^T \quad (14.15)$$

while its skew-symmetric component  $\mathbf{W}^s$  in the spin tensor. This is commonly associated with the material rotation rate tensor  $\boldsymbol{\Theta}^s = \dot{\mathbf{R}}^s (\mathbf{R}^s)^T$  [7], giving the angular velocity of the material. But the spin tensor differs, from  $\boldsymbol{\Theta}^s$  and is given by the

following expression:

$$\mathbf{W}^s = \frac{1}{2} [\mathbf{L}^s - (\mathbf{L}^s)^T] = \dot{\mathbf{R}}^s (\mathbf{R}^s)^T + \frac{1}{2} \mathbf{R}^s [\dot{\mathbf{U}}^s (\mathbf{U}^s)^{-1} - (\mathbf{U}^s)^{-1} \dot{\mathbf{U}}^s] (\mathbf{R}^s)^T \quad (14.16)$$

### 14.3 CONSTITUTIVE EQUATIONS

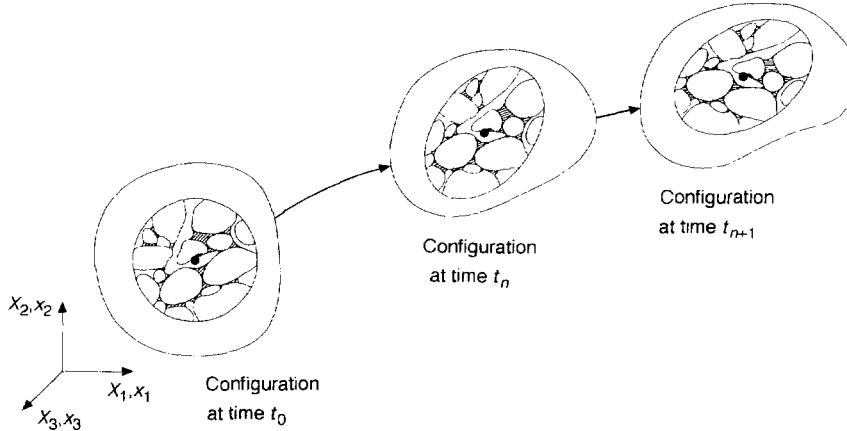
For the purpose of this chapter, we need to revisit the constitutive equations. It is assumed that, when we refer to the solid phase, the superscript  $s$  will no longer be required. When dealing with large deformation effects, which also involve large rotation, care must be taken in the material frame invariance of the constitutive law itself [6]. This can be obtained by expressing the hypoelastic relationship (Chapter 4) as a function of the objective fields. The constitutive law is here defined in terms of the objective Eulerian strain rate  $\mathbf{D}$  (14.15), and the associated objective stress rate measure defined by Jaumann, which is related to the rate of the Cauchy stress tensor by the non-objective spin tensor  $\mathbf{W}$  (14.16) via the well-known relationship

$${}^J\dot{\sigma} = \dot{\sigma} - \mathbf{W}\sigma + \sigma\mathbf{W} = \dot{\sigma} - \mathbf{W}\sigma - \sigma\mathbf{W}^T \quad (14.17)$$

Once the soil behaviour is defined in terms of the effective stress (equations 2.180 and 2.278) and the constitutive tangent tensor  $\mathbf{D}_T$  is introduced, the hypoelastic constitutive equation for the solid phase is

$${}^J\dot{\sigma}'' = \mathbf{D}_T[\mathbf{D}] \quad (14.18)$$

A more suitable stress rate measure would be the one introduced by Dienes [8] in order to eliminate the oscillating shear stress effects observed using the Jaumann stress rate [9, 10]. Dienes' stress rate differs from the Jaumann stress rate in that the material rotation



**Figure 14.1** Motion of body

rate  $\Theta$  is used instead of the non-objective spin tensor  $\mathbf{W}$ , which is also affected by pure stretching [11].

The Jaumann stress rate tensor can be correctly used as corotational measure associated with  $\mathbf{D}$  in an updated Lagrangian approach [7], if the strain increments are kept small enough during each step of the analysis. In this hypothesis,  $\mathbf{D}$  gives a suitable description of the strain rate, and  $\mathbf{W}$  accurately approximates the local angular velocity  $\Theta$ . In fact, in the reference configuration we have exactly that

$$\mathbf{F} = \mathbf{U} = \mathbf{R} = \mathbf{I} \quad \mathbf{D} = \dot{\mathbf{U}} \quad \mathbf{W} = \dot{\mathbf{R}} = \Theta \quad (14.19)$$

And at the end of the step it is true that

$$\mathbf{U} \cong \mathbf{I} \quad \mathbf{D} \cong \mathbf{R} \dot{\mathbf{U}} \mathbf{R}^T \quad \mathbf{W} \cong \Theta \quad (14.20)$$

An accurate large-deformation analysis can thus be developed by adopting Jaumann's measure for the stress rate plus the further condition that small deformations occur during each step. This formulation is advantageous from a computational viewpoint, since it can easily be implemented into a general finite element code as  $\mathbf{W}$  is immediately obtained once the deformation gradient is known. Alternatively the more correct approach of Dienes requires a polar decomposition to be performed at each stage of the analysis.

In an updated Lagrangian formulation, with small displacement gradients during each step, since we have  $\mathbf{F} \cong \mathbf{I}$  [6], we can write

$$\dot{\mathbf{E}} = \mathbf{F}^T \mathbf{D} \mathbf{F} \approx \mathbf{D} \quad (14.21)$$

and the incremental form of the constitutive law can be given in terms of the linear strain tensor increment, being  $d\mathbf{E} \cong d\boldsymbol{\epsilon}$  (equations 14.7 and 14.11).

This approach can also be used for elastoplastic analyses, and leads to a good approximation in the hypothesis of small elastic components of deformation. In fact, if we let  $\mathbf{X} \rightarrow \mathbf{x}$  be a given motion and if  $\mathbf{p}$  is the configuration corresponding to its irreversible component, i.e. the configuration reached by the body when unloaded, then  $\mathbf{X} \rightarrow \mathbf{x}_p$  and  $\mathbf{x}_p \rightarrow \mathbf{x}$  (Figure 14.2). The deformation gradient then has the following non-commutative representation for the plastic part and the elastic part [12]:

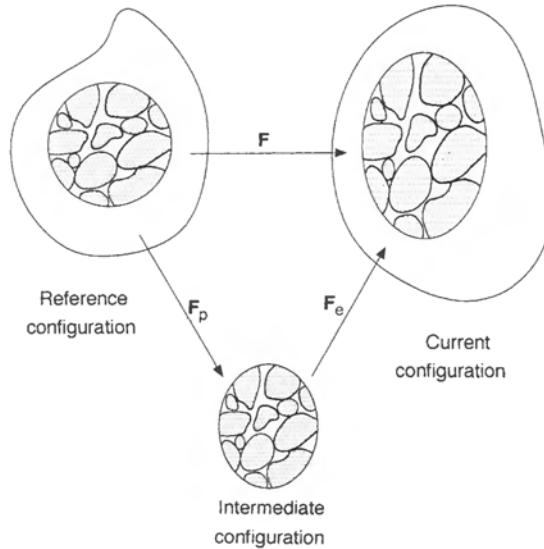
$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \frac{\partial \mathbf{x}}{\partial \mathbf{x}_p} \frac{\partial \mathbf{x}_p}{\partial \mathbf{X}} = \mathbf{F}_e \mathbf{F}_p \quad (14.22)$$

Since finite deformation measures are not expressed linearly in terms of displacements, the elastic and plastic components are generally not summable. The rate of deformation gradient is then as follows [13]:

$$\dot{\mathbf{F}} = \mathbf{L} \mathbf{F} \quad (14.23)$$

hence the general form of the velocity gradient is

$$\mathbf{L} = \dot{\mathbf{F}} \mathbf{F}^{-1} = \dot{\mathbf{F}}_e \mathbf{F}_e^{-1} + \mathbf{F}_e \dot{\mathbf{F}}_p \mathbf{F}_p^{-1} \mathbf{F}_e^{-1} = \mathbf{L}_e + \mathbf{F}_e \mathbf{L}_p \mathbf{F}_e^{-1} \quad (14.24)$$



**Figure 14.2** Kinematics of elastic-plastic deformation

If we choose a representation of the elastic part of the deformation that is independent of rigid motions ( $\mathbf{F}_e = \mathbf{V}_e$ ), then the deformation rate tensor is given by the symmetric part of the velocity gradient [13] in the form

$$\mathbf{D} = \mathbf{D}_e + \left( \frac{\mathbf{V}_e \mathbf{D}_p \mathbf{V}_e^{-1}}{2} + \frac{\mathbf{V}_e \mathbf{D}_p \mathbf{V}_e^{-1}}{2} \right) + \left( \frac{\mathbf{V}_e \mathbf{D}_p \mathbf{V}_e^{-1}}{2} - \frac{\mathbf{V}_e \mathbf{D}_p \mathbf{V}_e^{-1}}{2} \right) \quad (14.25)$$

Assuming small elastic strain components,  $\mathbf{V}_e \approx 1$ , leads to the desired summability [12]:

$$\mathbf{D} = \mathbf{D}_e + \mathbf{D}_p \quad (14.26)$$

Hence, in each step of an updated Lagrangian analysis with small strain increments, we have

$$d\boldsymbol{\varepsilon} \cong d\boldsymbol{\varepsilon}_e + d\boldsymbol{\varepsilon}_p \quad (14.27)$$

As far as the fluid phase is concerned, the constitutive equation governing the exchange of momentum between different phases can be expressed as a function of the average relative velocity  $\mathbf{v}^{\pi s}$ , once its frame invariance is proved. This variable depends on the volume fraction, which is objective, being a scalar, and it depends on the relative velocity  $\dot{\mathbf{r}}^{\pi s}$ , which is also frame indifferent. Its value, in any reference system moving relative to a chosen system with a rotational component given by the orthogonal tensor  $\mathbf{M}$ , is given by  $\mathbf{v}^{\pi s'} = \mathbf{M}\mathbf{v}^{\pi s}$  [3].

