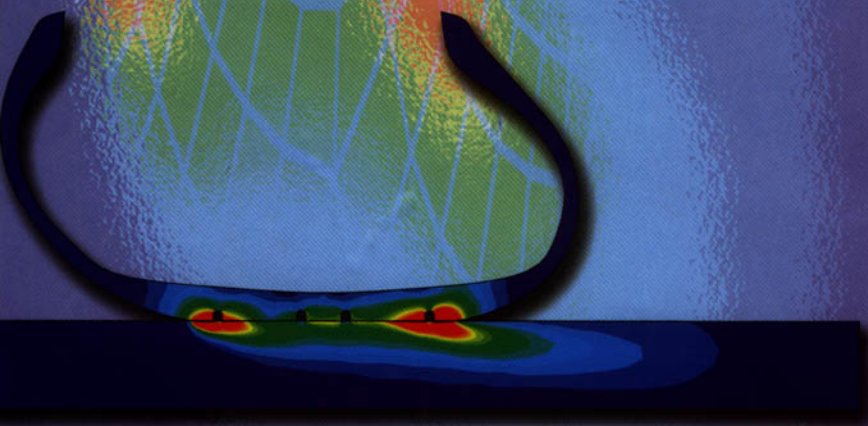


# Computational Contact Mechanics



**PETER WRIGGERS**

***Computational  
Contact Mechanics***

**This page intentionally left blank**

# ***Computational Contact Mechanics***

**P. Wriggers**

*University of Hannover, Germany*



JOHN WILEY & SONS, LTD



Copyright © 2002

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,  
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): [cs-books@wiley.co.uk](mailto:cs-books@wiley.co.uk)

Visit our Home Page on [www.wileyeurope.com](http://www.wileyeurope.com) or [www.wiley.com](http://www.wiley.com)

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 Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to [permreq@wiley.co.uk](mailto:permreq@wiley.co.uk), or faxed to (+44) 1243 770571.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

#### *Other Wiley Editorial Offices*

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

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

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

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

#### *Library of Congress Cataloging-in-Publication Data*

Wriggers, P.

Computational contact mechanics / P. Wriggers.

p. cm.

Includes bibliographical references and index.

ISBN 0-471-49680-4 (alk. paper)

1. Contact mechanics—Mathematical models. I. Title.

TA353.W75 2002

620.1'05—dc21

2002072091

#### *British Library Cataloguing in Publication Data*

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

ISBN 0-471-49680-4

Produced from  $\LaTeX$  Postscript files supplied by the author

Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire

This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

# Contents

<i>Preface</i>	<i>xiii</i>
<i>Introduction</i>	<i>xv</i>
<i>1 Introduction to Contact Mechanics</i>	<i>1</i>
<i>1.1 Contact in a Mass Spring System</i>	<i>1</i>
<i>1.1.1 A variational formulation</i>	<i>1</i>
<i>1.1.2 Lagrange multiplier method</i>	<i>3</i>
<i>1.1.3 Penalty method</i>	<i>4</i>
<i>1.2 Finite Element Analysis of the Contact of Two Bars</i>	<i>6</i>
<i>1.3 Thermo-mechanical Contact</i>	<i>9</i>
<i>1.4 Impact</i>	<i>13</i>
<i>2 Continuum Solid Mechanics and Weak Forms</i>	<i>19</i>
<i>2.1 Kinematics</i>	<i>19</i>
<i>2.1.1 Motion and deformation gradient</i>	<i>19</i>
<i>2.1.2 Strain measures</i>	<i>22</i>
<i>2.1.3 Transformation of vectors and tensors</i>	<i>24</i>
<i>2.1.4 Time derivatives</i>	<i>24</i>
<i>2.2 Balance Laws</i>	<i>26</i>
<i>2.2.1 Balance of mass</i>	<i>26</i>

2.2.2	<i>Local balance of momentum and moments of momentum</i>	26
2.2.3	<i>First law of thermodynamics</i>	27
2.2.4	<i>Transformation to the initial configuration, different stress tensors</i>	27
2.3	<i>Weak Form of Balance of Momentum, Variational Principles</i>	29
2.3.1	<i>Weak form of balance of momentum in the initial configuration</i>	29
2.3.2	<i>Spatial form of the weak formulation</i>	30
2.3.3	<i>Minimum of total potential energy</i>	30
2.4	<i>Constitutive Equations</i>	31
2.4.1	<i>Hyperelastic response function</i>	32
2.4.2	<i>Incremental constitutive tensor</i>	34
2.5	<i>Linearizations</i>	36
2.5.1	<i>Linearization of kinematical quantities</i>	39
2.5.2	<i>Linearization of constitutive equations</i>	40
2.5.3	<i>Linearization of the weak form</i>	40
2.5.4	<i>Linearization of a deformation dependent load</i>	43
3	<i>Contact Kinematics</i>	45
3.1	<i>Normal Contact of Three-dimensional Bodies</i>	46
3.2	<i>Tangential Contact of Three-dimensional Bodies</i>	49
3.2.1	<i>Stick condition</i>	50
3.2.2	<i>Slip condition</i>	50
3.3	<i>Variation of the Normal and Tangential Gap</i>	53
3.3.1	<i>Variation of normal gap</i>	53
3.3.2	<i>Variation of tangential gap</i>	54
4	<i>Constitutive Equations for Contact Interfaces</i>	55
4.1	<i>Normal Contact</i>	56
4.1.1	<i>Constraint formulation</i>	56
4.1.2	<i>Constitutive equations for normal contact</i>	58
4.2	<i>Tangential Contact</i>	61
4.2.1	<i>Stick as a constraint</i>	63
4.2.2	<i>Coulomb law</i>	63
4.2.3	<i>Regularization of the Coulomb law</i>	65
4.2.4	<i>Elasto-plastic analogy for friction</i>	66
4.2.5	<i>Friction laws for metal forming</i>	71

4.2.6	<i>Friction laws for rubber and polymers</i>	73
4.2.7	<i>Friction laws for concrete structures on soil</i>	76
4.2.8	<i>Friction laws from computational homogenization procedures</i>	79
4.3	<i>Lubrication</i>	83
4.4	<i>Adhesion</i>	86
4.5	<i>Decohesion</i>	88
4.6	<i>Wear</i>	89
4.7	<i>Fractal Contact Interfaces</i>	91
5	<i>Contact Boundary Value Problem and Weak Form</i>	93
5.1	<i>Frictionless Contact in Linear Elasticity</i>	93
5.2	<i>Frictionless Contact in Finite Deformations Problems</i>	97
5.3	<i>Treatment of Contact Constraints</i>	99
5.3.1	<i>Lagrange multiplier method</i>	100
5.3.2	<i>Penalty method</i>	101
5.3.3	<i>Direct constraint elimination</i>	104
5.3.4	<i>Constitutive equation in the interface</i>	105
5.3.5	<i>Nitsche method</i>	105
5.3.6	<i>Perturbed Lagrange formulation</i>	107
5.3.7	<i>Barrier method</i>	108
5.3.8	<i>Augmented Lagrange methods</i>	109
5.3.9	<i>Cross-constraint method</i>	111
5.4	<i>Comparison of Different Methods</i>	113
5.5	<i>Linearization of the Contact Contributions</i>	118
5.5.1	<i>Normal contact</i>	119
5.5.2	<i>Tangential contact</i>	121
5.5.3	<i>Special case of stick</i>	124
5.6	<i>Rolling Contact</i>	125
5.6.1	<i>Special reference frames for rolling contact</i>	125
5.6.2	<i>Strain measures</i>	127
5.6.3	<i>Weak Form</i>	128
5.6.4	<i>Constitutive equation</i>	129
5.6.5	<i>Contact kinematics</i>	130
6	<i>Discretization of the Continuum</i>	135
6.1	<i>Isoparametric Concept</i>	136
6.1.1	<i>Isoparametric interpolation functions</i>	139
6.1.2	<i>One-dimensional shape functions</i>	139

6.1.3	<i>Two-dimensional shape functions</i>	141
6.1.4	<i>Three-dimensional shape functions</i>	143
6.2	<i>Discretization of the Weak Forms</i>	145
6.2.1	<i>FE formulation of the weak form with regard to the initial configuration</i>	146
6.2.2	<i>Linearization of the weak form in the initial configuration</i>	150
6.2.3	<i>FE formulation of the weak form in the current configuration</i>	155
6.2.4	<i>Linearization of the weak form in the current configuration</i>	158
7	<i>Discretization, Small Deformation Contact</i>	161
7.1	<i>General Approach for Contact Discretization</i>	162
7.1.1	<i>Lagrange multiplier method</i>	162
7.1.2	<i>Penalty method</i>	165
7.2	<i>Node-to-Node Contact Element</i>	166
7.2.1	<i>Frictionless contact</i>	166
7.2.2	<i>Contact with friction</i>	170
7.3	<i>Isoparametric Discretization of the Contact Contribution</i>	172
7.3.1	<i>Examples for isoparametric contact elements</i>	176
7.4	<i>Discretization for Non-matching Meshes</i>	183
7.4.1	<i>Discretization with contact segments</i>	183
7.4.2	<i>Mortar method</i>	188
7.4.3	<i>Nitsche method</i>	194
8	<i>Discretization, Large Deformation Contact</i>	203
8.1	<i>Two-dimensional Node-to-Segment Contact Discretization</i>	204
8.2	<i>Alternative Discretization for the Two-dimensional NTS-Contact</i>	212
8.3	<i>Three-dimensional Contact Discretization</i>	218
8.3.1	<i>Node-to-surface contact element</i>	219
8.4	<i>Three-Node Master Segment for Frictionless Contact</i>	224
8.4.1	<i>Matrices for Node-To-Edge (NTE) elements</i>	227
8.4.2	<i>Matrices for Node-To-Node (NTV) elements</i>	229
8.5	<i>Smooth Contact Discretization</i>	230
8.5.1	<i>Hermite interpolation for frictionless contact</i>	231

8.5.2	<i>Bezier interpolation for frictionless contact</i>	237
8.5.3	<i>Bezier interpolation for frictional contact</i>	242
8.5.4	<i>Three-dimensional contact discretization</i>	250
8.6	<i>Numerical Examples</i>	253
8.6.1	<i>The sheet/plate rolling simulation</i>	253
8.6.2	<i>Simulation of a sliding and rolling wheel</i>	255
9	<i>Solution Algorithms</i>	259
9.1	<i>Contact Search</i>	261
9.1.1	<i>Spatial search, phase (I)</i>	264
9.1.2	<i>Contact detection, phase (II)</i>	266
9.2	<i>Solution Methods for Unconstrained Nonlinear Problems</i>	270
9.2.1	<i>Algorithms for time-independent problems</i>	271
9.2.2	<i>Algorithms for time-dependent problems</i>	273
9.3	<i>Global Solution Algorithms for Contact</i>	278
9.3.1	<i>Basic notation</i>	279
9.3.2	<i>Dual formulation</i>	282
9.3.3	<i>Penalty method</i>	284
9.3.4	<i>Lagrange multiplier method</i>	286
9.3.5	<i>Augmented Lagrange method, Uzawa algorithm</i>	287
9.3.6	<i>Partitioning method</i>	290
9.3.7	<i>SQP method</i>	294
9.3.8	<i>Active set method for quadratic program</i>	295
9.3.9	<i>Linear complementary problem</i>	297
9.3.10	<i>Contact algorithm of Dirichlet–Neumann type</i>	297
9.3.11	<i>Algorithm for dynamic contact</i>	299
9.4	<i>Global Algorithms for Friction</i>	301
9.5	<i>Local Integration of Constitutive Equations in the Contact Area</i>	305
9.5.1	<i>Evolution of adhesion</i>	306
9.5.2	<i>Friction laws</i>	306
10	<i>Thermo-mechanical Contact</i>	311
10.1	<i>Equations for the Continuum</i>	312
10.1.1	<i>Kinematical relations, multiplicative split</i>	312
10.1.2	<i>Thermoelastic constitutive law</i>	313

10.2	<i>Constitutive Equations for Thermo-mechanical Contact</i>	314
10.2.1	<i>Heat conductance through spots</i>	316
10.2.2	<i>Heat conductance through gas</i>	317
10.2.3	<i>Heat conductance by radiation</i>	319
10.3	<i>Initial Value Problem for Thermo-mechanical Contact</i>	320
10.4	<i>Weak Forms in Thermo-mechanical Analysis</i>	322
10.5	<i>Algorithmic Treatment</i>	323
10.6	<i>Discretization Techniques</i>	324
10.6.1	<i>Node-to-node contact element</i>	324
10.6.2	<i>Node-to-segment contact element</i>	327
10.7	<i>Examples</i>	330
10.7.1	<i>Heat transfer at finite deformations</i>	330
10.7.2	<i>Frictional heating at finite deformations</i>	332
11	<i>Beam Contact</i>	337
11.1	<i>Kinematics</i>	337
11.1.1	<i>Normal contact</i>	338
11.1.2	<i>Tangential contact</i>	341
11.2	<i>Variation of the Gap in Normal and Tangential Directions</i>	342
11.3	<i>Contact Contribution to Weak Form</i>	344
11.4	<i>Finite Element Formulation</i>	346
11.5	<i>Contact Search for Beams</i>	348
11.6	<i>Examples</i>	350
11.6.1	<i>Three beams in frictionless contact</i>	350
11.6.2	<i>Two beams in contact with friction</i>	351
12	<i>Adaptive Finite Element Methods for Contact Problems</i>	355
12.1	<i>Contact problem and discretization</i>	357
12.2	<i>Residual Based Error Estimator for Frictionless Contact</i>	358
12.3	<i>Error Indicator for Contact Based on Projection</i>	362
12.4	<i>Error Estimators based on Dual Principles</i>	365
12.4.1	<i>Displacement error control</i>	365
12.4.2	<i>Stress error control</i>	367
12.5	<i>Adaptive Mesh Refinement Strategy</i>	368
12.6	<i>Numerical Examples</i>	372
12.6.1	<i>Hertzian contact problem</i>	372

12.6.2	<i>Crossing tubes</i>	374
12.6.3	<i>Fractal interface</i>	375
12.7	<i>Error Indicator for Frictional Problems</i>	379
12.7.1	<i>Adaptive strategies</i>	381
12.7.2	<i>Transfer of history variables</i>	382
12.7.3	<i>Numerical examples</i>	384
13	<i>Computation of Critical Points with Contact Constraints</i>	389
13.1	<i>Inequality Constraints for Contact</i>	390
13.2	<i>Calculation of Stability Points</i>	391
13.3	<i>Extended System with Contact Constraints</i>	392
13.4	<i>Examples</i>	394
13.4.1	<i>Block pressing on arch</i>	394
13.4.2	<i>Two arches</i>	398
Appendix A	<i>Gauss integration rules</i>	401
A.1	<i>One-dimensional Integration</i>	401
A.2	<i>Two-dimensional Integration</i>	402
Appendix B	<i>Convective Coordinates</i>	405
Appendix C	<i>Parameter Identification for Friction Materials</i>	409
References		413
Index		437



# Acknowledgements

The author would like to thank the publishers below, who kindly allowed permission for the reproduction of their copyrighted figures.

Figures 1.5, 1.6, 1.7 and 1.8 from D. Dowson, *History of Tribology*, 1979, Longman; Figure 4.11 from Courtney-Pratt, J.S. and Eisner, E., *The effect of a tangential force on the contact of metallic bodies*, *Proceedings of the Royal Society of London*, 238-A, pp. 529-550, Copyright 1957; Figure 4.17 from Manfred Klüppel, Rubber Friction on Self-Affine Road Tracks, *Rubber Chemistry and Technology*, Vol. 73, pp 578-606. Reproduced with permission from *Rubber Chemistry and Technology*. Copyright © 2000, Rubber Division, American Chemical Society, Inc.; Figures 11.3, 11.4, 11.6, 11.7, 11.9, 11.10 and 11.11 from Litewka, P. and Wriggers, P., *Contact between 3D-beams with rectangular cross section*. *International Journal for Numerical Methods in Engineering*, Vol. 53: 9, Copyright 2001. Reproduced by permission of John Wiley & Sons, Ltd; Figures 12.7, 12.8, 12.9 and 12.10 from Stein, E. *et al. Error-controlled Adaptive Finite Element Methods in Solid and Structural Mechanics*, Copyright 2002. Reproduced by permission of John Wiley & Sons, Ltd; Figures 12.16–12.20 from Hu *et al.*, *Adaptive finite element analysis of fractal interfaces in contact problems*, reprinted from *Computer Methods in Applied Mechanics and Engineering*, Vol. 182, 2000, pp 17-38, Copyright 2000, with permission from Elsevier Science.

# *Preface*

Contact mechanics has its application in many engineering problems. No one can walk without frictional contact, and no car would move for the same reason. Hence contact mechanics has, from an engineering point of view, a long history, beginning in ancient Egypt with the movement of large stone blocks, over first experimental contributions from leading scientists like LEONARDO DA VINCI and COULOMB, to today's computational methods. In the past contact conditions were often modelled in engineering analysis by more simple boundary conditions since analytical solutions were not present for real world applications. In such cases, one investigated contact as a local problem using the stress and strain fields stemming from the analysis which was performed for the entire structure. With the rapidly increasing power of modern computers, more and more numerical simulations in engineering can include contact constraints directly, which make the problems nonlinear.

This book is an account of the modern theory of nonlinear continuum mechanics and its application to contact problems, as well as of modern simulation techniques for contact problems using the finite element method. The latter includes a variety of discretization techniques for small and large deformation contact. Algorithms play another prominent role when robust and efficient techniques have to be designed for contact simulations. Finally, adaptive methods based on error controlled finite element analysis and mesh adaption techniques are of great interest for the reliable numerical solution of contact problems. Nevertheless, all numerical models need a strong backup provided by modern continuum mechanics and its constitutive theory, which is applied in this book to the development of interface laws for normal and frictional contact.

The present text can be viewed as a textbook which is basically self-contained. It is written for students at graduate level and engineers who have to simulate contact problems in practical applications and wish to understand the theoretical and algorithmic background of modern finite element systems. The organization of the book is straightforward. After an introductory chapter which discusses relevant contact formulations in a simple matter, there follows a chapter which provides the continuum mechanics background. The special geometrical relations needed to set up the contact constraints and constitutive equations valid at the contact interface are then discussed in detail without going into a numerical treatment. The topic of computational contact is then described in depth in the next chapters, providing different formulations, algorithms and discretization techniques which have been established so far. Here solid and beam contact is considered, as well as contact of unstable systems and thermomechanical contact. The algorithmic side includes, besides a broad range of solution methods, adaptive discretization techniques for contact analysis. However, it can be concluded for the present that there exists nothing which can be called *the robust method* for all different types of contact simulations. This actually also holds for other simulations, including nonlinearities. However, especially due to the fact that such a method does not exist, it is necessary to discuss those methods which are *on the market* in the light of good or bad behaviour.

It is finally a pleasure to thank many people who have assisted me in writing the book, and who were always available in the last twenty years for deep discussions on computational contact mechanics, including the related formulations of continuum mechanics and implementation issues. This scientific collaboration often resulted in joint work in which new papers or reports were written. In particular, I should like to mention my PhD students Anna Haraldsson, Henning Braess, Michael Imhof, Joze Korelc, Lovre Krstulovic-Opara, Tilmann Raible, Albrecht Rieger, Oliver Scherf and Holger Tschöpe. But I have also to include colleagues who worked and still work with me on issues of computational contact mechanics: Christian Miehe, Bahram Nour-Omid, the late Panos Panagiotopoulos, Karl Schweizerhof, the late Juan Simo, Bob Taylor, Giorgio Zavarise and Tarek Zohdi.

Furthermore, I would like to express my appreciation to Bob Taylor, Giorgio Zavarise and Tarek Zohdi, who read early parts of the manuscript and helped with their constructive comments and criticisms to improve the text. I would also like to thank Elke Behrend, who together with Tilmann Raible, drew most of the figures in the text. Last but not least, I would like to thank John Wiley, especially Jan de Landtsheer, for patience regarding the delivery of the manuscript and for the good collaboration during the last years.

P. WRIGGERS

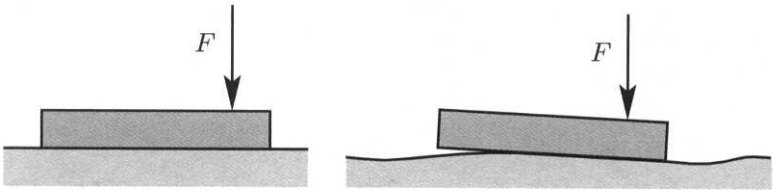
Hannover, January 2002

# Introduction

Boundary value problems involving contact are of great importance in industry related to mechanical and civil engineering, but also in environmental and medical applications. Virtually all movements on this planet involve contact and friction, like simple walking or running, driving of cars, riding bicycles or steaming of trains. If friction were not present (see movement on ice), all these motions would not be possible. Also, the area in which a foot, a tyre or a wheel interacts with the soil, the road or the rail is not known *a priori*, leading to a nonlinear boundary value problem for these simple everyday tasks.

Due to the nonlinear nature of contact mechanics, such problems in the past were often approximated by special assumptions within the design process. Due to the rapid improvement of modern computer technology, one can today apply the tools of computational mechanics to simulate applications which include contact mechanisms numerically. This can be done to an accuracy which is sufficient for design purposes. However, even now most of the standard finite element software is not fully capable of solving contact problems, including friction, with robust algorithms. Hence there is still a challenge for the finite element society to design efficient and robust methods for computational contact mechanics.

The range of application in contact mechanics starts with relatively simple problems like foundations in civil engineering, where the lift off of the foundation from the soil due to eccentric forces acting on a building are considered (see Figure I.1). Furthermore, foundations including piles as supporting members are of interest. Also, classical bearing problems of steel constructions, the connecting of structural members by bolts or screws or the impact of cars against building structures are areas in

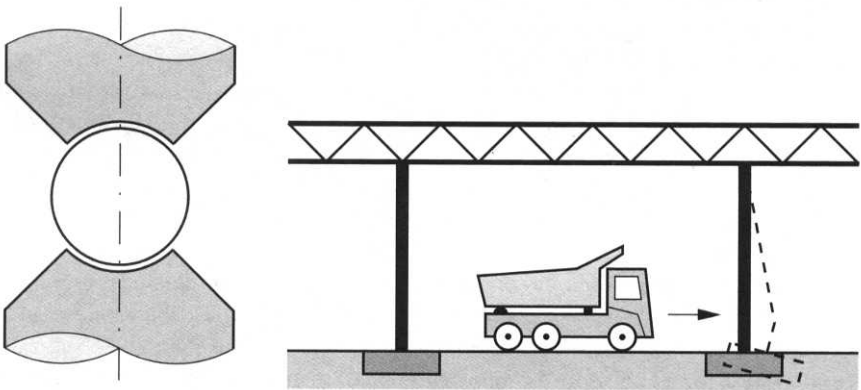


**Fig. 1.1** Contact problems: foundation.

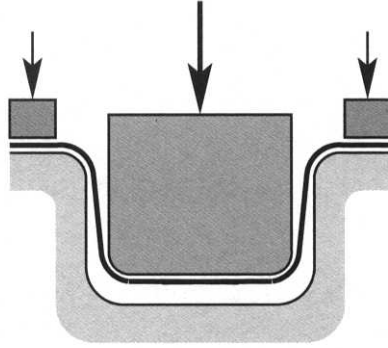
which contact analysis enters the design process in civil engineering (see Figure 1.2). Most of these problems can usually be treated by the assumption of small strains, however due to the nature of contact problems with the *a priori* unknown contact area, all applications are nonlinear and need special algorithms.

Applications of contact mechanics in mechanical engineering include the design of gears and metal forming processes, like sheet metal or bulk forming (see Figure 1.3). The latter problems depict large deformations within the sheet. Furthermore, drilling problems, crash analysis of cars, rolling contact of car tyres or railroad wheels are relevant technical applications of contact in mechanical engineering. Other applications are related to biomechanics where human joints or the implantation of teeth are of consideration. Here again, large deformation cannot be excluded in the analysis, and complicated nonlinear material models have to be applied for a successful numerical simulation.

Due to this variety, contact problems are today combined either with large elastic or inelastic deformations, including time-dependent responses. Hence a modern formulation within computational mechanics has to account for all these effects, leaving the linear theory as a special case. For most industrial applications, numerical



**Fig. 1.2** Contact problems: roller bearing and impact of a lorry.

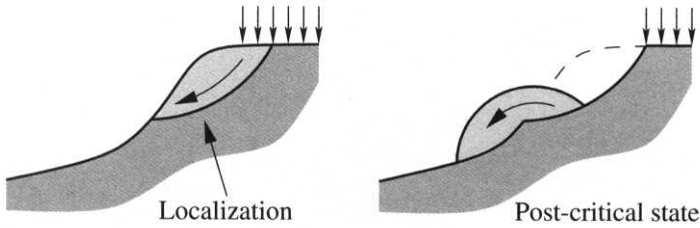


**Fig. I.3** Sheet metal forming.

methods have to be applied since the contacting bodies have complex geometries or undergo large deformations. Today we can distinguish several branches in computational contact mechanics which are applied to solve different classes of contact problems:

- Finite element methods, applied to problems undergoing small and large deformations, as well as in the elastic or inelastic range.
- Discrete element methods, used to compute problems in which up to  $10^8$  particles are coming into contact.
- Multi body systems, based on a description of the bodies as rigid ones. These systems are generally small, and can be applied to model the dynamic behaviour of engineering structures in which contact is also allowed.

Thermal coupling might need to be considered within contact analysis, cooling of electronic devices, heat removal within nuclear power plant vessels or thermal insulation of astronautic vehicles, where the mechanical response and the thermal conduction interacts in the contact area. When electronic devices are considered coupling with electro-magnetic field equations can be of interest. Even stability behaviour has to be linked to contact, like wrinkling arising in metal forming problems or the shearband formation in soils (see Figure I.4). The latter problem is also related to the simulation of avalanches. Here a contact formulation together with the correct modelling of the process in continuum mechanics can be used to compute the final position of a part of the avalanche which has sheared off. All together, Computational Contact Mechanics (CCM) has to cover topics from tribology, including friction, lubrication, adhesion and wear. One has to establish weak forms for finite deformation mechanics, coupling to other fields like thermal or electromagnetic fields, and to derive associated algorithms to solve the nonlinear boundary value problems, which include inequality constraints. Hence, CCM is an interdisciplinary area which needs input from tribolo-



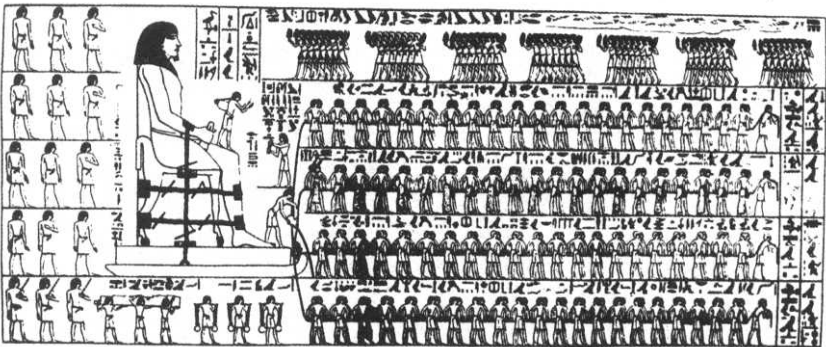
**Fig. 1.4** Shearband formation and collapse analysis in soils or avalanches.

gists, mathematicians, computer scientists and people from mechanics, together with people working in other fields like heat conduction or electromagnetism.

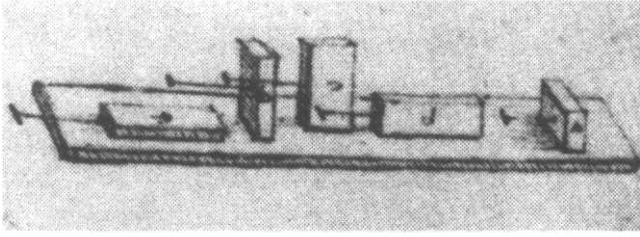
Here we will restrict ourselves mainly to finite element techniques for the treatment of contact problems, despite many other numerical schemes and analytical approaches which could be discussed as well. However, there are common formulations and algorithms, and also overlapping of the methods. These will be discussed in related chapters. Generally, an overview related to modern techniques applied in discrete element methods can be found in, for example Attig and Esser (1999) and for multi-body-systems with special relation to contact in Pfeiffer and Glocker (1996).

Before we provide a short summary of the topics covered in this book, a short historical overview on contact mechanics and computational contact mechanics is given.

**Historical remarks.** Due to this technical importance, a great number of researchers have investigated contact problems. In ancient Egypt people needed to move large stone blocks to build the pyramids, and thus had to overcome the frictional force associated with it. This is depicted in Figure I.5, where we can see that even in ancient Egypt people knew about the process of lubrication. There is a man



**Fig. 1.5** Stone block moved by Egyptian worker.



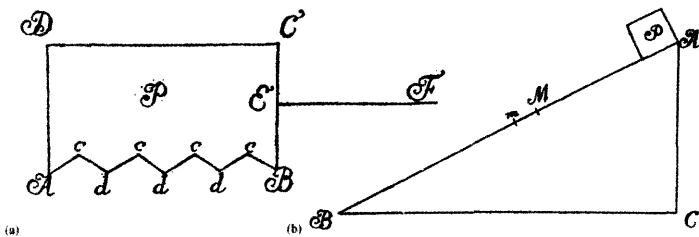
**Fig. I.6** DA VINCI's experiments.

standing on the sledge who pours a fluid onto the ground immediately in front of the sledge. Since friction occurs in many applications which are of technical importance, famous researchers in the past have investigated frictional contact problems, amongst them DA VINCI, who in the 15<sup>th</sup> century measured friction force and had already considered the influence of the contact area on the friction force using blocks with different contact area but the same weight (see Dowson (1979) and Figure I.6). He found that the friction force is proportional to the weight of the blocks, and is independent of the apparent contact area. Associated results are often attributed to Amontons (1699) neglecting the contribution of DA VINCI. When putting these findings in a formula one obtains the classical equation for friction (known as COULOMB's friction law), which every student in engineering learns during the first semesters of study:

$$F_T = \mu N \quad (\text{I.1})$$

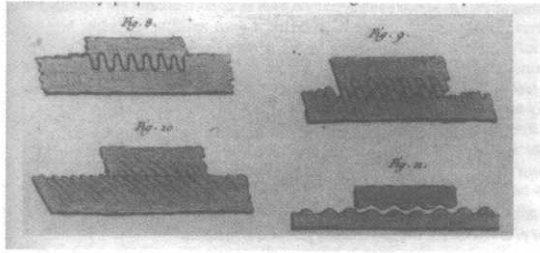
where  $F_T$  is the friction force,  $N$  is the normal force and  $\mu$  the coefficient of friction.

A first analysis from the mathematical point of view was carried out by EULER, who assumed triangular section asperities for the representation of surface roughness (Euler (1748b) and Euler (1748a)). His model is depicted in Figure I.7. He had already concluded from the solution of the equations of motion for a mass on a slope that the kinetic coefficient of friction has to be smaller than the static coefficient of friction. Actually, it was EULER who introduced the symbol  $\mu$  for the friction coefficient, which is the common symbol nowadays. A comprehensive experimental



**Fig. I.7** EULER's model for friction.





**Fig. 1.8** COULOMB'S model for rough surfaces.

study of frictional phenomena was later performed by Coulomb (1785); see Figure 1.8. He considered the following facts relating to friction: normal pressure, extent of surface area, materials and their surface coatings, ambient conditions (humidity, temperature and vacuum), and time dependency of friction force. These observations resulted in a formula for the frictional resistance to sliding of a body on a plane

$$F_T = A + \frac{N}{\mu^*}, \quad (1.2)$$

where  $F_T$  is the friction force,  $N$  is the normal force and  $\mu^*$  the inverse of the friction coefficient.  $A$  represents cohesion, an effect which was already described in Desaguliers (1725). The second term was attributed to a ploughing action within the interface. This result, today written as  $F_T = A + \mu N$ , is still acceptable, and is the basis for many developments of contact interface laws (see e.g. Tabor (1981)). Again COULOMB found that  $\mu$  is nearly independent of the normal force, the sliding velocity, the contact area (see also results from DA VINCI) and from the surface roughness. However,  $\mu$  depends strongly upon the material pairing in the contact interface. His further, remarkable results concerning the influence of the time of repose upon static friction are discussed in Dowson (1979).

Starting with the classical analytical work of Hertz (1882) the theory of elasticity was applied in contact mechanics. HERTZ investigated the elastic contact of two spheres and derived the pressure distribution in the contact area as well as the approach of the spheres under compression. However very few problems involving contact can be solved analytically. For an overview one may consult the books of Johnson (1985) or Timoshenko and Goodier (1970), and the references therein.

The finite element method developed together with the growing power of modern computers. Hence the first attempts to solve structural problems using finite elements were published in the late fifties (see Turner et al. (1956) or Argyris (1960)). After this, the literature grew enormously since there were many problems of industrial importance which could not be solved analytically. It then took another ten years for the first papers in which methods for the solution of contact problems with finite element methods appeared. As first contributions we list the work by Wilson and Parsons (1970) or Chan and Tuba (1971), which contain early treatments of contact using the geometrically linear theory. However, even at an earlier stage Wilkins (1964) devel-

oped the explicit HEMP-hydrocode which could deal with large strains, and included a simple contact model. Following this, the explicit codes DYNA2D and DYNA3D, as well as the implicit codes NIKE2D and NIKE3D, were developed at the Lawrence Livermore Laboratory by J. HALLQUIST, beginning in the mid-seventies. For the first time these codes provided the possibility to solve contact problems undergoing finite deformations on a large scale in an efficient way.

**Point of departure and connection of chapters.** The design of robust algorithms to treat contact problems efficiently within the finite element method needs input from different sources. These will be considered in the book, which also provides the physical and tribological background within the contact interface. Hence several chapters are devoted to theoretical aspects of continuum mechanics, contact kinematics and the constitutive behaviour in the contact interface. Other chapters contain discretization techniques for solids, and of course, for the contact interface. Furthermore, solution algorithms are discussed, as well as adaptive techniques for contact. Chapters dealing with special contact formulations or topics are also included to complete the treatment of contact problems. An interaction between the chapters will be denoted in the following more detailed description of the contents of the different chapters.

In the first introductory chapter, several contact problems and simple discretizations are treated to present the basic ideas and difficulties of contact mechanics, including coupled and impact problems. This chapter requires no further background besides standard engineering knowledge.

The second chapter is of a more general nature, and discusses the underlying theoretical background for finite deformation solid mechanics, including kinematics, weak forms, linearizations and simple hyperelastic constitutive equations. This chapter is needed to understand the following chapters regarding the kinematics of large deformation contact, and the associated weak formulations. It can be skipped if the reader is familiar with these formulations.

The third chapter discusses contact kinematics from the continuum mechanics point of view. The formulations stated in this chapter are the basis for the derivations in later chapters.

The physical background of the constitutive behaviour in the contact interface is considered in the fourth chapter. This section can be read on its own with a classical background in engineering. It contains material regarding normal and frictional contact for different material pairings, as well as basic formulations for lubrication, adhesion and wear.

The boundary value problem for frictionless and frictional contact is stated in Chapter 5. This also contains different methods on how the contact constraints can be incorporated in the weak forms needed for finite element analysis. This chapter is based on the formulations presented in Chapters 2 and 3. This chapter also contains a section on the treatment of rolling contact based on an Arbitrary LAGRANGIAN EULERIAN (ALE) formulation for stationary and non-stationary processes.

The discretization of solids in contact is derived in Chapter 6 on the basis of the theoretical formulations included in Chapter 2. This chapter is only concerned with

the continuum part of the bodies and hence can be skipped if the reader is familiar with this subject.

The discretization of the contact interfaces is described in Chapters 7 and 8 for linear and nonlinear geometry, respectively. Here interpolation functions and matrix formulations are given for two- and three-dimensional applications. Also, smooth interpolations are introduced to obtain more robust methods for arbitrary contact geometries. Furthermore, new techniques such as mortar or NITSCHKE interpolations are discussed in Chapter 7 which can be used for non-matching meshes. This chapter is based on the material derived in Chapters 2, 3, 4 and 6.

Solution methods for contact problems are contained in Chapter 9. Here different methods of algorithmic treatment are considered for the solution of contact boundary value problems which are defined in the weak sense in Chapter 5. Furthermore, search algorithms for contact are discussed for different applications with respect to global and local search.

In Chapter 10 we treat the coupled thermo-mechanical problem of contact. This chapter is concerned with the heat transfer at the contact interface, which depends upon the mechanical response. Furthermore, the associated finite element discretization for small and finite deformations and the algorithmic treatment of the coupled problem is considered. The contents of this chapter is based on formulations derived in Chapters 2, 3, 4, 7 and 8.