The symmetric permeability tensor  $\mathbf{k}^\pi$  depends on the orientation [14], changing for a rotation  $\mathbf{M}$  to

$$\mathbf{k}^{\pi'} = \mathbf{M}^T \mathbf{k}^\pi \mathbf{M} \quad (14.28)$$

If a finite deformation field arises, the permeability tensor should be further updated if it is assumed to be function of the void ratio  $e$  (Chapter 4). The void ratio  $e_0$  in the initial configuration is

$$e_0 = \frac{dv_0^f}{dv_0^s} = \frac{dv_0 - dv_0^s}{dv_0^s} = \frac{n_0}{1 - n_0} \quad (14.29)$$

which in the actual configuration becomes

$$e = \frac{n}{1 - n} = \frac{dv}{dv^s} - 1 = \frac{dv}{dv_0} \frac{dv_0}{dv_0^s} \frac{dv_0^s}{dv^s} - 1 = J(1 + e_0)J_s^{-1} - 1 \quad (14.30)$$

In an updated Lagrangian approach where the time-step increments are small enough to give an insignificant contribution from the second-order strain components during each step, the volume strain increment is satisfactorily approximated by the trace of the linear strain tensor ( $J \approx 1 + \text{tr } \boldsymbol{\epsilon}$ ) [4]. The current void ratio can then be evaluated by neglecting the specific contribution of grain deformation ( $J_s \approx 1$ ) to the overall volume strain, giving

$$e = (1 + \text{tr } \boldsymbol{\epsilon})(1 + e_0) - 1 = e_0 + (1 + e_0) \text{tr } \boldsymbol{\epsilon} \quad (14.31)$$

## 14.4 GOVERNING EQUATIONS AND THEIR WEAK FORM

Because we have chosen an updated Lagrangian reference system, we can neglect the convective terms in the balance equations. In terms of the appropriate total stress form, the equilibrium equation (2.262) is given by

$$-\rho \mathbf{a} - n S_w \rho^w \mathbf{a}^{ws} + \text{div } \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \quad (14.32)$$

where

$$\rho = (1 - n)\rho + n S_w \rho^w \quad (14.33)$$

After neglecting phase change (i.e.  $\dot{m} = 0$ ) and the gradient of water density, and assuming atmospheric pressure is the reference pressure, the mass balance equation for the water phase (2.253) becomes

$$\begin{aligned} & \left( \frac{(\alpha - n)}{K_s} S_w^2 + \frac{n S_w}{K_w} \right) p^w + \alpha S_w \text{div } \mathbf{v}^s + \left( \frac{(\alpha - n)}{K_s} \rho^w S_w + n \right) \dot{S}_w \\ & + \text{div} \left\{ \mathbf{k} \frac{k^{rw}}{\mu^w} [-\text{grad } p^w + \rho^w (\mathbf{g} - \mathbf{a}^w)] \right\} = 0 \end{aligned} \quad (14.34)$$

The associated boundary conditions are the traction boundary condition for the stresses:

$$\mathbf{t} = \bar{\mathbf{t}} \quad \text{on } \Gamma_u^q \quad (14.35)$$

and the flux boundary conditions for water:

$$q^w = \rho^w \mathbf{k} \frac{k^{rw}}{\mu^w} [-\nabla p^w + \rho^w(\mathbf{g} - \mathbf{a}^w)] \cdot \mathbf{n} \quad \text{on } \Gamma_w^q \quad (14.36)$$

The Dirichlet boundary conditions are of the standard form, i.e.

$$\begin{aligned} \mathbf{u} &= \hat{\mathbf{u}} && \text{on } \Gamma_u \\ \mathbf{v} &= \hat{\mathbf{v}} && \text{on } \Gamma_v \\ p^w &= \hat{p}^w && \text{on } \Gamma_w \end{aligned} \quad (14.37)$$

The initial conditions specify the full field of displacement and fluid pressure values at  $t = 0$ . It can immediately be verified that the weighted residual method, as applied in Section 3.4 to the equilibrium equation, yields the virtual power principle [15]. Since the virtual power equation (virtual velocities are used instead of virtual displacements) is advantageous for establishing the tangent stiffness matrix, we may write this equation directly to obtain an integral statement. The inner product  $(\boldsymbol{\sigma} : \boldsymbol{D})$  is therefore integrated over the body volume, where  $\boldsymbol{D}$  (virtual Eulerian strain rate tensor) is derived from the virtual velocity field  $\mathbf{v}$ , which satisfies kinematic compatibility and is unrelated to the real velocity field  $\mathbf{v}$ . The virtual power equation, after introduction of the effective stress, is given as follows:

$$\begin{aligned} &\int_{\Omega} [\boldsymbol{\sigma}'' - \mathbf{I}\alpha S_w p^w] : \boldsymbol{D} d\Omega - \int_{\Omega} \rho \mathbf{g} \cdot \mathbf{v} d\Omega - \int_{\Gamma_u^q} \bar{\mathbf{t}} \cdot \mathbf{v} da \\ &- \int_{\Omega} \rho \mathbf{a} \cdot \mathbf{v} d\Omega - \int_{\Omega} n \rho^w S_w \mathbf{a}^{ws} \cdot \mathbf{v} d\Omega = 0 \end{aligned} \quad (14.38)$$

Similarly, the weak form of the continuity equation for water (14.34) is obtained by multiplying this equation with the virtual pressure variation  $\delta p^w$ , which has first-derivative continuity and satisfies the natural boundary conditions. Finally, Green–Gauss's theorem is applied and yields

$$\begin{aligned} &\int_{\Omega} \delta p^w \left\{ \left( \frac{(\alpha - n)}{K_s} S_w^2 + \frac{n S_w}{K_w} \right) \dot{p}^w + \alpha S_w \operatorname{div} \mathbf{v}^s + \left( \frac{(\alpha - n)}{K_s} \rho^w S_w + n \right) \dot{S}_w \right\} d\Omega \\ &+ \int_{\Omega} \delta p^w \left[ \nabla (\delta p^w) \cdot \mathbf{k} \frac{k^{rw}}{\mu^w} [-\nabla p^w + \rho^w(\mathbf{g} - \mathbf{a}^w)] \right] d\Omega + \int_{\Gamma_w^q} \delta p^w \frac{q^w}{\rho^w} d\Gamma = 0 \end{aligned} \quad (14.39)$$

This procedure is a straightforward application of the virtual force principle and yields the same results as the weighted residual statement of Chapter 3.

## 14.5 THE RATE FORM OF STRESS POWER

The rate form of stress power [6] will now be introduced as it is useful in obtaining the tangent stiffness matrix, subsequently used in the iterative solution procedure of Section 14.7. The constitutive equations in Section 14.3 have been written in a rate form, hence they cannot be introduced directly into the virtual velocity equations.

The rate form of the stress power  $\dot{\Pi}$  is given by

$$\dot{\Pi} = \frac{D}{Dt} \int_{\Omega} \boldsymbol{\sigma} : \mathbf{D} d\Omega = \int_{\Omega} \left( \dot{\boldsymbol{\sigma}} : \mathbf{D} + \boldsymbol{\sigma} : \dot{\mathbf{D}} + \boldsymbol{\sigma} : \mathbf{D} \frac{\dot{J}}{J} \right) d\Omega \quad (14.40)$$

In this equation, because of the symmetry of the Cauchy stress tensor,  $\dot{\mathbf{D}}$  can be replaced by the rate of the virtual velocity gradient  $\dot{\mathbf{L}}$ , which from (14.24) is given as

$$\dot{\mathbf{L}} = \frac{d}{dt} \left( \frac{\partial \mathbf{v}}{\partial \mathbf{X}} \frac{\partial \mathbf{X}}{\partial \mathbf{x}} \right) = \frac{\partial \mathbf{v}}{\partial \mathbf{X}} \frac{d(\mathbf{F}^{-1})}{dt} = - \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \dot{\mathbf{F}} \mathbf{F}^{-1} - \mathbf{L} \mathbf{L} \quad (14.41)$$

We have taken into account that both the placement  $\mathbf{X}$  and the virtual velocity  $\mathbf{v}$  are not functions of time and the rate of the deformation gradient is given by (14.23). Using  $J = J \operatorname{tr} \mathbf{D}$  [3], equation (14.40) can be written as follows:

$$\dot{\Pi} = \int_{\Omega} (\dot{\boldsymbol{\sigma}} : \mathbf{D} - \boldsymbol{\sigma} : \mathbf{L} \mathbf{L} + (\operatorname{tr} \mathbf{D}) \boldsymbol{\sigma} : \mathbf{D}) d\Omega \quad (14.42)$$

The Jaumann stress rate (14.17) is now introduced. Since  $\boldsymbol{\sigma}$  and  $\mathbf{D}$  are symmetric tensors ( $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$ ) and  $\mathbf{D} = \mathbf{D}^T$ ) and  $\mathbf{W}$  is a skew tensor ( $\mathbf{W} = -\mathbf{W}^T$ ), we have

$$(\boldsymbol{\sigma} \mathbf{W}) : \mathbf{D} = (\boldsymbol{\sigma} \mathbf{W})^T : \mathbf{D}^T = (-\mathbf{W} \boldsymbol{\sigma}) : \mathbf{D} = -[(\mathbf{L} - \mathbf{D}) \boldsymbol{\sigma}] : \mathbf{D} = (\mathbf{D} \boldsymbol{\sigma}) : \mathbf{D} - (\mathbf{D} \mathbf{L}) : \boldsymbol{\sigma} \quad (14.43)$$

The Eulerian rate of stress power can finally be written as

$$\dot{\Pi} = \int_{\Omega} [(\mathbf{J} \dot{\boldsymbol{\sigma}} - 2\mathbf{D} \boldsymbol{\sigma}) : \mathbf{D} + \boldsymbol{\sigma} : \mathbf{L}^T \mathbf{L} + (\operatorname{tr} \mathbf{D}) \boldsymbol{\sigma} : \mathbf{D}] d\Omega \quad (14.44)$$

This equation could also be derived from the total Lagrangian rate equation for the particular case of instantaneously coincident reference and current configurations [16]. It is adopted in the updated Lagrangian formulation used in this chapter.

## 14.6 FINITE ELEMENT DISCRETISATION

Before carrying out the finite element discretisation, some further simplifications will be introduced. If high-frequency dynamic phenomena are not considered, all the terms linked to the relative component of fluid acceleration  $\mathbf{a}^{ws}$  can be ignored. This assumption is usually acceptable even in earthquake analyses [1].

Although these simplifications may significantly reduce the computational effort, they have a minimal influence on the results. The number of principal variables can be reduced to four (three displacements and the water pressure), then the system to be solved consists of the weak form of the momentum balance equation for the whole multiphase medium:

$$\int_{\Omega} [\boldsymbol{\sigma}'' - \mathbf{I}\alpha S_w p^w] : \mathbf{D} d\Omega - \int_{\Gamma_u} \mathbf{t} \cdot \mathbf{v} d\Gamma - \int_{\Omega} \rho \mathbf{g} \cdot \mathbf{v} d\Omega - \int_{\Omega} \rho \mathbf{a} \cdot \mathbf{v} d\Omega = 0 \quad (14.45)$$

and the weak form of the continuity equation for water:

$$\begin{aligned} & \int_{\Omega} \left[ \delta p^w \left( \alpha S_w \operatorname{div} \mathbf{v} + \frac{1}{Q^*} \dot{p}^w \right) + \operatorname{grad} (\delta p^w) \cdot \mathbf{k} \frac{k^{rw}}{\mu^w} [-\operatorname{grad} p^w + \rho^w (\mathbf{g} - \mathbf{a}^w)] \right] d\Omega \\ & + \int_{\Gamma_w} \delta p^w \frac{q^w}{\rho^w} d\Gamma = 0 \end{aligned} \quad (14.46)$$

where

$$\frac{1}{Q^*} = C_s + \frac{nS_w}{K_w} + \frac{(\alpha - n)}{K_s} S_w \left( S_w + \frac{C_s}{n} p^w \right) \quad (14.47)$$

Furthermore, the effective stress principle is valid as given by (3.65):

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}'' - \mathbf{I}\alpha S_w p^w \quad (14.48)$$

and from (2.181) the constitutive equation becomes

$$d\boldsymbol{\sigma}'' = \mathbf{D}_T [d\boldsymbol{\epsilon} - d\boldsymbol{\epsilon}_0] + d\Omega \boldsymbol{\sigma}'' - \boldsymbol{\sigma}'' d\Omega \quad (14.49)$$

because for small strain increments in each step, then  $d\boldsymbol{\epsilon} = \mathbf{D} dt$  (from 14.21) and  $d\Omega = \mathbf{W} dt$  (from 14.11 and 14.16). Finally, we need the capillary pressure relationship in (3.67) and (3.68).

#### 14.6.1 Spatial discretisation

For the spatial discretisation, we shall adopt the vector notation of Chapter 3. The unknown field variables  $\mathbf{u}$  and  $p^w$  are expressed in the whole domain as functions of the nodal values,  $\bar{\mathbf{u}}$  and  $\bar{p}^w$ , and also by the global shape function matrices,  $\mathbf{N}_u$  and  $\mathbf{N}_p$

$$\mathbf{u} = \mathbf{N}_u \bar{\mathbf{u}} \quad p^w = \mathbf{N}_p \bar{p}^w \quad (14.50)$$

This is also true for virtual displacement and pressure values.

The strain operator  $\mathbf{B}_0$  relates the strain rate vector with the vector of nodal velocities  $\bar{\mathbf{v}}$ :

$$\mathbf{D} = \mathbf{B}_0 \bar{\mathbf{v}} \quad (14.51)$$

And with reference to the last converged configuration in the updated Lagrangian approach we have

$$\mathbf{B}_{0i}^T = \begin{bmatrix} \frac{\partial N_{u_i}}{\partial x_1} & 0 & 0 & 0 & \frac{\partial N_{u_i}}{\partial x_3} & \frac{\partial N_{u_i}}{\partial x_2} \\ 0 & \frac{\partial N_{u_i}}{\partial x_2} & 0 & \frac{\partial N_{u_i}}{\partial x_3} & 0 & \frac{\partial N_{u_i}}{\partial x_1} \\ 0 & 0 & \frac{\partial N_{u_i}}{\partial x_3} & \frac{\partial N_{u_i}}{\partial x_2} & \frac{\partial N_{u_i}}{\partial x_1} & 0 \end{bmatrix} \quad (14.52)$$

When field variables are expressed as functions of the shape function matrices, the equilibrium equation (14.45) becomes

$$\int_{\Omega} \mathbf{B}_0^T \boldsymbol{\sigma}'' d\Omega - \mathbf{Q} \bar{\mathbf{p}}^w + \mathbf{M} \ddot{\mathbf{u}} = \mathbf{f}^u \quad (14.53)$$

where  $\mathbf{Q}$  is the coupling matrix:

$$\mathbf{Q} = \int_{\Omega} \mathbf{B}_0^T S_w \mathbf{m} \mathbf{N}_p d\Omega \quad (14.54)$$

$\mathbf{M}$  is the mass matrix:

$$\mathbf{M} = \int_{\Omega} \mathbf{N}_u^T [\rho^s(1-n) + \rho^w n S_w] \mathbf{N}_u d\Omega \quad (14.55)$$

and  $\mathbf{f}^u$  is the external load vector:

$$\mathbf{f}^u = \int_{\Omega} \mathbf{N}_u^T [\rho^s(1-n) + \rho^w n S_w] \mathbf{g} d\Omega + \int_{\Gamma} \mathbf{N}_u^T \bar{\mathbf{t}} d\Gamma \quad (14.56)$$

Except for one additional matrix, the discretisation of the continuity equation is the same as in Section 3.9 and it yields the following form:

$$\mathbf{H} \bar{\mathbf{p}}^w + \mathbf{G}^* \ddot{\mathbf{u}} + \mathbf{Q}^T \dot{\mathbf{u}} + \mathbf{S} \dot{\mathbf{p}}^w = \mathbf{f}^p \quad (14.57)$$

where  $\mathbf{H}$  is the permeability matrix:

$$\mathbf{H} = \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rw}}{\mu^w} \nabla \mathbf{N}_p d\Omega \quad (14.58)$$

$\mathbf{G}^*$  is the dynamic seepage matrix:

$$\mathbf{G}^* = \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rw}}{\mu^w} \rho^w \mathbf{N}_u d\Omega \quad (14.59)$$

$\mathbf{S}$  is the compressibility matrix:

$$\mathbf{S} = \int_{\Omega} \mathbf{N}_p^T \frac{1}{Q^w} \mathbf{N}_p d\Omega \quad (14.60)$$

and  $\mathbf{f}^p$  is the flow vector:

$$\mathbf{f}^p = \int_{\Omega} (\nabla \mathbf{N}_p)^T \frac{\mathbf{k} k^{rw}}{\mu^w} p^w \mathbf{g} d\Omega - \int_{\Gamma_w} \mathbf{N}_p^T \frac{q^w}{p^w} d\Gamma \quad (14.61)$$

Since the effect of the dynamic seepage forcing term is negligible [1], the coupled system at time  $t_{n+1}$  can be written as

$$\begin{cases} \mathbf{M}_{n+1} \ddot{\mathbf{u}}_{n+1} + \bar{\mathbf{P}}_{n+1} - \mathbf{Q}_{n+1} \bar{\mathbf{p}}_{n+1} = \mathbf{f}^u_{n+1} \\ \mathbf{Q}_{n+1} \dot{\bar{\mathbf{u}}}_{n+1} + \mathbf{H}_{n+1} \bar{\mathbf{p}}_{n+1} + \mathbf{S}_{n+1} \dot{\bar{\mathbf{p}}}_{n+1} = \mathbf{f}^p_{n+1} \end{cases} \quad (14.62)$$

where  $\bar{\mathbf{P}}_{n+1}$  is the equivalent nodal force vector.

#### 14.6.2 Discretisation in time and solution procedure

In the equilibrium equation there appear second-order time derivatives, and for the time discretisation a slightly different method is adopted than the method in Chapter 3. The system of equations (14.62) has been written at a time level  $t_{n+1}$ . A relation is employed that links the variables at  $t_{n+1}$  with their values at  $t_n$ . For this purpose the Newmark scheme is adopted, with the lowest allowable order for each variable: second order for displacements and first order for pressure

$$\begin{aligned} \dot{\bar{\mathbf{u}}}_{n+1} &= \dot{\bar{\mathbf{u}}}_n + \ddot{\bar{\mathbf{u}}}_n \Delta t + \beta_1 \Delta \ddot{\bar{\mathbf{u}}}_n \Delta t = \dot{\bar{\mathbf{u}}}_{n+1}^p + \beta_1 \Delta \ddot{\bar{\mathbf{u}}}_n \Delta t \\ \bar{\mathbf{u}}_{n+1} &= \bar{\mathbf{u}}_n + \dot{\bar{\mathbf{u}}}_n \Delta t + \frac{\ddot{\bar{\mathbf{u}}}_n \Delta t}{2} + \frac{\beta_2 \Delta \ddot{\bar{\mathbf{u}}}_n \Delta t^2}{2} \\ \bar{\mathbf{p}}_{n+1} &= \bar{\mathbf{p}}_n + \dot{\bar{\mathbf{p}}}_n \Delta t + \theta \Delta \dot{\bar{\mathbf{p}}}_n \Delta t \end{aligned} \quad (14.63)$$

Insertion of (14.63) into (14.62) leads to the following expression for the coupled system:

$$\begin{cases} \Psi_{n+1}^u = \mathbf{M}_{n+1} \Delta \ddot{\bar{\mathbf{u}}}_n + \bar{\mathbf{P}}_{n+1} - \mathbf{Q}_{n+1} \theta \Delta t \Delta \dot{\bar{\mathbf{p}}}_n - \mathbf{F}_{n+1}^u = 0 \\ \Psi_{n+1}^p = \mathbf{Q}_{n+1}^T \beta_1 \Delta t \Delta \ddot{\bar{\mathbf{u}}}_n + \mathbf{H}_{n+1} \theta \Delta t \Delta \dot{\bar{\mathbf{p}}}_n + \mathbf{S}_{n+1} \Delta \dot{\bar{\mathbf{p}}}_n - \mathbf{F}_{n+1}^p = 0 \end{cases} \quad (14.64)$$

with

$$\mathbf{F}_{n+1}^u = \mathbf{f}_{n+1}^u - \mathbf{M}_{n+1} \ddot{\bar{\mathbf{u}}}_n + \mathbf{Q}_{n+1} (\bar{\mathbf{p}}_n + \Delta t \dot{\bar{\mathbf{p}}}_n) \quad (14.65)$$

$$\mathbf{F}_{n+1}^p = \mathbf{f}_{n+1}^p - \mathbf{Q}_{n+1}^T (\dot{\bar{\mathbf{u}}}_n + \Delta t \ddot{\bar{\mathbf{u}}}_n) - \mathbf{H}_{n+1} (\bar{\mathbf{p}}_n + \Delta t \dot{\bar{\mathbf{p}}}_n) - \mathbf{S}_{n+1} \dot{\bar{\mathbf{p}}}_n \quad (14.66)$$

At the beginning of each time step, the equivalent nodal force vector  $\bar{\mathbf{P}}_{n+1}$  must be evaluated by integration of the constitutive law, the stress field at the previous step being known. The relative permeability and specific capacity values must also be updated, as well as the void ratio and the absolute permeability (if it varies with the void ratio).