The contact of beam elements is of interest in, for example, the micro-mechanical modelling of woven fabrics. Since the formulations do not fit completely into the general scope, all relevant equations – from the continuous formulation to the finite element discretization – are developed for the beam contact in Chapter 11. Knowledge of the background provided in Chapters 3, 5, 6 and 9 is necessary to understand the derivations.

Adaptive methods for contact problems which are necessary to control the errors inherited in the finite element method are described in Chapter 12. The objective of adaptive techniques is to obtain a mesh which is optimal in the sense that the computational costs involved are minimal under the constraint that the error in the finite element solution is below a certain limit. In general, adaptive methods rely on error indicators and error estimators, which can be computed *a priori* or *a posteriori*. In Chapter 12 an overview over different techniques is given, including different error estimators and indicators. Again, the basic formulations of the solid and the contact constraints from Chapters 2, 3, 6, 7, 8 and 9 are required.

Stability problems which include contact constraints are discussed in Chapter 13. These problems arise in, for example, sheet metal forming, but can also occur in civil engineering applications like the drilling of deep holes. Here the associated algorithms are stated based on the formulations given in Chapters 5 and 9.

# 1

---

## *Introduction to Contact Mechanics*

To introduce the basic methodology and difficulties related to contact mechanics, some simple contact problems will be discussed in this chapter. These are one-dimensional examples undergoing static, thermal or dynamic contact.

### **1.1 CONTACT IN A MASS SPRING SYSTEM**

#### **1.1.1 A variational formulation**

Let us consider a contact problem consisting of a point mass  $m$  under gravitational load which is supported by a spring with stiffness  $k$ . The deflection of the point mass  $m$  is restricted by a rigid plane, see Figure 1.1. The energy for this system can be written as

$$\Pi(u) = \frac{1}{2} k u^2 - m g u. \quad (1.1)$$

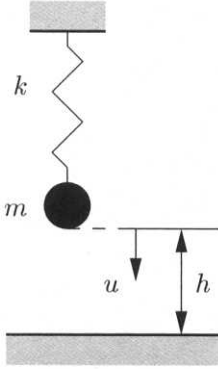
If we do not place any restriction on the displacement  $u$ , then we can compute the extremum of (1.1) by variation, leading to

$$\delta \Pi(u) = k u \delta u - m g \delta u = 0. \quad (1.2)$$

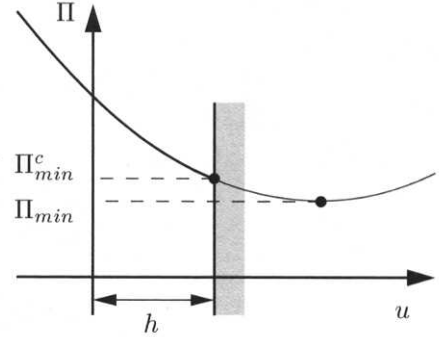
Since the second variation of  $\Pi$  yields  $\delta^2 \Pi = k$ , the extremum of (1.1) is a minimum at  $u = \frac{m g}{k}$ . This is depicted in Figure 1.2, in which the energy of the mass spring system is plotted.

The restriction of the motion of the mass by a rigid support can be described by

$$c(u) = h - u \geq 0, \quad (1.3)$$



**Fig. 1.1** Point mass supported by spring.



**Fig. 1.2** Energy of the mass spring system.

which excludes penetration as an inequality constraint. For  $c(u) > 0$  one has a gap between point mass and rigid support. For  $c(u) = 0$  the gap is closed.

Note that the variation  $\delta u$  is restricted at the contact surface; from (1.3) one obtains  $\delta u \leq 0$ , which means that the virtual displacement has to fulfil the constraint and can only point in the upward direction. The use of this variation in the variational form (1.2) yields an inequality

$$k u \delta u - m g \delta u \geq 0 \quad (1.4)$$

in which the greater sign follows from the fact that the force  $m g$  is greater than the spring force  $k h$  in the case of contact, and that the variation is  $\delta u < 0$  at the rigid support. Equation (1.4) is called a *variational inequality*. Due to the restriction of the solution space by the constraint condition (1.3) the solution of (1.1) is not at the minimum point associated with  $\Pi_{min}$ , but at the point associated with  $\Pi_{min}^c$ , which denotes the minimal energy within the admissible solution space, see Figure 1.2.

Often, instead of the variation  $\delta u$ , one uses the difference between a test function  $v$  and the solution  $u$ :  $\delta u = v - u$ . The test function has to fulfil the condition  $v - h \leq 0$  at the contact point, as also does the solution  $u$ . With the test function  $v$ , (1.2) can be written as

$$k u (v - u) - m g (v - u) = 0. \quad (1.5)$$

Since  $m g > k u$  at the contact point, we have with  $v - h \leq 0$

$$k u (v - h) \geq m g (v - h). \quad (1.6)$$

In both cases, inequality (1.3) which constrains the displacement  $u$  leads to variational inequalities which characterize the solution of  $u$ . These variational inequalities cannot be directly applied to solve the contact problem. For this one has to construct special methods. Some frequently used methods are discussed in the following sections.

Once the point mass contacts the rigid surface, a reaction force  $f_R$  appears. In classical contact mechanics, we assume that the reaction force between rigid surface and point mass is negative, hence the contact pressure can only be compression.

Such assumption excludes adhesion forces in the contact interface and leads to the restriction

$$f_R \leq 0. \quad (1.7)$$

This means that either we have a compression state ( $f_R < 0$ ) or an inactive reaction force ( $f_R = 0$ ).

Summarizing, one has to distinguish two cases within a contact problem where the motion is constrained by (1.3):

1. The spring stiffness is sufficiently large enough that the point mass does not touch the rigid surface. In this case, the following conditions are valid:

$$c(u) > 0 \quad \text{and} \quad f_R = 0. \quad (1.8)$$

2. The data of the system are such that the point mass comes into contact with the rigid support. In that case conditions

$$c(u) = 0 \quad \text{and} \quad f_R < 0 \quad (1.9)$$

hold.

Both cases can be combined in the statement

$$c(u) \geq 0, \quad f_R \leq 0 \quad \text{and} \quad f_R c(u) = 0 \quad (1.10)$$

which are known as HERTZ–SIGNORINI–MOREAU conditions in contact mechanics. Such conditions coincide with KUHN–TUCKER complementary conditions in the theory of optimization.

### 1.1.2 Lagrange multiplier method

The solution of a contact problem in which the motion is constrained by an inequality (1.3) can be obtained using the method of LAGRANGE multipliers. For this we assume that a constraint is active, which means condition (1.9) is fulfilled by the solution. Therefore, the LAGRANGE multiplier method adds to the energy of the system (1.1) a term which contains the constraint and yields

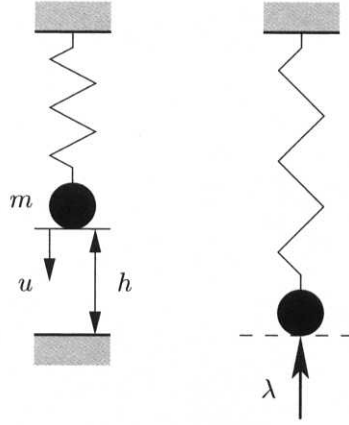
$$\Pi(u, \lambda) = \frac{1}{2} k u^2 - m g u + \lambda c(u). \quad (1.11)$$

A comparison with (1.10) shows that the LAGRANGE multiplier  $\lambda$  is equivalent to the reaction force  $f_R$ . The variation of (1.11) leads to two equations, since  $\delta u$  and  $\delta \lambda$  can be varied independently:

$$k u \delta u - m g \delta u - \lambda \delta u = 0, \quad (1.12)$$

$$c(u) \delta \lambda = 0. \quad (1.13)$$

The first equation represents the equilibrium for the point mass including the reaction force when it touches the rigid surface (see also Figure 1.3), and the second equation



**Fig. 1.3** Point mass supported by a spring and free body diagram for the LAGRANGE multiplier method.

states the fulfillment of the kinematical constraint equation (1.3) for contact:  $u = h$ . Due to that, the variation is no longer restricted, and one can solve for LAGRANGE multiplier  $\lambda$

$$\lambda = k h - m g = f_R . \quad (1.14)$$

However condition (1.7) still has to be checked and fulfilled by the solution (1.14). If this condition is not met, and hence an adhesion force is computed, then the assumption of contact no longer holds. This means the inequality constraint is inactive and the correct solution can be computed from (1.2) as  $u = \frac{m g}{k}$ ; furthermore, the reaction force or LAGRANGE multiplier is zero.

### 1.1.3 Penalty method

Another well known method which is often applied in finite element analysis of contact problems is the penalty approach. Here for an active constraint one adds a penalty term to the energy (1.1) as follows:

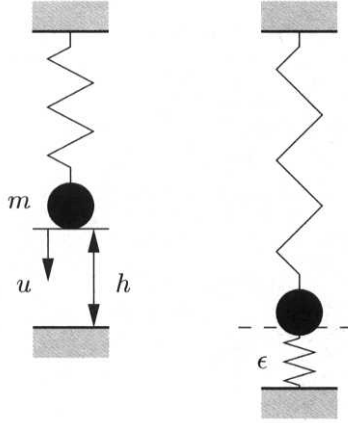
$$\Pi(u) = \frac{1}{2} k u^2 - m g u + \frac{1}{2} \epsilon [c(u)]^2 \quad \text{with } \epsilon > 0 . \quad (1.15)$$

As can be seen in Figure 1.4, the penalty parameter  $\epsilon$  can be interpreted as a spring stiffness in the contact interface between point mass and rigid support. This is due to the fact that the energy of the penalty term has the same structure as the potential energy of a simple spring. The variation of (1.15) yields

$$k u \delta u - m g \delta u - \epsilon c(u) \delta u = 0 , \quad (1.16)$$

from which the solution

$$u = (m g + \epsilon h) / (k + \epsilon) \quad (1.17)$$



**Fig. 1.4** Point mass supported by a spring and a penalty spring due to the penalty term.

can be derived. The value of the constraint equation is then

$$c(u) = h - u = \frac{k h - m g}{k + \epsilon}. \quad (1.18)$$

Since  $m g \geq k h$  in the case of contact, equation (1.18) means that a penetration of the point mass into the rigid support occurs, which is physically equivalent to a compression of the spring, see Figure 1.4. Note that the penetration depends upon the penalty parameter. The constraint equation is only fulfilled in the limit  $\epsilon \rightarrow \infty \implies c(u) \rightarrow 0$ . Hence, in the penalty method we can distinguish two limiting cases:

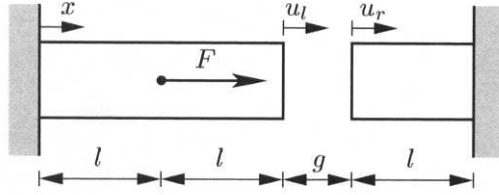
1.  $\epsilon \rightarrow \infty \implies u - h \rightarrow 0$ , which means that one approaches the correct solution for very large penalty parameters. Intuitively, this is clear since that means the penalty spring stiffness is very large, and hence only very small penetration occurs.
2.  $\epsilon \rightarrow 0$  represents the unconstrained solution, and thus is only valid for inactive constraints. In the case of contact, a solution with a very small penalty parameter  $\epsilon$  leads to a high penetration, see (1.18).

The reaction force for a penalty method is computed (see (1.16)) from  $\lambda = \epsilon c(u)$ . For this example, one arrives with (1.18) at

$$\lambda = f_R = \epsilon c(u) = \frac{\epsilon}{k + \epsilon} (k h - m g), \quad (1.19)$$

which in the limit  $\epsilon \rightarrow \infty$  yields the correct solution obtained with the LAGRANGE multiplier method, see (1.14).





**Fig. 1.5** System of two bars and loading.

## 1.2 FINITE ELEMENT ANALYSIS OF THE CONTACT OF TWO BARS

This example shows that even for a system which is built from two simple bars with geometrically linear and elastic behaviour a nonlinear response curve occurs in the case of contact. This is due to the change of stiffness within the contact process.

The potential energy of a bar loaded by  $i$  point loads is given by

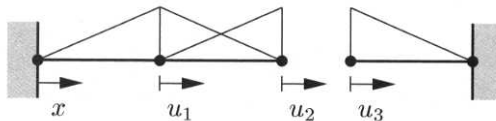
$$\Pi = \frac{1}{2} \int_{(l)} EA [u'(x)]^2 dx - \sum_i F_i u(x_i) \quad (1.20)$$

when distributed forces along the bar are neglected.  $EA$  denotes the axial stiffness,  $u(x)$  is the displacement of the bar and  $F_i$  describes a point load at point  $x_i$ . The problem depicted in Figure 1.5 shows a system consisting of two bars which are separated by a gap  $g$ . When the force  $F$ , acting at  $x = l$ , is large enough the gap will close. We assume that a penetration of bar 1 into bar 2 is impossible. Due to Figure 1.5, this yields the constraint equation

$$u_l - u_r \leq g. \quad (1.21)$$

For  $u_l - u_r < g$  no contact occurs, whereas contact takes place for  $u_l - u_r = g$ .

This system is discretized using three finite elements, two for the left bar and one for the right bar. Linear shape functions are chosen (see Figure 1.6) which already fulfil the boundary conditions at the left and right end of the structure, see also Figure 1.5. The explicit form of the shape functions and their derivatives is given within the



**Fig. 1.6** Shape functions.

elements as

$$\begin{aligned} 0 \leq x \leq l & : u(x) = \frac{x}{l} u_1 & u'(x) = \frac{u_1}{l}, \\ l < x \leq 2l & : u(x) = (2 - \frac{x}{l}) u_1 + (\frac{x}{l} - 1) u_2 & u'(x) = -\frac{u_1}{l} + \frac{u_2}{l}, \\ 2l < x \leq 3l & : u(x) = (3 - \frac{x}{l}) u_3 & u'(x) = -\frac{u_3}{l}. \end{aligned} \quad (1.22)$$

By inserting these interpolations into (1.20), the discretized form of the potential energy can then be derived by integration, leading for the bar system to

$$\Pi = \frac{1}{2} \frac{EA}{l} [u_1^2 + (u_2 - u_1)^2 + u_3^2] - F u_1. \quad (1.23)$$

The variation of  $\Pi$  yields

$$\delta\Pi = \frac{EA}{l} [u_1 \delta u_1 + (u_2 - u_1) (\delta u_2 - \delta u_1) + u_3 \delta u_3] - F \delta u_1 = 0. \quad (1.24)$$

The constraint condition (1.21) is now given by  $u_2 - u_3 \leq g$ :

- i) For  $u_2 - u_3 < g$  displacement  $u_3 = 0$  and no contact occurs. One says that the constraint equation is not active, since the gap is open. In this case, the solution follows directly from (1.24), which has the matrix form

$$\langle \delta u_1, \delta u_2, \delta u_3 \rangle \left\{ \begin{array}{c} \frac{EA}{l} (2u_1 - u_2) - F \\ \frac{EA}{l} (u_2 - u_1) \\ \frac{EA}{l} u_3 \end{array} \right\} = 0, \quad (1.25)$$

leading for arbitrary virtual displacements  $\delta u_i$  to the equation system

$$\frac{EA}{l} \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix} = \begin{Bmatrix} F \\ 0 \\ 0 \end{Bmatrix} \quad (1.26)$$

with the solution

$$u_1 = u_2 = \frac{Fl}{EA}, \quad u_3 = 0. \quad (1.27)$$

- ii) In case the load is increased such that  $F > EA \frac{g}{l}$ , contact occurs and the constraint  $u_2 - u_3 = g$  has to be fulfilled. Now the gap is closed, hence the constraint is active. The solution will be computed using the LAGRANGE multiplier method. As already shown in Section 1.1.2, one then has to add the constraint to the potential energy multiplied by the LAGRANGE multiplier. This yields

$$\Pi_{LM} = \Pi + \lambda g = \Pi + \lambda (g + u_3 - u_2). \quad (1.28)$$

Hence, the variation can be written in case of contact with (1.24) as

$$\delta\Pi_{LM} = \delta\Pi + \lambda (\delta u_3 - \delta u_2) + \delta\lambda (g + u_3 - u_2) = 0, \quad (1.29)$$

where the second term is associated with the reaction force (LAGRANGE multiplier) in the gap. The third term denotes the fulfillment of the constraint equation. The matrix form of (1.29) is given by

$$\langle \delta u_1, \delta u_2, \delta u_3, \delta \lambda \rangle \left\{ \begin{array}{c} \frac{EA}{l} (2u_1 - u_2) - F \\ \frac{EA}{l} (u_2 - u_1) - \lambda \\ \frac{EA}{l} u_3 + \lambda \\ g + u_3 - u_2 \end{array} \right\} = 0, \quad (1.30)$$

leading for arbitrary virtual displacements  $\delta u_i$  and the virtual LAGRANGE multiplier  $\delta \lambda$  to the equation system

$$\begin{bmatrix} 2\frac{EA}{l} & -\frac{EA}{l} & 0 & 0 \\ -\frac{EA}{l} & \frac{EA}{l} & 0 & -1 \\ 0 & 0 & \frac{EA}{l} & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \\ \lambda \end{Bmatrix} = \begin{Bmatrix} F \\ 0 \\ 0 \\ -g \end{Bmatrix}. \quad (1.31)$$

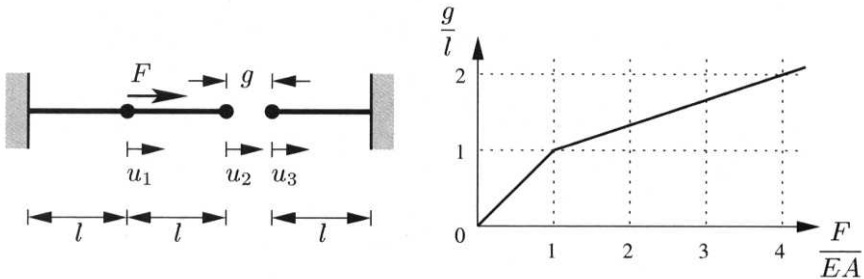
The solution of this system for  $u_2$  and  $\lambda$  leads to

$$u_2 = \frac{1}{3} \left( 2g + \frac{Fl}{EA} \right) \text{ and } \lambda = \frac{1}{3} \left( EA \frac{g}{l} - F \right). \quad (1.32)$$

Observe that the LAGRANGE multiplier fulfils condition (1.7), since  $F > (EA g)/l$  when the gap is closed, see also (1.27). From (1.32) one can now compute the dependency between load  $F$  and displacement  $u_2$ :

$$F = EA \left( 3 \frac{u_2}{l} - 2 \frac{g}{l} \right). \quad (1.33)$$

Figure 1.7 depicts the nonlinear load-deflection curve for the complete analysis. It is clear that the stiffness of the bar system increases when contact occurs; this can be observed from the fact that the load has to be three times as big to obtain the same increment to the displacement when the gap is closed as in the case when the gap is open.



**Fig. 1.7** Finite element discretization and load-deflection curve.

Condition (1.32) includes, for  $u_2 = g$ , the limiting case of the initiation of contact in which  $\lambda = 0$ . This case can also be obtained from (1.27).

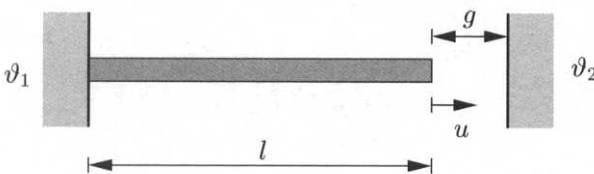
### Implications for numerical methods

- Generally, one can observe that in the case of contact two different states of the structural system are possible. One is related to the open gap, see (i), the other to the closed gap, see (ii). Both cases were solved using a different equation system, which means that the topology of the structure changes due to contact. This points out one of the difficulties when solving contact problems: the system matrix changes its size (or non-zero form) with active or inactive constraint equations. As will be seen in later chapters, this can also include a change of the topology when one finite element node moves during the deformation process from one element to another.
- Furthermore, we have the choice between different methods for the treatment of contact problems, including the LAGRANGE multiplier or the penalty formulation. The former introduces additional variables in the system, but does fulfil the constraint equation correctly; the latter leads to non-physical penetration, but has no additional variables. So both methods have advantages and disadvantages, which will be discussed in later chapters in detail, together with techniques to overcome the problems discussed above.

## 1.3 THERMO-MECHANICAL CONTACT

Contact can occur in a coupled thermo-mechanical analysis when two bodies have different temperatures. To show some of the main effects, the following example of a bar which contacts a rigid wall is investigated.

We can consider a problem as specified in Figure 1.8. The bar is fixed at the left end and heated at that point with a temperature of  $\vartheta_1$ . On the other side there is a gap between the end of the bar and the rigid wall which has temperature  $\vartheta_2$ . Hence we have to distinguish two situations: contact of the bar with the wall ( $\vartheta(l) = \vartheta_2$ ), and the open gap ( $\vartheta(l) = 0$ ). The material properties of the bar are given by the axial stiffness  $EA$  and the coefficient of heat transfer  $\alpha_T$ . This system will be analyzed under the assumption of steady state solution, thus time dependent solutions will not



**Fig. 1.8** Contact of a bar due to thermal heating.

be considered. Furthermore, the mechanical constitutive properties are assumed to be independent of the temperature. For this we can write the following equations for the mechanical and thermal problem:

### Mechanical problem

$$\begin{aligned}
 \text{Kinematics:} \quad & \varepsilon_{el} = \frac{du}{dx} - \alpha_T (\vartheta(x) - \vartheta_0) \\
 \text{Equilibrium:} \quad & \frac{d\sigma}{dx} = 0 \\
 \text{Constitutive equation:} \quad & \sigma = E \varepsilon_{el}
 \end{aligned} \tag{1.34}$$

with a given reference temperature  $\vartheta_0$ , the stress  $\sigma$ , and the elastic strain  $\varepsilon_{el}$  in the  $x$ -direction and YOUNG'S modulus  $E$ .

### Heat conduction

$$\begin{aligned}
 \text{Heat balance:} \quad & -\frac{dq}{dx} = 0 \\
 \text{Constitutive equation:} \quad & q = -k \frac{d\vartheta}{dx}
 \end{aligned} \tag{1.35}$$

where  $q$  is the heat flux,  $\vartheta$  is the temperature and  $k$  is the thermal conductivity in FOURIER'S law. Note that the assumption of steady state solutions has been made, and no internal heat will be generated in the bar.

The differential equation which governs the mechanical behaviour of the bar results from equations (1.34):

$$E \frac{d}{dx} \left[ \frac{du}{dx} - \alpha_T (\vartheta(x) - \vartheta_0) \right] = 0 \tag{1.36}$$

in the same way, from (1.35) one derives

$$\frac{d^2 \vartheta}{dx^2} = 0. \tag{1.37}$$

The mechanical and thermal problems are decoupled in the sense that the heat equation does not depend upon the mechanical quantities. So one can always solve for the thermal field  $\vartheta(x)$  independently of the mechanical field. Coupling is present in the case of finite deformations, and when dissipative processes like friction or plasticity have to be considered. This will be discussed in detail in later chapters.

Within the analysis one has to distinguish two different solution states. In the first the gap is still open, and in the second the gap is closed. This is the standard situation when contact is present (see also the previous section).

- i) Gap open (inactive constraint): in this situation no contact has been made. Hence the inequality  $u < g$  is valid, together with the fact that no contact pressure occurs. From (1.37) there follows a constant temperature distribution

in the bar with  $\vartheta(x) = \vartheta_1$ . Furthermore, the bar is stress free. Thus for elongation of the bar, the solution yields

$$u = \alpha_T (\vartheta_1 - \vartheta_0) l. \quad (1.38)$$

Observe, that the gap closes for a temperature of

$$\vartheta_g = \vartheta_0 + \frac{g}{\alpha_T l}. \quad (1.39)$$

- ii) Gap closed (active constraint): for a temperature which is larger than  $\vartheta_g$  the gap is closed. In that case, from (1.37) one obtains a linear temperature distribution along the bar  $\vartheta(x) = \vartheta_1 + (\vartheta_2 - \vartheta_1) \frac{x}{l}$  when a perfect conductance is assumed in the contact point. For the displacement the condition  $u = g$  holds, and hence the stress follows with (1.34) from

$$\begin{aligned} g &= \int_0^l \left\{ \frac{\sigma}{E} + \alpha_t \left[ (\vartheta_1 - \vartheta_0) + (\vartheta_2 - \vartheta_1) \frac{x}{l} \right] \right\} dx, \\ \sigma &= E \left[ \frac{g}{l} - \frac{\alpha_T}{2} ((\vartheta_1 - 2\vartheta_0) + \vartheta_2) \right]. \end{aligned} \quad (1.40)$$

As long as the second expression is now larger than the first term, a negative stress occurs in the bar, hence the contact stress is also negative and condition (1.9) is fulfilled.

However, if the temperature  $\vartheta_2$  is such that the second term in (1.40) is smaller than the first term, then a positive stress occurs, which means that the gap opens up again. This results in an on-off contact state, since after opening the temperature in the bar again changes to the constant value  $\vartheta_1$ , leading to contact. Hence the solution is no longer stable. Since such a response has never been observed in experiments, one has to reformulate the problem in such a way that this instability does not occur. One method which yields a unique solution introduces a pressure-dependent heat conduction  $h(\sigma)$  at the contact point. Such constitutive response can also be derived from micromechanical observations, e.g. see Section 10.2. A simple relation is given by

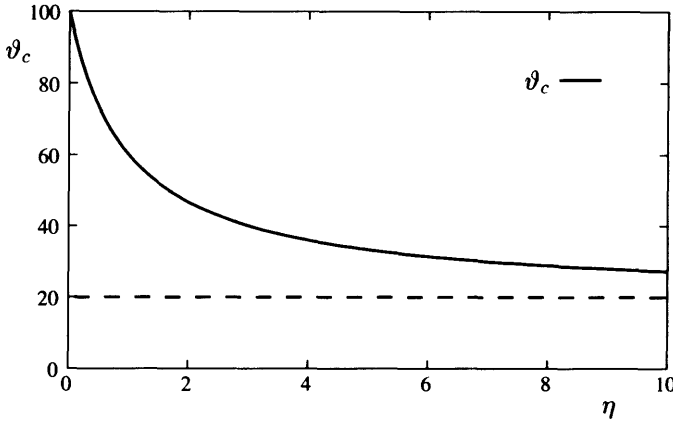
$$h(\sigma) = h_c \left( \frac{\sigma}{\bar{H}} \right)^\beta \quad (1.41)$$

with the thermal conductivity  $h_c$  in the contact point, the hardness of the material  $\bar{H}$  and a positive exponent  $\beta$  which has to be determined from experiments. The heat conduction in the contact interface is then given by

$$q_c = h(\sigma) (\vartheta_c - \vartheta_2), \quad (1.42)$$

where  $\vartheta_c$  is the contact temperature. Since  $q_c = -k \frac{d\vartheta}{dx}$ , from (1.41) and (1.42) one obtains the differential equation

$$-k \frac{d\vartheta}{dx} = h(\sigma) (\vartheta_c - \vartheta_2), \quad (1.43)$$



**Fig. 1.9** Pressure-dependent contact temperature  $\vartheta_c$ .

which has the solution

$$\vartheta(x) = \int \frac{h(\sigma)}{k} (\vartheta_2 - \vartheta_c) dx + C. \quad (1.44)$$

Evaluation of this equation at the contact point  $c$  by considering the boundary condition  $\vartheta(0) = \vartheta_1$  yields

$$\vartheta_c = \frac{\vartheta_1 + \eta \vartheta_2}{1 + \eta} \quad \text{with } \eta = \frac{h(\sigma) l}{k}, \quad (1.45)$$

which means that there is a jump in the temperature at the contact point, since  $\vartheta_c \neq \vartheta_2$ . The contact temperature is depicted in Figure 1.9 as a function of the dimensionless parameter  $\eta$  which includes the pressure dependency (large  $\eta$  means a higher contact pressure). The curve in Figure 1.9 is plotted for the values  $\vartheta_1 = 100K$  and  $\vartheta_2 = 20K$ . The limit cases are  $\eta = 0 \Rightarrow \vartheta_c = \vartheta_1$  and  $\eta \rightarrow \infty \Rightarrow \vartheta_c \rightarrow \vartheta_2$ , as can be seen in Figure 1.9. Hence for small contact pressures, almost no heat is conducted through the surface. Due to the possibility of incorporating a temperature jump at the boundary, the solution of the thermo-mechanical contact problem is stable. However, the solution for the contact stress now has to be computed from a nonlinear equation, which follows from the condition  $u = g$  with (1.34) and (1.45):

$$\begin{aligned} g &= \int_0^l \left\{ \frac{\sigma}{E} + \alpha_T \left[ \vartheta_1 + (\vartheta_c - \vartheta_1) \frac{x}{l} \right] \right\} dx, \\ \Rightarrow 0 &= -\frac{g}{l} + \frac{\sigma}{E} + \frac{\alpha_T}{2} \frac{2\vartheta_1 + \eta(\vartheta_1 + \vartheta_2)}{1 + \eta}. \end{aligned} \quad (1.46)$$

where  $\eta$  is a function of  $\sigma$  defined in (1.45) and (1.41).

## Implications for numerical methods

- In the case of thermo-mechanical contact, in general one has to solve a system of coupled field equations which leads to algorithms for different problem classes when non-stationary processes are involved, since the time dependency of the heat conduction equation is first order, and second order for the equations of the solid.
- In the contact zone a pressure-dependent constitutive equation is needed to avoid instability. This means that one has to use a finite element discretization technique for contact, which yields the contact pressure and not a contact force.

## 1.4 IMPACT

When two bodies which have different velocities come into contact an impact occurs. Within an impact analysis one is interested in the velocities of the bodies after impact and in the impact force as a function of time.

Here a one-dimensional example is discussed in which a bar of length  $l_1$  impacts another bar of length  $l_2$ , see Figure 1.10. Both bars have the same material properties  $EA_1 = EA_2 = EA$  and densities  $\rho_1 = \rho_2 = \rho$ . The left bar has an initial velocity of  $v_{01}$ , whereas the right bar is at rest.

The solution of this problem can be derived from the one-dimensional wave equation

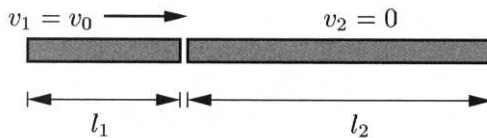
$$EA \frac{\partial^2 u}{\partial x^2} = -\rho A \frac{\partial^2 u}{\partial t^2}. \quad (1.47)$$

Furthermore, one has to fulfil the initial and boundary conditions of the problem stated in Figure 1.10 and the standard contact conditions (1.10), which describe that no penetration can occur at the contact point, and also that the contact force has to be a compression force.

The solution of (1.47) is given by

$$u(t) = f\langle x - ct \rangle + g\langle x + ct \rangle \quad \text{with} \quad c = \sqrt{\frac{E}{\rho}}, \quad (1.48)$$

where  $c$  denotes the speed of wave travelling in the bars. Function  $f$  corresponds to a wave travelling in the  $x$ -direction of the bar, while  $g$  is associated with a wave



**Fig. 1.10** Longitudinal impact of two bars.



travelling in the opposite direction. By differentiation of both sides of (1.48) with respect to  $x$  and  $t$ , one derives

$$\varepsilon = \frac{\partial u}{\partial x} = f'\langle x - ct \rangle + g'\langle x + ct \rangle, \quad (1.49)$$

$$v = \frac{\partial u}{\partial t} = c [-f'\langle x - ct \rangle + g'\langle x + ct \rangle], \quad (1.50)$$

with  $f' = \frac{\partial f}{\partial x}$  and  $g' = \frac{\partial g}{\partial x}$ . The identity

$$\frac{\partial u}{\partial t} = c \frac{\partial u}{\partial x} \quad (1.51)$$

follows by comparing the last two relations. Using (1.51), one can conclude that the normal stress  $\sigma$  in the bar is given by

$$\sigma = E\varepsilon = \frac{E}{c} \frac{\partial u}{\partial t} = \rho c \frac{\partial u}{\partial t}, \quad (1.52)$$

which shows that there is a linear relationship between the stress at any point in the bar and the particle velocity. Hence when a wave travels with speed  $c$  along the bar, there is also a stress pulse which travels with the same velocity. When such a pulse reaches the free end of the bar, one can compute the behaviour of the pulse from the condition that the end of the bar has to be stress free. This leads, with (1.50), to the condition

$$\sigma = E u'(t) = 0 = E (f'\langle x - ct \rangle + g'\langle x + ct \rangle) \quad \forall t, \quad (1.53)$$

from which a relation between  $f'$  and  $g'$  follows for the free ends at  $x = 0$  and  $x = l_1 + l_2$ ,

$$f'\langle x - ct \rangle = -g'\langle x + ct \rangle \quad \forall t. \quad (1.54)$$

Thus a reflection occurs at the free ends with equal amplitude in the stress pulse but with opposite velocity. Furthermore, the initial conditions can be stated for the impact of two bars described in Figure 1.10 as:

$$\begin{aligned} v = \frac{\partial u}{\partial t} \Big|_{t=0} &= c [-f'\langle x \rangle + g'\langle x \rangle] = v_0 \text{ for } 0 \leq x \leq l_1, \\ v = \frac{\partial u}{\partial t} \Big|_{t=0} &= c [-f'\langle x \rangle + g'\langle x \rangle] = 0 \text{ for } l_1 < x \leq l_1 + l_2, \\ \sigma = E \frac{\partial u}{\partial x} \Big|_{t=0} &= f'\langle x \rangle + g'\langle x \rangle = 0 \text{ for } 0 \leq x \leq l_1 + l_2. \end{aligned} \quad (1.55)$$

From these conditions follow the initial values of  $f'$  and  $g'$  as

$$\begin{aligned} f'\langle x \rangle &= -\frac{v_0}{2c} & g'\langle x \rangle &= \frac{v_0}{2c} & \text{for } 0 \leq x \leq l_1, \\ f'\langle x \rangle &= 0 & g'\langle x \rangle &= 0 & \text{for } l_1 < x \leq l_1 + l_2. \end{aligned} \quad (1.56)$$

The problem stated in Figure 1.10 can be solved with the relations stated above. Since the left bar has an initial velocity of  $v_0$ , one has a distribution of  $f'$  and  $g'$ , as shown in Figure 1.11 for  $t = 0$ . These are associated with two waves, one travelling in the  $x$ -direction and the other in the opposite direction. Several states of the waves are depicted at certain times in Figure 1.11 for the case in which bar 2 has length  $l_2 = 2l_1$ . Figure 1.11 also shows the stress distribution due to the travelling strain waves. The two bodies remain in contact until time  $T_{imp} = 4l_1/c$ , which corresponds to the time at which the reflected wave in bar 2 arrives at the contact point. Since the first bar is stress free, this wave encounters a free end and hence does not enter bar 1, but reflects due to the stress free boundary condition. After that time, the bars are no longer in contact. The final velocity of bar 1 after impact is  $v_{1e} = 0$ , and for bar 2 the velocity is then  $v_{2e} = v_0/2$ . As can be seen in Figure 1.11, there is still an oscillation due to the travelling stress wave in bar 2, whereas bar 1 is at rest. If one assumes that both bars are made of steel ( $E = 2.1 \cdot 10^8 \text{ kN/m}^2$ ,  $\rho = 7.85 \cdot 10^3 \text{ kg/m}^3$ ), and that  $l_1 = 1 \text{ m}$ , then the wave speed is  $c = \sqrt{E/\rho} = 5172 \text{ m/s}$ , and hence the impact time is  $T_{imp} = 4/5172 = 7.73 \cdot 10^{-4} \text{ s}$ . If the initial velocity is chosen to be  $5 \text{ m/s}$ , a stress amplitude of  $\sigma = \rho c v_0/2$  follows from equation (1.52). This leads in this example to a stress of  $\sigma = 7.85 \cdot 10^3 \cdot 5172 \cdot 2.5 = 10.2 \cdot 10^4 \text{ kN/m}^2$ , which represents 42% of the yield stress ( $\sigma_Y = 24 \cdot 10^4$ ) of a standard steel.

It is interesting to note that the classical impact theory for rigid bodies yields, under the assumption of an elastic impact, the final velocities  $v_{1e} = -v_0/3$  and  $v_{2e} = 2v_0/3$ , which are different when compared to the wave solution above. This is due to the oscillations remaining in bar 2 after impact, which is, as also the impact time, neglected in the case of rigid body impact.