The non-linear coupled system (14.64) is solved by an iterative procedure. If a Newton–Raphson scheme is adopted to linearise the problem (Chapter 4), the Jacobian matrix of transformation  $\mathbf{J}$ , at the  $i$ th-iteration is

$$\mathbf{J} = \frac{\partial \Psi}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}_i} = \begin{bmatrix} \frac{\partial \Psi''}{\partial (\Delta \ddot{\mathbf{u}})} & \frac{\partial \Psi''}{\partial (\Delta \dot{\mathbf{p}})} \\ \frac{\partial \Psi''}{\partial (\Delta \ddot{\mathbf{u}})} & \frac{\partial \Psi''}{\partial (\Delta \dot{\mathbf{p}})} \end{bmatrix} = \begin{bmatrix} \mathbf{M} + \frac{1}{2} \mathbf{K}_T \beta_2 \Delta t^2 & -\mathbf{Q} \theta \Delta t \\ \mathbf{Q}^T \beta_1 \Delta t & \mathbf{H} \theta \Delta t + \mathbf{S} \end{bmatrix} \quad (14.67)$$

In a small-strain situation  $\mathbf{K}_T$  is given by (3.44) and (4.30). In a finite-strain situation the Eulerian rate of stress power (14.42) can be expressed in matrix form as

$$\dot{\Pi} = \int_V [\mathbf{D}^T \mathbf{J} \dot{\boldsymbol{\sigma}}'' - \mathbf{D}^T \boldsymbol{\sigma}_{d1} \mathbf{D} + \mathbf{L}^T \dot{\boldsymbol{\sigma}} \mathbf{L} + \mathbf{D}^T \boldsymbol{\sigma}_{d2} \mathbf{D}] d\Omega \quad (14.68)$$

The matrices  $\boldsymbol{\sigma}_{d1}$  and  $\boldsymbol{\sigma}_{d2}$  are given below. The discretised form of the Eulerian rate of stress power, bearing in mind (14.52), is given as

$$\dot{\Pi} = \mathbf{K}_T \bar{\mathbf{v}} \quad (14.69)$$

The tangent stiffness matrix  $\mathbf{K}_T$  is then given by

$$\mathbf{K}_T = \int_V [\mathbf{B}_0^T \tilde{\mathbf{C}} \mathbf{B}_0 + \mathbf{G}_r^T \hat{\boldsymbol{\sigma}} \mathbf{G}_r] d\Omega = \mathbf{K} + \mathbf{K}_\sigma \quad (14.70)$$

The tangent stiffness matrix is therefore equal to the sum of two terms. The first one corresponds to the small-strain stiffness matrix

$$\mathbf{K} = \int_V \mathbf{B}_0^T \tilde{\mathbf{C}} \mathbf{B}_0 d\Omega \quad (14.71)$$

but is referred to the constitutive matrix  $\tilde{\mathbf{C}}$ , modified by the Cauchy stresses at time  $t$ , as

$$\tilde{\mathbf{C}} = \mathbf{D}_T - \boldsymbol{\sigma}_{d1} + \boldsymbol{\sigma}_{d2} = \mathbf{D}_T - \boldsymbol{\sigma}_d \quad (14.72)$$

where

$$\boldsymbol{\sigma}_{d1} = \begin{bmatrix} 2\sigma''_{11} & 0 & 0 & 0 & \sigma''_{13} & \sigma''_{12} \\ 0 & 2\sigma''_{22} & 0 & \sigma''_{23} & 0 & \sigma''_{21} \\ 0 & 0 & 2\sigma''_{33} & \sigma''_{32} & \sigma''_{31} & 0 \\ 0 & \sigma''_{32} & \sigma''_{23} & \bar{\sigma}_{23} & \sigma''_{21} & \sigma''_{31} \\ \sigma''_{31} & 0 & \sigma''_{13} & \sigma''_{12} & \bar{\sigma}_{13} & \sigma''_{32} \\ \sigma''_{21} & \sigma''_{12} & 0 & \sigma''_{13} & \sigma''_{23} & \bar{\sigma}_{13} \end{bmatrix} \quad \text{with } \bar{\sigma}_{ij} = \frac{\sigma''_{ii} + \sigma''_{jj}}{2} \quad (14.73)$$

$$\boldsymbol{\sigma}_{d2} = \begin{bmatrix} \sigma''_{11} & \sigma''_{11} & \sigma''_{11} & 0 & 0 & 0 \\ \sigma''_{22} & \sigma''_{22} & \sigma''_{22} & 0 & 0 & 0 \\ \sigma''_{33} & \sigma''_{33} & \sigma''_{33} & 0 & 0 & 0 \\ \sigma''_{23} & \sigma''_{23} & \sigma''_{23} & 0 & 0 & 0 \\ \sigma''_{13} & \sigma''_{13} & \sigma''_{13} & 0 & 0 & 0 \\ \sigma''_{12} & \sigma''_{12} & \sigma''_{12} & 0 & 0 & 0 \end{bmatrix} \quad (14.74)$$

The second term is called the initial stress stiffness matrix  $\mathbf{K}_\sigma$ , and is equal to

$$\mathbf{K}_\sigma = \int_V \mathbf{G}_r^T \hat{\boldsymbol{\sigma}} \mathbf{G}_r d\Omega \quad (14.75)$$

where

$$\hat{\boldsymbol{\sigma}} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \hat{\sigma}_{13} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} & \hat{\sigma}_{23} \\ \hat{\sigma}_{31} & \hat{\sigma}_{32} & \hat{\sigma}_{33} \end{bmatrix} \quad \text{with} \quad \hat{\sigma}_{ij} = \sigma''_{ii} \mathbf{I} \quad (14.76)$$

and

$$\mathbf{G}_r^T = \left[ \frac{\partial N_{u_i}}{\partial X_1} \mathbf{I} \quad \frac{\partial N_{u_i}}{\partial X_2} \mathbf{I} \quad \frac{\partial N_{u_i}}{\partial X_3} \mathbf{I} \right] \quad (14.77)$$

Since the matrix  $\boldsymbol{\sigma}_{d2}$  is not symmetric,  $\mathbf{K}_T$  is not symmetric, but this is usually not important. When elastic behaviour only is considered, the elements of the constitutive matrix  $\mathbf{D}_T$  have the same order of magnitude as the elastic moduli, so the elements of  $\boldsymbol{\sigma}_d$  are small in comparison. However, they must be included in any plastic analysis, when using materials with a small amount of strain hardening, since they can be of the same order even for small deformation gradients [16].

The system to be solved can then be written in the following form, which is symmetric providing the tangent stiffness matrix is itself symmetric:

$$\begin{bmatrix} \mathbf{M} + \frac{1}{2} \mathbf{K}_T \beta_2 \Delta t^2 & -\mathbf{Q} \theta \Delta t \\ -\mathbf{Q}^T \theta \Delta t & -\frac{\theta}{\beta_1} (\mathbf{H} \theta \Delta t + \mathbf{S}) \end{bmatrix} \begin{bmatrix} \Delta \ddot{\mathbf{u}} \\ \Delta \dot{\mathbf{p}} \end{bmatrix} = - \begin{bmatrix} -\Psi^u \\ -\frac{\theta}{\beta_1} \Psi^p \end{bmatrix} \quad (14.78)$$

Since the Newton–Raphson method requires the Jacobian matrix to be evaluated and inverted at each iteration, other modified schemes may also be used to achieve convergence with less computational effort. In particular, secant updates, like the Davidon method and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method, are advantageous in non-linear analyses [17].

Stability and accuracy analyses for dynamic problems can then be carried out, as outlined in Chapter 3. For an extensive outline of recent developments, the reader is referred to Zienkiewicz *et al.* [17]. It can be shown that unconditional stability of the

solution procedure for linear problems requires that the parameters  $\beta_1$ ,  $\beta_2$  and  $\theta$  (chosen in the range [0–1] are as follows:

$$\beta_2 \geq \beta_1 \geq 0.5 \quad \text{and} \quad \theta \geq 0.5 \quad (14.79)$$

The general solution scheme also incorporates special cases without requiring any modification and for no loss of computational efficiency. Thus, for quasi-static problems or long-term phenomena, e.g. soil consolidation, any inertial effects are negligible and in (14.53) and (14.57) all acceleration terms are omitted and the mass matrix is set to zero. In pure statics the time variable is retained, but is used merely as an artificial variable allowing load incrementation. If the solution of either static or dynamic undrained analyses is required, then the permeability and compressibility matrices are set to zero.

## 14.7 EXAMPLES

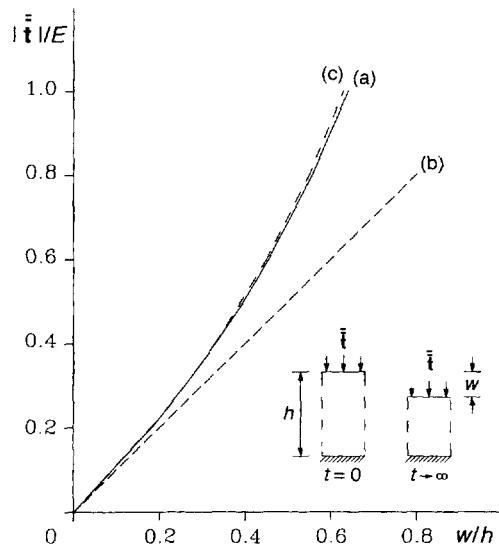
Several examples are shown in this section, which demonstrate a few of the many problems that can be solved with the model discussed in this chapter. These examples refer to finite-strain and small-strain situations; the first two examples compare the two assumptions, and where necessary, they amplify the importance of finite-strain analysis. Rather advanced problems, such as strain localisation, are also addressed in this section. The examples were solved by the finite element code Swandyne [1–5]. This code assumes that pore pressures are positive if in tension, opposite to our earlier convention. Bear this in mind, when interpreting pore pressure values in the graphs.

### 14.7.1 Finite-strain and small-strain fully saturated consolidation

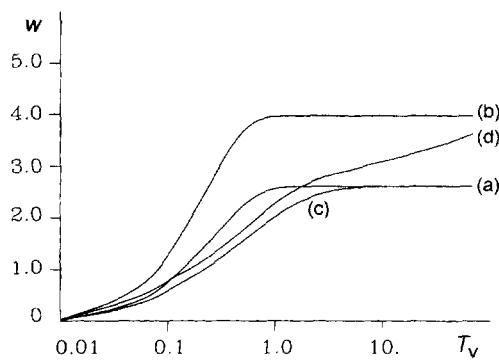
The first example [3,4] refers to the consolidation analysis of a 10 m deep ground, fully saturated by water and indefinitely extending in the horizontal direction. A step load is applied at the top surface, with drainage being allowed only via this surface. The problem, which can be considered as one-dimensional with each vertical section as a plane of symmetry, is modelled as a saturated soil column in plane-strain conditions. The following boundary conditions are imposed for the displacement field: all nodes are horizontally constrained and the nodes at the bottom level are also vertically fixed. The pressure at the top level is assumed to be atmospheric and the lateral surfaces are impermeable. Different load levels are considered, with a maximum intensity of the uniform distributed load being equal to the Young's modulus of the ground. Porosity  $n = 0.3$ , isotropic permeability  $k = 0.01 \text{ m s}^{-1}$ , elastic modulus  $E = 1 \text{ GPa}$  and Poisson's ratio  $\nu = 0$ .

The logarithmic theoretical relationship between the applied load and the final displacement is represented in Figure 14.3. The computational results (curve a) are in very good agreement with this theoretical behaviour (curve c). Curve b represents the results of a geometrically linear analysis.

A linear dependency of the intrinsic permeability on the value of the void ratio is also considered, varying from the initial value to zero when the porosity becomes zero. Figure 14.4 plots the maximum vertical displacement versus the normalized time  $T_v$  for



**Figure 14.3** Vertical settlement versus load level: (a) numerical finite deformation results; (b) linear analysis response; (c) theoretical solution in the finite deformation regime



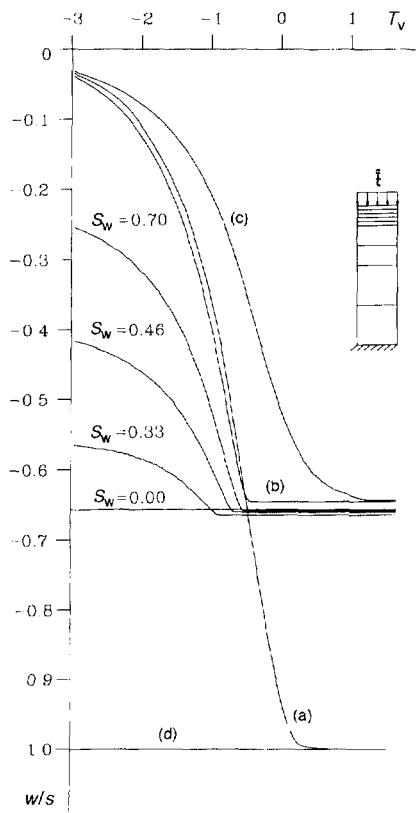
**Figure 14.4** Vertical settlement versus normalised time: (a) load =  $0.3 E$ , constant permeability; (b) load =  $0.5 E$ , constant permeability; (c) load =  $0.3 E$ , variable permeability; (d) load =  $0.5 E$ , variable permeability

different load levels ( $T_v = c_v t$ , and  $c_v = kE/\gamma^w h^2$  is the time factor, where  $\gamma^w$  is the specific weight of the pore fluid and  $h$  is the height of the soil column). The degree of consolidation with a constant and a variable permeability are described respectively by curves a and c for the load level equal to  $0.3E$  and by curves b and d for the load level of

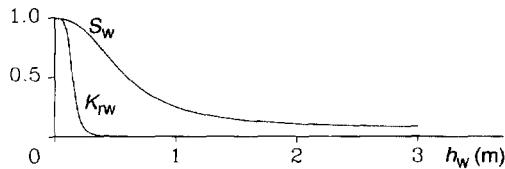
$0.5E$ . Notice the significant increase of time (in logarithmic scale on the figure) necessary to reach full consolidation.

#### 14.7.2 Finite-strain and small-strain partially saturated consolidation

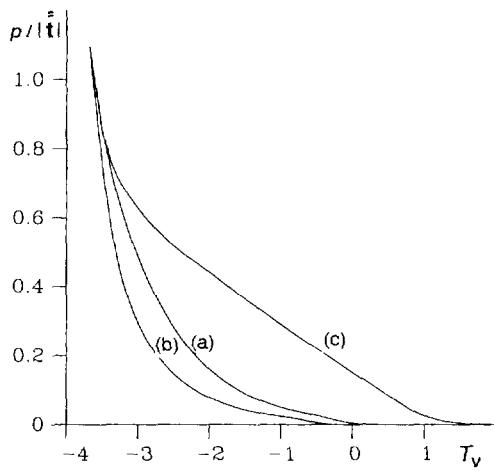
A second example [3,4] is where partial saturation conditions are also considered. A one-dimensional column, 7 m high, is modelled in accordance with Kim *et al.* [18] using plane-strain elements, adopting a material with porosity  $n = 0.5$ , isotropic permeability  $k = 4 \times 10^{-7} \text{ m s}^{-1}$  elastic modulus  $E = 6 \text{ GPa}$  and Poisson's ratio  $\nu = 0.4$ . A uniformly distributed load is applied in the vertical direction up to a maximum value of 10 GPa.



**Figure 14.5** Normalised vertical settlement versus normalised time for the illustrated model: (a) small deformation analysis, fully saturated; (b) large deformation analysis, constant permeability, fully saturated; (c) large deformation analysis, variable permeability, fully saturated; (d) small deformation analysis, partially saturated



**Figure 14.6** Saturation ( $S_w$ ) and relative permeability ( $K_{rw}$ ) versus hydraulic load

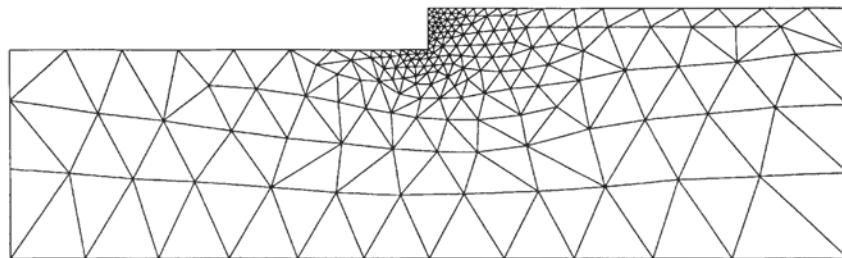


**Figure 14.7** Transient water pressure at the top, normalised with respect to the applied load under fully saturated conditions: (a) small deformation analysis; (b) large deformation analysis, constant permeability; (c) large deformation analysis, variable permeability

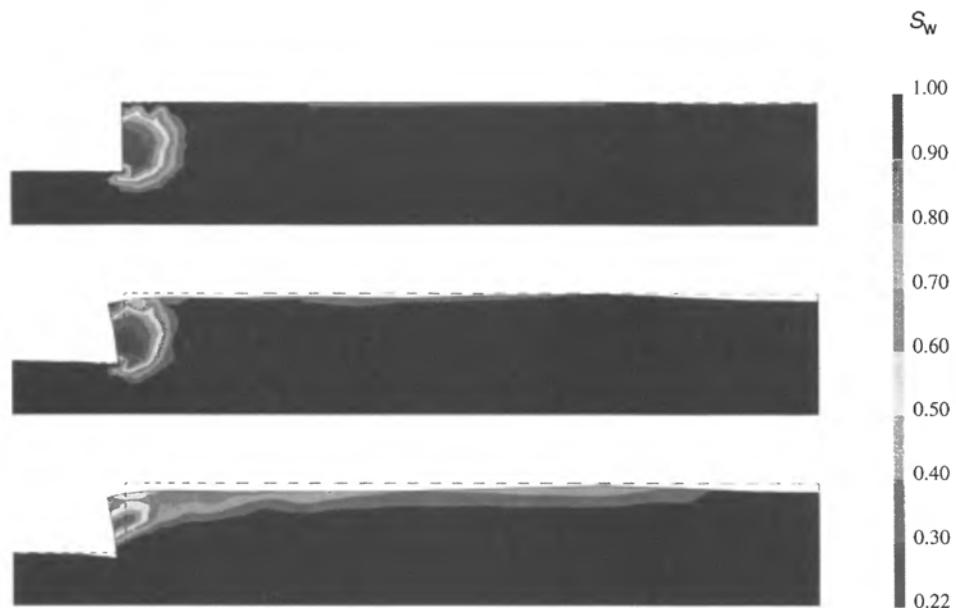
Figure 14.5 plots the vertical settlement, normalised with respect to the corresponding values of Terzaghi's theory, versus normalised time for the case of fully saturated conditions for small deformations (curve a) and large deformations (curve b). The large deformation analysis was also performed for a permeability which had a linear dependence on the void ratio (curve c). The large deformation results for three different homogeneous partially saturated initial conditions are also given. These partially saturated initial conditions were imposed by assigning the initial capillary pressure distribution corresponding to the saturation degree via the capillary pressure relationship of Figure 14.6. The small deformation analysis, with air pressure equal to the atmospheric pressure, gives the final results at the first time steps (curve d), whereas the large deformation analysis leads to a real consolidation process. The dissipation of water pressure for the fully saturated case at the upper element is shown in Figure 14.7.

#### 14.7.3 Slope under seismic behaviour: finite strains

A further example [4.5] refers to the elastoplastic large-strain behaviour of a 9.15 m vertical slope subjected to gravitational loads during the first 25 s of the El-Centro



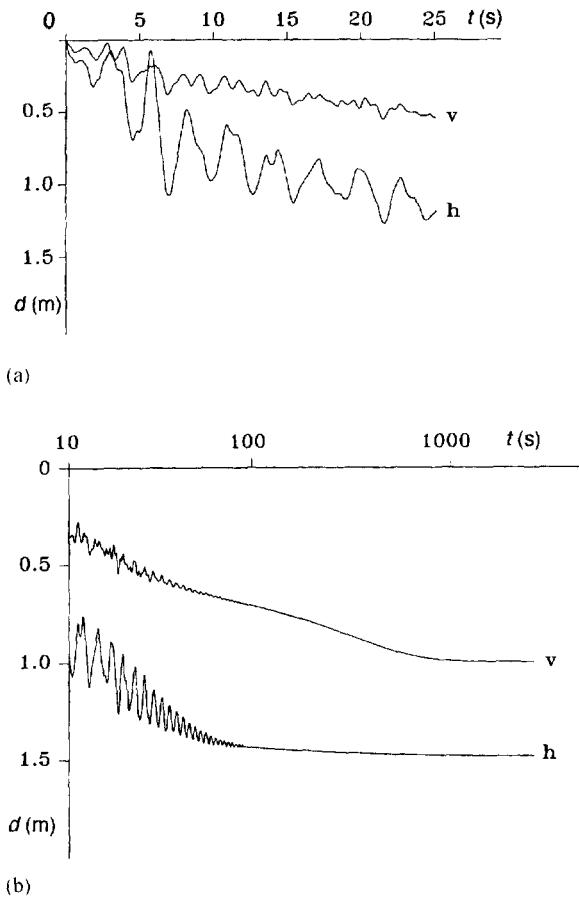
**Figure 14.8** Initial configuration



**Figure 14.9** Saturation distribution over deformed configuration of the area close to the vertical slope at three different steps of the analysis

earthquake (N–S component of horizontal acceleration) and during the following consolidation phase. The fully saturated domain, after excavation, is modelled by linear triangular finite elements. Figure 14.8 shows the initial configuration.