Another possibility to solve the wave equation (1.47) is by separation of variables. Using

$$u(x, t) = v(x) \tau(t) \quad (1.57)$$

one derives

$$c^2 \frac{v''}{v} = \frac{\ddot{\tau}}{\tau}, \quad (1.58)$$

which has the solution

$$\begin{aligned} v(x) &= A \cos kx + B \sin kx \\ \tau(x) &= a \cos \omega t + b \sin \omega t \end{aligned} \quad k^2 = \frac{\omega^2}{c^2}. \quad (1.59)$$

For the bar system with free ends, one obtains with the boundary conditions  $\sigma(0) = \sigma(3l_1) = 0$  the equation  $\sin k(3l_1) = 0$ , which has the eigenvalues  $k_n(3l_1) = n\pi$  for  $n = 1, 2, 3, \dots$ . The related eigenfunctions are

$$\varphi_n(x) = \cos \frac{n\pi}{3l_1} x, \quad (1.60)$$

which with (1.59) yield the solution

$$u(x, t) = \sum_{n=1,2,3,\dots}^{\infty} \left[ a_n \cos c \frac{n\pi}{3l_1} t + b_n \sin c \frac{n\pi}{3l_1} t \right] \cos \frac{n\pi}{3l_1} x, \quad (1.61)$$

which has to be adjusted to the initial conditions. These then lead to a *FOURIER* series representation of the initial conditions in terms of *sin* and *cos* functions. Here details will be omitted. They can be found in standard textbooks. However, note that there are two possibilities for solving the impact problem. The latter has the inconvenience that an overshooting can occur which is a high oscillatory result near the wave fronts.

### **Implications for numerical methods**

- As shown above for impact problems, the impact time is very short and the stresses generated are high. Hence, the numerical methods to solve impact problems have to include nonlinear material behaviour and have to be designed for short time responses.
- Due to the possibility of high oscillatory responses near wave fronts, one has to be careful when constructing algorithms for impact problems, in the sense that one should not destroy the wave front characteristics within the numerical scheme.

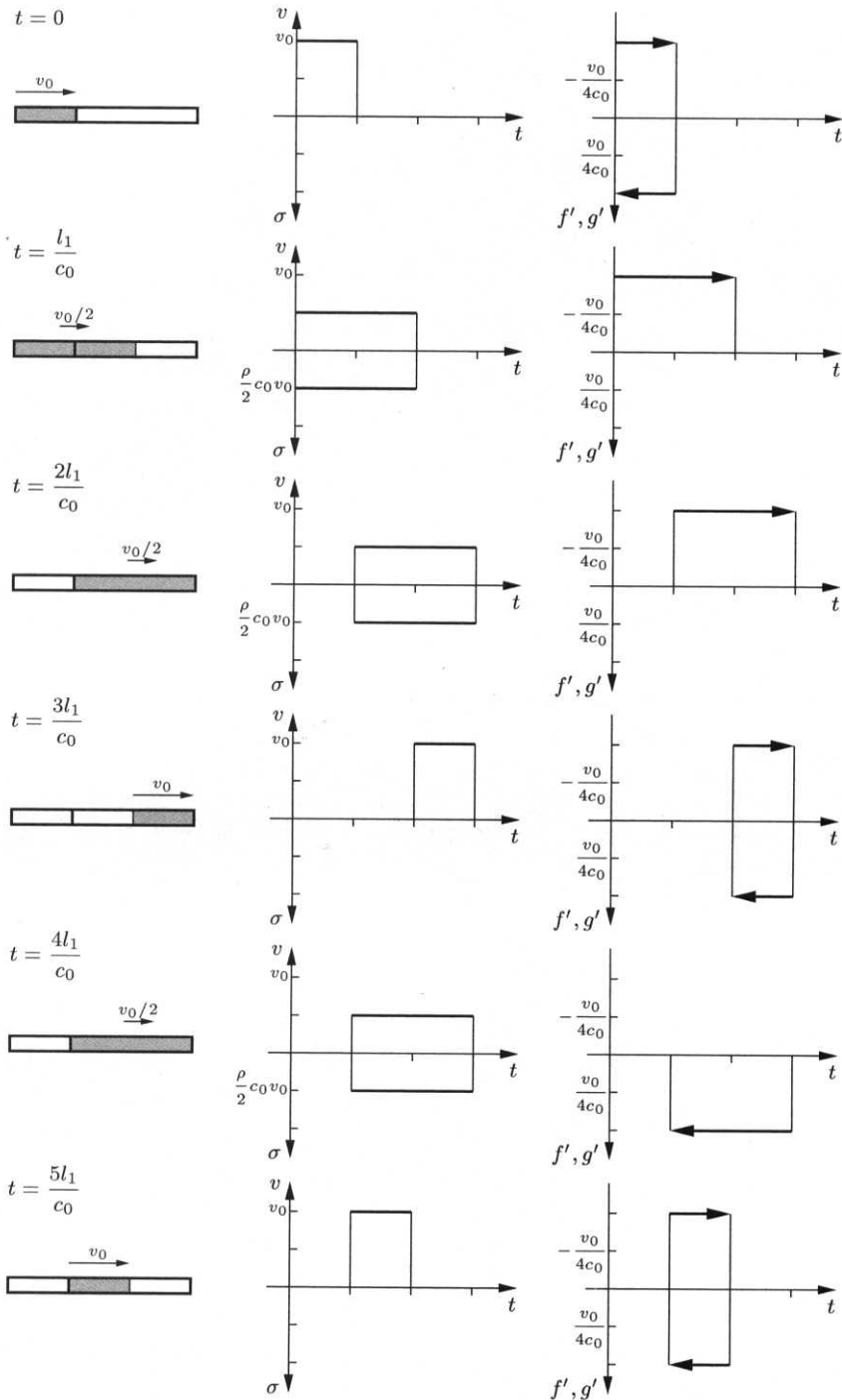


Fig. 1.11 Wave solution for bar impact.

**This page intentionally left blank**

# 2

---

## *Continuum Solid Mechanics and Weak Forms*

The deformation of solids is generally described by the kinematic relations, the equations of balance and the constitutive equations. This chapter summarizes the main equations which govern the deformation of solids. For a detailed treatment of this subject the reader should consult the literature, e.g. the standard books of Eringen (1967), Malvern (1969), Truesdell and Noll (1965), Truesdell and Toupin (1960), Ogden (1984) or Chadwick (1999).

### 2.1 KINEMATICS

#### 2.1.1 Motion and deformation gradient

In this section we discuss the motion and deformation of continua. A body  $\mathcal{B}$  can be described by a set of points which are in a region of the EUCLIDEAN space  $\mathbb{E}^3$ . A configuration of  $\mathcal{B}$  is then a one-to-one mapping  $\varphi: \mathcal{B} \rightarrow \mathbb{E}^3$ , which places the particles of  $\mathcal{B}$  in  $\mathbb{E}^3$ . The position of a particle  $X$  of  $\mathcal{B}$  in the configuration  $\varphi$  is defined by  $\mathbf{x} = \varphi(X)$ . The placement of the body  $\mathcal{B}$  is described by  $\varphi(\mathcal{B}) = \{\varphi(X) \mid X \in \mathcal{B}\}$  and therefore be denoted as configuration  $\varphi(\mathcal{B})$  of body  $\mathcal{B}$ .

The motion of body  $\mathcal{B}$  is then a temporally parametric series of configurations  $\varphi_t: \mathcal{B} \rightarrow \mathbb{E}^3$ . For the position of the particle  $X$  at time  $t \in \mathbb{R}^+$  we have

$$\mathbf{x} = \varphi_t(X) = \varphi(X, t). \quad (2.1)$$

This equation describes a curve in  $\mathbb{E}^3$  for the particle  $X$ .  $\mathbf{X} = \varphi_0(X)$  defines the reference configuration of body  $\mathcal{B}$ , where  $\mathbf{X}$  is the position of particle  $X$  in this

configuration. With (2.1) we have

$$\mathbf{x} = \varphi(\varphi_0^{-1}(\mathbf{X}), t). \quad (2.2)$$

**REMARK 2.1:** *Body  $\mathcal{B}$  does not have to assume the reference configuration at any time. Since the reference configuration can be chosen in an arbitrary way, it is often selected to coincide with the initial configuration. Later, in the sections regarding finite element discretizations, it will be advantageous to use a special reference configuration, e.g. for isoparametric elements, which can be handled in a simple yet somewhat artificial manner.*

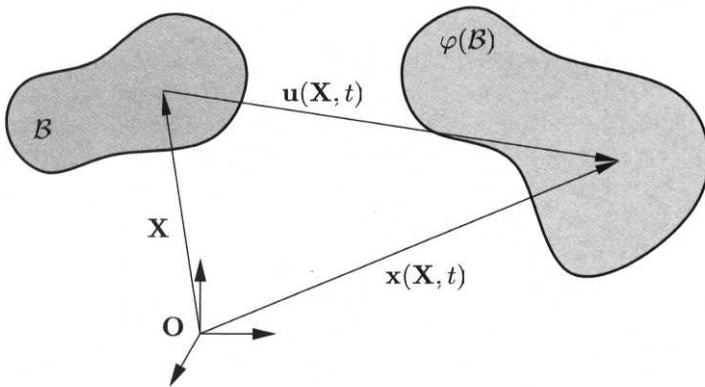
For practical applications we do not need to differentiate between  $\mathbf{X}$  and  $X$ . This simplifies the notation, and we can write (2.2) as

$$\mathbf{x} = \varphi(\mathbf{X}, t), \quad (2.3)$$

where  $\mathbf{X}$  depicts the position of particle  $X$  in the reference configuration  $\mathcal{B}$ . With this, the positions  $\mathbf{x}$  and  $\mathbf{X}$  are described as vectors in  $\mathbb{E}^3$  with respect to the origin  $\mathbf{O}$ , as shown in Figure 2.1. The point  $X$  is denoted in the reference configuration by the position vector  $\mathbf{X} = X_A \mathbf{E}_A$ . Here  $\mathbf{E}_A$  defines an orthogonal basis in the reference configuration with origin  $\mathbf{O}$ . Therefore (2.3) can be written in components:

$$x_i = \varphi_i(X_A, t). \quad (2.4)$$

In the following by indices in capital letters we will denote components of vectors and tensors if these refer to the basis  $\mathbf{E}_A$  of the reference configuration.  $X_A$  are the LAGRANGE coordinates of the particle  $X$ . Small letters are used for indices which refer to the basis  $\mathbf{e}_i$  of the spatial or current configuration. The quantities  $x_i$  denote the spatial coordinates of  $X$ . To simplify notation, we employ an orthogonal Cartesian basis. This coincides with the finite element method, since isoparametric interpolations are always defined in an orthogonal Cartesian basis. The change to arbitrary curvilinear coordinates is a purely technical matter.



**Fig. 2.1** Configurations of body  $\mathcal{B}$ .

The equations of mechanics of continua can be formulated with respect to the deformed or undeformed configuration of a body  $\mathcal{B}$ . From the theoretical standpoint, there is no difference whether the equations refer to the current or the reference configuration of the body. However, one should consider the implications due to physical modelling, as in plasticity. When formulating numerical methods for continua, considerable differences in efficiency can occur when the equations are related to either the spatial or the reference configuration. Thus we will define strain measures with respect to both configurations. Within this discussion, we denote by small letters tensors which refer to the current configuration  $\varphi(\mathcal{B})$ , and we use capital letters for the reference configuration  $\mathcal{B}$ .

To describe the deformation process locally, we introduce the deformation gradient  $\mathbf{F}$  which maps tangent vectors of the reference configuration to tangent vectors in the spatial configuration. It is a tensor which associates to a material line element  $d\mathbf{X}$  in  $\mathcal{B}$  the line element  $d\mathbf{x}$  in  $\varphi(\mathcal{B})$ :

$$d\mathbf{x} = \mathbf{F} d\mathbf{X}. \quad (2.5)$$

The components of the deformation gradient follow from the direct notation  $\mathbf{F} = \partial \mathbf{x} / \partial \mathbf{X}$  as partial derivatives  $\partial x_i / \partial X_A = x_{i,A}$ . With (2.3) and (2.4) we obtain

$$\mathbf{F} = \text{Grad } \varphi(\mathbf{X}, t) = F_{iA} \mathbf{e}_i \otimes \mathbf{E}_A = \frac{\partial x_i}{\partial X_A} \mathbf{e}_i \otimes \mathbf{E}_A. \quad (2.6)$$

Since the gradient (2.6) is a linear operator, the local transformation (2.5) is also linear. To preserve the continuous structure in  $\mathcal{B}$  during the deformation, the mapping (2.5) has to be one-to-one, i.e.  $\mathbf{F}$  cannot be singular. This is equivalent to the condition

$$J = \det \mathbf{F} \neq 0, \quad (2.7)$$

where  $J$  defines the JACOBIAN determinant. Furthermore, to exclude self-penetration of the body,  $J$  has to be greater than 0. Thus its inverse exists, which is denoted by  $\mathbf{F}^{-1}$ . With this we can invert equation (2.5):

$$d\mathbf{X} = \mathbf{F}^{-1} d\mathbf{x}. \quad (2.8)$$

The inverse of the deformation gradient has the following form:

$$\mathbf{F}^{-1} = (F_{iA})^{-1} \mathbf{E}_A \otimes \mathbf{e}_i \quad \text{with} \quad (F_{iA})^{-1} = \left( \frac{\partial x_i}{\partial X_A} \right)^{-1} = \frac{\partial X_A}{\partial x_i}, \quad (2.9)$$

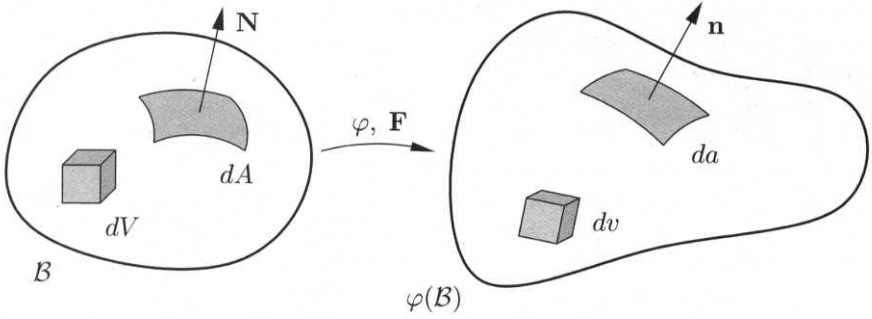
where  $\mathbf{X} = \varphi^{-1}(\mathbf{x})$ .

It is well known that the deformation gradient  $\mathbf{F}$  can be decomposed by the polar decomposition theorem into a stretching and a rotational part, e.g. see Malvern (1969):

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}, \quad F_{iA} = R_{iB} U_{BA} = V_{ik} R_{kA}, \quad (2.10)$$

where  $\mathbf{U}$  is the right stretch tensor with a basis in the reference configuration, and  $\mathbf{V}$  is the left stretch tensor which is an object in the current configuration. The rotation tensor  $\mathbf{R}$  is a two-field tensor which connects both configurations.





**Fig. 2.2** Transformation between area and volume elements.

Once the deformation gradient  $\mathbf{F}$  is known, transformations of area and volume elements between  $\mathcal{B}$  and  $\varphi(\mathcal{B})$  can be derived. The transformation of area elements between  $\mathcal{B}$  and  $\varphi(\mathcal{B})$  is given by the formula due to NANSON (see e.g. Ogden (1984), pp. 88):

$$\boxed{d\mathbf{a} = \mathbf{n} da = J \mathbf{F}^{-T} \mathbf{N} dA = J \mathbf{F}^{-T} d\mathbf{A}}. \quad (2.11)$$

In this equation  $\mathbf{n}$  is the normal to the surface of  $\varphi(\mathcal{B})$  and  $\mathbf{N}$  denotes the normal to the surface of  $\mathcal{B}$  (see Figure 2.2).  $J$  is the JACOBI determinant defined in (2.7) and  $da$  (respectively  $dA$ ) are the surface elements in the associated configuration. For the transformation of volume elements from the reference to the spatial configuration, we have

$$\boxed{dv = J dV}. \quad (2.12)$$

With the introduction of a displacement vector  $\mathbf{u}(\mathbf{X}, t)$  as the difference in position vectors of a point in the reference and current configurations,

$$\mathbf{u}(\mathbf{X}, t) = \varphi(\mathbf{X}, t) - \mathbf{X}, \quad (2.13)$$

we can write for the deformation gradient (2.6)

$$\mathbf{F} = \text{Grad} [\mathbf{X} + \mathbf{u}(\mathbf{X}, t)] = \mathbf{1} + \text{Grad} \mathbf{u} = \mathbf{1} + \mathbf{H}, \quad (2.14)$$

where  $\mathbf{H} = \text{Grad} \mathbf{u}$  is the displacement gradient with respect to  $\mathbf{X}$ .

### 2.1.2 Strain measures

In this section we describe different strain measures which will be applied later. One of the most common strain measures is the right CAUCHY–GREEN tensor  $\mathbf{C}$ , which refers to the initial configuration  $\mathcal{B}$

$$\boxed{\mathbf{C} = \mathbf{F}^T \mathbf{F}}. \quad (2.15)$$

Since this strain measure is not zero at the initial state (there we have  $\mathbf{F} = \mathbf{1} \Rightarrow \mathbf{C} = \mathbf{1}$ ), it is convenient to introduce the GREEN-LAGRANGIAN strain tensor  $\mathbf{E}$  which refers to the initial configuration  $\mathcal{B}$

$$\mathbf{E} = \frac{1}{2} (\mathbf{F}^T \mathbf{F} - \mathbf{1}) = \frac{1}{2} (\mathbf{C} - \mathbf{1}), \quad E_{AB} = \frac{1}{2} (F_{iA} F_{iB} - \delta_{AB}). \quad (2.16)$$

In (2.16)  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$  is the positive definite right CAUCHY-GREEN tensor which expresses the square of the infinitesimal line element  $d\mathbf{x}$  via the material line element  $d\mathbf{X}$  by  $d\mathbf{x} \cdot d\mathbf{x} = d\mathbf{X} \cdot \mathbf{C} d\mathbf{X}$ . Thus the strain  $\mathbf{E}$  is the difference of the square of the line elements in  $\mathcal{B}$  and  $\varphi(\mathcal{B})$ . Furthermore, we have  $\mathbf{C} = \mathbf{U}^T \mathbf{U} = \mathbf{U}^2$ .

The spectral decomposition of  $\mathbf{C}$  can often be advantageous when hyperelastic constitutive equations have to be formulated. We have

$$\mathbf{C} = \sum_{i=1}^3 \lambda_i^2 \mathbf{N}_i \otimes \mathbf{N}_i, \quad (2.17)$$

where  $\lambda_i$  and  $\mathbf{N}_i$  follow from the eigenvalue problem

$$(\mathbf{C} - \lambda_i^2) \mathbf{N}_i = \mathbf{0}. \quad (2.18)$$

Based on the spectral decomposition (2.17), we define more general strain measures:

$$\mathbf{E}^\alpha = \frac{1}{\alpha} \sum_{i=1}^3 (\lambda_i^\alpha - 1) \mathbf{N}_i \otimes \mathbf{N}_i \quad \text{and} \quad \mathbf{e}^\alpha = \frac{1}{\alpha} \sum_{i=1}^3 (\lambda_i^\alpha - 1) \mathbf{n}_i \otimes \mathbf{n}_i. \quad (2.19)$$

$\mathbf{E}^\alpha$  refers to the reference configuration, and  $\mathbf{e}^\alpha$  has its bases in the current configuration. As special cases we obtain the GREEN-LAGRANGIAN strain tensor

$$\mathbf{E} = \sum_{i=1}^3 \frac{1}{2} (\lambda_i^2 - 1) \mathbf{N}_i \otimes \mathbf{N}_i. \quad (2.20)$$

With respect to the current configuration, the ALMANSI strain tensor

$$\mathbf{e} = \frac{1}{2} (\mathbf{1} - \mathbf{b}^{-1}), \quad e_{ik} = \frac{1}{2} [\delta_{ik} - (F_{iA})^{-1} (F_{kA})^{-1}] \quad (2.21)$$

is often applied. It is defined with the left CAUCHY-GREEN tensor  $\mathbf{b} = \mathbf{F} \mathbf{F}^T$ . The ALMANSI strain tensor is connected to the GREEN-LAGRANGE strain tensor via the following transformation:

$$\mathbf{E} = \mathbf{F}^T \mathbf{e} \mathbf{F}, \quad (2.22)$$

which can easily be verified with (2.21).

### 2.1.3 Transformation of vectors and tensors

Since we know the transformation between differential elements in the reference and current configuration, we can also transform vectors or tensors from reference to current configuration, and vice versa. This stems from the fact that the base vectors can be viewed as differential line elements.

If we transform a quantity from the current to the initial configuration, we call this transformation a *pull back* operation, see Marsden and Hughes (1983). A transformation in the other direction is considered as a *push forward* operation.

For the gradient of a scalar field  $G(\mathbf{X}) = g(\mathbf{x}) = g[\varphi(\mathbf{X})]$ , we have

$$\text{Grad } G = \mathbf{F}^T \text{grad } g \iff \frac{\partial G}{\partial X_A} = \frac{\partial g}{\partial x_i} \frac{\partial x_i}{\partial X_A}, \quad (2.23)$$

$$\text{grad } g = \mathbf{F}^{-T} \text{Grad } G. \quad (2.24)$$

In an analogous way, for the gradient of a vector field  $\mathbf{W}(\mathbf{X}) = \mathbf{w}(\mathbf{x}) = \mathbf{w}[\varphi(\mathbf{X})]$  we obtain

$$\boxed{\text{Grad } \mathbf{W} = \text{grad } \mathbf{w} \mathbf{F} \iff \text{grad } \mathbf{w} = \text{Grad } \mathbf{W} \mathbf{F}^{-1}}. \quad (2.25)$$

An application of these general results is given by the computation of the deformation gradient in terms of the displacement field  $\mathbf{u}[\varphi(\mathbf{X})]$ . Using (2.14) and (2.25) yields

$$\begin{aligned} \mathbf{F} &= \mathbf{1} + \text{Grad } \mathbf{u}, \\ \mathbf{1} &= \mathbf{F}^{-1} + \text{Grad } \mathbf{u} \mathbf{F}^{-1}, \\ \implies \mathbf{F}^{-1} &= \mathbf{1} - \text{grad } \mathbf{u}. \end{aligned} \quad (2.26)$$

Thus the inverse of the deformation gradient can be computed from the displacements that refers to the current configuration.

**REMARK 2.2:** *In the case of small deformations, the GREEN-LAGRANGIAN strain tensor  $\mathbf{E}$  can be written in terms of the displacement field. Since the deformation gradient can be reformulated as  $\mathbf{F} = \text{Grad } \mathbf{x} = \text{Grad } \mathbf{X} + \text{Grad } \mathbf{u} = \mathbf{1} + \text{grad } \mathbf{u}$ , by neglecting the nonlinear terms, one obtains from (2.16)*

$$\boldsymbol{\epsilon}(\mathbf{u}) = \frac{1}{2} (\text{Grad } \mathbf{u} + \text{Grad}^T \mathbf{u}). \quad (2.27)$$

### 2.1.4 Time derivatives

The dependency of the deformation  $\varphi(\mathbf{X}, t)$  on the time must be considered in non-linear problems when either the constitutive relations are time- or history-dependent, as in the case of friction, or if the complete process is time-dependent, like an impact problem. Here we compute the time derivatives of kinematical quantities.

The velocity of a material point in the reference configuration is defined by the material time derivative

$$\mathbf{v}(\mathbf{X}, t) = \frac{\partial \varphi}{\partial t}(\mathbf{X}, t) = \dot{\varphi}(\mathbf{X}, t). \quad (2.28)$$

In the current configuration, we write for the velocity  $\hat{\mathbf{v}}$  of a particle, which is a point  $\mathbf{x}$  at time  $t$  in  $\varphi(B)$ ,

$$\hat{\mathbf{v}}(\mathbf{x}, t) = \hat{\mathbf{v}}(\varphi(\mathbf{X}, t), t) = \mathbf{v}(\mathbf{x}, t). \quad (2.29)$$

In an analogous way, we obtain the acceleration by differentiation of the velocity

$$\mathbf{a} = \ddot{\varphi}(\mathbf{X}, t) = \dot{\mathbf{v}}(\mathbf{X}, t). \quad (2.30)$$

Using this definition, the acceleration with respect to the current configuration yields, with (2.29) and the chain rule,

$$\hat{\mathbf{a}} = \dot{\hat{\mathbf{v}}} = \frac{\partial}{\partial t} [\hat{\mathbf{v}}(\varphi(\mathbf{X}, t), t)] = \frac{\partial \hat{\mathbf{v}}}{\partial t} + \text{grad } \hat{\mathbf{v}} \hat{\mathbf{v}}. \quad (2.31)$$

The first term is known as the *local derivative* the second term is the *convective part* of the time derivative. The local time derivative is computed by fixing the spatial position. Time derivative (2.31) must be applied to Eulerian descriptions of motions, which is mostly the case in fluid mechanics.

The time derivative of the deformation gradient  $\mathbf{F}$ , with (2.6), (2.28) and (2.25), yields

$$\dot{\mathbf{F}} = \text{Grad } \dot{\varphi}(\mathbf{X}, t) = \text{Grad } \mathbf{v} = \text{grad } \hat{\mathbf{v}} \mathbf{F}. \quad (2.32)$$

The spatial velocity gradient in (2.32) is often described by  $\mathbf{l}$ . With (2.32) we can define the spatial velocity gradient by  $\mathbf{F}$ :

$$\mathbf{l} = \dot{\mathbf{F}} \mathbf{F}^{-1}. \quad (2.33)$$

Equation (2.32) can now be applied to compute the time derivative of the GREEN-LAGRANGIAN strain tensor (2.16):

$$\dot{\mathbf{E}} = \frac{1}{2} (\dot{\mathbf{F}}^T \mathbf{F} + \mathbf{F}^T \dot{\mathbf{F}}). \quad (2.34)$$

The time derivative of  $\mathbf{E}$  can be rewritten with the last relation in (2.32):

$$\dot{\mathbf{E}} = \mathbf{F}^T \frac{1}{2} (\mathbf{l} + \mathbf{l}^T) \mathbf{F} = \mathbf{F}^T \mathbf{d} \mathbf{F}. \quad (2.35)$$

This equation has a structure similar to (2.22), and hence it denotes a *pull back* of the symmetrical spatial velocity gradient

$$\mathbf{d} = \frac{1}{2} (\mathbf{l} + \mathbf{l}^T) \quad (2.36)$$

to the initial configuration.

## 2.2 BALANCE LAWS .

The partial differential equations which represent the local balance laws of continuum mechanics are summarized in this section. For a detailed derivation, see (Malvern (1969), Chap. 5) or Marsden and Hughes (1983) for example.

### 2.2.1 Balance of mass

The balance of mass  $m$  of a body is given by the relation

$$\boxed{m = \int_B \rho_0 dV = \int_{\varphi(B)} \rho dV = \text{const.}} , \quad (2.37)$$

where  $\rho_0$  is the density in the initial configuration and  $\rho$  the density in the current configuration. Within the LAGRANGE description of a motion, we can conclude, assuming sufficient smoothness, that  $\rho_0 = J \rho$ . This equation yields a relation for the volume elements in the initial and current configurations

$$dv = \frac{\rho_0}{\rho} dV = J dV . \quad (2.38)$$

### 2.2.2 Local balance of momentum and moments of momentum

The local balance of momentum with respect to a volume element in the current configuration  $\varphi(B)$  can be written as

$$\boxed{\text{div } \boldsymbol{\sigma} + \rho \bar{\mathbf{b}} = \rho \dot{\mathbf{v}}; \quad \sigma_{ik,i} + \rho \bar{b}_k = \rho \dot{v}_k} . \quad (2.39)$$

In this equation  $\boldsymbol{\sigma}$  denotes the CAUCHY stress tensor. In (2.39)  $\rho \bar{\mathbf{b}}$  defines the volume or body force (e.g. due to gravitation).  $\rho \dot{\mathbf{v}}$  is the inertia force term, which can be neglected in the case of static analysis. Furthermore, we have the CAUCHY theorem, which relates the stress vector  $\mathbf{t}$  to the surface normal vector  $\mathbf{n}$  by

$$\mathbf{t} = \boldsymbol{\sigma}^T \mathbf{n}, \quad t_i = \sigma_{ik} n_k, \quad \left\{ \begin{matrix} t_1 \\ t_2 \\ t_3 \end{matrix} \right\} = \begin{bmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} \\ \sigma_{12} & \sigma_{22} & \sigma_{32} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \left\{ \begin{matrix} n_1 \\ n_2 \\ n_3 \end{matrix} \right\} . \quad (2.40)$$

This relation has been stated here in direct notation, index and matrix notation.

The local balance of angular of momentum in the absence of micropolar stresses, which is usually the case in non-magnetic materials (e.g. see Truesdell and Toupin (1960)), yields

$$\boxed{\boldsymbol{\sigma} = \boldsymbol{\sigma}^T, \quad \sigma_{ik} = \sigma_{ki}} , \quad (2.41)$$

which dictates the symmetry of the CAUCHY stress tensor.

### 2.2.3 First law of thermodynamics

A further balance law postulates the conservation of energy of a thermodynamical process. This is the first law of thermodynamics. This law states that the change of total energy  $E$  is induced by mechanical power  $P$  and heat transfer  $Q$  into the system:  $\dot{E} = P + Q$ . Within continuum mechanics, ignoring magnetism, etc., the mechanical power is defined by

$$P = \frac{d}{dt} \int_{\varphi(B)} \frac{1}{2} \rho \mathbf{v} \cdot \mathbf{v} dv + \int_{\varphi(B)} \boldsymbol{\sigma} \cdot \mathbf{d} dv, \quad (2.42)$$

and thus by the material time derivative of the kinetic energy and the stress power of the CAUCHY stress tensor and the symmetrical spatial velocity gradient  $\mathbf{d}$ ,  $\boldsymbol{\sigma} \cdot \mathbf{d}$ , which contributes to the internal energy. The heat input into the system

$$Q = - \int_{\varphi(\partial B)} \mathbf{q} \cdot \mathbf{n} da + \int_{\varphi(B)} \rho r dv \quad (2.43)$$

has two sources: the heat transfer through the surface of the body, described by the heat flux vector  $\mathbf{q}$ ; and the surface normal  $\mathbf{n}$ , and an internal heat source  $r$ . The total energy consists of the kinetic energy  $K = \int_{\varphi(B)} \frac{1}{2} \rho \mathbf{v} \cdot \mathbf{v} dv$  and the internal energy  $U = \int_{\varphi(B)} \rho u dv$  ( $u$  is the specific internal energy). Introducing all relations into the equation  $\dot{E} = P + Q$  yields, after some transformations, the local form of the first law of thermodynamics:

$$\boxed{\rho \dot{u} = \boldsymbol{\sigma} \cdot \mathbf{d} + \rho r - \operatorname{div} \mathbf{q}, \quad \rho \dot{u} = \sigma_{ik} d_{ik} + \rho r - q_{i,i}}. \quad (2.44)$$

In this equation the term  $\boldsymbol{\sigma} \cdot \mathbf{d}$  denotes the specific stress power.

In the framework of the constitutive theory, the free HELMHOLTZ energy is often introduced, which is defined by

$$\psi = u - \eta \theta, \quad (2.45)$$

where  $\eta$  is the entropy of the system and  $\theta$  the absolute temperature. With this transformation the first law of thermodynamics can be written as

$$\rho \dot{\psi} = \boldsymbol{\sigma} \cdot \mathbf{d} + \rho r - \operatorname{div} \mathbf{q} - \dot{\eta} \theta - \eta \dot{\theta}. \quad (2.46)$$

### 2.2.4 Transformation to the initial configuration, different stress tensors

Equations (2.39) and (2.41) refer to the current configuration. Often one needs a formulation of these equations in quantities which are related to the initial configuration  $B$ . For this transformation, also often called *pull back*, we define more stress tensors,

which follow from the equivalence of a force which is defined in  $\mathcal{B}$  and  $\varphi(\mathcal{B})$ :

$$\int_{\partial\varphi(\mathcal{B})} \boldsymbol{\sigma} \mathbf{n} \, da = \int_{\partial\mathcal{B}} \boldsymbol{\sigma} J \mathbf{F}^{-T} \mathbf{N} \, dA = \int_{\partial\mathcal{B}} \mathbf{P} \mathbf{N} \, dA. \quad (2.47)$$

This relation defines the first PIOLA-KIRCHHOFF stress tensor  $\mathbf{P}$ . We have the transformation

$$\mathbf{P} = J \boldsymbol{\sigma} \mathbf{F}^{-T} \quad P_{Ak} = J \sigma_{ik} (F_{iA})^{-1} \quad (2.48)$$

between the CAUCHY and the first PIOLA-KIRCHHOFF stresses, which are the actual stresses in terms of the area of the initial configuration. Since in equation (2.48) the spatial quantity  $\boldsymbol{\sigma}$  is only multiplied on one side by  $\mathbf{F} \mathbf{P}$ , it is a so-called *two field tensor* where one base vector lies in  $\mathcal{B}$  and the other in  $\varphi(\mathcal{B})$ . After some manipulation, we can now transform the local balance of momentum (2.39) to the reference configuration

$$\text{DIV } \mathbf{P} + \rho_0 \bar{\mathbf{b}} = \rho_0 \dot{\mathbf{v}}. \quad (2.49)$$

However, when using (2.48) in the balance of angular of momentum (2.41), we see that the PIOLA-KIRCHHOFF stress tensor is general in nonsymmetric:  $\mathbf{P} \mathbf{F}^T = \mathbf{F} \mathbf{P}^T$ .

A symmetric stress tensor which is defined with regard to the reference configuration is the second PIOLA-KIRCHHOFF stress tensor, which follows from the complete *pull back* of the CAUCHY stress tensor to the reference configuration  $\mathcal{B}$ :

$$\begin{aligned} \mathbf{S} &= \mathbf{F}^{-1} \mathbf{P} = J \mathbf{F}^{-1} \boldsymbol{\sigma} \mathbf{F}^{-T}, \\ S_{AB} &= (F_{Ai})^{-1} P_{Bi} = J (F_{Ai})^{-1} \sigma_{ik} (F_{kB})^{-1}. \end{aligned} \quad (2.50)$$

$\mathbf{S}$  does not represent an experimentally measurable stress. However, it is an essential stress measure that plays a prominent role in the constitutive theory. It is "work conjugated" (duality paired) with the GREEN-LAGRANGIAN strain tensor (2.16).

Besides the CAUCHY stress tensor  $\boldsymbol{\sigma}$ , the KIRCHHOFF stress tensor  $\boldsymbol{\tau}$  is often employed, which is defined as the *push forward* of the second PIOLA-KIRCHHOFF stress tensor  $\mathbf{S}$  to the current configuration

$$\boldsymbol{\tau} = \mathbf{F} \mathbf{S} \mathbf{F}^T, \quad \boldsymbol{\tau} = J \boldsymbol{\sigma}. \quad (2.51)$$

The transformation of the first law of thermodynamics (2.44) to the initial configuration yields, with (2.35),

$$J \boldsymbol{\sigma} \cdot \mathbf{d} = (\mathbf{F} \mathbf{S} \mathbf{F}^T) \cdot (\mathbf{F}^{-T} \dot{\mathbf{E}} \mathbf{F}^{-1}) = \mathbf{S} \cdot \dot{\mathbf{E}}. \quad (2.52)$$

Furthermore, for (2.38) we have

$$\rho_0 \dot{U} = \mathbf{S} \cdot \dot{\mathbf{E}} - \text{Div } \mathbf{Q} + \rho_0 R, \quad \rho_0 \dot{U} = S_{AB} \dot{E}_{AB} - Q_{A,A} + \rho_0 R, \quad (2.53)$$

where the internal energy,  $U$ , the internal heat source,  $R$ , and the heat flux,  $\mathbf{Q}$ , are referred to the initial configuration. With (2.51) or (2.16) we can also state the stress power (2.52) as

$$\mathbf{S} \cdot \dot{\mathbf{E}} = \frac{1}{2} \mathbf{S} \cdot \dot{\mathbf{C}} = \boldsymbol{\tau} \cdot \mathbf{d} \quad (2.54)$$

## 2.3 WEAK FORM OF BALANCE OF MOMENTUM, VARIATIONAL PRINCIPLES

For the solution of boundary value problems stemming from the continuum, we shall employ numerical methods based on variational formulations. Thus we need associated formulations, which are given in the next sections.

### 2.3.1 Weak form of balance of momentum in the initial configuration

The principle of virtual work is an equivalent formulation of the balance of momentum which often – due to its reduced regularity requirements – is called the weak form of equilibrium. Since no constitutive equations enter *a priori* the weak form, it is valid for all problem classes, including plasticity, friction or non-conservative loading. The derivation of the weak form starts from the local equilibrium equation ( $\text{Div } \mathbf{P} + \rho_0 \bar{\mathbf{b}} = \rho_0 \dot{\mathbf{v}}$ ), which is multiplied by a vector valued function  $\boldsymbol{\eta} = \{\boldsymbol{\eta} \mid \boldsymbol{\eta} = \mathbf{0} \text{ on } \partial B_u\}$  – often called a virtual displacement or test function. Integration over the volume of the body under consideration yields

$$\int_B \text{Div } \mathbf{P} \cdot \boldsymbol{\eta} dV + \int_B \rho_0 (\bar{\mathbf{b}} - \dot{\mathbf{v}}) \cdot \boldsymbol{\eta} dV = 0. \quad (2.55)$$

Partial integration of the first term and use of the divergence theorems leads, with the boundary conditions, to the weak form of

$$G(\boldsymbol{\varphi}, \boldsymbol{\eta}) = \int_B \mathbf{P} \cdot \text{Grad } \boldsymbol{\eta} dV - \int_B \rho_0 (\bar{\mathbf{b}} - \dot{\mathbf{v}}) \cdot \boldsymbol{\eta} dV - \int_{\partial B_\sigma} \bar{\mathbf{t}} \cdot \boldsymbol{\eta} dA = 0. \quad (2.56)$$

The gradient of  $\boldsymbol{\eta}$  can also be viewed as a virtual variation  $\delta \mathbf{F}$  of the deformation gradient

$$\delta \mathbf{F} = \left. \frac{d}{d\epsilon} [\mathbf{F}(\mathbf{x} + \epsilon \boldsymbol{\eta})] \right|_{\epsilon=0}. \quad (2.57)$$

In (2.56) we can exchange the first PIOLA–KIRCHHOFF stress tensor with  $\mathbf{P} = \mathbf{F}\mathbf{S}$  by the second PIOLA–KIRCHHOFF stress tensor:

$$\mathbf{P} \cdot \text{Grad } \boldsymbol{\eta} = \mathbf{S} \cdot \mathbf{F}^T \text{Grad } \boldsymbol{\eta} = \mathbf{S} \cdot \frac{1}{2} (\mathbf{F}^T \text{Grad } \boldsymbol{\eta} + \text{Grad}^T \boldsymbol{\eta} \mathbf{F}) = \mathbf{S} \cdot \delta \mathbf{E}, \quad (2.58)$$

where the variation of the GREEN–LAGRANGIAN strain tensor, computed according to (2.57), has been used. Note that  $\delta \mathbf{E} = \frac{1}{2} \delta \mathbf{C}$ . In (2.58) one makes use of the symmetry of  $\mathbf{S}$  so that the antisymmetric part of  $\mathbf{F}^T \text{Grad } \boldsymbol{\eta}$  disappears in the scalar product. With (2.58) we can rewrite (2.56) as

$$\boxed{G(\boldsymbol{\varphi}, \boldsymbol{\eta}) = \int_B \mathbf{S} \cdot \delta \mathbf{E} dV - \int_B \rho_0 (\bar{\mathbf{b}} - \dot{\mathbf{v}}) \cdot \boldsymbol{\eta} dV - \int_{\partial B_\sigma} \bar{\mathbf{t}} \cdot \boldsymbol{\eta} dA = 0.} \quad (2.59)$$



The first term in (2.59) denotes the virtual internal work or the stress divergence, and the last two terms contain the virtual work of the external forces. This equation can be written in index notation as follows:

$$G(\varphi, \eta) = \int_B S_{AB} \delta E_{AB} dV - \int_B \rho_0 (\bar{b}_A - \dot{v}_A) \eta_A dV - \int_{\partial B_\sigma} \bar{t}_A \eta_A dA = 0. \quad (2.60)$$

We note the equivalence of the strong form, (2.49), and the weak form, (2.59), provided the solution is smooth enough.

### 2.3.2 Spatial form of the weak formulation

The transformation of the weak form (2.56) to the current configuration follows by pure geometrical operations. For this purpose we need to transform the associated tensors by *push forward* operations to the current configuration  $\varphi(B)$ . With the transformation of the first PIOLA-KIRCHHOFF stress tensor to the CAUCHY stress tensor (see (2.48)),  $\sigma = \frac{1}{J} \mathbf{P} \mathbf{F}^T$ , and by using (2.25) we derive

$$\mathbf{P} \cdot \text{Grad } \eta = J \sigma \mathbf{F}^{-T} \cdot \text{Grad } \eta = J \sigma \cdot \text{Grad } \eta \mathbf{F}^{-1} = J \sigma \cdot \text{grad } \eta.$$

Furthermore, as  $dv = J dV$  and thus  $\rho = \rho_0 J$  is valid, we can transform the weak form (2.56) into the current configuration:

$$g(\varphi, \eta) = \int_{\varphi(B)} \sigma \cdot \text{grad } \eta dv - \int_{\varphi(B)} \rho (\bar{\mathbf{b}} - \dot{\mathbf{v}}) \cdot \eta dv - \int_{\varphi(\partial B_\sigma)} \bar{\mathbf{t}} \cdot \eta da = 0. \quad (2.61)$$

In this equation the result from (2.47) has been used to transform the stress vector  $\bar{\mathbf{t}}$  into  $\varphi(B)$ . Symmetry of the CAUCHY stress tensor enables us to replace the spatial gradient of  $\eta$  by its symmetric part. Hence, with the definition

$$\nabla^S \eta = \frac{1}{2} (\text{grad } \eta + \text{grad}^T \eta), \quad (2.62)$$

it follows that

$$g(\varphi, \eta) = \int_{\varphi(B)} \sigma \cdot \nabla^S \eta dv - \int_{\varphi(B)} \rho (\bar{\mathbf{b}} - \dot{\mathbf{v}}) \cdot \eta dv - \int_{\varphi(\partial B_\sigma)} \bar{\mathbf{t}} \cdot \eta da = 0. \quad (2.63)$$

This relation has exactly the same structure as the principle of virtual work in the geometrically linear theory. The difference, however, is that all integrals, stresses and gradients have to be computed with respect to the current coordinates, which reflects the nonlinearity of (2.63).

### 2.3.3 Minimum of total potential energy

In the case of a hyperelastic material there exists the strain energy function  $W$  (see Section 2.4.1) which describes the elastic energy stored in a body  $B$ . Based on

this function, the classical minimum principle of the total elastic potential can be formulated. For this also one has to consider the potential energy of the forces applied. We assume that these forces are conservative (meaning they are path-independent). Neglecting dynamical effects, we obtain

$$\Pi(\varphi) = \int_B [W(\mathbf{C}) - \rho_0 \bar{\mathbf{b}} \cdot \varphi] dV - \int_{\partial B_\sigma} \bar{\mathbf{t}} \cdot \varphi dA \Rightarrow MIN. \quad (2.64)$$

Out of all possible deformation states  $\varphi$ , the one which minimizes  $\Pi$  fulfils the equilibrium equations. The minimum can be computed by a variation of (2.64). It is related to the weak form (2.59). This can be shown by applying the directional derivative, which leads to the so-called first variation of  $\Pi$ :

$$\delta\Pi = D\Pi(\varphi) \cdot \eta = \left. \frac{d}{d\alpha} \Pi(\varphi + \alpha\eta) \right|_{\alpha=0}. \quad (2.65)$$

In explicit form we obtain

$$D\Pi(\varphi) \cdot \eta = \int_B \left[ \frac{\partial W}{\partial \mathbf{C}} \cdot \delta\mathbf{C} - \rho_0 \bar{\mathbf{b}} \cdot \eta \right] dV - \int_{\partial B_\sigma} \bar{\mathbf{t}} \cdot \eta dA = G(\varphi, \eta) = 0. \quad (2.66)$$

The variation of the right CAUCHY–GREEN tensor  $\delta\mathbf{C}$  can easily be expressed in terms of the GREEN–LAGRANGE strain tensor:  $2\delta\mathbf{C} = \delta\mathbf{E}$ , see also (2.58). The partial derivative of  $W$  with respect to  $\mathbf{C}$  leads, with  $2\partial W / \partial \mathbf{C} = \mathbf{S}$ , to the 2. PIOLA–KIRCHHOFF stress tensor, see Eq. (2.67) in Section 2.4.1. Hence (2.66) is equivalent to the weak form (2.59).

The construction of a minimal principle is important in several respects, since it enables investigations regarding the existence and uniqueness of solutions. Furthermore, special solution methods can be developed on the basis of a minimal principle which are efficient and reliable.

## 2.4 CONSTITUTIVE EQUATIONS

Since contact takes place at the interface between bodies, the constitutive laws for the bodies coming into contact which describe the material behaviour within the bodies can be arbitrary, and do not affect the main formulation of contact problems. However it is clear that the physical properties of the surfaces of the bodies are influenced by the general constitutive behaviour. Thus, to include a nonlinear constitutive equation valid for large deformations, we discuss finite elasticity. Of course, we can consider more complicated constitutive relations which can also be of inelastic nature, but this is not the aim of this book and we refer to Desai and Siriwardane (1984), Lubliner (1990), Khan and Huang (1995) or Simo and Hughes (1998) for example.

### 2.4.1 Hyperelastic response function

Throughout this section we briefly discuss hyperelastic constitutive relations. For more detailed information, see Ogden (1984) for example. These can be applied to describe the constitutive behaviour of rubber or foam for instance. In the case of small deformations, these constitutive equations reduce to the classical HOOKE's law of linear elasticity.

The constitutive equation or response function for the second PIOLA-KIRCHHOFF stress is, in the case of a hyperelastic material, given by the partial derivative of the strain energy  $W$  function with respect to the right CAUCHY-GREEN tensor, e.g. see Ogden (1984),

$$\mathbf{S} = 2 \frac{\partial W(\mathbf{C}, \mathbf{X})}{\partial \mathbf{C}}, \quad S_{AB} = 2 \frac{\partial W(\mathbf{C}, \mathbf{X})}{\partial C_{AB}}. \quad (2.67)$$

This response function represents a constitutive relation which fulfils the requirements of frame indifference, and hence is objective. In the case of a homogeneous material, the strain energy  $W$  does not depend upon  $\mathbf{X}$ . Here we restrict ourselves to homogeneous isotropic materials. Thereafter, the strain energy function can be specialized, and is represented by an isotropic tensor function

$$W(\mathbf{C}) = W(I_C, II_C, III_C). \quad (2.68)$$

The second PIOLA-KIRCHHOFF stresses now follow with (2.67) by using the chain rule

$$\mathbf{S} = 2 \left[ \left( \frac{\partial W}{\partial I_C} + I_C \frac{\partial W}{\partial II_C} \right) \mathbf{1} - \frac{\partial W}{\partial II_C} \mathbf{C} + III_C \frac{\partial W}{\partial III_C} \mathbf{C}^{-1} \right]. \quad (2.69)$$

Within this equation, the following results for the derivative of invariants with respect to tensors have been used:

$$\frac{\partial I_C}{\partial \mathbf{C}} = \mathbf{1}, \quad \frac{\partial II_C}{\partial \mathbf{C}} = I_C \mathbf{1} - \mathbf{C}, \quad \frac{\partial III_C}{\partial \mathbf{C}} = III_C \mathbf{C}^{-1}. \quad (2.70)$$

For the special choice of the strain energy function  $W$ , we obtain the simplest possible response function, which is known as compressible NEO-HOOKIAN material. We choose

$$W(I_C, J) = g(J) + \frac{1}{2} \mu (I_C - 3). \quad (2.71)$$

For compressible materials, function  $g(J)$  in (2.71) has to be convex. Furthermore, the following growth conditions must hold:

$$\lim_{J \rightarrow +\infty} W \rightarrow \infty \quad \text{and} \quad \lim_{J \rightarrow 0} W \rightarrow -\infty. \quad (2.72)$$

These conditions are equivalent with the conditions that the stress for a deformed body whose volume goes to zero has to go to  $-\infty$ , and for a deformed body whose

volume goes to  $+\infty$  the stress also has to go to  $+\infty$ . These growth conditions are fulfilled when the compressible part  $g(J)$  is chosen, according to, Ciarlet (1988), as

$$g(J) = c(J^2 - 1) - d \ln J - \mu \ln J \quad \text{with} \quad c > 0, d > 0. \quad (2.73)$$

The response function of the NEO-HOOKIAN material (2.71) now follows with (2.69), and for the second PIOLA-KIRCHHOFF stress tensor yields

$$\begin{aligned} \mathbf{S} &= \frac{\Lambda}{2} (J^2 - 1) \mathbf{C}^{-1} + \mu (\mathbf{1} - \mathbf{C}^{-1}), \\ S_{AB} &= \frac{\Lambda}{2} (J^2 - 1) (C_{AB})^{-1} + \mu [\delta_{AB} - (C_{AB})^{-1}], \end{aligned} \quad (2.74)$$

where the constants  $c$  and  $d$  have been chosen as  $c = \Lambda/4$  and  $d = \Lambda/2$ . The material constants  $\Lambda$  and  $\mu$  are the LAMÉ constants, which have to be determined by experiments.

Note that with Definition (2.51), the KIRCHHOFF stress can be written in terms of quantities define in the initial configuration:

$$\boldsymbol{\tau} = 2 \mathbf{F} \frac{\partial W(\mathbf{C})}{\partial \mathbf{C}} \mathbf{F}^T, \quad \tau_{ik} = 2 F_{iA} \frac{\partial W(\mathbf{C})}{\partial C_{AB}} F_{kB}.$$

From this form, the KIRCHHOFF stress is given with (2.69) by

$$\boldsymbol{\tau} = 2 \left[ \left( \frac{\partial W}{\partial I_C} + I_C \frac{\partial W}{\partial II_C} \right) \mathbf{F} \mathbf{F}^T - \frac{\partial W}{\partial II_C} \mathbf{F} \mathbf{C} \mathbf{F}^T + III_C \frac{\partial W}{\partial III_C} \mathbf{F} \mathbf{C}^{-1} \mathbf{F}^T \right].$$

Since the invariants of  $\mathbf{C}$  and  $\mathbf{b}$  are equal using  $\mathbf{F} \mathbf{C}^{-1} \mathbf{F}^T = \mathbf{1}$  one derives

$$\boldsymbol{\tau} = 2 \left[ \left( \frac{\partial W}{\partial I_b} + I_b \frac{\partial W}{\partial II_b} \right) \mathbf{b} - \frac{\partial W}{\partial II_b} \mathbf{b}^2 + III_b \frac{\partial W}{\partial III_b} \mathbf{1} \right].$$

Comparing this result to (2.69), the KIRCHHOFF stresses can also be derived from

$$\boxed{\boldsymbol{\tau} = 2 \mathbf{b} \frac{\partial \psi(\mathbf{b})}{\partial \mathbf{b}}} \quad (2.75)$$

directly in term of spatial quantities.

Equation (2.74) can also be transformed directly into the current configuration by the standard *push forward* operations. We note that the CAUCHY stress tensor is related, via  $\boldsymbol{\sigma} = J^{-1} \mathbf{F} \mathbf{S} \mathbf{F}^T$ , to the second PIOLA-KIRCHHOFF stresses, see (2.50). With this, after some manipulation we obtain

$$\boldsymbol{\sigma} = \frac{\Lambda}{2J} (J^2 - 1) \mathbf{1} + \frac{\mu}{J} (\mathbf{b} - \mathbf{1}), \quad (2.76)$$

$$\sigma_{ik} = \frac{\Lambda}{2J} (J^2 - 1) \delta_{ik} + \frac{\mu}{J} (b_{ik} - \delta_{ik}). \quad (2.77)$$

### 2.4.2 Incremental constitutive tensor

To derive the incremental constitutive tensor we have to compute the rate of the response function (2.67). Thus, the response function must be differentiated with respect to time. This leads to

$$\dot{\mathbf{S}} = 2 \frac{\partial^2 W}{\partial \mathbf{C} \partial \mathbf{C}} [\dot{\mathbf{C}}], \quad (2.78)$$

and hence to an incremental relation between the rate of the second PIOLA-KIRCHHOFF stress tensor  $\mathbf{S}$  and the right CAUCHY-GREEN tensor  $\mathbf{C}$ . With the definition of a fourth order incremental constitutive tensor

$$\mathbb{C} = 4 \frac{\partial^2 W}{\partial \mathbf{C} \partial \mathbf{C}}, \quad \mathbb{C}_{ABCD} = 4 \frac{\partial^2 W}{\partial C_{AB} \partial C_{CD}}, \quad (2.79)$$

for (2.78) we obtain

$$\dot{\mathbf{S}} = \mathbb{C} \left[ \frac{1}{2} \dot{\mathbf{C}} \right], \quad \dot{S}_{AB} = \mathbb{C}_{ABCD} \frac{1}{2} \dot{C}_{CD}. \quad (2.80)$$

The *push forward* of equation (2.80) to the current configuration yields, with the LIE-derivative of the KIRCHHOFF stress tensor,

$$(\mathcal{L}_v \boldsymbol{\tau})_{ik} = F_{iA} \dot{S}_{AB} F_{kB}, \quad (2.81)$$

and with the time derivative of the right CAUCHY-GREEN tensors, see (2.16) and (2.35),

$$\dot{C}_{CD} = 2 F_{lC} d_{lm} F_{mD} \quad (2.82)$$

the relation

$$(\mathcal{L}_v \boldsymbol{\tau})_{ik} = F_{iA} F_{lC} F_{mD} F_{kB} \mathbb{C}_{ABCD} d_{lm}. \quad (2.83)$$

$\mathbf{d}$  is the symmetrical spatial velocity gradient, see (2.36). Since in (2.83) each base vector of the incremental constitutive tensor  $\mathbb{C}$  is transformed by  $\mathbf{F}$ , we can define a spatial incremental constitutive tensor  $\mathfrak{c}$  as

$$\mathfrak{c}_{iklm} = F_{iA} F_{lC} F_{mD} F_{kB} \mathbb{C}_{ABCD}. \quad (2.84)$$

This leads to the compact form of: (2.83)

$$(\mathcal{L}_v \boldsymbol{\tau})_{ik} = \mathfrak{c}_{iklm} d_{lm}, \quad \mathcal{L}_v \boldsymbol{\tau} = \mathfrak{c} [\mathbf{d}]. \quad (2.85)$$

Thereafter, we derive the incremental constitutive tensor for the constitutive equations (2.74) and its *push forward* to the current configuration. The response function (2.74) depends upon the deformation via the inverse of the right CAUCHY-GREEN tensor and its determinant:  $J = \sqrt{\text{III}_{\mathbf{C}}}$ . Thus for the computation of  $\mathbb{C}$  using (2.79), the derivatives of  $J$  and  $\mathbf{C}^{-1}$  with respect to  $\mathbf{C}$  have to be computed.

With (2.63) the derivative of the JACOBIAN is

$$\frac{\partial J}{\partial \mathbf{C}} = \frac{1}{2} J \mathbf{C}^{-1}. \quad (2.86)$$

The derivative of  $\mathbf{C}^{-1}$  follows from relation,  $\frac{\partial}{\partial C_{CD}} [C_{AM} C_{MB}^{-1}] = 0$ , as

$$\frac{\partial C_{AB}^{-1}}{\partial C_{CD}} = -C_{AC}^{-1} C_{BD}^{-1}. \quad (2.87)$$

Since  $\mathbf{C}$  is symmetric we only need the symmetrical part of (2.87), and introduce the fourth order tensor  $\mathbb{I}_{C^{-1}}$  which has the index notation

$$\mathbb{I}_{C^{-1}ABCD} = \frac{1}{2} (C_{AC}^{-1} C_{BD}^{-1} + C_{AD}^{-1} C_{BC}^{-1}). \quad (2.88)$$

With these preliminaries the constitutive tensor can be derived. After some algebraic manipulations, we obtain

$$\begin{aligned} \mathbf{C} &= \Lambda J^2 \mathbf{C}^{-1} \otimes \mathbf{C}^{-1} + [2\mu - \Lambda(J^2 - 1)] \mathbb{I}_{C^{-1}}, \\ \mathbb{C}_{ABCD} &= \Lambda J^2 C_{AB}^{-1} C_{CD}^{-1} + [2\mu - \Lambda(J^2 - 1)] \mathbb{I}_{C^{-1}ABCD}. \end{aligned} \quad (2.89)$$

Transformation of the incremental constitutive tensor  $\mathbb{C}$  to the current configuration yields, with (2.84) and

$$C_{AC}^{-1} C_{BD}^{-1} = F_{pA}^{-1} F_{pC}^{-1} F_{qB}^{-1} F_{qD}^{-1} F_{iA} F_{lC} F_{mD} F_{kB} = \delta_{pi} \delta_{pl} \delta_{qk} \delta_{qm} = \delta_{il} \delta_{km},$$

the incremental constitutive tensor in  $\varphi(B)$ :

$$\begin{aligned} \mathfrak{C} &= \Lambda J^2 \mathbf{1} \otimes \mathbf{1} + [2\mu - \Lambda(J^2 - 1)] \mathbb{I}, \\ \mathfrak{C}_{iklm} &= \Lambda J^2 \delta_{ik} \delta_{lm} + [2\mu - \Lambda(J^2 - 1)] \mathbb{I}_{iklm}, \end{aligned} \quad (2.90)$$

where  $\mathbf{1}$  is the second order unit tensor and  $\mathbb{I}$  is a fourth order unit tensor. Both tensors are related to the current configuration. Tensor  $\mathbb{I}$  has, in index notation, the form

$$\mathbb{I}_{iklm} = \frac{1}{2} (\delta_{il} \delta_{km} + \delta_{im} \delta_{kl}). \quad (2.91)$$

For a formulation of the equations of elasticity for numerical treatment within the method of finite elements, it is preferable to have the matrix representation of equation (2.85). For this purpose the components of the LIE derivative of the KIRCHHOFF stresses and the symmetrical spatial velocity gradient  $\mathbf{d}$  are represented in vector form. In that case, the incremental constitutive tensor (2.90) is a matrix which can be used to compute the incremental KIRCHHOFF stresses once  $\mathbf{d}$  is known,

$$\mathcal{L}_v \boldsymbol{\tau} = \mathbf{D} \mathbf{d}, \quad (2.92)$$

or explicitly

$$\begin{Bmatrix} \mathcal{L}_v \tau_{11} \\ \mathcal{L}_v \tau_{22} \\ \mathcal{L}_v \tau_{33} \\ \mathcal{L}_v \tau_{12} \\ \mathcal{L}_v \tau_{23} \\ \mathcal{L}_v \tau_{31} \end{Bmatrix} = \begin{bmatrix} 2\mu + \Lambda & \Lambda J^2 & \Lambda J^2 & 0 & 0 & 0 \\ \Lambda J^2 & 2\mu + \Lambda & \Lambda J^2 & 0 & 0 & 0 \\ \Lambda J^2 & \Lambda J^2 & 2\mu + \Lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha \end{bmatrix} \begin{Bmatrix} d_{11} \\ d_{22} \\ d_{33} \\ 2d_{12} \\ 2d_{23} \\ 2d_{31} \end{Bmatrix}$$

with  $\alpha = \mu - \frac{1}{2} \Lambda (J^2 - 1)$ . (2.93)

The deformation gradient  $\mathbf{F}$  is equal to  $\mathbf{1}$  in the undeformed initial configuration. Hence, also  $\mathbf{C}^{-1} = \mathbf{1}$  and  $J = 1$ . Thus, when the incremental constitutive tensor in (2.89) is evaluated at the undeformed state in the initial configuration, we obtain

$$\mathbf{C}_0 = \Lambda \mathbf{1} \otimes \mathbf{1} + 2\mu \mathbf{I}. \quad (2.94)$$

This equation also follows directly from (2.90) since for  $\mathbf{F} = \mathbf{1}$  the initial and current configuration coincide. The constitutive tensor  $\mathbf{C}_0$  is identical to the elasticity tensor of the geometrical linear theory of elasticity. Its matrix form is

$$\sigma = \mathbf{D}_0 \epsilon, \quad (2.95)$$

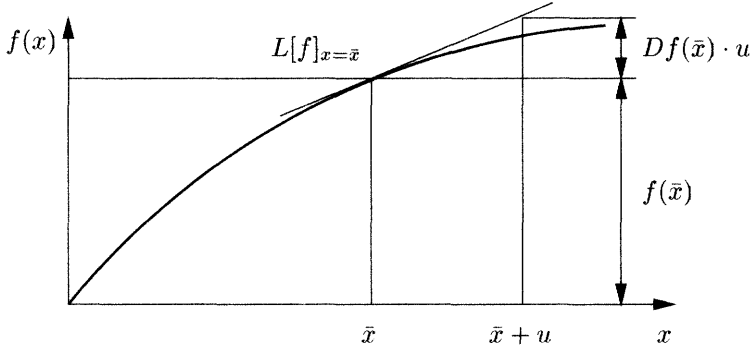
or explicitly,

$$\begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{31} \end{Bmatrix} = \begin{bmatrix} 2\mu + \Lambda & \Lambda & \Lambda & 0 & 0 & 0 \\ \Lambda & 2\mu + \Lambda & \Lambda & 0 & 0 & 0 \\ \Lambda & \Lambda & 2\mu + \Lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \begin{Bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ 2\epsilon_{12} \\ 2\epsilon_{23} \\ 2\epsilon_{31} \end{Bmatrix}. \quad (2.96)$$

## 2.5 LINEARIZATIONS

Different phenomena lead to nonlinearities in continuum mechanics. There are geometrical nonlinearities, nonlinearities stemming from the constitutive equations, or nonlinearities due to unilateral boundary conditions, as in contact. Linearization of the mathematical models is necessary once the associated initial or boundary value problems have to be solved. Especially for numerical methods like the finite element method, it has been proven that NEWTON's method is a very efficient solution algorithm for nonlinear continuum problems.

Thus it is necessary to have a mathematical tool which allows us to compute linearizations of nonlinear continuum problems. The purpose of this section is to provide these mathematical tools, and to apply them to kinematical relations, to constitutive equations and to the weak forms. Mathematical details are omitted, but they can be found for example in Marsden and Hughes (1983).



**Fig. 2.3** Linearization of  $f$  at  $\bar{x}$ .

The idea of the linearization process will first be described by means of an example. Let us introduce a scalar valued function  $f$  which is continuous and has continuous first derivatives ( $C^1$ -continuity). Under this assumption it is possible to express  $f$  by a TAYLOR series expansion at  $x$ :

$$f(\bar{x} + u) = \bar{f} + \bar{D}f \cdot u + R. \quad (2.97)$$

In this equation the following notation has been used:  $\bar{f} = f(\bar{x})$  and  $\bar{D}f = Df(\bar{x})$ . The operator  $D$  denotes the derivative of  $f$  with respect to  $x$ . The “ $\cdot$ ” is in this case a simple multiplication.  $u$  is an increment and the residual term  $R = R(u)$  has the property  $\lim_{u \rightarrow 0} \frac{R}{|u|} \rightarrow 0$ . Figure 2.3 depicts the geometrical interpretation of equation (2.97). With  $u$  being the independent variable and  $\bar{x}$  a fixed coordinate in (2.97), the tangent to the curve described by  $f$  at  $\bar{x}$  is

$$f(u) = \bar{f} + \bar{D}f \cdot u \quad (2.98)$$

which touches the curve in  $(\bar{x}, \bar{f})$ . The linear part of  $f(x)$  in  $x = \bar{x}$  defines the linearization

$$L[f]_{x=\bar{x}} \equiv f(u). \quad (2.99)$$

This result for the one-dimensional case can be extended to scalar valued functions in three dimensions. The  $f$  is a function of  $(\mathbf{x})$ . The TAYLOR series expansion is then

$$f(\bar{\mathbf{x}} + \mathbf{u}) = \bar{f} + \bar{D}f \cdot \mathbf{u} + R. \quad (2.100)$$

Here  $\bar{\mathbf{x}}$  is a point in three-dimensional space, and  $\mathbf{u}$  is a vector with its origin in  $\bar{\mathbf{x}}$ . We obtain, with the definitions,

$$\bar{f} = f(\bar{\mathbf{x}}) \quad \text{and} \quad \bar{D}f = Df(\bar{\mathbf{x}}) = \left. \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\bar{\mathbf{x}}}, \quad (2.101)$$

where  $\bar{D}f$  denotes the gradient vector of  $f$  at  $\bar{\mathbf{x}}$ . Equation (2.100) can be written as

$$f(\bar{\mathbf{x}} + \mathbf{u}) = \bar{f} + \text{Grad } f(\bar{\mathbf{x}}) \cdot \mathbf{u} + R. \quad (2.102)$$



The product “.” in (2.102) is now a scalar product between two vector. Now the directional derivative of  $f$  is computed at  $\bar{\mathbf{x}}$  in the direction of  $\mathbf{u}$ . The directional derivative is defined by

$$\left. \frac{d}{d\epsilon} [f(\bar{\mathbf{x}} + \epsilon \mathbf{u})] \right|_{\epsilon=0},$$

where  $\epsilon$  is a scalar parameter. Since  $\bar{\mathbf{x}} + \epsilon \mathbf{u}$  is a line in the three-dimensional space, the directional derivative measures the increment of  $f$  in the direction of this line in  $\bar{\mathbf{x}}$ . The computation of the directional derivative follows with the chain rule as

$$\left. \frac{d}{d\epsilon} [f(\bar{\mathbf{x}} + \epsilon \mathbf{u})] \right|_{\epsilon=0} = \left[ \frac{\partial f(\bar{\mathbf{x}} + \epsilon \mathbf{u})}{\partial \mathbf{x}} \cdot \frac{\partial (\bar{\mathbf{x}} + \epsilon \mathbf{u})}{\partial \epsilon} \right]_{\epsilon=0} = \frac{\partial f(\bar{\mathbf{x}})}{\partial \mathbf{x}} \cdot \mathbf{u}.$$

A comparison yields that the directional derivative

$$\left. \frac{d}{d\epsilon} [f(\bar{\mathbf{x}} + \epsilon \mathbf{u})] \right|_{\epsilon=0} = \bar{D}f \cdot \mathbf{u}$$

is in coincidence with the tangent to  $f$  in  $\bar{\mathbf{x}}$ . Thus the linear part of  $f$  at  $\bar{\mathbf{x}}$  is given by the value of  $f$  and the directional derivative at  $\bar{\mathbf{x}}$ . The directional derivative is a linear operator, hence rules for standard derivatives like the product rule apply.

The directional derivative for infinite dimensional function spaces is given by a formal application of the foregoing results. Hence, one can consider the following  $C^1$ -mapping,  $\mathbf{G} : \mathcal{E} \rightarrow \mathcal{F}$ , where  $\bar{\mathbf{x}}, \mathbf{u}$  are points in the associated space:

$$\mathbf{G}(\bar{\mathbf{x}} + \mathbf{u}) = \bar{\mathbf{G}} + \bar{D}\mathbf{G} \cdot \mathbf{u} + \mathbf{R}. \quad (2.103)$$

Here now the “.” is the inner product of the elements characterizing the associated space. Again, the directional derivative is

$$\left. \frac{d}{d\epsilon} [\mathbf{G}(\bar{\mathbf{x}} + \epsilon \mathbf{u})] \right|_{\epsilon=0} = \bar{D}\mathbf{G} \cdot \mathbf{u}. \quad (2.104)$$

Thus the linear part of the mapping at  $\bar{\mathbf{x}}$  is

$$\mathbf{L}[\mathbf{G}]_{x=\bar{\mathbf{x}}} = \bar{\mathbf{G}} + \bar{D}\mathbf{G} \cdot \mathbf{u}. \quad (2.105)$$

Here elements describing the space under consideration can be scalars, vectors or tensors.

To simplify notation, the directional derivative  $\bar{D}\mathbf{G} \cdot \mathbf{u}$  will also be written as  $\Delta \bar{\mathbf{G}}$ . Here the bar denotes the evaluation of  $\mathbf{G}$  at point  $\bar{\mathbf{x}}$ .

Tensors which refer to the current configuration are linearized by first performing a *pull back* transformation, see (B.7), to the reference configuration. There the linearization is computed according to the rules stated above, and then the result is transformed back to the reference configuration (*push forward* operation). Note that the *pull back* and *push forward* operations depend upon the description of the tensors, e.g. a covariant tensor has a different *pull back* than a contravariant tensor. Thus

for tensors  $\boldsymbol{\tau}$  which refer to a covariant base (e.g. stress tensors), the directional derivative has the form

$$D \boldsymbol{\tau} \cdot \mathbf{u} = \mathbf{F} \{ D [ \mathbf{F}^{-1} \boldsymbol{\tau} \mathbf{F}^{-T} ] \cdot \mathbf{u} \} \mathbf{F}^T. \quad (2.106)$$

In an analogous way, a tensor which refers to a contravariant base like a strain tensor has the directional derivative

$$D \mathbf{e} \cdot \mathbf{u} = \mathbf{F}^{-T} \{ D [ \mathbf{F}^T \mathbf{e} \mathbf{F} ] \cdot \mathbf{u} \} \mathbf{F}^{-1}. \quad (2.107)$$

### 2.5.1 Linearization of kinematical quantities

The linearization of strain measures is described in this section in detail for strain measures referring to the initial and the current configuration.

The first strain measure which will be discussed is the GREEN-LAGRANGIAN strain tensor (2.16). The linear part is given with (2.105) as

$$\mathbf{L} [\mathbf{E}]_{\varphi=\bar{\varphi}} = \bar{\mathbf{E}} + \bar{D} \mathbf{E} \cdot \mathbf{u} = \bar{\mathbf{E}} + \Delta \bar{\mathbf{E}}, \quad (2.108)$$

where the directional derivative  $\bar{D} \mathbf{E} \cdot \mathbf{u} = \Delta \bar{\mathbf{E}}$  has to be computed according to (2.104)

$$\begin{aligned} \bar{D} \mathbf{E} \cdot \mathbf{u} &= \left. \frac{d}{d\epsilon} \left[ \frac{1}{2} \mathbf{F}^T (\bar{\boldsymbol{\varphi}} + \epsilon \mathbf{u}) \mathbf{F} (\bar{\boldsymbol{\varphi}} + \epsilon \mathbf{u}) - \mathbf{1} \right] \right|_{\epsilon=0} \\ \Delta \bar{\mathbf{E}} &= \frac{1}{2} [\bar{\mathbf{F}}^T \text{Grad} \mathbf{u} + \text{Grad}^T \mathbf{u} \bar{\mathbf{F}}]. \end{aligned} \quad (2.109)$$

This result is linear in  $\mathbf{u}$  and depends upon the deformation at  $\bar{\boldsymbol{\varphi}}$ , which is represented by  $\bar{\mathbf{F}}$ . The computation of (2.109) at  $\boldsymbol{\varphi} = \mathbf{X}$  yields the linear strain tensor

$$\mathbf{L} [\mathbf{E}]_{\varphi=\mathbf{X}} = \mathbf{0} + \frac{1}{2} [\text{Grad} \mathbf{u} + \text{Grad}^T \mathbf{u}]. \quad (2.110)$$

As a next strain measure the ALMANSI strain tensor,  $\mathbf{e} = \frac{1}{2} (\mathbf{1} - \mathbf{b}^{-1})$ , is considered, which refers to the current configuration. First one has to compute the *pull back* of  $\mathbf{e}$  using (2.22), and then apply the directional derivative. This result is then *pushed forward* to the current configuration

$$\begin{aligned} D \mathbf{e} \cdot \mathbf{u} &= \bar{\mathbf{F}}^{-T} \{ D \mathbf{E} \cdot \mathbf{u} \} \bar{\mathbf{F}}^{-1} = \frac{1}{2} (\text{Grad} \mathbf{u} \bar{\mathbf{F}}^{-1} + \bar{\mathbf{F}}^{-T} \text{Grad}^T \mathbf{u}) \\ &= \frac{1}{2} (\overline{\text{grad}} \mathbf{u} + \overline{\text{grad}}^T \mathbf{u}) = \bar{\nabla}^S \Delta \mathbf{u}. \end{aligned} \quad (2.111)$$

Comparison of this result with (2.109) shows that

$$\Delta \bar{\mathbf{E}} = \bar{\mathbf{F}}^T \bar{\nabla}^S \Delta \mathbf{u} \bar{\mathbf{F}}, \quad (2.112)$$

and thus the linearization of the ALMANSI strain tensor leads to the same structure as shown in equation (2.35) for the time derivative of the GREEN-LAGRANGIAN strain tensor.

### 2.5.2 Linearization of constitutive equations

Linearization of the constitutive equations can be computed for hyperelastic response functions in an analogous way as that for the time derivatives. For inelastic constitutive equations, the linearization depends upon the algorithm which is used to integrate the evolution equations, and thus the linearization can only be computed once the integration algorithm is known. Here we only consider hyperelastic constitutive equations.