The material characteristics of the soil, saturation relationships included, are assumed with reference to the Pastor–Zienkiewicz model of the clay core of the San Fernando Dam [5]. In order to impose initial fully saturated conditions, with the free surface at the top nodes of the domain, an undrained analysis is first carried out with horizontal constraints at the vertical slope nodes. Then the seismic excitation is applied at the bottom nodes while the horizontal constraints are released and the soil is allowed to desaturate.



**Figure 14.10** Vertical ( $v$ ) and horizontal ( $h$ ) displacements versus time for the top point of the slope: (a) seismic phase and (b) consolidation phase

On the left and right vertical sides of the domain, the horizontal displacements are fixed and a hydrostatic pressure distribution is assigned with an atmospheric value at the two corresponding top nodes.

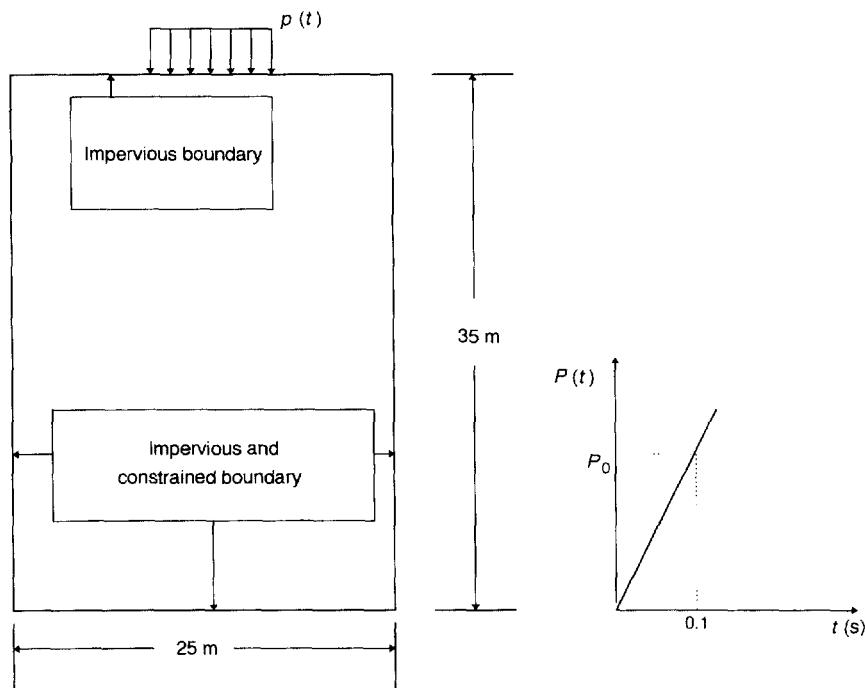
Figure 14.9 represents the saturation distribution at different times over the corresponding deformed configuration in the area close to the slope, whereas Figure 14.10 shows the time history of the vertical and horizontal displacements at the nodes above the vertical slope.

#### 14.7.4 Dynamic strain localisation

Strain localisation occurs when the deformation in a continuum is localised in well-defined but narrow zones of intense straining. It occurs in many engineering materials,

including single-phase solids like metals and multiphase geomaterials. In the localisation zone the material behaviour is inelastic, elsewhere it generally remains elastic. The trigger to form localisation zones is the occurrence of inhomogeneities in the material and stress concentrations. Stress concentrations can occur in dynamic strain localisation by the interaction of incoming and reflected waves, hence localisation may be initialised.

Extensive laboratory experiments have been carried out on geomaterials [19,20]. Laboratory tests have also been made on dense and loose sands [21]. For undrained samples of dense sand, strain localisation was observed only when there developed pore water tractions (with reference to atmospheric pressure) of the order of 80 kPa. Cavitation of the pore fluid also took place. This is not the case in samples of medium-loose sands, where no negative water pressures were observed. Rice [22] used a mathematical model to study the amount of dilatant strengthening in a fully saturated dilating frictional material. Rudnicki [23] analysed a fluid-saturated rock mass with an embedded weakened layer. Vardoulakis [24] showed the importance of including inertia terms in the stability analysis for undrained simple-shear deformation of water-saturated



**Figure 14.11** Geomaterial characteristics and the load function of the first example

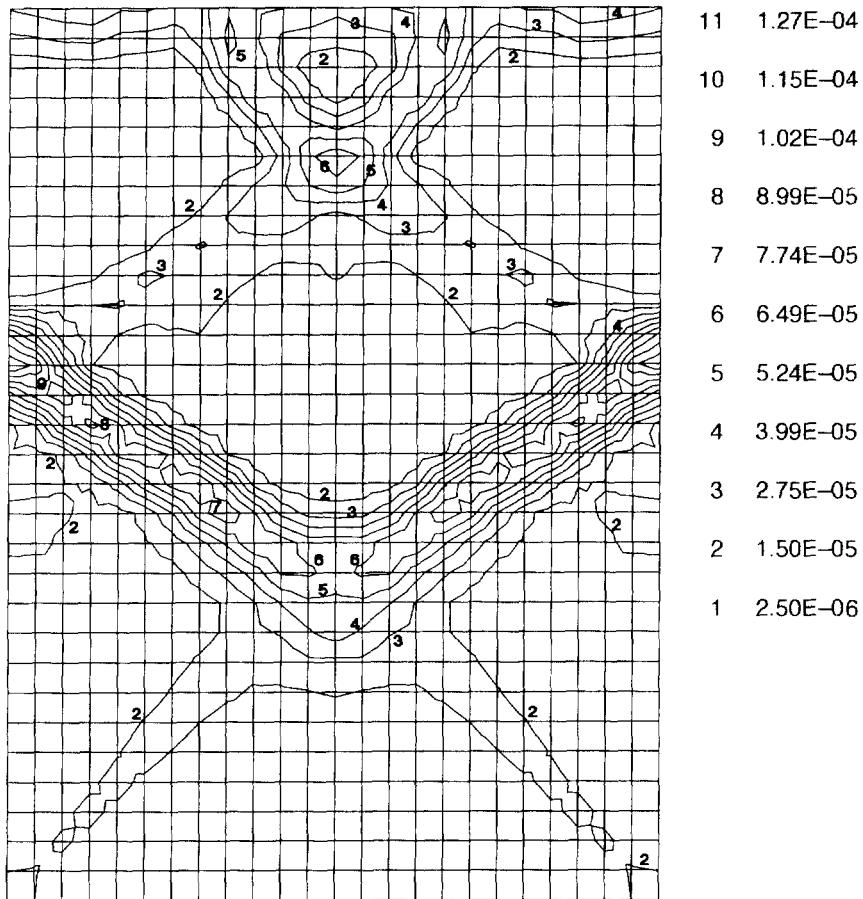
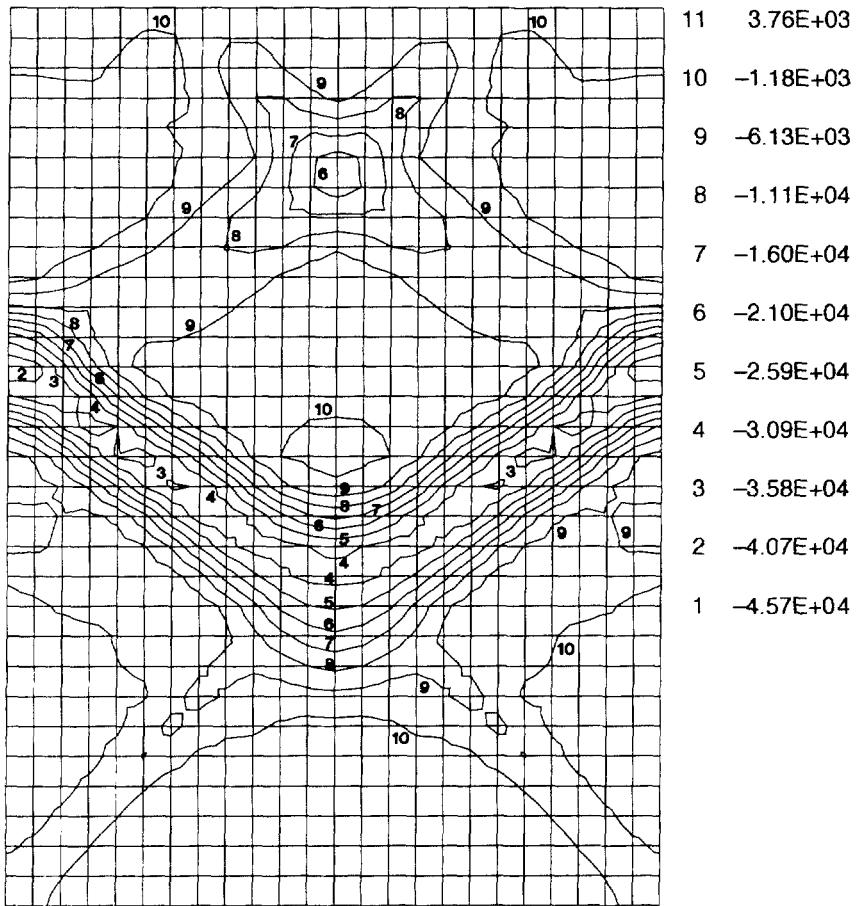


Figure 14.12 Contours of effective plastic strain at  $t = 0.112\text{ s}$

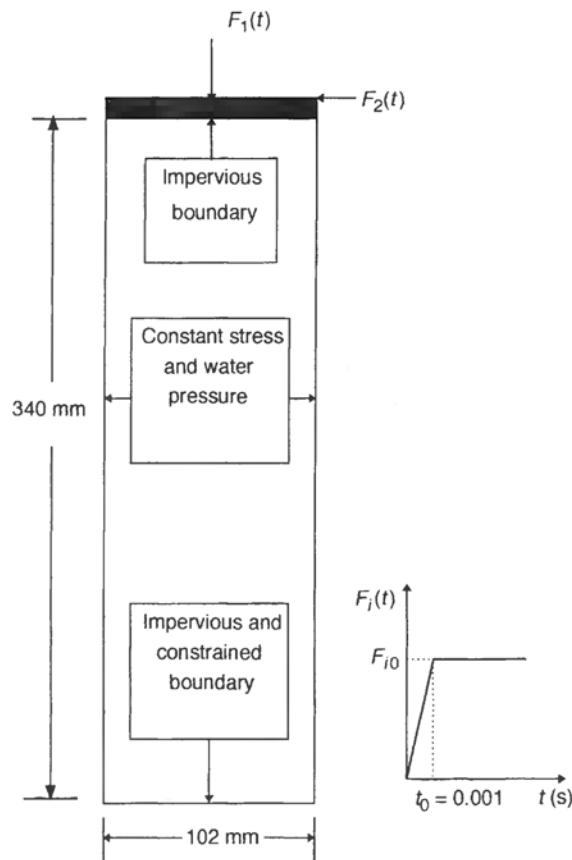
granular soil, at least when the material was contracting. Loret and Prevost [25] and Schrefler *et al.* [26,27] used numerical models to study the shear band dominated processes in fully saturated soil masses using the dynamic strain localisation theory.