The hyperelastic constitutive equation (2.67) describes the response function for the second PIOLA–KIRCHHOFF stress tensor depending on the right CAUCHY–GREEN tensor. According to (2.105) its linearization yields

$$\begin{aligned} \mathbf{L}[\mathbf{S}]_{\varphi=\bar{\varphi}} &= \bar{\mathbf{S}} + \bar{\mathbf{D}} \mathbf{S} \cdot \mathbf{u} = \bar{\mathbf{S}} + \Delta \bar{\mathbf{S}} \\ &= \bar{\mathbf{S}} + \left. \frac{\partial \mathbf{S}}{\partial \mathbf{C}} \right|_{\varphi=\bar{\varphi}} [\bar{\mathbf{D}} \mathbf{C} \cdot \mathbf{u}]. \end{aligned} \quad (2.113)$$

This result can be reformulated with (2.79) and (2.109) as

$$\mathbf{L}[\mathbf{S}]_{\varphi=\bar{\varphi}} = \bar{\mathbf{S}} + \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}]. \quad (2.114)$$

A comparison with (2.113) yields

$$\Delta \bar{\mathbf{S}} = \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}]. \quad (2.115)$$

Relation (2.115) has the same structure as the incremental constitutive equation (2.80). Only the time derivatives have to be replaced by the directional derivatives.

### 2.5.3 Linearization of the weak form

Solution of nonlinear boundary value problems can in general only be obtained by approximate methods. Many of these methods, like the finite element method, are based on the variational formulation of the field equations, given for instance by the weak form or principle of virtual work, hence equations (2.56) or (2.59) provide the starting point for a numerical method. For the solution of these nonlinear equations an iterative scheme has to be developed, since the discretization of the weak form results in a nonlinear system of algebraic equations.

Among many possible iterative algorithms, NEWTON's method has been proven to often be the most efficient scheme, since it exhibits quadratic convergence near the solution point. Within NEWTON's method a correction of the solution is achieved by the TAYLOR series expansion of the nonlinear equation set at a point where the approximated solution is already known. The necessary linearization can be computed with the aid of the directional derivative.

The linearization of the weak form is first derived with respect to the initial configuration, which is based on equation (2.56). We assume that the linearization is computed at a deformation state  $\bar{\varphi}$  at which the body under investigation is in equilibrium.

The linear part of the weak form is

$$L[G]_{\varphi=\bar{\varphi}} = G(\bar{\varphi}, \boldsymbol{\eta}) + DG(\bar{\varphi}, \boldsymbol{\eta}) \cdot \Delta \mathbf{u}. \quad (2.116)$$

$G(\bar{\varphi}, \boldsymbol{\eta})$  is equal to (2.58), only  $\varphi$  is exchanged by the state  $\bar{\varphi}$ . The directional derivative of  $G$ , needed to compute the linearization, has only to be applied to the first term in (2.58) when the assumption of conservative loading is made

$$DG(\bar{\varphi}, \boldsymbol{\eta}) \cdot \Delta \mathbf{u} = \int_B [D\mathbf{P}(\bar{\varphi}) \cdot \Delta \mathbf{u}] \cdot \text{Grad } \boldsymbol{\eta} dV; \quad (2.117)$$

all other terms do not depend upon the deformation. The linearization of the first PIOLA-KIRCHHOFF stress tensor yields, with  $\mathbf{P} = \mathbf{F}\mathbf{S}$ ,

$$DG(\bar{\varphi}, \boldsymbol{\eta}) \cdot \Delta \mathbf{u} = \int_B \{ \text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} + \bar{\mathbf{F}} [D\mathbf{S}(\bar{\varphi}) \cdot \Delta \mathbf{u}] \} \cdot \text{Grad } \boldsymbol{\eta} dV. \quad (2.118)$$

Quantities labelled with a bar have to be evaluated at  $\bar{\varphi}$ . For linearization of the second PIOLA-KIRCHHOFF stresses, equation (2.115) can be used. This leads to

$$D\mathbf{S}(\bar{\varphi}) \cdot \Delta \mathbf{u} = \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}], \quad (2.119)$$

where the last term is the linearization of the GREEN-LAGRANGIAN strain tensor  $\mathbf{E}$  at  $\bar{\varphi}$ , see (2.109). The incremental elasticity tensor  $\mathcal{C}_R$  which is evaluated with respect to the reference configuration is given, with (2.79), by

$$\bar{\mathbf{C}} = 4 \left. \frac{\partial^2 W}{\partial \mathbf{C} \partial \mathbf{C}} \right|_{\varphi=\bar{\varphi}} \quad (2.120)$$

at  $\bar{\varphi}$ .

Inserting equation (2.120) into (2.118) completes the linearization:

$$DG(\bar{\varphi}, \boldsymbol{\eta}) \cdot \Delta \mathbf{u} = \int_B \{ \text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} + \bar{\mathbf{F}} \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}] \} \cdot \text{Grad } \boldsymbol{\eta} dV. \quad (2.121)$$

Note that also  $\bar{\mathbf{C}}$  has to be computed at  $\bar{\varphi}$ . By making use of the trace operation and by considering symmetry of  $\bar{\mathbf{C}}$ , a compact form of (2.121) can be obtained:

$$\boxed{DG(\bar{\varphi}, \boldsymbol{\eta}) \cdot \Delta \mathbf{u} = \int_B \{ \text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} \cdot \text{Grad } \boldsymbol{\eta} + \delta \bar{\mathbf{E}} \cdot \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}] \} dV}. \quad (2.122)$$

Note the symmetry of the linearization with respect to  $\boldsymbol{\eta}$  and  $\Delta \mathbf{u}$ . The first term in (2.122) is the so-called *geometrical matrix* or *initial stress matrix*. The second term contains the initial deformations which occur in the incremental constitutive tensor  $\bar{\mathbf{C}}$ , the variation of the GREEN-LAGRANGIAN strains  $\delta \bar{\mathbf{E}} = \frac{1}{2}(\bar{\mathbf{F}}^T \text{Grad } \boldsymbol{\eta} + \text{Grad}^T \boldsymbol{\eta} \bar{\mathbf{F}})$

and its linearization  $\Delta \bar{\mathbf{E}} = \frac{1}{2}(\bar{\mathbf{F}}^T \text{Grad } \Delta \mathbf{u} + \text{Grad}^T \Delta \mathbf{u} \bar{\mathbf{F}})$ . Equation (2.122) is given in index notation as

$$DG(\bar{\varphi}, \eta)_A \Delta u_A = \int_B \{ \Delta u_{A,B} \bar{S}_{BC} \eta_{A,C} + \delta \bar{E}_{AB} \bar{\mathbf{C}}_{ABCD} \Delta \bar{E}_{CD} \} dV. \quad (2.123)$$

With the last equations, all relations with respect to the initial and current configurations, are known, which have to be applied within an iterative solution procedure, e.g. NEWTON's method. Thus the basis for discretization using the finite element method for nonlinear problems in solid mechanics is known.

The linearization of the weak form, defined in quantities of the current configuration, follows by *push forward* of linearization (2.122) to the already obtained deformations state  $\bar{\varphi}$ . With the *push forward*  $\bar{\nabla}^S \Delta \mathbf{u}$  of the GREEN-LAGRANGIAN strain tensor, the second term in (2.122) can be re-written as

$$\int_B \bar{\nabla}^S \eta \cdot \bar{\mathbf{c}} [\bar{\nabla}^S \Delta \mathbf{u}] dV.$$

The fourth order tensor  $\bar{\mathbf{c}}$  follows from  $\bar{\mathbf{C}}$  by the transformation (2.79).

The first term in (2.122) can be transformed directly with  $\bar{\tau} = \bar{\mathbf{F}} \bar{\mathbf{S}} \bar{\mathbf{F}}^T$ , and thus expressed in KIRCHHOFF stresses:

$$\text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} \cdot \text{Grad } \eta = \bar{\mathbf{F}} \text{Grad } \Delta \mathbf{u} \bar{\mathbf{F}}^{-1} \bar{\tau} \bar{\mathbf{F}}^{-1} \cdot \text{Grad } \eta = \overline{\text{grad}} \Delta \mathbf{u} \bar{\tau} \cdot \overline{\text{grad}} \eta. \quad (2.124)$$

With these transformations, the linearization in terms of quantities at the current configuration state  $\bar{\varphi}$  is

$$Dg(\bar{\varphi}, \eta) \cdot \Delta \mathbf{u} = \int_B \{ \overline{\text{grad}} \Delta \mathbf{u} \bar{\tau} \cdot \overline{\text{grad}} \eta + \bar{\nabla}^S \eta \cdot \bar{\mathbf{c}} [\bar{\nabla}^S \Delta \mathbf{u}] \} dV. \quad (2.125)$$

With  $d\bar{v} = \bar{J} dV$  integral (2.125) can be transformed into the current configuration  $\bar{\varphi}$ . For this purpose, we use the CAUCHY stress tensor  $\bar{\sigma} = \frac{1}{\bar{J}} \bar{\tau}$ , and define the incremental constitutive tensor

$$\hat{\bar{\mathbf{c}}} = \frac{1}{\bar{J}} \bar{\mathbf{c}} \quad (2.126)$$

such that the final result

$$Dg(\bar{\varphi}, \eta) \cdot \Delta \mathbf{u} = \int_{\bar{\varphi}(B)} \{ \overline{\text{grad}} \Delta \mathbf{u} \bar{\sigma} \cdot \overline{\text{grad}} \eta + \nabla^S \eta \cdot \hat{\bar{\mathbf{c}}} [\bar{\nabla}^S \Delta \mathbf{u}] \} dv \quad (2.127)$$

follows. Equation (2.127) in the literature is also known as the *updated Lagrange* formulation e.g. see Bathe et al. (1975), since the deformation state  $\bar{\varphi}$  is always updated during the nonlinear incremental solution procedure.

With the last equations all relations with respect to the current configuration are known. These have to be applied within an iterative solution procedure, e.g. NEWTON's method. Thus the basis for discretization using the finite element method for nonlinear problems in solid mechanics is known.

### 2.5.4 Linearization of a deformation dependent load

The description of a pressure load stemming from gases or fluids without internal friction leads to a surface load which depends upon the current deformation state. The stress vector  $\mathbf{t}$  is then given in terms of the pressure  $p$  and the surface normal  $\mathbf{n}$  by  $\mathbf{t} = p \mathbf{n}$ . This leads in the weak form (2.63) to the additional term

$$g(\varphi, \eta) + g_p(\varphi, \eta) = g(\varphi, \eta) + \int_{\varphi(\partial B_p)} p \mathbf{n} \cdot \boldsymbol{\eta} da. \quad (2.128)$$

For this term the linearization is derived next.

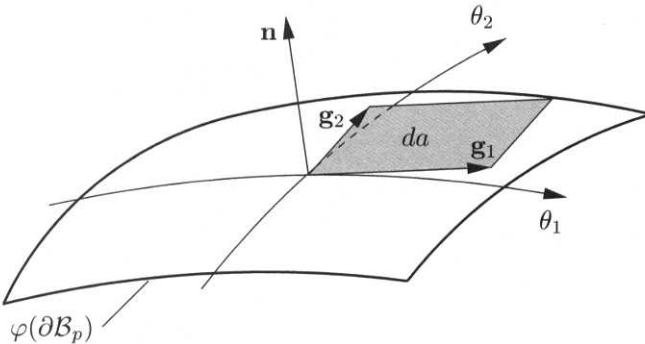
Again, to compute the linearization, it is preferable to perform a *pull back* operation and to refer (2.128) to the initial configuration. Two methods are possible. The first relies on the transformation of the surface normal by (2.11). This yields the expression  $\int_B p J \mathbf{F}^{-T} \mathbf{N} \cdot \boldsymbol{\eta} dA$  where linearization is complicated. Simpler is the second method in which the normal vector  $\mathbf{n}$  is expressed by the cross product of the tangent vectors which are tangent to the convective coordinates of the surface of the body, see Figure 2.4.

The normal unit vector follows with the tangent vectors, as defined in Figure 2.4,  $\mathbf{g}_\alpha$  ( $\alpha = 1, 2$ ) as

$$\mathbf{n} = \frac{\mathbf{g}_1 \times \mathbf{g}_2}{\|\mathbf{g}_1 \times \mathbf{g}_2\|}.$$

The tangent vectors can be computed from the deformation state using equation (B.4):  $\mathbf{g}_\alpha = \varphi_{,\alpha}$ . The area element  $da$  can be expressed by  $da = \|\mathbf{g}_1 \times \mathbf{g}_2\| d\theta_1 d\theta_2$  in terms of the tangent vectors with respect to the convective coordinates. Based on these relations, the virtual work for pressure loading is

$$g_p(\varphi, \eta) = \int_{(\theta_1)} \int_{(\theta_2)} p(\varphi_{,1} \times \varphi_{,2}) \cdot \boldsymbol{\eta} d\theta_1 d\theta_2. \quad (2.129)$$



**Fig. 2.4** Pressure dependent surface loads.

With the introduction of the displacement field, the tangent vectors have the form  $\varphi_{,\alpha} = (\mathbf{X} + \mathbf{u})_{,\alpha}$ . Hence the linearization of (2.129) yields

$$D g_p(\varphi, \eta) \cdot \Delta \mathbf{u} = \int_{(\theta_1)} \int_{(\theta_2)} p (\Delta \mathbf{u}_{,1} \times \varphi_{,2} + \varphi_{,1} \times \Delta \mathbf{u}_{,2}) \cdot \boldsymbol{\eta} d\theta_1 d\theta_2, \quad (2.130)$$

when  $p$  itself is independent of the deformation state. The linearization refers to the convected coordinates. It can be *pushed forward* to the current configuration, leading to

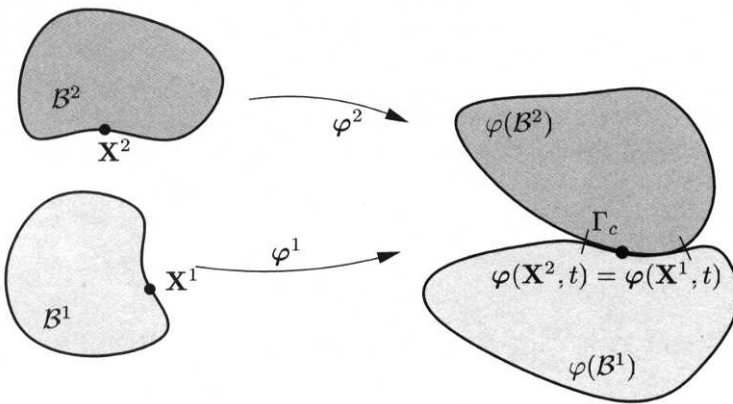
$$D g_p(\varphi, \eta) \cdot \Delta \mathbf{u} = \int_{\varphi(\partial B_p)} p \frac{\Delta \mathbf{u}_{,1} \times \varphi_{,2} + \varphi_{,1} \times \Delta \mathbf{u}_{,2}}{\|\varphi_{,1} \times \varphi_{,2}\|} \cdot \boldsymbol{\eta} da. \quad (2.131)$$

With this the linearization of a deformation-dependent pressure load, see (2.128), has been derived. More theoretical considerations with regard to the nonconservative nature of deformation dependent loads can be found in Sewell (1967), Bufler (1984), Ogden (1984) or Simo et al. (1991).

# 3

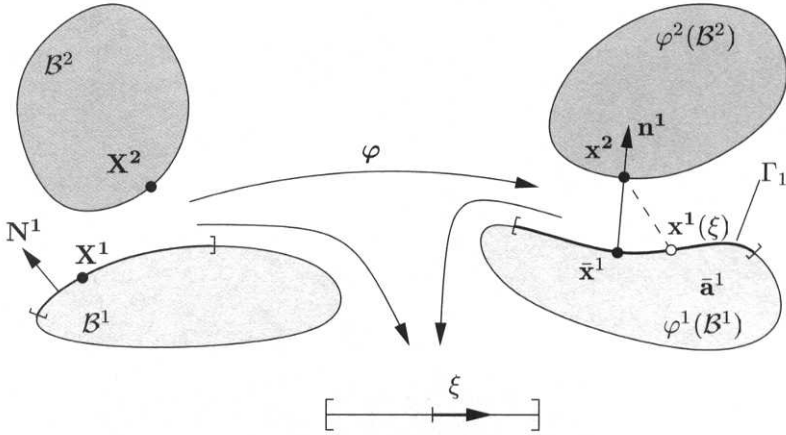
## Contact Kinematics

Many technical contact problems involve large deformations of the bodies that are in contact. Thus we will formulate all contact relations for finite deformations, so we look at problems where two or more bodies  $B^\alpha$  approach each other during a finite deformation process and come into contact on parts of their boundaries denoted by  $\Gamma_c$ , see Figure 3.1. We observe that two points,  $\mathbf{X}^1$  and  $\mathbf{X}^2$ , in the initial configuration of the bodies which are distinct can occupy the same position in the current configuration,  $\varphi(\mathbf{X}^2) = \varphi(\mathbf{X}^1)$ , within the deformation process. Hence contact conditions have to be formulated with respect to the current configuration. In general, two steps have



**Fig. 3.1** Finite deformation contact.





**Fig. 3.2** Deformed configuration of bodies  $B^\alpha$ , minimum distance.

to be performed to detect whether or not contact takes place. These are the global search for contact and the set-up of local kinematical relations which are needed to formulate the contact constraints. Here we focus on the latter; search algorithms will be discussed in Section 9.1.

In a large deformation, continuum-based formulation of contact kinematics, the distance between the bodies in contact is minimized, as can be found for the classical non-penetration condition in Curnier and Alart (1988) for example.

In the case that a small penetration due to the approach of the two bodies in contact has to be allowed, the contact kinematics are developed in Wriggers and Miehe (1992). This non-penetration function also plays a significant role for the definition of the tangential velocity in the contact interface, which is needed to formulate frictional problems, e.g. see Simo and Laursen (1992), Wriggers and Miehe (1992), Laursen and Simo (1993b), or Curnier et al. (1995).

Let us consider two elastic bodies  $B^\alpha$ ,  $\alpha = 1, 2$ , each occupying the bounded domain  $\Omega^\alpha \subset R^3$ . The boundary  $\Gamma^\alpha$  of a body  $B^\alpha$  consists of three parts:  $\Gamma_\sigma^\alpha$  with prescribed surface loads,  $\Gamma_u^\alpha$  with prescribed displacements, and  $\Gamma_c^\alpha$ , where the two bodies  $B^1$  and  $B^2$  come into contact. In the contact area, we have to formulate the constraint equations or the approach function for normal contact, as well as the kinematical relations for the tangential contact.

### 3.1 NORMAL CONTACT OF THREE-DIMENSIONAL BODIES

Assume that two bodies come into contact. In that case, the non-penetration condition is given by

$$(\mathbf{x}^2 - \mathbf{x}^1) \cdot \mathbf{n}^1 \geq 0, \quad (3.1)$$

see Figure 3.2.  $\mathbf{x}^\alpha$  denotes the coordinates of the current configuration  $\varphi(\mathcal{B}^\alpha)$  of body  $\mathcal{B}^\alpha$ :  $\mathbf{x}^\alpha = \mathbf{X}^\alpha + \mathbf{u}^\alpha$ , where  $\mathbf{X}^\alpha$  is related to the initial configuration  $\mathcal{B}^\alpha$  and  $\mathbf{u}^\alpha$  is the displacement field, see also Appendix B. The normal vector  $\mathbf{n}^1$  is associated with body  $\mathcal{B}^1$ . By assuming that the contact boundary describes, at least locally, a convex region, we can relate to every point  $\mathbf{x}^2$  on  $\Gamma^2$  a point  $\bar{\mathbf{x}}^1 = \mathbf{x}^1(\bar{\xi})$  on  $\Gamma^1$  via the minimum distance problem

$$\hat{d}^1(\xi^1, \xi^2) = \|\mathbf{x}^2 - \bar{\mathbf{x}}^1\| = \min_{\mathbf{x}^1 \in \Gamma^1} \|\mathbf{x}^2 - \mathbf{x}^1(\xi)\|, \quad (3.2)$$

see Fig 3.2b for the two-dimensional case. The distance can then be used to define the gap or penetration between the two bodies.  $\xi = (\xi^1, \xi^2)$  denotes the parameterization of the boundary  $\Gamma^1$  via convective coordinates, e.g. see Wriggers and Miehe (1992), Wriggers and Miehe (1994) or Laursen and Simo (1993b). The point  $\bar{\mathbf{x}}^1$  is computed from the necessary condition for the minimum of the distance function (3.2)

$$\frac{d}{d\xi^\alpha} \hat{d}^1(\xi^1, \xi^2) = \frac{\mathbf{x}^2 - \hat{\mathbf{x}}^1(\xi^1, \xi^2)}{\|\mathbf{x}^2 - \hat{\mathbf{x}}^1(\xi^1, \xi^2)\|} \cdot \hat{\mathbf{x}}_\alpha^1(\xi^1, \xi^2) = 0. \quad (3.3)$$

The solution of (3.3) requires orthogonality of the first and second terms. Since  $\hat{\mathbf{x}}_\alpha^1(\xi^1, \xi^2)$  is the tangent vector  $\mathbf{a}_\alpha^1$ , the first term must have the same direction as the normal vector  $\mathbf{n}^1$  at the minimum point  $(\bar{\xi}^1, \bar{\xi}^2)$ . Thus we have the condition  $-\mathbf{n}^1(\bar{\xi}^1, \bar{\xi}^2) \cdot \mathbf{a}_\alpha^1(\bar{\xi}^1, \bar{\xi}^2) = 0$ , which means that the current master point  $\hat{\mathbf{x}}^1(\bar{\xi}^1, \bar{\xi}^2)$  is the orthogonal projection of a given slave point  $\mathbf{x}^2$  onto the current master surface  $\varphi_t^1(\Gamma_c^1)$ .

Here and in the following, we denote by a bar over a quantity its evaluation at the minimal distance point  $(\bar{\xi}^1, \bar{\xi}^2)$ , which means that these values denote the solution point of (3.3). Thus

$$\bar{\mathbf{n}}^1 = \frac{(\bar{\mathbf{a}}_1^1 \times \bar{\mathbf{a}}_2^1)}{\|\bar{\mathbf{a}}_1^1 \times \bar{\mathbf{a}}_2^1\|} \quad (3.4)$$

is the outward unit normal on the current master surface at the master point, where  $\bar{\mathbf{a}}_\alpha^1$  are tangent vectors at  $\hat{\mathbf{x}}^1(\bar{\xi}^1, \bar{\xi}^2)$ . Note that the normal can also be defined using (3.3). The result is

$$\bar{\mathbf{n}}^1 = -\frac{\mathbf{x}^2 - \hat{\mathbf{x}}^1(\xi^1, \xi^2)}{\|\mathbf{x}^2 - \hat{\mathbf{x}}^1(\xi^1, \xi^2)\|}. \quad (3.5)$$

The application of this definition is more convenient in special cases; however it can only be used in relation with the penalty method, and it has the problem that  $\bar{\mathbf{n}}^1$  is not defined for  $\|\mathbf{x}^2 - \hat{\mathbf{x}}^1(\xi^1, \xi^2)\| = 0$ .

Once the point  $\bar{\mathbf{x}}^1$  is known, we can define either the inequality constraint of the non-penetration condition

$$\boxed{g_N = (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \bar{\mathbf{n}}^1 \geq 0}, \quad (3.6)$$

or a penetration function

$$\boxed{g_N^- = \begin{cases} (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \bar{\mathbf{n}}^1 & \text{if } (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \bar{\mathbf{n}}^1 < 0 \\ 0 & \text{otherwise} \end{cases}} \quad (3.7)$$

The latter defines the magnitude of penetration of one body into the other and has to be used in conjunction with the penalty method.

**REMARK 3.1:** It can happen that the distance function is locally non-differentiable. In such a case the projection point  $\hat{\mathbf{x}}^1(\xi^1, \xi^2)$  is not unique. Examples of points which have a non-unique projection are shown in Figure 3.3. Usually the discontinuity depicted in Figure 3.3a is not of great influence in practical applications, since in a next iteration step when the point gets closer to the contact surface, the discontinuity disappears. A criterion which provides a measure to detect whether the projection (3.3) yields a unique solution or not was derived in Curnier et al. (1995). Using the transformation of line elements which lie in the tangent plane of the master and slave surface  $d\hat{\mathbf{x}}^1 = \mathbf{P}_\perp d\mathbf{x}^2$ , for the area elements on master and slave surface one obtains

$$\bar{\mathbf{n}}^1 da^1 = \mathbf{P}_\perp \mathbf{n}^2 da^2. \quad (3.8)$$

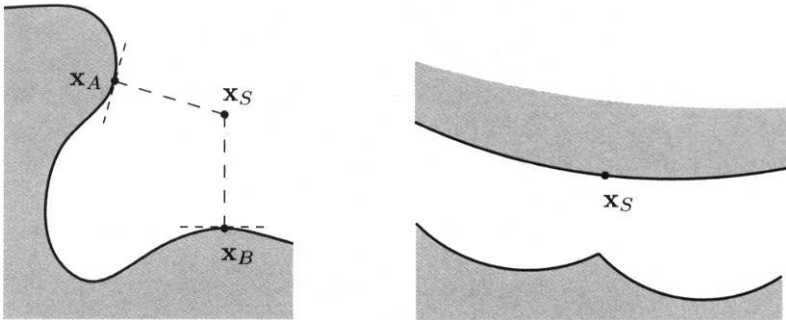
The ratio of these area element is now given by (see Curnier et al. (1995)),

$$\bar{j} = \frac{da^1}{da^2} = \frac{-\mathbf{n}^2 \cdot \bar{\mathbf{n}}^1}{|1 - (\kappa_1 + \kappa_2) g_N + \kappa_1 \kappa_2 g_n^2|}, \quad (3.9)$$

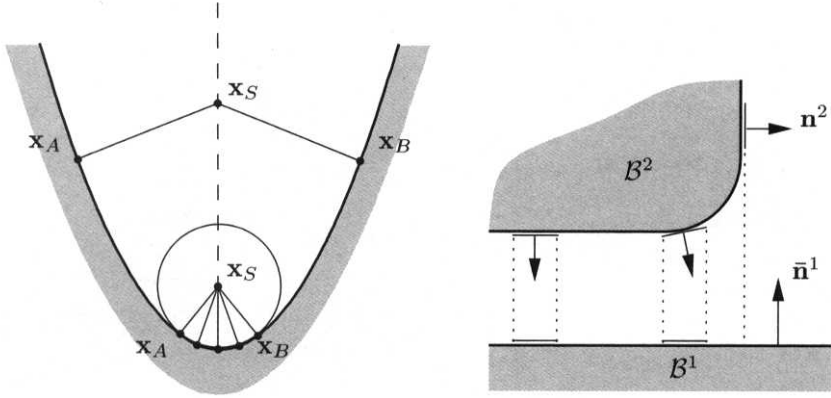
where we can also write  $\bar{j} = \|\mathbf{P}_\perp \mathbf{n}^2\|$ .  $\kappa_1$  and  $\kappa_2$  describe the principle curvature of the master surface in point  $\hat{\mathbf{x}}^1(\xi^1, \xi^2)$ . The projection is unique when  $\bar{j} < \infty$ , which means that the point  $\mathbf{x}^2$  should not be located at a point of focus which is the cutting point of two normal vectors of equal length (see Figure 3.4a) or it should not coincide with the center of curvature of the master surface. One can show (see Thorpe (1979)) that  $\hat{\mathbf{x}}^1(\xi^1, \xi^2)$  is associated with a local minimum if no point of focus lies on the line between  $\mathbf{x}^2$  and  $\hat{\mathbf{x}}^1(\xi^1, \xi^2)$ . A global minimum of (3.3) is given for

$$|g_N| < \frac{1}{\max(|\kappa_1|, |\kappa_2|)}, \quad (3.10)$$

which means that the distance between the slave node and master surface is smaller than the smallest radius of curvature for each point of the master surface. Furthermore, to obtain a projection (3.8) which is bijective one has to require that  $\bar{j} > 0$ , which means that the condition  $-\mathbf{n}^2 \cdot \bar{\mathbf{n}}^1 > 0$  holds (see Figure 3.4b) for the normal vectors.



**Fig. 3.3** Non-differentiable distance functions.



**Fig. 3.4** Uniqueness and regularity for minimum distance problem.

The case shown in Figure 3.3b is of more practical importance. This is because often, low order interpolations are applied for the finite element discretization, which use a linear function within the element to describe the deformed boundary. Hence there is always a discontinuity in the normal field when point  $\mathbf{x}^2$  moves from one element to the next. This problem will be discussed in more detail in Section 9.3.

In the case of geometrically linear kinematics, it is advantageous to write the inequality constraint as

$$(\mathbf{u}^2 - \bar{\mathbf{u}}^1) \cdot \bar{\mathbf{n}}^1 + g_X \geq 0, \quad (3.11)$$

where  $\mathbf{u}^\alpha$  are the displacements of the bodies  $B^\alpha$ . The initial gap  $g$  between the two bodies is given by  $g_X = (\mathbf{X}^2 - \bar{\mathbf{X}}^1) \cdot \bar{\mathbf{n}}^1$ . In view of the penalty formulation which will be applied to solve the contact problems, we introduce a penetration function as follows:

$$u_N^- = \begin{cases} (\mathbf{u}^2 - \bar{\mathbf{u}}^1) \cdot \bar{\mathbf{n}}^1 + g_X & \text{if } (\mathbf{u}^2 - \bar{\mathbf{u}}^1) \cdot \bar{\mathbf{n}}^1 + g_X < 0 \\ 0 & \text{otherwise.} \end{cases} \quad (3.12)$$

Functions  $g_N^-$  and  $u_N^-$  indicate a penetration of one body into the other, and show in which parts of  $\Gamma^\alpha$  the constraint equations, preventing penetration, have to be activated. Thus (3.7) or (3.12) can be used to determine the contact area  $\Gamma_c^\alpha \subseteq \Gamma^\alpha$ .

**REMARK 3.2:** In the case of contact between a rigid surface and a deformable body, equations (3.7) and (3.12) also hold. In this case,  $\bar{\mathbf{u}}^1 \equiv \mathbf{0}$  holds and  $\bar{\mathbf{n}}^1$  is the normal of the rigid body.

### 3.2 TANGENTIAL CONTACT OF THREE-DIMENSIONAL BODIES

In the tangential direction of the contact interface one has generally to distinguish two cases. The first is the so-called 'stick state' in which a point which is in contact is not

allowed to move in a tangential direction. The second case is *sliding*, which means that a point moves in a tangential direction in the contact interface. This movement can of course only be described by the relative deformation in the interface.

### 3.2.1 Stick condition

In the first case a mathematical condition for stick is needed. Such a condition can be derived from projection (3.3). It is clear that a point which sticks to another body does not move in a tangential direction, and hence the computed values for the convective coordinates  $(\xi^1, \xi^2)$  do not change for this point during the motion  $\dot{\xi}^\alpha = 0$ . Therefore, the condition

$$\mathbf{g}_T = g_{T\alpha} \bar{\mathbf{a}}^\alpha = \mathbf{0} \quad \text{with} \quad g_{T\alpha} = (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \bar{\mathbf{a}}_\alpha^1 \quad (3.13)$$

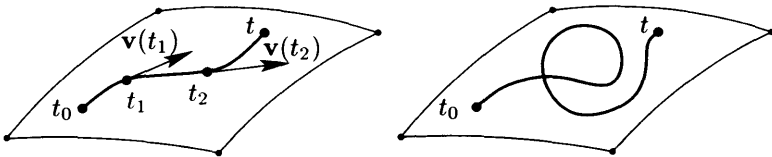
can be formulated. Here  $\mathbf{g}_T$  denotes the relative displacement in a tangential direction, which has to be zero. Note that in the stick case, one does not have to distinguish between the normal and tangential directions. Thus in case the normal gap is closed, see (3.6),  $g_N = (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \bar{\mathbf{n}}^1 = 0$ ; then one can instead combine conditions  $g_N = 0$  and  $\mathbf{g}_T = \mathbf{0}$  to the more simple condition

$$\mathbf{x}^2 - \bar{\mathbf{x}}^1 = \mathbf{0}. \quad (3.14)$$

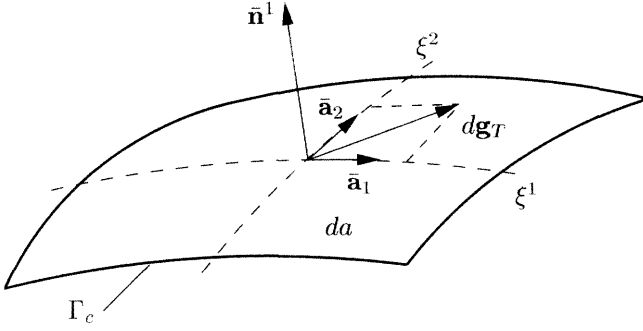
The implication of this choice with respect to numerical implementation of contact is discussed in Remark 5.2 and Section 8.2.

### 3.2.2 Slip condition

The tangential relative slip between two bodies is related to the change of the point  $\mathbf{x}^2$  relative to the projection  $\bar{\mathbf{x}}^1$ . This means that the solution point  $\bar{\xi} = (\bar{\xi}^1, \bar{\xi}^2)$  which has been obtained via the minimal distance problem (3.2) will move on the master surface. The sketch in Figure 3.5a depicts the path of the point  $\mathbf{x}^2$  along on the master surface beginning at  $t_0$  and ending at time  $t$ . Furthermore, the velocity  $\mathbf{v}$  of point  $\mathbf{x}^2$  relative to the master surface is shown at times  $t_1$  and  $t_2$ . Note that the path of a point  $\mathbf{x}^2$  on the master surface is not known *a priori*. It could even cross its own line, as depicted in Figure 3.5b. Thus, during our calculations we cannot assume anything regarding the path. Hence in a frictional sliding situation, one has to integrate the



**Fig. 3.5** Path of the point  $\mathbf{x}^2$  relative to the master surface



**Fig. 3.6** Increment of the frictional path.

relative velocities to obtain the path of  $\mathbf{x}^2$  on the master surface. Details regarding the continuum formulation are stated below.

First we state the tangential relative displacement of a point  $\mathbf{x}^2$  on the contact surface, which is defined in terms of body  $B^1$ . We compute the path of point  $\mathbf{x}^2$  on  $\Gamma^1$  from

$$d\mathbf{g}_T = \bar{\mathbf{a}}_\alpha^1 d\xi^\alpha = \bar{\mathbf{x}}_{,\alpha}^1 d\xi^\alpha, \quad (3.15)$$

see Figure 3.6. From (3.15) the length of the frictional path can be computed with  $d\mathbf{g}_T = \|\mathbf{d}\mathbf{g}_T\|$  and  $d\xi^\alpha = \dot{\xi}^\alpha dt$  as

$$g_T = \int_{t_0}^t \|\dot{\xi}^\alpha \bar{\mathbf{x}}_{,\alpha}^1\| dt = \int_{t_0}^t \sqrt{\dot{\xi}^\alpha \dot{\xi}^\beta a_{\alpha\beta}} dt, \quad (3.16)$$

where  $t$  is the time which is used to parameterize the path of point  $\mathbf{x}^2$ . For the evaluation of (3.16), we have to know the time derivative of  $\dot{\xi}^\alpha$ . This can be computed from the relation

$$[\mathbf{x}^2 - \bar{\mathbf{x}}^1] \cdot \bar{\mathbf{a}}_\alpha^1 = 0, \quad (3.17)$$

which is valid at the contact point, since the difference  $\mathbf{x}^2 - \bar{\mathbf{x}}^1$  is normal to the contact surface, and  $\bar{\mathbf{a}}_\alpha^1$  denotes the tangent vector to the surface  $\Gamma^1$  at the minimal distance point, see Figure 3.2b. This yields the following result:

$$\frac{d}{dt} [\mathbf{x}^2 - \bar{\mathbf{x}}^1(\xi^1, \xi^2)] \cdot \bar{\mathbf{a}}_\alpha^1 = [\mathbf{v}^2 - \bar{\mathbf{v}}^1 - \bar{\mathbf{a}}_\beta^1 \dot{\xi}^\beta] \cdot \bar{\mathbf{a}}_\alpha^1 + [\mathbf{x}^2 - \bar{\mathbf{x}}^1] \cdot \dot{\bar{\mathbf{a}}}_\alpha^1 = 0. \quad (3.18)$$

With  $\dot{\bar{\mathbf{a}}}_\alpha^1 = \bar{\mathbf{v}}_{,\alpha}^1 + \hat{\mathbf{x}}_{,\alpha\beta}^1 \dot{\xi}^\beta$  we obtain  $\dot{\xi}^\beta$  from (3.18), which leads to the following system of equations:

$$\bar{H}_{\alpha\beta} \dot{\xi}^\beta = \bar{R}_\alpha, \quad (3.19)$$

with

$$\begin{aligned}\bar{H}_{\alpha\beta} &= [\bar{a}_{\alpha\beta} + g_N \bar{\mathbf{x}}_{,\alpha\beta}^1 \cdot \bar{\mathbf{n}}^1] \\ \bar{R}_\alpha &= [\mathbf{v}_2^1 - \bar{\mathbf{v}}^1] \cdot \bar{\mathbf{a}}_\alpha^1 + [\mathbf{x}^2 - \bar{\mathbf{x}}^1] \cdot \bar{\mathbf{v}}_{,\alpha}^1 \\ &= [\mathbf{v}^2 - \bar{\mathbf{v}}^1] \cdot \bar{\mathbf{a}}_\alpha^1 + g_N \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{v}}_{,\alpha}^1.\end{aligned}\quad (3.20)$$

Well known results from differential geometry of surfaces introduce  $\bar{a}_{\alpha\beta}$  as the metric tensor and  $\bar{\mathbf{x}}_{,\alpha\beta}^1(\bar{\xi}^1, \bar{\xi}^2) \cdot \bar{\mathbf{n}}^1$  as the curvature tensor  $\bar{b}_{\alpha\beta}$ , see also (B.15). Thus we can rewrite  $\bar{H}_{\alpha\beta} = [\bar{a}_{\alpha\beta} + g_N \bar{b}_{\alpha\beta}]$ . Denoting  $\bar{H}^{\alpha\beta} = (\bar{H}_{\alpha\beta})^{-1}$ , we can now explicitly solve for the relative velocity  $\dot{\xi} = \{\bar{\xi}^1, \bar{\xi}^2\}$ ,

$$\dot{\xi}^\beta = \bar{H}^{\alpha\beta} [(\mathbf{v}^2 - \bar{\mathbf{v}}^1) \cdot \bar{\mathbf{a}}_\alpha^1 + g_N \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{v}}_{,\alpha}^1]. \quad (3.21)$$

Using these results we define as the second important kinematical function the tangential relative velocity function on the current surface  $\varphi^2(\Gamma_c^2)$  by setting

$$\boxed{\mathcal{L}_v \mathbf{g}_T := \dot{\xi}^\alpha \bar{\mathbf{a}}_\alpha}. \quad (3.22)$$

Equation (3.22) determines the evolution of the tangential slip  $\mathbf{g}_T$ , which enters as a local kinematical the constitutive function for the contact tangential stress. In the following chapter, we will abbreviate the LIE derivative  $\mathcal{L}_v \mathbf{g}_T$ , denoting the relative tangential velocity by the term  $\dot{\mathbf{g}}_T$  to simplify notation.

**REMARK 3.3:**

1. Note that the second terms on the right-hand side of (3.19) depends upon the penetration  $g_N$ . Thus in the case of a strong enforcement of the non-penetration condition ( $g_N = 0$ ) with LAGRANGE multipliers, this term vanishes. The evolution  $\mathcal{L}_v$  in (3.22) is then given by the projection of the spatial velocities  $\mathbf{v}^2$  and  $\hat{\mathbf{v}}^1(\bar{\xi})$ , evaluated at the contact points onto the tangential direction of the contact surface at the master point

$$\mathcal{L}_v \mathbf{g}_T := \bar{\mathbf{P}}_T [\mathbf{v}^1 - \hat{\mathbf{v}}^2(\bar{\xi}^1, \bar{\xi}^2)], \quad \text{with} \quad \bar{\mathbf{P}}_T = \bar{\mathbf{a}}_\alpha^2 \otimes \bar{\mathbf{a}}^{2\alpha}. \quad (3.23)$$

2. If the deformed contact surface is flat, then the curvature tensor  $\bar{b}_{\alpha\beta}$  is zero. This is always the case for a surface discretization by three node triangular elements.
3. Note that the (a priori objective) Lie derivative of the tangential vector  $\mathbf{g}_T$  has the representation  $\mathcal{L}_v \mathbf{g}_T = \mathbf{F}^2 \{ \frac{d}{dt} [\mathbf{F}^{2-1}(\mathbf{g}_T)] \} = \dot{\xi}^\alpha \bar{\mathbf{a}}_\alpha$  based on the deformation gradient  $\mathbf{F}^2$  of the master surface defined above. Thus (3.22) represents an evolution equation for the objective rate  $\mathcal{L}_v \mathbf{g}_T$  of the tangential vector introduced above.
4. In the case of no relative movement in a tangential direction (stick condition), we have  $\mathcal{L}_v \mathbf{g}_T = \mathbf{g}_T = \mathbf{0}$ .
5. In the geometrically linear case, from (3.19) and (3.21) we obtain

$$\frac{d}{dt} \{ [\mathbf{x}^1 - \hat{\mathbf{x}}^2(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{A}}_\alpha^2 \} = [\mathbf{v}^1 - \hat{\mathbf{v}}^2(\bar{\xi}^1, \bar{\xi}^2) - \bar{\mathbf{A}}_\beta^2 \dot{\xi}^\beta] \cdot \bar{\mathbf{A}}_\alpha^2,$$

which yields

$$\bar{\mathbf{A}}_{\alpha\beta}^2 \dot{\xi}^\beta = [\mathbf{v}^1 - \hat{\mathbf{v}}^2(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{A}}_\alpha^2.$$

The terms multiplied by  $g_N$  can be neglected. Thus  $\dot{\xi}^\beta$  is given by the projection of the difference velocity of the two bodies at the contact point on the tangent direction of the undeformed surface. From the last equation, we can deduce the relative tangential velocity at the contact point:  $\dot{\mathbf{g}}_T = \dot{\xi}^\beta \bar{\mathbf{A}}_\alpha^2$ .

For a penetration  $g_N < 0$  we have to take into account the second term in (3.21) and the scaling factors  $\bar{H}_{\alpha\beta}$ , both consequences of the time dependence of  $\bar{\mathbf{a}}_\alpha^2$ .

For the two-dimensional contact we can specify the result in (3.19), which then yields

$$\dot{\xi} = \frac{1}{\bar{a}_{11} + g_N \bar{b}_{11}} \{ [\mathbf{v}^2 - \bar{\mathbf{v}}^1] \cdot \bar{\mathbf{x}}^1_{,\xi} + g_N \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{v}}^1_{,\xi} \}, \quad (3.24)$$

where  $\bar{a}_{11} = \mathbf{x}^1_{,\xi}(\bar{\xi}) \cdot \mathbf{x}^1_{,\xi}(\bar{\xi})$  describes the metric and  $b_{11} = \mathbf{x}^1_{,\xi\xi}(\bar{\xi}) \cdot \bar{\mathbf{n}}^1$  the curvature of the boundary. The vectors  $\mathbf{v}^\alpha$  denote the velocities at  $\mathbf{x}^\alpha$ . Knowing the change of the coordinate  $\bar{\xi}$ , we can define the relative tangential velocity as

$$\mathcal{L}_v \mathbf{g}_T := \dot{\xi} \bar{\mathbf{x}}^1_{,\xi}(\bar{\xi}). \quad (3.25)$$

In that case, for the total sliding distance (3.16) we obtain

$$g_T = \int_{t_0}^t \|\dot{\xi}^1 \bar{\mathbf{x}}^1_{,\xi}\| dt = \int_{\xi_0}^{\xi} \sqrt{\bar{a}_{11}} d\xi. \quad (3.26)$$

The upper limit  $\xi$  in this integral still depends upon the deformation, which is clear from the implicit definition of  $\xi$  via (3.3). However, one can transform (3.26) into an integral with fixed limits by introducing a different parameterization, e.g. as in the case when the isoparametric map is used in a finite element context.

### 3.3 VARIATION OF THE NORMAL AND TANGENTIAL GAP

The variation of the geometrical contact quantities are needed in the weak forms for contact. In this section the variation of the gap in the normal contact direction and the variation of the tangential relative displacement in the contact interface are derived for the different cases discussed in the previous sections.

#### 3.3.1 Variation of normal gap

The first case discussed is the variation of the gap function in the normal direction for the contact between three-dimensional solids.

The variation of the normal gap follows from (3.6) as

$$\delta g_N = \delta \{ [\mathbf{x}^2 - \mathbf{x}^1(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{n}}^1(\bar{\xi}^1, \bar{\xi}^2) \}. \quad (3.27)$$

To compute the variation (3.27) explicitly, we have to consider all terms which depend upon the deformation. In the case of contact, we then have to take into account



the projection of point  $\mathbf{x}^2$  onto the master surface parameterized by the convective coordinates  $\xi^1$  and  $\xi^2$ . This leads to

$$\delta g_N = [\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1 - \bar{\mathbf{x}}_{,\alpha}^1 \delta \xi^\alpha] \cdot \bar{\mathbf{n}}_1 + [\mathbf{x}^2 - \bar{\mathbf{x}}^1] \cdot \delta \bar{\mathbf{n}}^1, \quad (3.28)$$

where we have set  $\boldsymbol{\eta}^\alpha = \delta \mathbf{x}^\alpha$ . Equation (3.28) simplifies due the fact that  $\bar{\mathbf{x}}_{,\alpha}^1 \cdot \bar{\mathbf{n}}^1 = 0$ . Furthermore we have  $\bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{n}}^1 = 0$ , see Appendix B. With the definition of the normal (3.5), this eliminates the last term in (3.28). Hence we obtain the result

$$\boxed{\delta g_N = [\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1] \cdot \bar{\mathbf{n}}^1}. \quad (3.29)$$

Note, however, that we have to start from (3.28) if we want to derive the linearization of the variation of the gap function (3.27).

### 3.3.2 Variation of tangential gap

The variation of the tangential slip can be stated in the same way as the time derivative (relative tangential slip velocity) was computed in (3.21) and (3.22). We obtain

$$\boxed{\delta \mathbf{g}_T = \delta \bar{\xi}^\alpha \bar{\mathbf{a}}_\alpha^1}, \quad (3.30)$$

with the variation of  $\xi^\alpha$

$$\delta \bar{\xi}^\alpha = \bar{H}^{\alpha\beta} [(\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot \bar{\mathbf{a}}_\beta^1 + g_N \bar{\mathbf{n}}^1 \cdot \bar{\boldsymbol{\eta}}_{,\beta}^1]. \quad (3.31)$$

The latter relation follows simply from (3.21) by replacing the velocities  $\mathbf{v}$  by the test function  $\boldsymbol{\eta}$  with  $\bar{H}^{\alpha\beta} = [\bar{a}_{\alpha\beta} + g_N \bar{b}_{\alpha\beta}]^{-1}$ .

In the two-dimensional case, these equations simplify. The variation of  $\xi$  can be computed using (3.24) according to (3.31). This leads to

$$\boxed{\delta \bar{\xi} = \frac{1}{\bar{a}_{11} + g_N \bar{b}_{11}} \{ [\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1] \cdot \bar{\mathbf{x}}_{,\xi}^1 + g_N \bar{\mathbf{n}}^1 \cdot \bar{\boldsymbol{\eta}}_{,\xi}^1 \}}. \quad (3.32)$$

# 4

---

## *Constitutive Equations for Contact Interfaces*

The design of technical systems depends greatly upon the knowledge of the contact behaviour in interfaces which connect different parts of the system. Such systems are very general, examples being cars, printing or copy machines, human joints or implants, unfolding space structures, robots, micro machines or base isolation systems for buildings to protect against earthquakes.

Related to the precision which is needed to resolve the mechanical behaviour in the contact interface, different approaches have been established over the centuries to model the mechanical behaviour in the contact area.

Two main techniques can be followed to impose contact conditions in the normal direction. These are the formulation of the non-penetration condition as a purely geometrical constraint (the normal contact stresses then follow from the constraint equation), and the development of an elastic or elasto-plastic constitutive laws for the micromechanical approach within the contact area, which yields a response function for the normal contact stresses. Such constitutive equations are often derived from statistical models.

For the tangential direction, one has the same situation as for normal contact when stick in the contact interface is considered. Again, either a geometrical constraint equation can be formulated, or a constitutive law for the tangential relative micro displacements between the contacting bodies can be applied. For tangential sliding between bodies, one always has to derive a constitutive equation for friction which can be stated in the form of an evolution equation. Usually the frictional evolution depends on different parameters, like normal force, relative tangential velocity, temperature or total slip distance.

## 4.1 NORMAL CONTACT

The next two sections discuss the formulation of normal contact as a constraint for non-penetration, or by introducing constitutive equations which describe the approach of the contacting surfaces. In the case of frictionless contact, one of the models proposed for normal contact is sufficient within the analysis. In the case of friction, the tangential stress also has to be considered, see Section 4.2.

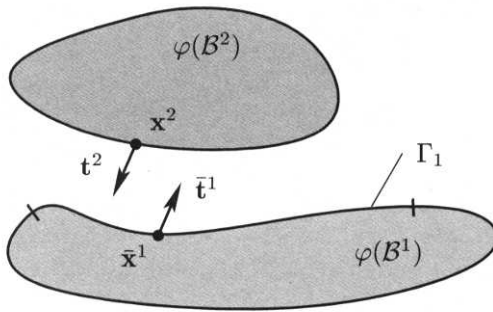
### 4.1.1 Constraint formulation

The formulation which treats normal contact as a unilateral constraint problem is in general used when it is most essential to enforce the geometrical constraints in a correct way (one could coin the term “low contact precision” for such a situation). In this case, it is not necessary to specify constitutive relations in the contact interface. Therefore, the normal contact pressure cannot be computed from a constitutive equation, but is then obtained as a reaction in the contact area, and hence can be deduced from the constraint equations, see Section 5.3.1. This procedure is the classical way to formulate contact constraints, thus numerous researchers have used this strategy, e.g. see Johnson (1985) or Kikuchi and Oden (1988).

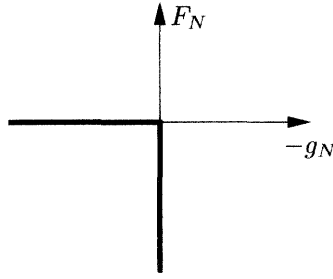
The mathematical condition for non-penetration is stated as  $g_N \geq 0$ , see (3.6), which precludes the penetration of body  $B^1$  into body  $B^2$ . Contact takes place when  $g_N$  is equal to zero. In this case, the associated normal component  $p_N^1$  of the stress vector

$$\mathbf{t}^1 = \boldsymbol{\sigma}^1 \bar{\mathbf{n}}^1 = p_N^1 \bar{\mathbf{n}}^1 + t_T^{1\beta} \bar{\mathbf{a}}_\beta^1 \quad (4.1)$$

in the contact interface must be non-zero. The stress vector acts on both surfaces (see Figure 4.1) obeying the action-reaction principle:  $\mathbf{t}^1(\bar{\xi}^1, \bar{\xi}^2) = -\mathbf{t}^2$  in the contact point  $\bar{\mathbf{x}}^1$ . We have  $p_N = p_N^1 = p_N^2 < 0$ , since adhesive stresses will not be allowed in the contact interface throughout our considerations. Note that the tangential stress  $t_T^{1\beta}$  is zero in the case of frictionless contact. For contact one has the conditions  $g_N = 0$  and  $p_N < 0$ . If there is a gap between the bodies, then the relations  $g_N > 0$



**Fig. 4.1** Stresses in the contact interface.



**Fig. 4.2** Contact force versus normal gap.

and  $p_N = 0$  hold. This leads to the statements

$$\boxed{g_N \geq 0, \quad p_N \leq 0, \quad p_N g_N = 0}, \quad (4.2)$$

which are known as HERTZ–SIGNORINI–MOREAU conditions for frictionless contact. These conditions provide the basis to treat frictionless contact problems in the context of constraint optimization. In MOREAU (1974) this law is expressed by subgradients (see also CURNIER (1999)). In optimization theory such a set of equations is called a KUHN–TUCKER condition (sometimes known as a KUHN–TUCKER–KARUSH condition, since it was recently discovered that KARUSH developed them first in 1938). As can be seen from Figure 4.2, equations (4.2) lead to a non-smooth contact law for the normal contact pressure.

So far stresses related to the current configuration have been introduced. Hence CAUCHY's theorem was formulated in terms of the CAUCHY stresses  $\sigma$ , see Eq. (4.1). In the same way, one can also introduce the nominal stress or first PIOLA–KIRCHHOFF stress  $\mathbf{P}$  this leads to two different representations of the stress vector

$$\mathbf{t} = \sigma \mathbf{n} \quad \text{or} \quad \mathbf{T} = \mathbf{P} \mathbf{N}. \quad (4.3)$$

With NANSON's formula (2.11) and the transformation (2.48), one concludes that  $\mathbf{T} dA = \mathbf{t} da$ . Hence by introducing the JACOBIAN for the area  $j = da / dA$ , one obtains

$$\mathbf{T} = j \mathbf{t} \quad \text{with} \quad j = J \|\mathbf{F}^{-T} \mathbf{N}\|. \quad (4.4)$$

The normal and tangential components follow from

$$\mathbf{t} = \mathbb{P} \mathbf{t} + (\mathbb{I} - \mathbb{P}) \mathbf{t}, \quad (4.5)$$

with the projection tensor  $\mathbb{P} = \bar{\mathbf{n}}^1 \otimes \bar{\mathbf{n}}^1$  and the fourth order unit tensor  $\mathbf{t} = \mathbb{I} \mathbf{t}$ . Note that  $\mathbb{P}$  has the following properties:  $\mathbb{P}^2 = \mathbb{P}$  and  $\mathbb{P} (\mathbb{I} - \mathbb{P}) = \mathbb{O}$ . With (4.4) and (4.1), the relations

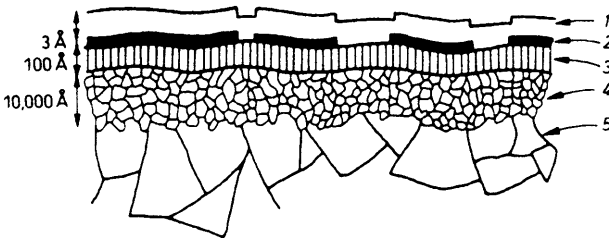
$$T_N = j p_N = j \bar{\mathbf{n}}^1 \cdot \mathbb{P} \mathbf{t} \quad \text{and} \quad \mathbf{T}_T = j \mathbf{t}_T = j (\mathbb{I} - \mathbb{P}) \mathbf{t} \quad (4.6)$$

hold. Both formulations, using the nominal stresses or the CAUCHY stresses, are equivalent when the transformations derived above are introduced in a consistent way into constitutive equations or weak forms.

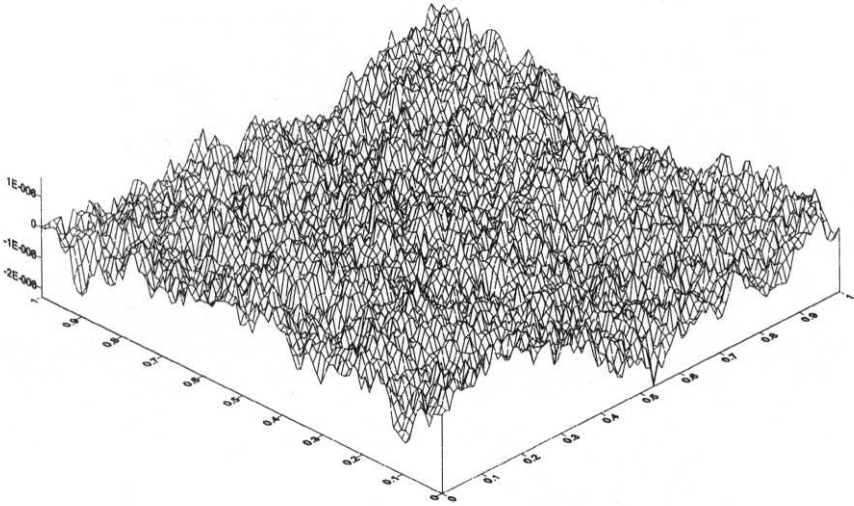
#### 4.1.2 Constitutive equations for normal contact

There exist several contact problems where knowledge of the micromechanical approach of the surfaces to each other is essential for a proper treatment of the physical phenomena. In that case an interface compliance is needed for these problems with “high contact precision”. Constitutive equations which include deformation of the micro structure can be developed for normal contact by investigating the micromechanical behaviour within the contact surface. Associated models have been developed based on experiments, e.g. see Greenwood and Williamson (1966) or Kragelsky et al. (1982). The micromechanical behaviour depends in general upon material parameters like hardness and on geometrical parameters like surface roughness. However, depending on the materials of the bodies being in contact other parameters also have to be considered when all effects stemming from the micromechanics of the surfaces have to be included. This can be observed from Figure 4.3, which shows the complex layer structure of a solid at its boundary. It consists for example of a contaminant layer (1), an absorbed gas layer (2), an oxide film (3), a work-hardened layer (4), and finally, the metal substrate (5). Hence the real micromechanical phenomena are very complex: it is possible to initiate even chemical reactions in the interface by extremely high local pressures at the asperities stemming from applied mechanical forces. However, such phenomena will be neglected here.

A micromechanical view of a rough surface is shown in Figure 4.4, which clearly depicts different randomly distributed asperities. It is obvious that contacting bodies at the asperities of the rough surfaces come in contact where high local stresses occur. Thus, early investigations assumed plastic deformations when the interface laws were derived. This is correct when the loads are only applied once. For repeated loading one can assume that the surface is flattened due to initial plastic deformations and hence after this pre-deformation can bear loads elastically, e.g. see Mikic (1971). Thus the models which are used try only to capture the most important phenomena,



**Fig. 4.3** Surface layers on a workpiece; schematic view from Rabinowicz (1995).



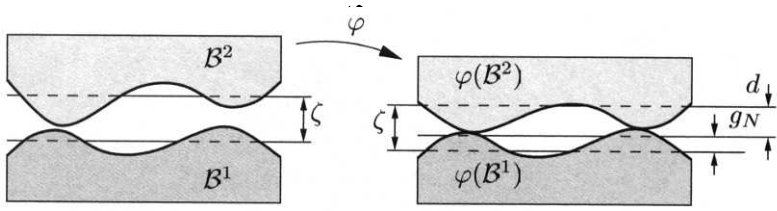
**Fig. 4.4** : Scan of a rough micro-mechanical surface geometry.

and assume either an elastic or a plastic deformation of the asperities that are actually in contact at the interface.

The derivation of constitutive contact equations for the approach of two rough surfaces involves two main steps. First one has to find a mathematical description of the surface geometry by statistical, e.g. see Greenwood and Williamson (1966) or lately also by fractal models e.g. see Majumdar and Bhushan (1990) and Section 4.7. Secondly a model which describes the mechanical behaviour of one summit of the rough surface under loading has to be introduced. Such a method leads to a contact law for normal contact of the form (see Willner and Gaul (1997))

$$p_N(\gamma) = \frac{1}{h^2} \int_0^\infty \int_0^\infty N_i(\gamma) P(\zeta_\sigma, \kappa_\sigma) d\kappa d\zeta. \quad (4.7)$$

Here  $\zeta_\sigma = z / \sigma_z$  is the normalized asperity height, normalized by the rms-height  $\sigma_z$ . The height  $z$  is measured on a regular grid with spacing  $h$  (see Greenwood and Williamson (1966)). The curvature  $k_s$  stems then from a finite difference approximation, and yields the mean curvature by  $\kappa_\sigma = k_s / \sigma_k$ , normalized by the rms-curvature  $\sigma_k$ .  $P(\zeta_\sigma, \kappa_\sigma)$  is the probability distribution of a joint.  $N_i(\gamma)$  is the normal contact force related to one summit, with  $\gamma = g_N / \sigma_z$  being the gap function normalized by the rms-height. The contact pressure related to the apparent contact area follows from equation (4.7), when all input data are known from measurements.



**Fig. 4.5** Physical approach in  $\Gamma_c$ , undeformed and deformed situations.

In most applications it is sufficient to formulate the constitutive relation for the apparent contact pressure like

$$p_N = f(d) \quad \text{or} \quad d = h(p_N), \quad (4.8)$$

where  $f$  and  $h$  are nonlinear functions of the current mean plane distance  $d$  or the contact pressure  $p_N$ , respectively.

In the case of contact, the current mean plane distance  $d$  is related to the geometrical approach  $g_N$ , see (3.6), via the relation

$$g_N = \zeta - d. \quad (4.9)$$

$\zeta$  is the initial mean plane distance in the contact area  $\Gamma_c$ , which is defined as the mean plane distance of the two surfaces when the two surfaces just touch each other in the initial configuration, see Figure 4.5.

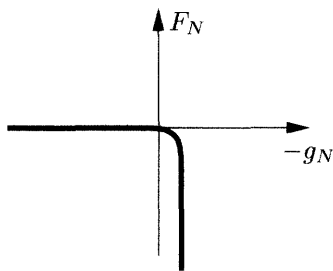
Most of the interface laws can be written in the form (4.8). Out of many different possibilities, two constitutive equations for normal pressure in the contact area will be stated besides the general relation (4.7). The first was developed in Zavarise (1991) and Zavarise et al. (1992a), and is based on a statistical model of the micro-geometry proposed by Song and Yovanovich (1987). This constitutive relation for the approach of both surfaces yields an exponential law of the form

$$p_N = c_3 e^{-c_4 d^2}. \quad (4.10)$$

A more detailed description of this interface law shows the dependency of  $c_3$  and  $c_4$  on the geometrical and material parameters of the surface:

$$p_N = \frac{c_1 \left(1617646.152 \frac{\sigma}{m}\right)^{c_2}}{5.589^{1+0.0711 c_2}} \exp \left[ -\frac{1 + 0.0711 c_2}{(1.363\sigma)^2} d^2 \right]. \quad (4.11)$$

$c_1$  and  $c_2$  are mechanical constants which express the nonlinear distribution of the surface hardness,  $\sigma$  and  $m$  are statistical parameters of the surface profile, representing the RMS surface roughness and the mean absolute asperity slope. As can be seen, many constitutive parameters are needed to describe the approach of two rough surfaces within the contact area. Thus, these models are only used when really needed,



**Fig. 4.6** Contact compliance law.

like in thermo-mechanical contact where the approach enters the constitutive relation for the heat flux, see Section 10.2.

Another law for the contact pressure has been given, based on experimental investigations, by Kragelsky et al. (1982). These authors formulated the following constitutive equation for the contact pressure:

$$p_N = c_N (g_N)^n \quad (4.12)$$

in terms of the gap function  $g_N$  which is defined in (3.6).  $c_N$  and  $n$  are constitutive parameters which have to be determined by experiments. For metals  $n$  is usually in the range  $2 \leq n \leq 3.33$ . Thus the simplest possibility with  $n = 1$  is excluded, which would coincide with a standard penalty method, see Section 5.3.2.

Both interface laws could be viewed as nonlinear springs (see Figure 4.6) which yield a regularization of the HERTZ–SIGNORINI–MOREAU conditions in (4.2).

Often, elasto-plastic deformation has to be considered in a realistic micromechanical model. Hence inelastic behaviour can be applied in the derivation of the statistical models discussed above. Thereafter, a potential from which the contact pressure follows via partial derivation with respect to the normal gap cannot be constructed.

The constitutive equations (4.10) and (4.12) represent very high stiffness, since the approach of both surfaces is very small and limited once all asperities are flattened. This fact leads to ill-conditioning when numerical solutions have to be obtained with (4.10) or (4.12), e.g. by the finite element method. A way in which one can treat such models numerically is described in Section 5.3.8.

## 4.2 TANGENTIAL CONTACT

The interfacial behaviour related to frictional response is very important in most technical applications. The science related to this topic is called *tribology*, which is mainly concerned with technical systems (e.g. breaks or bearings) where friction is present. Tribology covers topics like adhesion, friction, wear, lubrication, thermal contact or electric contact. Also from the economical point of view, tribology is an important science, hence it can be applied to estimate the time until a mechanical part



looses its usefulness. Examples for such processes are the wear of oil well drills, the wear of artificial joints in biomechanics, the wear of metal cutting machines or the wear of railroad tracks.

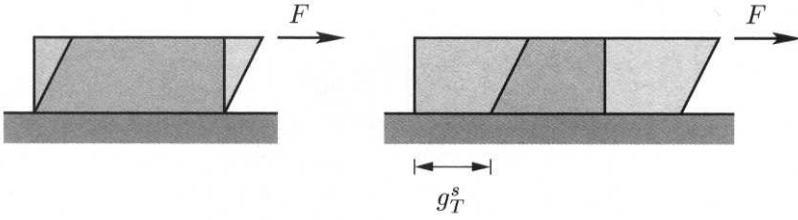
Sometimes it is desirable to maximize friction such as for the contact between a car tyre and the road during braking, or for the foundations of buildings when wind forces are acting. However, often it is important to reduce friction such as during skiing or ice-skating. Since friction is directly linked to wear, it is necessary to decrease friction in production processes, in e.g. the heads of computer disks, in bearings, in all moving parts of engines, etc.

Despite the fact that friction has been investigated for a long time, starting with LEONARDO DA VINCI, AMONTONS and COULOMB, many frictional phenomena have to-date not been completely understood. This is due to the fact that the frictional behaviour on the surface takes place at the atomic level with, e.g. an interaction of chemical, electro-magnetic and mechanical processes. Thus, some researchers tried to formulate a third body in the interface which has a special properties, and is only present at the moment of the tangential mechanical loading, e.g. see Kragelsky (1956). Also, the structure of surfaces that are in contact is complicated, since a typical metallic surface consists of several layers which influence friction. In general, the friction coefficient depends upon the normal pressure, the relative tangential velocity, the surface roughness and the temperature, to name but a few of the most relevant parameters. For a recent extensive overview regarding the physical understanding of friction see Persson (2000).

We will restrict ourselves to more simple formulations which yield constitutive equations for frictional contact for dry friction. Lubrication which is essential in many technical applications since it reduces friction and wear will not be considered here in detail, however a short introduction is given in section 4.3.

The most frequently used constitutive equation is the classical law of COULOMB. However, other frictional laws are available which take into account local, micro-mechanical phenomena within the contact interface, e.g. see Woo and Thomas (1980). An overview with relation to numerical modelling may be found in Oden and Martins (1986).

The main governing phenomena for dry friction are adhesion of the surfaces and ploughing of the asperities. For the physical background, see Tabor (1981). During the last few years, frictional phenomena have also been considered within the framework of the theory of plasticity. This leads to non-associative slip rules, for which different relations have been proposed by for instance Bowden and Tabor (1964) and Michalowski and Mroz (1978). Further discussion on the theory of friction is contained in Curnier (1984). Laws which investigate the non-local character of friction can be found in Oden and Pires (1983a) and Oden and Pires (1983b). Constitutive equations for friction with respect to metal forming have a long history, and are discussed in Tabor (1981); modern treatments with respect to micromechanical behaviour are present in, for example Anand (1993), Stupkiewicz and Mroz (1999) or Stupkiewicz (2001). At the contact interface the response in the tangential direction can be divided into two different actions. In the first, no tangential relative displacement occurs in the contact zone under a loading due to, for example a force  $F$ . This



**Fig. 4.7** Stick or slip in the contact area.

behaviour is called stick (see the left part of Figure 4.7). The second action due to the force  $F$  is associated with a relative tangential movement  $g_T^s$  in the contact interface, which denotes the so-called slip (see the right part of Figure 4.7).

#### 4.2.1 Stick as a constraint

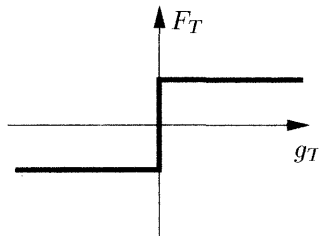
Stick is equivalent to the case in which the relative tangential velocity is zero. Hence, the stick condition can be obtained from (3.22) as

$$\dot{\mathbf{g}}_T = \mathbf{0} \Leftrightarrow \mathbf{g}_T = \mathbf{0}. \quad (4.13)$$

This condition is formulated in the current configuration, and thus in general imposes a nonlinear constraint equation on the motion in the contact interface. Associated with this constraint is a LAGRANGE multiplier,  $\lambda_T$ , which denotes the reaction due to the constraint (4.13). There exist many possibilities to enforce the constraint condition (4.13). These will be discussed in detail in Section 5.3.

#### 4.2.2 Coulomb law

Once the tangential forces are above a certain limit (see Figure 4.8) then the contacting surfaces no longer stick to each other, but move relative to each other. This relative tangential movement is called *sliding*, and classically is described by the law of



**Fig. 4.8** COULOMB'S friction law.

COULOMB. It takes the form

$$\mathbf{t}_T = -\mu |p_N| \frac{\dot{\mathbf{g}}_T}{\|\dot{\mathbf{g}}_T\|} \quad \text{if } \|\mathbf{t}_T\| > \mu |p_N|, \quad (4.14)$$

where  $\mu$  is the sliding friction coefficient. This coefficient is constant in the classical COULOMB law. It depends upon the material pairing of the solids in contact. Table 4.1 contains some values for different materials.

It can be seen from the bandwidth of the friction coefficient for a specific material pairing that this coefficient must be influenced by other physical and geometrical circumstances. In general, the friction coefficient depends upon different parameters like the surface roughness, the relative sliding velocity  $\dot{\mathbf{g}}_T$  between the contacting bodies, the contact normal pressure  $p_N$  or the temperature  $\theta$ , e.g. see Section 4.2.7. If such effects are introduced, one obtains a variant of the COULOMB law with a variable coefficient  $\mu = \mu(\dot{\mathbf{g}}_T, p_N, \theta)$ .

One such heuristic friction law which incorporates the relative sliding velocity  $\dot{\mathbf{g}}_T$  in the expression for the friction coefficient is given by

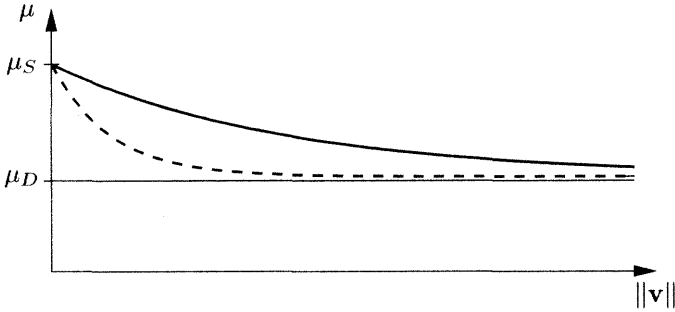
$$\mu(\dot{\mathbf{g}}_T) = \mu_D + (\mu_S - \mu_D) e^{-c \|\dot{\mathbf{g}}_T\|}. \quad (4.15)$$

It depends upon three constitutive parameters  $\mu_S$ ,  $\mu_D$  and  $c$ . Here for zero sliding velocity, the friction coefficient assumes the static value  $\mu_S$ . For large velocities the dynamic friction coefficient  $\mu_D$  is approached (see also Figure 4.9). The additional constitutive parameter  $c$  describes how fast the static coefficient approaches the dynamic one, see the two different curves in Figure 4.9, which stem from different values of  $c$ .

For many applications in which the surface roughness is not too large or too smooth, the friction coefficient is independent from the surface roughness and hence from the

**Table 4.1** Friction coefficient for different material pairings.

Material pairing	Friction coefficient $\mu$
concrete–concrete	0.5 — 1.0
concrete–sand	0.35 — 0.6
concrete–steel	0.2 — 0.4
metal–wood	0.3 — 0.65
rubber–steel	0.15 — 0.65
steel–steel	0.2 — 0.8
steel–teflon	0.04 — 0.06
steel–concrete	0.2 — 0.4
steel–ice	0.015 — 0.03
wood–steel	0.5 — 1.2
wood–wood	0.4 — 1.0



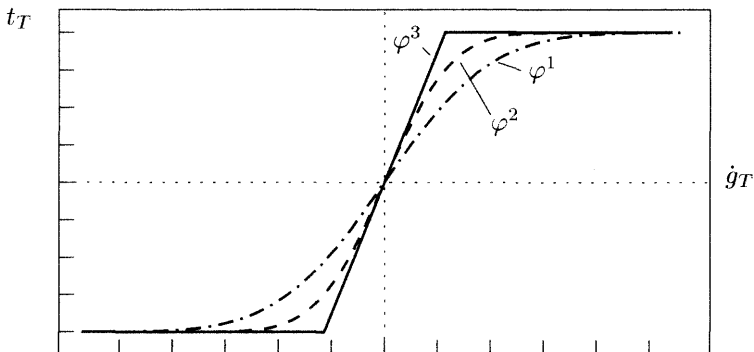
**Fig. 4.9** Velocity-dependent COULOMB'S friction law.

real contact area in the interface (see below). If, furthermore, also the sliding velocity is neither too large nor too small, then the friction force is proportional to the normal force, leading to a constant friction coefficient  $\mu$ .