There is a difference in localisation between single-phase materials and fluid-saturated geomaterials. Single-phase strain-softening materials have no internal length scale [28]. When softening occurs, the domain splits into an elliptic part with an imaginary wave speed (standing wave) and a hyperbolic part where waves can propagate. Because of the inability of the standing wave to propagate, the localisation zone has zero thickness with no energy consumption, which is against experimental



**Figure 14.13** Contours of excess pore water pressure at  $t = 0.112$  s

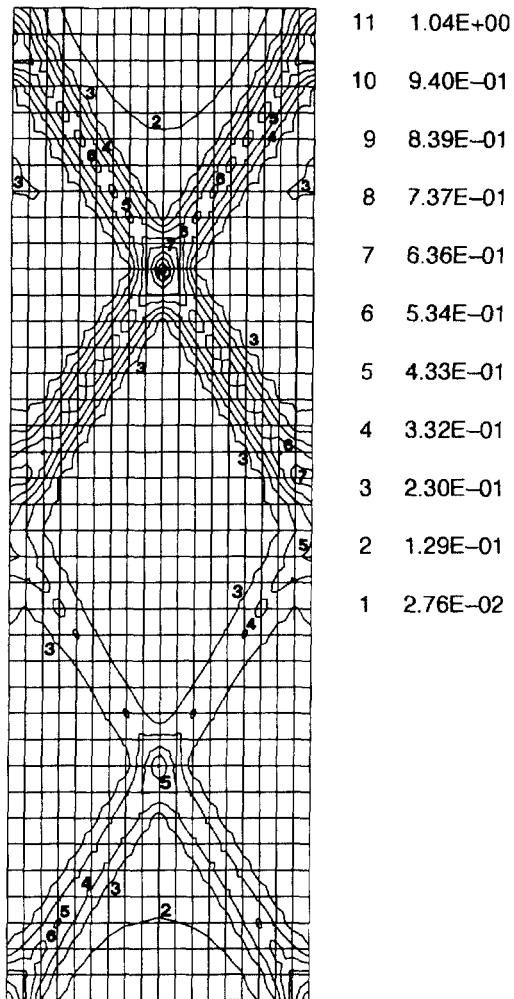
evidence. Finite element models attempt to simulate this phenomenon; when the mesh is refined, the shear band width decreases, creating a pathological mesh dependence of the results. Multiphase materials behave differently; for fully and partially saturated porous media with a softening solid-phase behaviour, in the case of axial waves, there exists a domain of parameters for which the dynamic equations do not lose their hyperbolic character at the onset of strain softening, so a solution with a real wave speed may be obtained [29]. This domain and its associated internal length depend strongly on the permeability. No such weakly stable domain exists for shear waves [22,29]. These results explain why, in the models of Schrefler and coworkers [26,27], a mesh -



**Figure 14.14** Geomaterial characteristics and load function

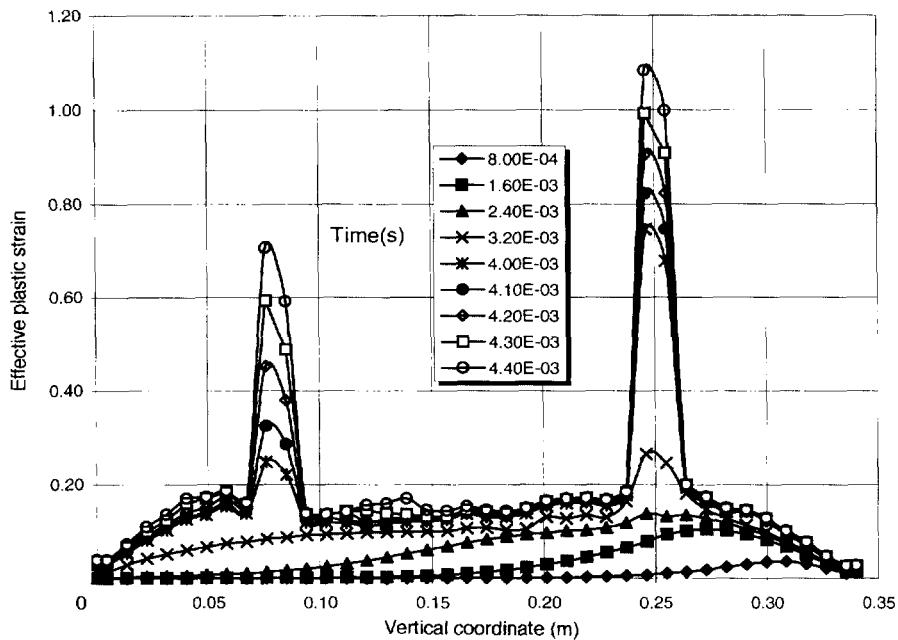
independent value of the equivalent plastic strain in the shear band was observed, whereas the shear band width was still weakly mesh dependent. Schrefler and coworkers [26,27] give examples of dynamic strain localisation with a softening solid-phase behaviour, using a Mohr–Coulomb constitutive relationship with cohesion softening. We will use the Pastor–Zienkiewicz model, as described in Chapter 4.

First, we take the cross-section of a geological formation (Figure 14.11) [25,27]. The load is applied at the central part of the top surface and the incremental load function is also shown in Figure 14.11 ( $P_0 = 2.04 \text{ kPa}$ ). The self-weight and the hydrostatic pressure distribution are taken into account as initial conditions for the Pastor–Zienkiewicz model. The permeability for the saturated condition is  $k = 0.25 \text{ mm s}^{-1}$ ;



**Figure 14.15** Contours of effective plastic strain for the symmetric load case at  $4.4 \times 10^{-3}$  s

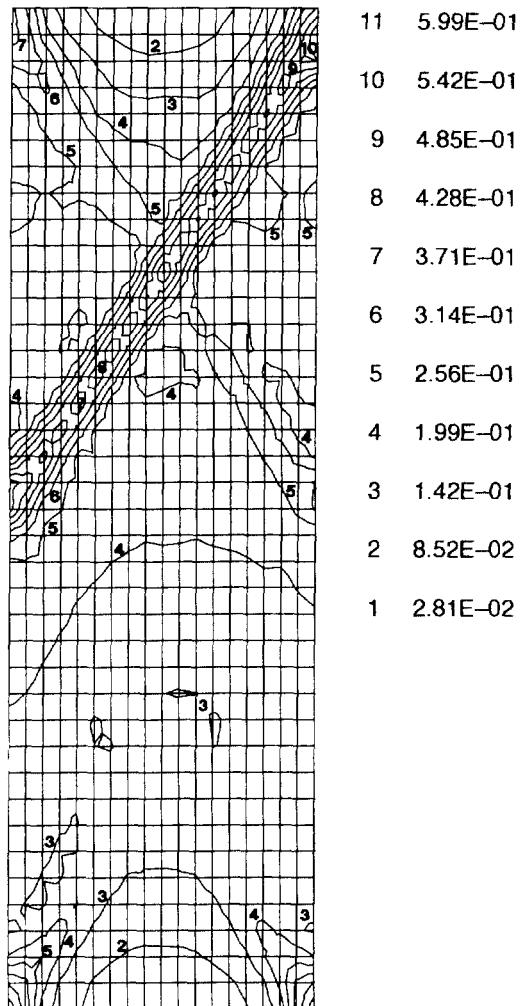
the other material parameters are  $E = 285$  MPa,  $\nu = 0.4285$ ,  $K_w = 0.20$  GPa,  $K_s = 6.78$  GPa,  $n = 0.20$ ,  $\rho_s = 2000 \text{ kg m}^{-3}$ ,  $\rho_w = 1000 \text{ kg m}^{-3}$ . The Pastor-Zienkiewicz generalised constitutive sand model is adopted for medium-loose sand with  $M_f = 0.4$ ,  $H_{u0} = 600$ ,  $\alpha_g = 0.45$ ,  $\beta_1 = 0.2$ ,  $M_g = 1.5$ ,  $\gamma_u = 2.0$ ,  $\alpha_f = 0.45$ ,  $H_0 = 350$ ,  $\gamma_{DM} = 2.0$ ,  $\beta_0 = 4.2$ ,  $K_{evo}/p_0 = 35\,000$ ,  $K_{eso}/p_0 = 52\,500$ .



**Figure 14.16** Development of effective plastic strain in the cross-section  $x = 0.051$  m for the symmetric load case

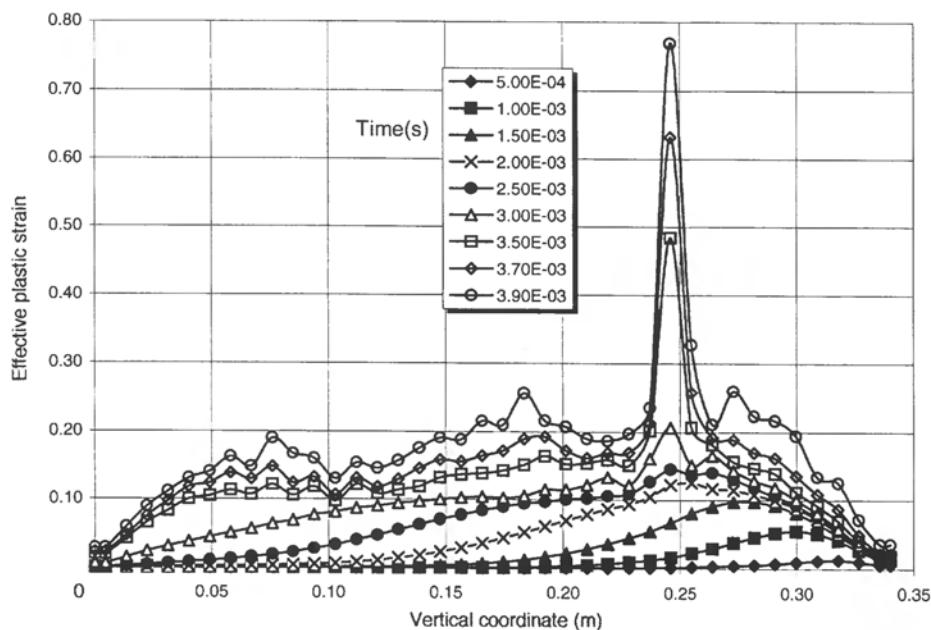
The shear band formation commences at the top surface near the loading area, because in the assumed model the sand stiffness increases in the vertical direction. The shear bands are then reflected at the lateral rigid surfaces and continue to propagate towards the direction of the bottom surface. The shear bands after 0.112 s are shown in Figure 14.12. The excess water pressure results at time 0.112 s are presented in Figure 14.13 and no water traction (with reference to atmospheric pressure) is observed in the shear bands.