#### 4.2.3 Regularization of the Coulomb law

Another possibility to formulate tangential constitutive equations in the contact interface is given by a regularization of the stick-slip behaviour. Such a formulation is used to avoid the non-differentiability of COULOMB'S law at the onset of sliding, see Figure 4.8. Associated models were introduced by Oden and Pires (1983b); see also Raous (1999). These are based on a functional form which yields a smooth transition from stick to slip, as depicted in Figure 4.10. The explicit forms are given for two-dimensional problems by

$$t_T = -\mu \varphi^i(\dot{g}_T) |p_N|, \quad (4.16)$$



**Fig. 4.10** Regularization of COULOMB'S friction law.

where the function  $\varphi^i$  describing the smooth transition from sticking to sliding is defined by:

1. a square root regularization

$$\varphi^1 = \frac{\dot{g}_T}{\sqrt{\dot{g}_T^2 + \varepsilon^2}}, \quad (4.17)$$

2. a hyperbolic tangent regularization

$$\varphi^2 = \tanh \left( \frac{\dot{g}_T}{\varepsilon} \right), \quad (4.18)$$

3. or a piecewise polynomial regularization

$$\varphi^3 = \begin{cases} -1 & , \text{if } \dot{g}_T < -\varepsilon \\ \frac{\dot{g}_T}{2\varepsilon} & , \text{if } -\varepsilon \leq \dot{g}_T \leq \varepsilon \\ 1 & , \text{if } \dot{g}_T > \varepsilon \end{cases}. \quad (4.19)$$

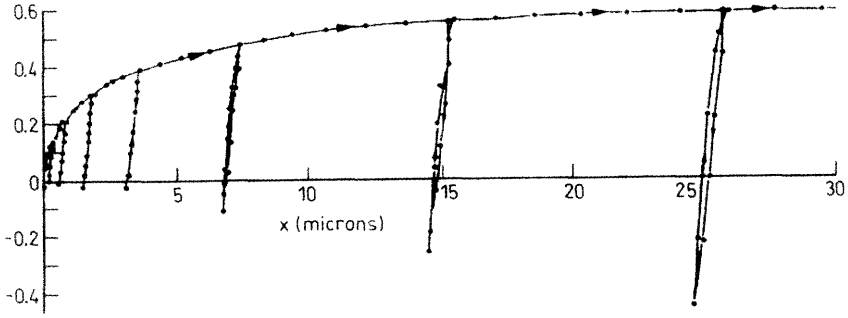
The scalar parameter  $\varepsilon$  denotes in all cases, the regularization variable, which for  $\varepsilon \rightarrow 0$  as the limit case yields the classical COULOMB law.

These regularized constitutive interface laws physically have the drawback that they only describe the stick-slip motion in an approximate fashion. For a value of the parameter  $\varepsilon$  that is too large, such a model might not be able to predict real stick-slip motions. On the other hand, due to the differentiability of formulations (1) and (2), these models lead to numerical algorithms which are simpler and more robust. The piecewise polynomial regularization can be used to construct an algorithm for stick-slip motion which is similar to well known algorithms from the theory of plasticity (see also the next section).

#### 4.2.4 Elasto-plastic analogy for friction

COULOMB'S law and other constitutive equations for friction can also be formulated in the framework of elasto-plasticity. This has been investigated by several authors who developed different constitutive equations for frictional problems, e.g. see Michalowski and Mroz (1978) or Curnier (1984). A treatment of frictional interface laws in terms of non-associated plasticity has been considered within a finite element formulation by Wriggers (1987), Giannokopoulos (1989), Wriggers et al. (1990) or Laursen and Simo (1993b).

The reason to formulate elasto-plastic constitutive relations in the contact interface stems from two sources. One is associated with the wish to regularize COULOMB'S law. The other reason has its source in experimental observations, as shown in Figure 4.11. These experiments were carried out by Courtney-Pratt and Eisner (1957) with metallic surfaces (platinum-to-platinum), and suggest the use of classical elasto-plastic relations which split the tangential motion into a elastic or adhesive (stick) part



**Fig. 4.11** Friction coefficient versus tangential movement; experiments from Courtney-Pratt and Eisner (1957).

and a plastic or slip part. Experiments for steel and copper are discussed in Anand (1993) which depict the same behaviour.

As depicted in Figure 4.12, in general one can assume softening (1) as well as hardening (2) for the constitutive behaviour of friction, once a threshold value  $\hat{t}_T$  is passed. This is also in accordance with experiments (e.g. see Figure 4.11 for hardening) or it stems from the fact that the sliding coefficient of friction is lower than the coefficient describing stick, which results in softening.

The key idea of the elasto-plastic approach is a split of the tangential slip  $\mathbf{g}_T$  into an elastic (stick or adhesive) part  $\mathbf{g}_T^e$  and a plastic (slip) part  $\mathbf{g}_T^s$ , see Figure 4.12:

$$\boxed{\mathbf{g}_T = \mathbf{g}_T^e + \mathbf{g}_T^s} . \quad (4.20)$$

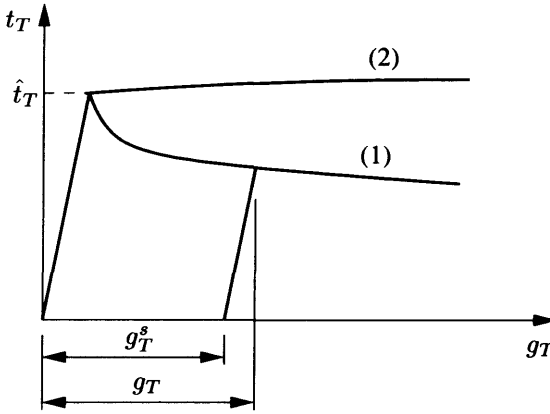
This split can be viewed as a regularization step of the frictional law. However, one also can find a physical interpretation, in the sense that elastic tangential micro displacements  $\mathbf{g}_T^e$  occur in the contact interface. These stem from the elastic behaviour of the asperities in the real contact area. Hence, they model an elastic stick behaviour since the associated deformations vanish once the loading is removed from the system. The simplest possible model for the elastic part of the tangential contact is an isotropic linear elastic relation, which yields

$$\mathbf{t}_T = c_T \mathbf{g}_T^e , \quad (4.21)$$

where  $c_T$  is the elastic constant. However, one can also think of anisotropic elastic constitutive behaviour leading to

$$\mathbf{t}_T = \mathbf{C}_T \mathbf{g}_T^e , \quad (4.22)$$

with the constitutive tensor  $\mathbf{C}_T$ . The directions of anisotropy have to be determined from experiments by taking into account the micro-structure of the contact surfaces, see Figure 4.4.



**Fig. 4.12** Kinematical split of slip and elastic stick.

The plastic frictional tangential slip  $g_T^s$  is governed by a constitutive evolution equation which can be formally derived by using standard concepts of elasto-plasticity theory. Let us introduce for the derivation of a slip rule, in analogy to the derivation of a plastic flow rule, the dissipation

$$\mathcal{D}^s = \mathbf{t}_T \cdot \dot{\mathbf{g}}_T^s \geq 0 \quad (4.23)$$

due to the plastic slip. Now consider an elastic domain

$$\mathbb{E}_t = \{ \mathbf{t}_T \in \mathbb{R}^2 \mid f_s(\mathbf{t}_T) \leq 0 \} \quad (4.24)$$

in the space of the contact tangential stress which is bounded by the function  $\hat{f}_s$ . A simple expression for  $f_s$  in (4.24) is given by

$$f_s(\mathbf{t}_T) = \|\mathbf{t}_T\| - \mu p_N \leq 0, \quad (4.25)$$

which is the plastic slip criterion (see figure 4.13(a) for a given contact pressure  $p_N$ , and is equivalent to classical COULOMB'S law. The frictional coefficient  $\mu$  is a material parameter.

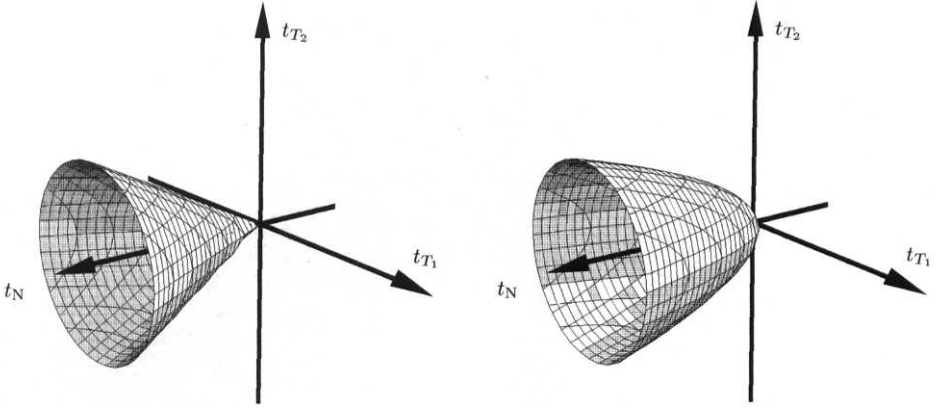
The evolution equation for the plastic slip can now be computed from the maximum dissipation principle, well known from the theory of plasticity, e.g. see Lubliner (1985). The derivation is done here assuming the simple slip criterion (4.25), however the methods also hold for more advanced slip criteria discussed below.

Holding  $p_N$  fixed, from the so-called maximum dissipation principle one obtains

$$(\mathbf{t}_T - \mathbf{t}_T^*) \cdot \dot{\mathbf{g}}_T^s \geq 0 \quad \forall \mathbf{t}_T^* \in \mathbb{E}_t \quad (4.26)$$

the constitutive evolution equation for the plastic slip

$$\boxed{\dot{\mathbf{g}}_T^s = \dot{\gamma} \frac{\partial f_s(\mathbf{t}_T)}{\partial \mathbf{t}_T} = \dot{\gamma} \mathbf{n}_T} \quad \text{with} \quad \mathbf{n}_T = \frac{\mathbf{t}_T}{\|\mathbf{t}_T\|}, \quad (4.27)$$



**Fig. 4.13** (a) Coulomb frictional cone (b) more general slip surface.

which denotes the normality rule for fixed contact pressure. The left-hand side of (4.27) denotes the relative tangential velocity due to slip. Additionally, we have the loading-unloading conditions in KUHN-TUCKER form:

$$\dot{\gamma} \geq 0, \quad f_s(\mathbf{t}_T) \leq 0, \quad \dot{\gamma} f_s(\mathbf{t}_T) = 0, \quad (4.28)$$

which determine the plastic parameter  $\dot{\gamma}$ .

Another slip criterion has been formulated in Wriggers et al. (1990) which additionally takes into account the pressure dependency of the tangential response. Here the form  $\mu = \tau_0 / p_r + \beta$  proposed by Tabor (1981) for the frictional parameter is assumed, where  $\tau_0$  and  $\beta$  are constitutive parameters describing a model with linear varying shear strength of the interfacial material due to the true contact pressure. The true pressure  $p_r$  is related to the true contact area  $A_r = \sum_i A_{ri}$  (real contact area due to the contact of the asperities in the contact interface) whereas the pressure  $p_N$  is associated with the nominal contact area  $A_a$ , see Figure 4.14.

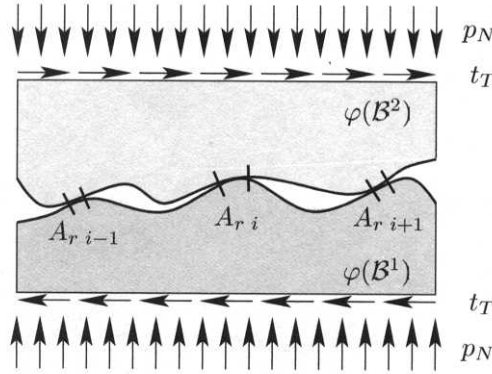
Woo and Thomas (1980) have formulated a relation for metals between the true and the nominal area based on experimental observations:

$$\frac{A_r}{A_a} = \left( \frac{|p_N|}{A_a H} \right)^n, \quad n = \frac{5}{6}, \quad (4.29)$$

with the hardness  $H$  of the material. With these relations one arrives at the following slip criterion (e.g. see Figure 4.13(b)):

$$f_s(\mathbf{t}_T, p_N) = \|\mathbf{t}_T\| - \alpha |p_N|^n - \beta |p_N| \leq 0, \quad \alpha = \frac{A_a \tau_0}{(A_a H)^n}. \quad (4.30)$$





**Fig. 4.14** Contact interface, micromechanical view.

Note that the choice of one of the slip criteria (4.25) or (4.30) has to be made with regard to experimental data within the contact interface; of course, other slip criteria are possible.

Again we assume that the tangential plastic slip  $\mathbf{g}_T^s$  is governed by a constitutive evolution equation which can be derived using standard concepts of the theory of elasto-plasticity. Within this framework we can formulate a plastic slip criterion

$$f_s(\mathbf{t}_T, p_N, \theta, g_v) \leq 0 \quad (4.31)$$

for a given contact pressure  $p_N$  with material parameter  $\mu$ , a given temperature and a hardening/softening variable like the effective slip  $g_v$ . This slip criterion function can be specialized as follows:

$$f_s(\mathbf{t}_T, p_N, \theta, g_v) = \|\mathbf{t}_T\| - \hat{g}_s(p_N, \theta, g_v) \leq 0, \quad (4.32)$$

with the special case of classical COULOMB'S model, see (4.25).

Inequality (4.32) can also be expressed in terms of the nominal stresses in the contact interface. This results, with (4.6), in

$$f_s(\mathbf{T}_T, T_N, \theta, g_v) = \|\mathbf{T}_T\| - j \hat{g}_s(j^{-1} T_N, \theta, g_v) \leq 0. \quad (4.33)$$

It can be seen that when function  $\hat{g}_s$  is a homogeneous function of its argument, as in COULOMB'S law (4.25), then  $\hat{g}_s = \hat{G}_s$ . Thus for COULOMB one can write

$$f_s(\mathbf{t}_T, p_N) = \|\mathbf{t}_T\| - \mu p_N \leq 0 \Leftrightarrow F_s(\mathbf{T}_T, T_N) = \|\mathbf{T}_T\| - \mu T_N \leq 0. \quad (4.34)$$

In some applications (see Section 4.2.5) it makes sense to introduce constitutive equations for friction stresses which depend upon the deformation of the contacting bodies, and hence on the area JACOBIAN  $j$  defined in (4.4).

In all formulations derived so far, the constitutive evolution equation for the plastic or frictional slip can be stated in a general form of a slip rule for large deformations

in the contact zone as follows:

$$\begin{aligned}\dot{\mathbf{g}}_T^s &= \dot{\gamma} \frac{\partial f_s}{\partial \mathbf{t}_T} = \dot{\gamma} \mathbf{n}_T, & \text{with } \mathbf{n}_T &= \frac{\mathbf{t}_T}{\|\mathbf{t}_T\|}, \\ \dot{g}_v &= \dot{\gamma},\end{aligned}$$

where equation (4.35) describes the evolution of the effective slip, which is defined as

$$g_v = \int_0^t \|\dot{\mathbf{g}}_T(\tau)\| d\tau. \quad (4.35)$$

$\dot{\gamma}$  is a parameter which describes the magnitude of the plastic slip. Also, here we can set up the KUHN–TUCKER conditions, as before, see (4.28).

A generalization of friction models to non-isotropic behaviour is sometimes necessary if the surface in the interface has a special texture. Models for anisotropic friction can be found in Mroz and Stupkiewicz (1992) or He and Curnier (1993). Mroz and Stupkiewicz (1992) introduced a model which is an extension of the equations derived to-date. To simplify the notation they introduced an orthogonal cartesian reference frame in the tangent plane of the contact interface. With respect to that, it is now possible to describe the slip function by a super ellipse:

$$f_s(p_N, t_{Tx}, t_{Ty}) = \left( \left| \frac{t_{Tx}}{\mu_x} \right|^n + \left| \frac{t_{Ty}}{\mu_y} \right|^n \right)^{1/n} - p_N \leq 0, \quad (4.36)$$

where  $t_{Tx}$  and  $t_{Ty}$  denote the tangential stresses in the  $x$ - and  $y$ -directions and  $n$  is a parameter which generates the form of the ellipse. Note that this model assumes a special case of anisotropic behaviour which has different friction parameters with respect to the cartesian co-ordinate axes.

According to (4.35), the evolution equation is now given by two equations for the  $x$ - and  $y$ -directions,

$$\dot{g}_{Tx}^s = \dot{\gamma} \frac{\partial f_s}{\partial t_{Tx}}, \quad \dot{g}_{Ty}^s = \dot{\gamma} \frac{\partial f_s}{\partial t_{Ty}}. \quad (4.37)$$

Again, the KUHN–TUCKER conditions can be set up to complete the formulation:

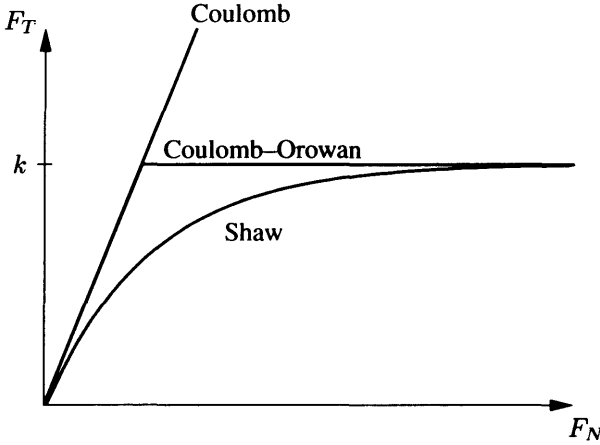
$$\dot{\gamma} \geq 0, \quad f_s(t_{Tx}, t_{Ty}) \leq 0, \quad \dot{\gamma} f_s(t_{Tx}, t_{Ty}) = 0. \quad (4.38)$$

The dissipation function, which is needed for instance in thermo-mechanical applications or problems where wear has to be computed, can also be stated; it yields

$$\mathcal{D}^s = \mathbf{t}_T \cdot \dot{\mathbf{g}}_T^s = t_{Tx} \dot{g}_{Tx}^s + t_{Ty} \dot{g}_{Ty}^s. \quad (4.39)$$

#### 4.2.5 Friction laws for metal forming

In case of metal forming, especially in bulk-forming, the use of COULOMB's friction law is very limited. This is because the prediction of the frictional stress related to



**Fig. 4.15** Different possibilities to limit the frictional stress.

sliding is much too high when applying this law. Hence, in most practical applications, a threshold value is introduced to limit the friction stress, see Figure 4.15.

The simplest approach to account for a limit stress is to change the slip function (4.25) to

$$f_s(\mathbf{t}_T) = \|\mathbf{t}_T\| - h \leq 0, \quad (4.40)$$

where  $h$  is a function which can be described in different ways. Two possible choices for function  $h$  which are relatively simple are stated in the following:

1. COULOMB-OROWAN law:

$$h = \min(\mu |p_N|, Y_0), \quad (4.41)$$

where the constitutive parameter  $Y_0$  denotes the elastic limit of the material, as can be seen in Figure 4.15.

2. SHAW law:

$$h = \beta Y_0 \quad \text{with} \quad \beta = A_r / A_a, \quad (4.42)$$

where  $A_r$  is the real and  $A_a$  the nominal contact area of the contact surface. This quotient can be computed from (4.29), for example. Thus the parameter  $\beta$  describes the flattening of the asperities depending on the normal pressure. The law of SHAW, compared to the COULOMB-OROWAN law, leads to a smooth function between the normal pressure and tangential stress. Hence, it is better suited for numerical purposes.

In the literature, other nonlinear relations are also introduced to describe a smooth transition from COULOMB's law to a constant limiting tangential stress for high contact pressures depending on the sliding stress and normal pressure. Constitutive

equations which include this feature can be derived from micro-mechanical models (e.g. see Wanheim et al. (1974), Avitzur and Nakamura (1986) or Anand (1993)). A commonly used model (see also Stupkiewicz (2001)) is provided by the relation

$$\mathbf{t}_T = \alpha Y_0 \tanh \left[ \frac{-\mu p_N}{\alpha Y_0} \right]. \quad (4.43)$$

Additionally to the friction coefficient  $\mu$  for low contact pressures, already introduced in (4.25), two more parameters appear.  $Y_0$  is the shear yield stress of the workpiece, and  $\alpha$  can be regarded as another friction coefficient which is valid in the high pressure range.

Furthermore, there are many applications in metal forming where a coating is used to improve the sliding conditions in the contact interface. This results in a decreased frictional resistance, and hence less work is needed to form a workpiece. However, it can be that the forming process is such that the surface of the workpiece can expand. Thereafter the coating can break, resulting in a change of the friction behaviour which cannot be neglected. In this case it is useful to assume that the friction coefficient  $\mu$  depends on the change of area in the contact interface  $\mu = \mu(j)$ , where  $j$  is the JACOBIAN connecting the area elements  $da = j dA$ . The formula is equivalent to equation (2.12), which describes the transformation of volume elements. Using NANSON's formula (2.11), one obtains

$$da = j dA \quad \text{with} \quad j = J \|\mathbf{F}^{-T} \mathbf{N}\|. \quad (4.44)$$

Now, as shown in Stupkiewicz (2001), one can formulate a relation for COULOMB's friction law as follows:

$$\begin{aligned} \mathbf{t}_T &= -\mu(j) |p_N| \frac{\dot{\mathbf{g}}_T}{\|\dot{\mathbf{g}}_T\|} \quad \text{with} \\ \mu(j) &= \frac{1}{2} (\mu_1 + \mu_2) + \frac{1}{2} (\mu_2 - \mu_1) \tanh \left[ \frac{j - j_0}{\bar{j}} \right], \end{aligned} \quad (4.45)$$

where now four constitutive parameters are introduced. These are the friction coefficients  $\mu_1$  for coating and  $\mu_2$  for broken coating, with  $\mu_2 > \mu_1$ , and the critical surface extension  $j_0$ , as well as the width of the transition zone  $\bar{j}$ . Further, more advanced models can be found in Stupkiewicz and Mroz (1999).

#### 4.2.6 Friction laws for rubber and polymers

Many technical components include rubber or polymer sealings. Also, rolling contact of car tyres on roads or rubber coated cylinders in printing machines have a need for the analysis of constitutive equations for friction. Since the behaviour of rubber and polymers is different from that of metals, different friction laws have to be considered. In general the micro-mechanical mechanism is different; a concise statement of the interface physics known so far for rubber materials can be found in Persson (2000).

Here only the main effects are considered, leading to a constitutive model which yields a frictional coefficient depending on the sliding velocity and the normal pressure.

In polymers the internal friction stemming from the visco-elastic behaviour of the solids contributes to the sliding friction. Hence the frictional coefficient depends upon the frequency  $\omega \sim v / r$ , where  $v$  is the sliding velocity and  $r$  denotes the diameter of the real contact area. When  $\omega \approx 2 / \tau$ , where  $\tau$  is the relaxation time of the polymer, then the friction coefficient is maximal. Thus one can design the polymer such that one achieves maximum friction by choosing the material so that its relaxation time is  $\tau = 2 / \omega$ , or for small internal friction one has to use a material whose relaxation time is  $\tau \gg 2 / \omega$  or  $\tau \ll 2 / \omega$ .

Besides the internal friction, adhesion is also a main contributor to the sliding friction of soft materials like rubber. However, this effect depends strongly upon the surface properties. If the surface is not completely clean, as in most technical applications, then this effect is reduced. For other polymers one observes that for small internal friction the sliding process takes place in a thin contamination layer. However, for large sliding velocities, melting of the polymer occurs at the surface due to low thermal conductivity and a low melting point of the polymers. This leads to a considerable decrease of the friction coefficient with increasing wear.

All considerations result in complex constitutive relations if one wants to take into account all of the effects. For practical purposes, the following forms for the friction coefficient are applied.

A velocity- and temperature-dependent form for the frictional coefficient

$$\mu(\theta, v_T) = \mu_0 + c_1 \theta [\ln v_T - \ln(c_2 \theta)] \quad (4.46)$$

was proposed in Rieger (1968). This constitutive model for the friction of rubber depends upon the sliding velocity  $v_T$  and the temperature  $\theta$ . It has three parameters which have to be determined by experiments. If the temperature can be neglected, such as when only small sliding velocities occur, then the constitutive relation

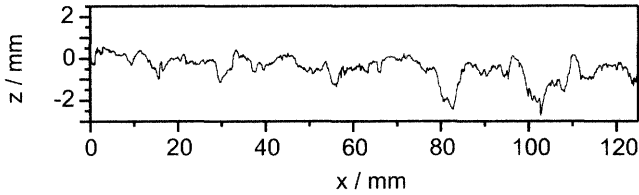
$$\mu(v_T) = \mu_0(p_N) + c_1 \ln \frac{v_T}{v_1} - c_2 \ln \frac{v_T}{v_2}, \quad (4.47)$$

which can be found in Nackenhurst (2000), can be used for a range of the sliding velocity from  $10^{-3}$  to 10 m/s. In this model the parameter  $\mu_0$ , depends upon the normal pressure  $p_N$ ,

$$\mu_0(p_N) = \mu_1 \left[ \frac{p_N}{p_0} \right]^\alpha. \quad (4.48)$$

Thus, this model needs seven parameters deduced from experiments.

Other frictional interface laws for contact between a road and a tyre were developed in Klüppel and Heinrich (2000). These authors base their model on the assumptions that the adhesion in the interface depends upon the surface free energy of the bulk rubber and the hysteretic energy loss due to the deformation of rubber by surface asperities. The latter effect is assumed to be the dominant factor of rubber friction (see also Persson (2000)). To understand the frictional dynamics of rubber stemming from stochastic excitations during sliding over multiple scale surfaces, the authors introduce self-affine surfaces and a visco-elastic model for the behaviour of the rubber.



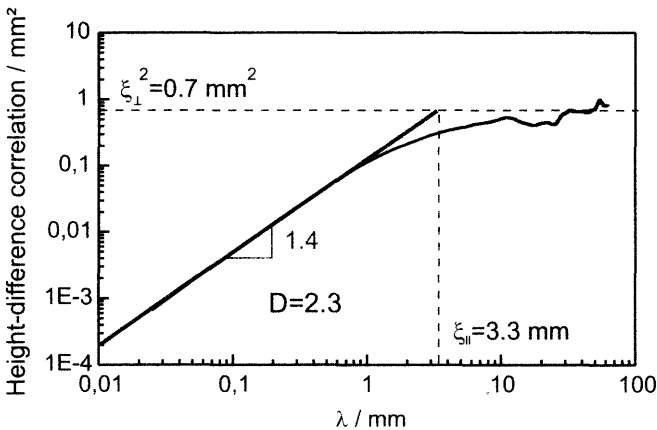
**Fig. 4.16** Profile of an asphalt road track from laser measurement, see Klüppel and Heinrich (2000).

The self-affine surfaces are described by

$$S_z(\lambda) = \left[ \frac{\lambda}{\xi_{\parallel}} \right]^{2H} \xi_{\perp}^2 \quad \text{for } \lambda < \xi_{\parallel}, \quad (4.49)$$

where  $\xi_{\parallel}$  and  $\xi_{\perp}$  are the normal and tangential correlation lengths of the rough surface. They denote the maximum length scales for which self-affinity is fulfilled.  $H = 3 - D$  is the HURST coefficient with the fractal dimension  $D$ . With (4.49) the surface is described by the set  $D$ ,  $\xi_{\parallel}$  and  $\xi_{\perp}$ . The values for the correlation lengths can be obtained from measurements, see Figures 4.16 and 4.17.

This, together with the visco-elastic model, results in a friction coefficient which depends upon the sliding velocity  $v_T$  and the real and imaginary elastic moduli,  $E'$



**Fig. 4.17** Resulting parameters  $D$ ,  $\xi_{\parallel}$  and  $\xi_{\perp}$  describing the surface, see Klüppel and Heinrich (2000).

and  $E^{\parallel}$ :

$$\mu(v_T) = \frac{1}{4} \left[ \frac{\xi_{\perp}}{\xi_{\parallel}} \right]^2 \frac{E^{\parallel}}{E^{\perp}} \arctan \left[ \frac{(v_1 - v_2) v_T}{v_T^2 + v_1 v_2} \right], \quad (4.50)$$

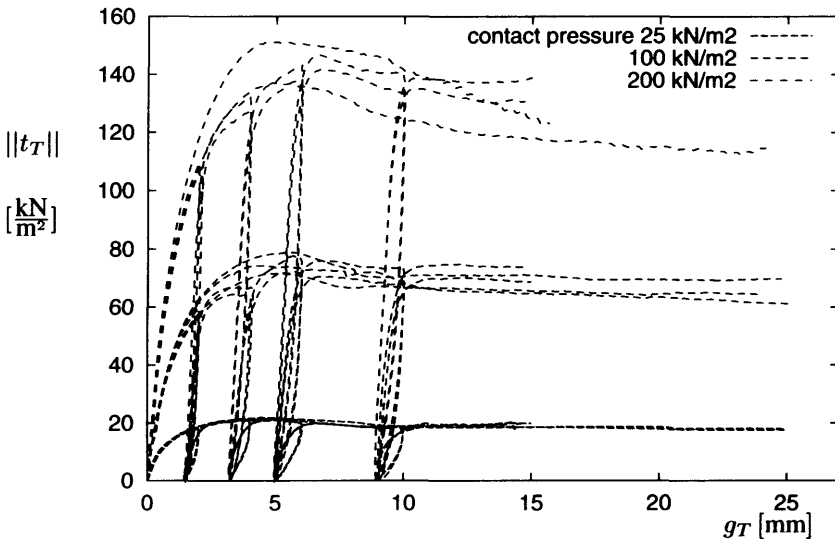
where the constants  $v_1$  and  $v_2$  depend upon the relaxation time and the correlation lengths. For more details, and the implications of the formulation with respect to the choice of fillers in the rubber material for the generation of specific friction coefficients, see Klüppel and Heinrich (2000).

#### 4.2.7 Friction laws for concrete structures on soil

In this section two possible friction laws which describe the interface behaviour of a concrete structure on soil are discussed. With regard to the soil material, we restrict ourselves to sand. The phenomenological frictional laws are developed based on results of experimental shear tests, see Figure 4.18. The plot shows a dependency of the frictional force on the relative tangential movement in the interface in a non-linear way. After a first elastic response, one observes elasto-plastic behaviour with softening. Also, hysteretic effects are present in the loading/unloading phases of the experiment.

The first material law to describe frictional behaviour between sand and concrete is developed following a model which was constructed for micro-sliding between metals, see Sellgren and Olofsson (1999). This derivation is based on the observation depicted in Figure 4.11 for metals which show a similar material behaviour.

No elasto-plastic approach (see section 4.2.4) is used. Here the frictional coefficient is introduced as a function of sliding distance and contact pressure. It thus



**Fig. 4.18** Results of shear tests, see Reul (2000).

changes throughout the motion. At each change of direction of motion the parameter of the frictional coefficient function is actualized, so that the friction law becomes *history dependent*. The following constitutive relation is considered for the tangential stresses:

$$\mathbf{t}_T = \hat{t}_N^{a_6} \mu^* \frac{\tilde{\mathbf{g}}_T}{\|\tilde{\mathbf{g}}_T\|}, \quad (4.51)$$

with the normalized contact pressure  $\hat{t}_N = \frac{p_N}{p_N^0}$  and the friction coefficient

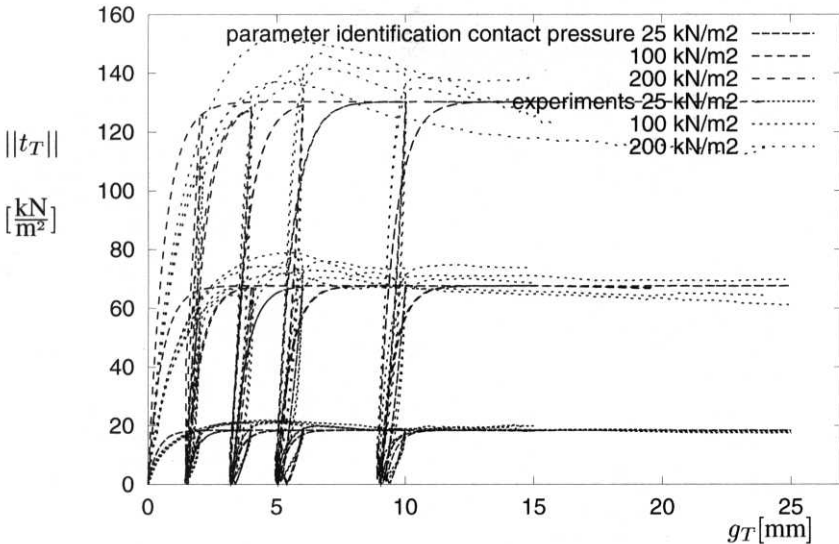
$$\mu^* = \left| -\mu_0 + (a_1 + \mu_0) \left[ 1 - \left( 1 - \frac{a_2 \delta}{(a_4 + \mu_0) \hat{t}_N^{a_5}} \right)^{a_3} \right] \right|. \quad (4.52)$$

The relative tangential displacement  $\tilde{\mathbf{g}}_T$  defines the direction of the frictional stress in (4.51), and  $\delta = \int \|\tilde{\mathbf{g}}_T\| dt$  denotes the accumulated sliding distance during the loading process, which is summed over all time increments  $\Delta t_i$ , leading to  $\delta = \sum_i \|\tilde{\mathbf{g}}_{T_i}\|$ .

At a reversal of the tangential motion the condition  $\tilde{\mathbf{g}}_{T_{n+1}} \cdot \tilde{\mathbf{g}}_{T_n} \leq 0$  holds, where the subscript  $n + 1$  denotes the current and  $n$  the previous time steps. At that stage, the following parameters are changed:

$$\begin{aligned} \delta &= 0 \\ \mu_0 &= \mu^*. \end{aligned} \quad (4.53)$$

In this approach a nonlinear dependency of the contact pressure on frictional stresses is considered in (4.51), which is different from the model proposed by Sellgren and Olofsson (1999). The model works for two- and three-dimensional problems, and



**Fig. 4.19** Fit of friction law (4.51) to experiments.



has six constitutive parameters  $a_1$  to  $a_6$  which have to be determined by experiments. This task has been performed for the experiments shown in Figure 4.18 using the parameter identification procedure described in appendix C. The fit of these data to the experiments yields the following parameters:  $a_1 = 0.891 \text{ kN/m}^2$ ,  $a_2 = 40 \text{ m/kN}$ ,  $a_3 = 95.193$ ,  $a_4 = 0.719 \text{ kN/m}^2$ ,  $a_5 = 0.226$  and  $a_6 = 0.914$ . The result of parameter identification using model (4.51) is depicted in Figure 4.19.

One observes that this model represents the hysteretic effects but not the decrease of the frictional force with respect to increasing relative tangential displacement. However, the dependency on the normal pressure is reproduced correctly.

Another formulation for an interface law between soil and concrete can be derived on the basis of the elasto-plastic approach, which was described in Section 4.2.4. The material law for the elastic depends upon the contact pressure as follows:

$$\mathbf{t}_T = \epsilon_T \hat{t}_N^{a_7} \mathbf{g}_T^e = \epsilon_T \hat{t}_N^{a_7} (\mathbf{g}_T - \mathbf{g}_T^p). \quad (4.54)$$

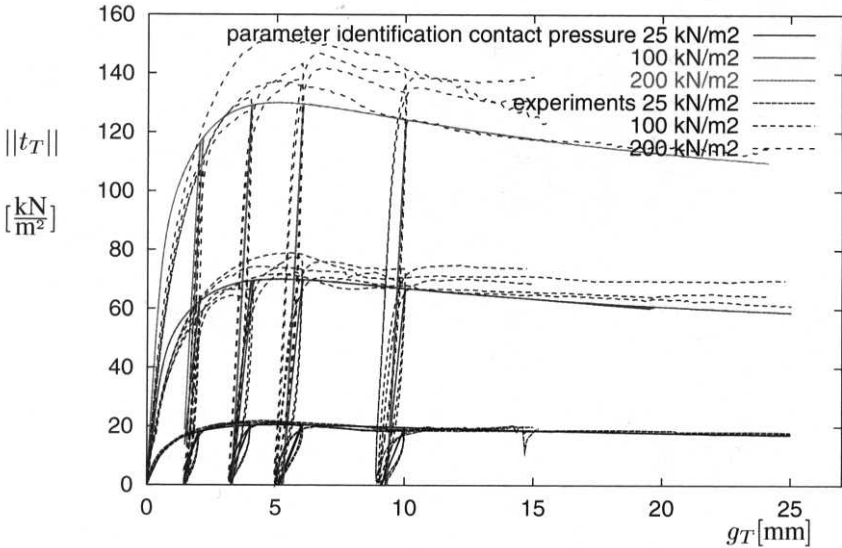
Again  $\hat{t}_N$  is a normalized contact pressure as introduced in (4.51). In this approach, the slip criterion  $f_s$  is described by slip criterion

$$f_s = \|\mathbf{t}_T\| - t_N^{a_6} \left[ \frac{a_1 w}{1 + a_2 w + a_3 w^2} + a_4 \arctan(a_5 w) \right] = \|\mathbf{t}_T\| - t_N^{a_6} \mu(w). \quad (4.55)$$

This formulation is analogous to the introduction of a hardening function which was used to describe the constitutive behaviour of sand, see Arslan (1980).