In the last example a dense sand sample was investigated with the Pastor–Zienkiewicz model. The geometrical data is the same as in Mokni [21] and is depicted in Figure 14.14. The load function is also shown where both symmetric and unsymmetric (small lateral load) loading cases are chosen to simulate the shear band formation ( $F_{10} = 204$  kN,  $F_{20} = 3.4$  kN). The basic material parameters are the same as in the first example, i.e. permeability,  $E$ ,  $K_w$ ,  $K_s$ ,  $n$ ,  $\rho^s$ ,  $\rho^w$ . The initial conditions, chosen from Mokni [21], are  $\sigma_{xx} = \sigma_{yy} = -100$  kPa,  $\sigma_{zz} = -130$  kPa,  $\sigma_{xy} = \sigma_{xz} = \sigma_{yz} = 0$  and  $p^w = -300$  kPa; the material parameters of dense sand for the Pastor–Zienkiewicz model are  $M_f = 1.0$ ,  $H_{n0} = 600$ ,  $\alpha_g = 0.45$ ,  $\beta_1 = 0.2$ ,  $M_g = 1.0$ ,  $\gamma_u = 2.0$ ,  $\alpha_f = 0.45$ ,  $H_0 = 350$ ,  $\gamma_{DM} = 2.0$ ,  $\beta_0 = 4.2$ ,  $K_{evo} = 11\,550$  kPa,  $K_{eso} = 22\,000$  kPa [30].

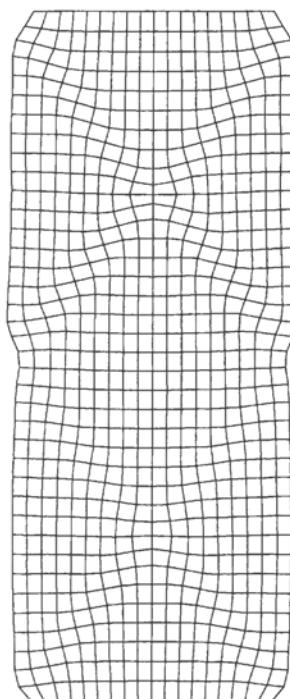


**Figure 14.17** Contours of effective plastic strain for the unsymmetric load case at  $3.9 \times 10^{-3}$  s

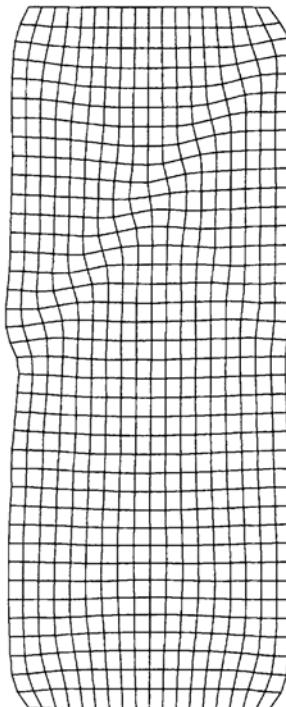
As in the laboratory test, the numerical simulation of the shear band starts from the top surface, near the loading area for the symmetric or unsymmetric load cases. The development of the shear band for symmetric and unsymmetric loading is shown in Figures 14.15 to 14.18, where the effective plastic strain is depicted. The unsymmetric shear band formation starts after a small disturbing load is applied and then propagates in a fixed direction controlled by this disturbing load. The symmetric shear bands



**Figure 14.18** Development of effective plastic strain in the cross-section  $x = 0.051\text{ m}$  for the unsymmetric load case



**Figure 14.19** Structure deformation for the symmetric case at  $4.4 \times 10^{-3}\text{ s}$



**Figure 14.20** Structure deformation for the unsymmetric case at  $3.9 \times 10^{-3}$  s

propagate from the top surface of the sample, and new shear bands are generated at the bottom surface after the loading wave hits the lower boundary. This was also observed in the laboratory test. Because of the uniform water pressure applied around the sample, the results of water pressure are not as regular as in the previous example. But water tractions are still obtained in the centre of the shear band; a value of over 80 kPa is obtained for symmetric and unsymmetric cases in the shear band hence cavitation occurs. A more advanced model is needed for cavitation modelling [27,30]. Figures 14.19 and 14.20 show the deformed shape when shear bands are formed.

## 14.8 CONCLUSIONS

This chapter has dealt with the problem of small-strain and finite-strain behaviour in porous media, static and dynamic, partially saturated and fully saturated. All these situations can be handled with the same approach used throughout this book. It has been assumed that the air pressure in the partially saturated zones remains constant and equal to the external ambient pressure, a common assumption in soil mechanics. Several examples – including advanced problems such as dynamic strain localisation in dense

and medium-dense sand samples and seismic behaviour of a desaturating vertical slope showed the breadth of applications for this model. A more extensive treatment of such soil dynamics problems can be found in Zienkiewicz *et al.* [17].

## REFERENCES

- Zienkiewicz, O. C., Chan, A. H. C., Pastor, M., Paul, D. K. and Shiomi, T. (1990) Static and dynamic behaviour of soils: a rational approach to quantitative solutions, I. Fully saturated problems. *Proc. R. Soc. Lond. A*, **429**, 285–309.
- Zienkiewicz, O. C., Xie, Y. M., Schrefler, B. A., Ledesma, A. and Bicanic, N. (1990) static and dynamic behaviour of soils: a rational approach to quantitative solutions, II. Semi-saturated problems. *Proc. R. Soc. Lond. A*, **429**, 311–21.
- Meroi, E. (1993) Comportamento non lineare per geometria di mezzi porosi parzialmente saturi. Doctoral thesis, University of Padua.
- Meroi, E. A., Schrefler, B. A. and Zienkiewicz, O. C. (1995) Large-strain static and dynamic semisaturated soil behaviour. *Int. J. Num. Anal. Meth. Geomech.*, **19**, 81–106.
- Meroi, E. A. and Schrefler, B. A. (1995) Large-strain analysis of porous media, in *Modern Issues in Non-Saturated Soils*, A Gens, P. Jouanna, B. A. Schrefler (ed.), Springer-Verlag, Wien, Ch. 9.
- L. E. Malvern, (1969) *Introduction to the Mechanics of a Continuous Medium*, Prentice Hall, Englewood Cliffs NJ.
- Cheng, W. F. and Tsui, Y. (1992) Limitations to the large-strain theory. *Int. J. Num. Meth. Engng.*, **33**, 101–14.
- Dienes J. K. (1979) On the analysis of rotation, stress rate in deforming bodies. *Acta Mech.*, **32**, 217–32.
- Dafalias Y. F. (1983) Corotational rates for kinematic hardening at large plastic deformations. *J. Appl. Mech.*, **50**, 561–65.
- Johnson, G. C. and Bammman D. J. (1984) A discussion of stress rates in finite deformation problems. *Int. J. Solids Struct.*, **8**, 725–37.
- Molenkamp, F. (1986) Limits to the Jaumann stress rate, *Int. J. Num. Anal. Meth. Geomech.*, **10**, 151–76.
- Lee, E. H. (1981) Some comments on elastic–plastic analysis. *Int. J. Solids Struct.*, **17**, 859–72.
- Lubarda, V. A. and Lee E. H. (1981) A correct definition of elastic and plastic deformation and its computational significance. *J. Appl. Mech.*, **48**, 35–40.
- Carter, J. P. Booker, J. R. and Small, J. C. (1979) The analysis of finite elasto-plastic consolidation. *Int. J. Num. Anal. Meth. Geomech.*, **3**, 107–29.
- Zienkiewicz, O. C. and Taylor, R. L. (1991) *The Finite Element Method*, McGraw-Hill, London.
- McMeeking R. M. and Rice, J. R. (1975) Finite element formulations for problems of large elastic–plastic deformation. *Int. J. Solids Struct.*, **11**, 601–16.
- Zienkiewicz, O. C., Chan, A. H., Pastor, M., Schrefler, B. A., and Shiomi, T. (1999) *Computational Geomechanics: With Special Reference to Earthquakes*, Wiley, Chichester.
- Kim, C. S., Lee, T. S., Advani, S. H. and Lee, J. H. (1993) Hygrothermomechanical evaluation of porous media under finite deformation, II. Model validations and field simulations. *Int. J. Num. Meth. Engng.*, **36**, 161–79.
- Lade, P. V. (1982) Localization effects in triaxial tests on sand, in *Proceedings of the IUTAM Symposium on Deformation and Failure of Granular Materials*, P. A. Vermeer and H. J. Luger, (eds), Balkema, Rotterdam, pp. 461–71.

20. Peters, J. F., Lade, P. V. and Bro, A. (1988) Shear band formation in triaxial and plane strain tests, in *Advanced Triaxial Testing of Soil and Rock*, ASTM STP 977, R. T. Donaghe, R. C. Chaney and M. L. Silver (eds), ASTM, Philadelphia PA, pp. 604–27.
21. Mokni, M. (1992) Relations entre déformations en masse et déformations localisées dans les matériaux granulaires. PhD thesis, Institut de Mécanique, Grenoble.
22. Rice, J. R. (1985) On the stability of dilatant hardening for saturated rock masses. *J. Geophys. Res.*, **80**, 1531–36.
23. Rudnicki, J. W. (1984) Effect of dilatant hardening on the development of concentrated shear-deformation in fissured rock masses. *J. Geophys. Res.*, **89**(B11), 9259–70.
24. Vardoulakis, I. (1986) Dynamic stability analysis of undrained simple shear on water-saturated granular soils. *Int. J. Num. Anal. Meth. Geomech.*, **10**, 177–90.
25. Loret, B. and Prevost, J. H. (1991) Dynamic strain localization in fluid-saturated porous media. *J. Engng Mech.*, **11**, 907–22.
26. Schrefler, B. A., Majorana, C. E. and Sanavia, L. (1995) Shear band localization in saturated porous media. *Arch. Mech.*, **47**, 577–99.
27. Schrefler, B. A., Sanavia, L. and Majorana, C. E. (1996) A multiphase medium model for localisation and postlocalisation simulation in geomaterials. *Mech. Cohesive-Frictional Mater. Struct.*, **1**, 95–114.
28. Sluys, L. J. (1992) Wave propagation, localisation and dispersion in softening solids. PhD thesis, Delft University, Netherlands.
29. Zhang, H. W., Sanavia, L. and Schrefler, B. (1999) Internal length scale in strain localization multiphase porous media. Forthcoming. To appear in *Mechanics of Cohesive-Frictional Materials and Structures*.
30. Sanavia, L., Zhang, H. W. and Schrefler, B. A. (1998) Strain localization modelling in saturated sand samples. *Int. J. Num. Meth. Fluids*, **27**, 1–4, 109–125.