The slip rule is given by

$$\dot{\mathbf{g}}_T^s = \lambda \frac{\partial f_s}{\partial \mathbf{t}_T} = \lambda \frac{\mathbf{t}_T}{\|\mathbf{t}_T\|}. \quad (4.56)$$



**Fig. 4.20** Fit of friction law (4.54) to experiments.

The evolution equation for hardening is described by

$$\dot{w} = \lambda \quad (4.57)$$

with the hardening parameter  $w$ .

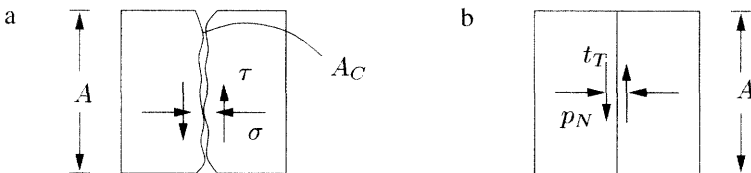
Again, the parameters were fitted for the experiments shown in Figure 4.18 using the parameter identification procedure described in Appendix C. This leads to the following set of parameters:  $a_1 = 16621. \text{ kN} / \text{ m}^3$ ,  $a_2 = 19817.1 / \text{ m}$ ,  $a_3 = 19739311 / \text{ m}^2$ ,  $a_4 = 0.4915 \text{ kN} / \text{ m}^2$ ,  $a_5 = 535.71 / \text{ m}$ ,  $a_6 = 0.89$  and  $a_7 = 0.99$ .

The result of parameter identification using model (4.54) is depicted in Figure 4.20. Now the hysteretic effects are not represented by this model, but the dependency of the frictional stress on the relative tangential displacement  $g_T$ , as well as the dependency on the normal pressure, is described correctly.

For both models the parameter identification procedure yields good results. This is not obvious, since the model is highly nonlinear and history-dependent due to the softening behaviour of the tangential stress with respect to the absolute relative displacement in the interface.

#### 4.2.8 Friction laws from computational homogenization procedures

Another method to derive contact compliances (see section 4.1.2) and friction interface laws is provided by a computational homogenization procedure. In this case a numerical simulation is performed under homogeneous boundary conditions for a Representative Volume Element (RVE). Such an RVE is depicted as a two-dimensional sketch in Figure 4.21 (a). The RVE has to include the real contact geometry with known profiles from the micro-mechanical surfaces. It has to be large enough to resolve the micro-structure of the surface geometries in the contact area. When the boundary conditions are such that only deformations are applied in the normal direction, then the computational homogenization procedure yields a homogenized contact law for the normal pressure  $p_N$ , see Figure 4.21(b). When additionally, relative tangential motions are described, then a friction law for  $t_T$  can be derived. Note that all computations have to be performed using three-dimensional models, since a real micro-geometry is always a two-dimensional surface. To arrive at a statistically representative homogenized constitutive equation for the contact interface, several com-



**Fig. 4.21** (a) Micro-structure, (b) homogenized model.

putations have to be performed with a different distribution of the micro-geometry until the standard deviation is below a given tolerance.

Still, for the numerical model which includes the micro-geometry, see Figure 4.21(a), one has to provide a general contact law for the contact stresses in the normal and tangential directions. In such micro-mechanical computation no friction is assumed here between the solids. The friction law then stems either from a constitutive relation describing adhesion in the contact area, and/or the elastic-plastic response of the solid which is related to ploughing.

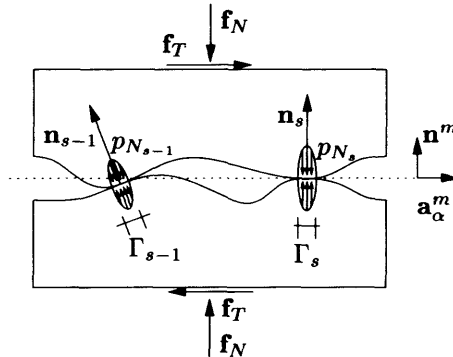
Using a micro-structure in the contact surface, as depicted in Figure 4.21 (a) (see also Figure 4.14) the normal contact stresses on the rough microstructural surface are computed. These stresses only occur in some parts  $\Gamma_s$  of the micro-asperities, see Figure 4.22. The sum over all areas  $\Gamma_s$  in which contact stresses  $p_{N_s} \mathbf{n}_s$  are present yields a resultant force on the entire contact surface. The contact stress can be represented with respect to the basis  $(\mathbf{a}_\alpha^m, \mathbf{n}^m)$  of the contact mean plane surface

$$\sum_s \int_{\Gamma_s} p_{N_s} \mathbf{n}_s d\Gamma = \mathbf{f}_N + \mathbf{f}_T = f_N \mathbf{n}^m + f_T^\alpha \mathbf{a}_\alpha^m. \quad (4.58)$$

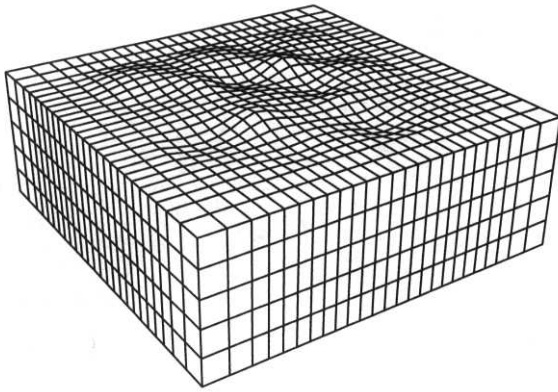
Averaging over the mean plane of the contact surface is then carried out. In the case depicted in Figure 4.22, the mean plane is straight. Hence the resulting mean contact stresses have the same direction

$$\begin{aligned} p_N^m \mathbf{n}^m &= \frac{1}{A} f_N \mathbf{n}^m \Rightarrow p_N^m = \frac{1}{A} f_N \\ t_T^{m\alpha} \mathbf{a}_\alpha^m &= \frac{1}{A} f_T^\alpha \mathbf{a}_\alpha^m \Rightarrow t_T^{m\alpha} = \frac{1}{A} f_T^\alpha \end{aligned} \quad (4.59)$$

Initial estimates regarding the structure of a frictional law can be drawn by evaluating these stresses. Therefore, one can derive a general form for the homogenized friction



**Fig. 4.22** Micro-structure of the contact zone.



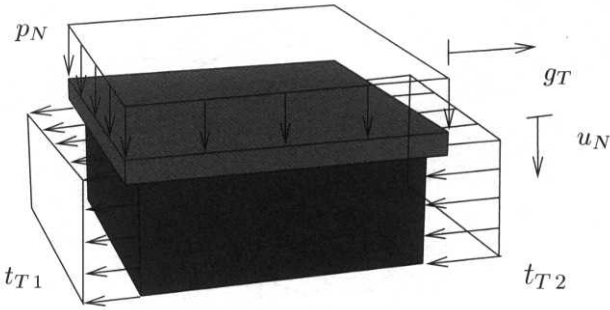
**Fig. 4.23** Finite element mesh and randomly generated surface.

law; some results of such procedures are reported in Haraldsson and Wriggers (2001) or Bandeira et al. (2002).

The general procedure to obtain friction laws from such computations is outlined below. The three steps have to be executed for each of the samples in a statistically representative set. In general, this means that around 100 different samples have to be solved:

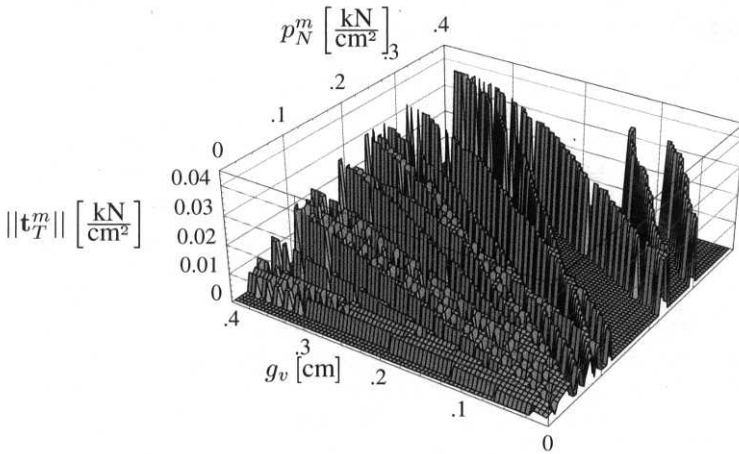
1. Discretization of two solid blocks with a randomly generated surface with the mean plane of the contact surface being  $A$ . A mesh stemming from such a procedure is presented for one of the solids in Figure 4.23. The micro-structure is modelled with  $b$ -splines in order to be able to discretize a wavy contact surface for the solids. In the example depicted in Figure 4.23, 25  $b$ -spline-nodes with randomly chosen heights were used to generate the asperities of the surface.
2. Application of a homogeneous deformation pattern at the boundary of the blocks in the normal and tangential directions, see Figure 4.24. To derive a friction law one has to perform several computations with different constant normal forces  $p_N A$ . Hence, a normal deformation  $u_N$  is first applied to obtain a certain normal force, which is then fixed throughout the computation. After that, the relative displacement  $g_T$  is applied incrementally in the tangential direction.
3. Incremental solution of the boundary value problem for a fixed normal force, e.g. fixed normal displacement  $u_N$ . Due to the applied tangential relative displacement  $g_T$ , the tangential stresses  $t_{T1}$  and  $t_{T2}$  are computed as reactions. From this result for normal and tangential reactions, one can then derive the homogenized contact stresses  $p_N$  and  $\mathbf{t}_T$  according to equation (4.59).

This procedure was used to derive a contact law for a concrete-soil interface; for details see Haraldsson and Wriggers (2001). One then obtains for different numerical

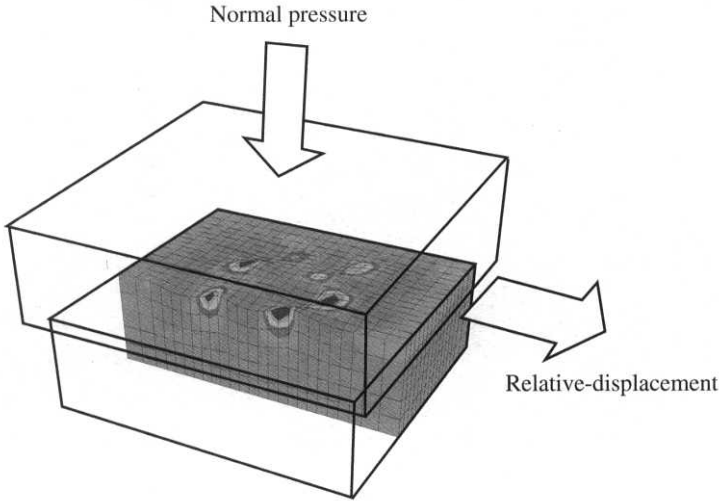


**Fig. 4.24** Movement of the block and resultant stresses  $t_{T1}$  and  $t_{T2}$ .

experiments a distribution of the tangential stress in relation to normal pressure  $p_N$  and total slip  $g_v$ , see Figure 4.25. Each line with values greater than zero for  $p_N^m$  stands for one numerical experiment. For the flat areas in between, no numerical experiments were performed; such numerical experiments still take some time, since one has to solve a geometrical and material nonlinear three-dimensional boundary value problem with many unknowns. However, the effort compared with real experiments is still small. Hence it makes a lot of sense to support experiments by additional numerical simulations in order to obtain a statistical representative result for the homogenized law. Also, one can vary the material parameters of the solids near the contact surface and hence can account additionally for work hardening in the case of metals. Furthermore the virtual testing procedure allows us to gain insight into the behaviour within the micro-structure of the contact interface. A view of the



**Fig. 4.25** Relation between mean normal pressure  $p_N^m$ , mean tangential stress  $t_T^m$  and slip  $g_v$ .



**Fig. 4.26** VON MISES stresses in the micro-structure.

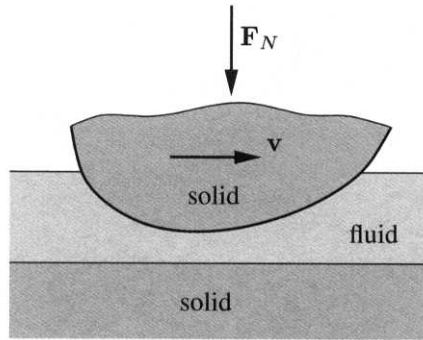
stress distribution in the micro-structure is depicted in Figure 4.26 for the model computation described above. The distribution of the VON MISES stresses shows clearly that the largest plastic deformations occur at the asperities, which is intuitively clear.

### 4.3 LUBRICATION

Lubrication is used to reduce friction in tribological systems like bearings, engines, gears, etc. This phenomenon is produced by, for example, a thin layer of fluid – the lubricant – between the solids in contact. Lubrication can also appear when a chemical reaction in the contact interface leads to a layer of contaminants which reduce friction; also, abrasive wear can decrease the friction force in the interface.

Even before the steam engine was developed, scientists were interested in the influence of lubrication when the friction between metals had to be reduced. Leibniz (1706) investigated already the difference between sliding and rolling friction and required special technical solutions to reduce friction. Later the industrial revolution demanded more insight, which finally lead to the formulation of the REYNOLDS equation for thin film lubrication and the STRIBECK curves, which can be used to distinguish mixed friction and pure hydrodynamic lubrication, see Figure 4.28. Here only lubrication due to the elasto-hydrodynamic effects of a fluid in the interface will be considered; for more detailed descriptions of other phenomena, see Polzer and Meissner (1983), Rabinowicz (1995) or Persson (2000).

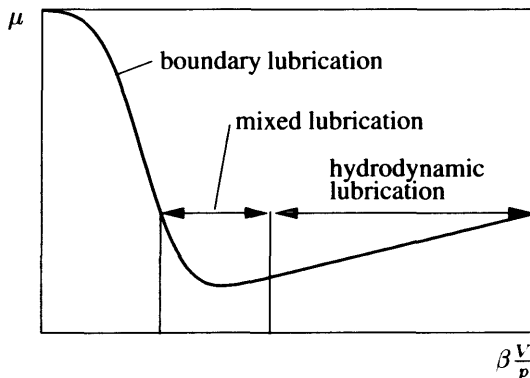
Lubrication only works when there is a certain relative velocity  $v$  between the solids such that the fluid layer does not disappear, see Figure 4.27. Such a process is called



**Fig. 4.27** Solid-fluid interaction in the case of lubrication.

hydrodynamic lubrication. If the velocity is too small, then the fluid will be squeezed out between the bodies leading to direct contact of the bodies, also called boundary lubrication. Experiments and theoretical predictions show that the frictional force is much higher for boundary lubrication than for hydrodynamic lubrication. Another observation is that the hydrodynamic lubrication depends upon the sliding velocity, leading to an increase in friction force for increasing velocity, see Figure 4.28.

There are two ways to introduce lubrication: one is the coupled treatment of the problem; the other is related to the fact that lubrication in general reduces the coefficient of friction. Hence it can be incorporated into the constitutive relations for friction, discussed in the previous sections, by a different choice of parameters. These have to be found from experiments or derived from equations governing lubrication, e.g. see Khonsari and Booser (2001).



**Fig. 4.28** Frictional stress versus normalized sliding velocity.

If lubrication is treated as a coupled problem, one has to state the relevant equations for the fluid film. These derive from the classical NAVIER-STOKES equations of fluid dynamics by introducing several simplifications inherent in the problem. It is well known that the flow in the interface is laminar for most cases, and it is also incompressible. With the further assumption that the nonlinear convective term in the NAVIER-STOKES equations can be neglected as well as inertia terms, the general equations reduce to the REYNOLDS equation,

$$\frac{\partial}{\partial x} \left( \frac{h_s^3}{\eta_s} \frac{\partial p_N}{\partial x} \right) + \frac{\partial}{\partial y} \left( \frac{h_s^3}{\eta_s} \frac{\partial p_N}{\partial y} \right) = 6 v_x \left( \frac{\partial p_N}{\partial x} \right). \quad (4.60)$$

This equation is valid for stationary processes when a constant relative sliding velocity  $v_x$  in the  $x$ -direction is present. The other variables are defined in Figure 4.29.  $h_s$  is the height of the gap, which can depend upon the deformations of the solids, and which then leads to a nonlinear coupled problem with the coupling terms being the contact pressure  $p_N$  and the deformation-dependent height  $h_s$ . Note that (4.60) is only valid for  $\delta_0 > 0$ , see Figure 4.29.  $\eta_s$  denotes the viscosity of the lubricant. This constitutive parameter depends in general upon the contact pressure  $p_N$ , the temperature  $\theta$  or the gap height  $h_s$ . An empirical formula which includes observations from experiments and takes the change due to  $h_s$  into account can be defined as

$$\eta_s(h_s) = \eta_s \left( \frac{h_s}{h_{s \max}} \right)^r, \quad 0 < r < 1. \quad (4.61)$$

Here  $r$  and  $\eta_s$  have to be adjusted via experimental data, and  $h_{s \max}$  follows from measurements of the problem geometry which can include micromechanical effects. A pressure-dependent viscosity coefficient follows from the function

$$\eta_s(p_N) = \eta_0 e^{\eta_1 p_N}, \quad (4.62)$$

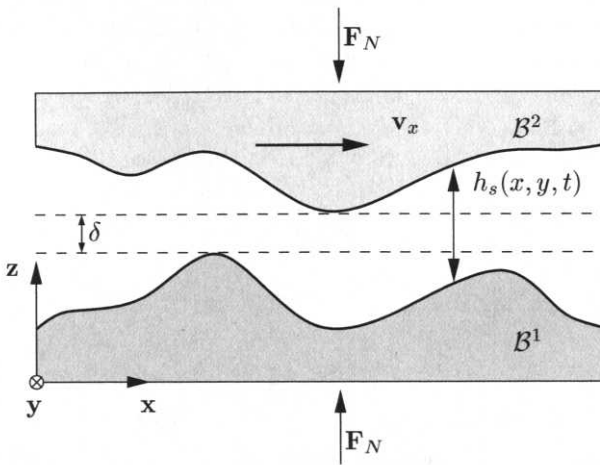


Fig. 4.29 Definition of lubrication interface.



with two constitutive parameters  $\eta_0$  and  $\eta_1$ .

Note that the gap height is usually very small, e.g.  $\sim 10^{-5}$  m. Hence, when the lubrication interface problem is treated in a coupled manner using the finite element method, one needs a high accuracy (very fine mesh) in the contact area to resolve these small deformations of the solid.

#### 4.4 ADHESION

If two clean plates (e.g. made of glass) are pressed together and then the load is reversed such that one tries to pull the plates apart, one can observe that the plates seem to be glued together. This effect is due to adhesion. This means that a positive normal stress has to be admitted in the contact interface ( $p_N > 0$ ). Furthermore, if the contact stresses are large enough, the adhesion will be broken in the contact interface, leading to a classical unilateral contact problem as defined in Section 5.1. Applications in which adhesive forces are present are a model of the micromechanics in contact interfaces (see the previous section or debonding processes in heterogeneous material, e.g. Wriggers et al. (1998), Zohdi and Wriggers (2000) or Zohdi and Wriggers (2001)). Furthermore, the analysis of tape which works due to adhesion can be mentioned.

Hence adhesion changes the unilateral contact problem and, as a result, a constitutive equation has to be formulated in the contact interface which is able to describe the transition from total adhesion to unilateral contact. Work in this direction was extensively pursued by the French school, e.g. see Fremond (1987). This author formulated adhesion by using a measure for the intensity of adhesion on the interface, which is similar to a description of damage mechanics by a damage function. Coupling of adhesion to friction was discussed in Raous et al. (2000) and Raous (1999) who also discussed the thermomechanical background and introduced a numerical model. However, there are also other models used to investigate the debonding processes in between matrix materials and particles, e.g. see Suresh et al. (1993).

We will not go into the details of the derivation of adhesive models, which can be found in Fremond (1987). Here only the resulting model will be discussed. For this purpose a condition, similar to (4.2), will be introduced, see Raous (1999),

$$-p_N + C_N g_N \beta^2 \geq 0, \quad g_N \geq 0, \quad (-p_N + C_N g_N \beta^2) g_N = 0, \quad (4.63)$$

where  $g_N$  is the gap function, see (3.6). The value  $\beta$  describes the intensity of adhesion:

$$\beta = \begin{cases} 1 & \text{total adhesion,} \\ 0 < \beta < 1 & \text{partial adhesion,} \\ 0 & \text{no adhesion.} \end{cases} \quad (4.64)$$

Furthermore  $C_N$  is a constitutive parameter which is associated with the stiffness in the normal direction of the interface. Still missing in this model is an evolution

equation for the intensity of adhesion. The evolution equation

$$\dot{\beta} = - \left[ \frac{1}{\eta} \left( w \frac{\partial h(\beta)}{\partial \beta} - C_N g_N^2 \beta \right) \right]^{1/p} \quad (4.65)$$

was proposed by Raous (1999). Here,  $\eta$  describes the viscosity of the adhesive interface.  $w$  is the limit of the debonding energy which acts as a threshold value: only when the force due to adhesion  $C_N g_N^2 \beta$  is greater than  $w \frac{\partial h(\beta)}{\partial \beta}$  is an evolution of the adhesion intensity possible.  $h$  is a given function which can be used to adjust the constitutive relation to experimental data. If one chooses  $h(\beta) = \beta$  and  $p = 1$ , (4.65) simplifies to

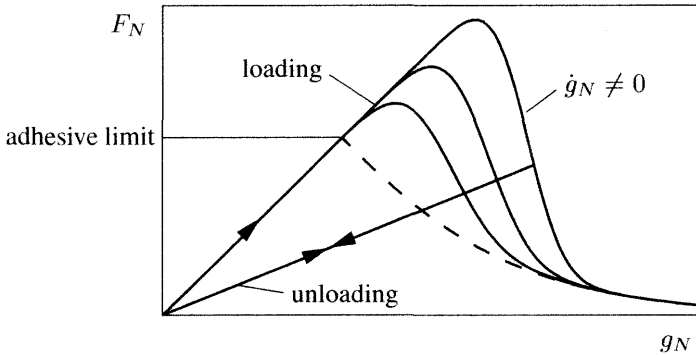
$$\dot{\beta} = - \frac{1}{\eta} (w - C_N g_N^2 \beta). \quad (4.66)$$

This equation is to be solved for the given constitutive parameters. The behaviour of the constitutive equation defined above is depicted in Figure 4.30. One observes that the constitutive equations leads to a stress response which is well known from continuum damage theory.

The incremental form of the adhesion law needed in a finite element algorithmic treatment is provided in Section 9.5.1.

The extension of the constitutive behaviour defined above to friction will also include adhesion in the stick/slip phase of the tangential movement. This leads to a set of equations which basically have the same structure as the adhesion model for normal contact:

$$\begin{aligned} f_s^a &= \|t_T - C_T g_T \beta^2\| - (p_N - C_N g_N \beta^2) \leq 0, \\ \dot{g}_T &= \dot{\gamma} \frac{\partial f_s^a}{\partial t_T}, \\ \dot{\beta} &= - \frac{1}{\eta} [w - (C_N g_N^2 + C_T \|g_T\|^2) \beta]. \end{aligned} \quad (4.67)$$



**Fig. 4.30** Adhesive normal stress

In this set of equations the only new constitutive parameter is  $C_T$ , which describes the elastic stiffness in the tangential direction of the adhesive zone. Also, only the simplified model (4.66) was used, but the model (4.65) can also be applied adequately.

#### 4.5 DECOHESION

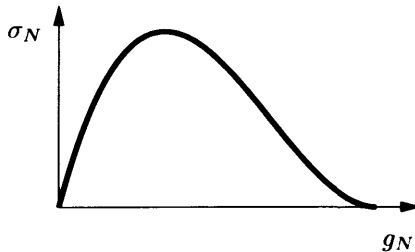
Another application, where special constitutive equations have to be formulated in the contact interface, is debonding. Debonding describes the loss of contact of e.g. micro-fibres or particles in a matrix material. Special constitutive equations have to be developed to simulate debonding processes. Out of many different possible formulations, e.g. see Tvergaard (1990) or Xu and Needleman (1993) a constitutive equation developed by Needleman (1990) for the two-dimensional case is reported, see Figure 4.31. Here normal stresses and the tangential stresses follow from an elastic strain energy function  $\Phi$  which relates the stresses in the interface to the gap opening  $g_N$  in the following way:

$$\sigma_A = \frac{\partial \Phi}{\partial g_A}, \quad (4.68)$$

where the subscript  $A$  stands for the direction normal or tangential to the interface. Using the strain energy function  $\Phi$  in Needleman (1990), one arrives at

$$\begin{aligned} \sigma_N &= C [\hat{g}_N (\hat{g}_N - 1)^2 + \alpha \hat{g}_T^2 (\hat{g}_N - 1)] , \\ \sigma_T &= \alpha C [\hat{g}_T (\hat{g}_N - 1)^2] . \end{aligned} \quad (4.69)$$

In this equation the normalized gap opening  $\hat{g}_N = g_N / \delta$  and the normalized relative tangential deformation  $\hat{g}_T = g_T / \delta$  have been introduced. The length  $\delta$  denotes the maximum gap opening which can be tolerated in the interface before complete debonding occurs. The constant  $C$  is related to the maximum stress carried by the interface  $C = 27/4 \sigma_{max}$ . In the case of  $g_T = 0$ , we obtain at  $\bar{g}_N = \frac{1}{3}$  as the maximum value the stress  $\sigma_{max}$ .



**Fig. 4.31** Adhesive normal stress for decohesion

## 4.6 WEAR

Wear among other events like breakage or corrosion is one of the causes leading to the fact that a part, an engine or a tool is no longer useful. There are many examples which underline this statement. One can think of a record player, even if such an object is almost history, where the record as well as the needle undergoes a wear process during use, which eventually can be observed easily just by listening to the quality of the sound. Another simple process from everyday life is writing. What one sees when using a pencil to write on a piece of paper is just the outcome of a wearing process. If these processes are connected with wear, then it is clear that wear means the removal of material from solid surfaces under mechanical action. In general, wear is related to sliding contact. However, the amount of material which is removed is quite small thus the wear process is not always easily detected in industrial applications where wear between metals has to be considered.

If studied in more detail, one observes that wear involves many different disciplines, e.g. material science, chemistry and applied mechanics. This is because besides the geometrical properties of surfaces and the frictional forces in the interface, surface chemistry and the material properties close to the surface also play a major role.

Basically, there are four different forms of wear mechanisms, which are discussed below:

- **Adhesive wear.** Fragments are pulled off one surface and adhere to the other during a sliding of smooth surfaces. If later these particles come off the surface, they can form loose wear particles.
- **Abrasive wear.** When a rough hard surface and a soft surface are in sliding contact, the hard surface ploughs grooves in the soft one. The material stemming from the grooves then forms the wear particles.
- **Corrosive wear.** This process occurs in a corrosive environment when a film protects the surfaces and is removed by a sliding action. In that case, the attack due to corrosion starts due to sliding contact.
- **Surface fatigue wear.** This happens if, during repeated sliding or rolling, the material undergoes many loading/unloading cycles which lead to cracks close to the surface. These might result in a breakup of the surface, and lead to the formation of wear particles.

These wear mechanisms cannot be discussed in this monograph in depth; for further details see Rabinowicz (1995), for example. From all of the above, it can be deduced that a wear process is complicated and can also involve different mechanisms at different stages of the process. Wear in general depends upon the properties of the material surfaces, the surface roughness, the sliding distance, the sliding velocity and the temperature. If one wants to describe wear mechanism through constitutive equations and evolution laws, then one has to determine which of the aforementioned effects play a major role. Depending on the softness of the surfaces and on the material

data of the particles which are worn off (debris), one might have to consider a third body (consisting of the debris) which acts in between the contacting interfaces.

The first constitutive equations which can be used to compute the volume of material lost by the wear process are very simple. A simple law for adhesive wear was proposed by Holm (1946) and Archard (1953). This HOLM-ARCHARD law has the form

$$V_{wear} = k_{ad} \frac{F_N g_T}{H}, \quad (4.70)$$

where  $k_{ad}$  is the so-called *wear coefficient*, which depends upon the materials in contact, and hence can be put in relation to the friction coefficient, see Rabinowicz (1995).  $H$  denotes the hardness of the surface which is worn away.  $F_N$  is the normal force and  $g_T$  the relative sliding distance between the materials. Equation (4.70) provides a linear relationship between wear volume  $V_{wear}$  and relative sliding distance  $g_T$  should the normal force  $F_N$  be kept constant. The wear coefficient  $k_{ad}$  has to be deduced from experimental data; for metals it is in the range of  $10^{-3}$  to  $10^{-8}$ .

Despite the fact that a completely different mechanism is associated with abrasive wear, one can apply the same constitutive equation as (4.70) to abrasive wear, see Rabinowicz (1995). It reads

$$V_{wear} = k_{abr} \frac{F_N g_T}{H}. \quad (4.71)$$

The only change is the abrasive wear coefficient  $k_{abr}$ , which physically represents a value that depends upon the average of the roughness angle, and ranges between  $10^{-2}$  to  $10^{-5}$ .

Due to the small amount of material which is worn off, no coupled analysis (concerning the change of mass or change of geometry) is needed when wear problems have to be solved. If laws (4.70) or (4.71) are used within an analysis, one only has to apply them in a post-processing phase. However, this means that the normal force is kept constant.

If this is not the case, as in contact problems in which deformation is involved, then the constitutive equations have to be refined. A thermodynamically consistent theoretical background for such models was provided by Zmitrowicz (1987a) and Zmitrowicz (1987b). Constitutive relations leading to evolution equations were also discussed in Curnier (1984) and application within the finite element method can be found in Neto et al. (1996), Strömberg et al. (1996), Strömberg (1997) or Agelet de Saracibar and Chiumenti (1999).

These authors indicated a connection between friction and wear effects. As a result of that, the mean rate of wear can be assumed to be proportional to the dissipation rate due to friction, see (4.23) or (4.39). The mean rate of wear itself is characterized by a function which represents the movement of the contacting surfaces in a normal direction, since wear removes volume and hence brings the surfaces closer together. The wear rate  $\dot{V}$  is then given by

$$\dot{V}_n = k_{wear} \mathcal{D}^s, \quad (4.72)$$

where  $k_{wear}$  is the wear parameter and  $D^s$  is the dissipation rate. In general, one has to assume different wear parameters for both contacting surfaces, which was omitted here. Using (4.72) leads to an algorithm for wear in which the wear rate is simply computed in each time or load increment of the frictional analysis.

In a more general situation, the wear parameter can depend upon the sliding velocity and on the temperature generated in the contact interface due to friction:  $k_{wear} = k_{wear}(\dot{\mathbf{g}}_T, \theta)$ . In the latter case, the wear computation needs a coupled thermo-mechanical treatment, e.g. see Chapter 10.

## 4.7 FRACTAL CONTACT INTERFACES

Rough contact surfaces can also be modelled by fractals. Here we shall state the general ideas which lead to such a model. Examples will be given in Section 12.6.3. Structures involving interfaces with fractal geometry are referred to as sequence of classical interfaces problems, which result from the consideration of the fractal interfaces as the unique *fixed point* or the *deterministic attractor* of a given Iterated Function System (IFS). On the interface itself, unilateral contact is assumed to hold.

The geometry of a fractal permits the accurate geometrical description of certain physical forms and of the figures and graphs resulting in physical, chemical and biomechanical processes, see e.g. Mandelbrot (1982), Takayasu (1990), Scholz (1989), Barnsley (1988) and Barnsley and Hurd (1993). One could mention here the forms of clouds and mountains, landscape and coastline geometry, fluvial system geometry, the distribution of craters in planets, etc.

In structural analysis and applied mechanics, we often have to deal with fractal domains  $\Omega$  and/or fractal boundaries  $\Gamma$ . These are, for example, the cases of the crack interfaces in natural bodies, the free surfaces and interfaces in fractured bones, metals and rocks, the geometry of metallic interfaces after sandblasting or meteoritic rain, the crashed interfaces in composite and granular materials, the geometry of fluvial systems, nervous cells and the geometry of plants, see e.g. Takayasu (1990), Scholz (1989) or Feder (1988).

The finite element theory on fractal interfaces in contact problems, based on the approach of Barnsley (1988), was investigated by Panagiotopoulos et al. (1992). He proposed an inequality fractal formulation under the assumption of the fractal interfaces obeying unilateral contact with friction conditions. The influence of fractal geometry on the mechanical quantities and on the displacement and stress fields of deformable bodies was discussed. The reliable numerical applications from rock mechanics, bone mechanics and fracture engineering were given. Also, certain methods for consideration of the fractal geometry in the FEM calculations were exploited, on the assumption that the fractal geometry does not change during the loading process. Here we apply this approach to define rough surfaces through fractals, see also Hu et al. (2000).

An important tool to the theory, which finds many applications to the mechanics of structure, is the fractal interpolation function. Suppose that in  $\mathbb{R}^2$ , for instance, we have a set of given data  $z_i = \{x_i, y_i\}$ ,  $i = 0, 1, \dots, N$ , where the maps are affine

transformations of the special structure

$$\{x, y\} \rightarrow w_i \begin{Bmatrix} x \\ y \end{Bmatrix} = \begin{bmatrix} a_i & 0 \\ c_i & d_i \end{bmatrix} \begin{Bmatrix} x \\ y \end{Bmatrix} + \begin{Bmatrix} e_i \\ f_i \end{Bmatrix}, \quad (4.73)$$

with the restrictions

$$w_i(z_0) = z_{i-1}, \quad w_i(z_N) = z_i \quad \text{for } i = 1, \dots, N. \quad (4.74)$$

In (4.73) one parameter is free. Choosing  $0 \leq d_i < 1$ , we obtain that

$$\begin{aligned} a_i &= \frac{(x_i - x_{(i-1)})}{(x_N - x_0)}, \quad e_i = \frac{(x_N x_{(i-1)} - x_0 x_i)}{(x_N - x_0)}, \\ c_i &= \frac{(y_i - y_{(i-1)})}{(x_N - x_0)} - d_i \frac{(y_N - y_0)}{(x_N - x_0)}, \\ f_i &= \frac{(x_N y_{(i-1)} - x_0 y_i)}{(x_N - x_0)} - d_i \frac{(x_N y_0 - y_N x_0)}{(x_N - x_0)}. \end{aligned} \quad (4.75)$$

Thus we have constructed an IFS from a given set of points  $z_i = \{x_i, y_i\}$ . If  $A$  denotes the attractor of this IFS, then  $A$  is the “fractal graph” of a continuous function  $f : [x_0, x_N] \rightarrow \mathbb{R}$  which interpolates the data  $\{x_i, y_i\}$ ,  $i = 0, 1, \dots, N$ . This function is called a “fractal interpolation function” and it is shown that it can be defined as the unique fixed point of a contractive transformation  $T$  which is defined by the relation

$$Tf(x) = c_i l_i^{-1}(x) + d_i f(l_i^{-1}(x)) + f_i, \quad (4.76)$$

where  $l_i^{-1}$  is the inverse of the linear mapping  $l_i(x) = a_i x + e_i$ .

Let us now assume that set  $A$  is called the deterministic attractor of IFS. Furthermore,  $\{\mathbb{R}^n; w_i\}$  is the deterministic fractal of the IFS considered. It can easily be shown that if the points  $x_0, \dots, x_N$  are equidistant, then the  $\dim A$ , denoted by  $D$ , is given by the formula

$$D = 1 + \frac{\ln \left( \sum_{i=1}^N |d_i| \right)}{\ln N} \quad (4.77)$$

if the points  $\{x_i, y_i\}$ ,  $i = 0, 1, \dots, N$  do not form a straight line (in this case  $D = 1$ ), and if  $\sum_{i=1}^N |d_i| > 1$ . If  $x_0, \dots, x_N$  are not equidistant, then  $f$  is the real solution of the equation

$$\sum_{i=1}^N |d_i| a_i^{D-1} = 1, \quad (4.78)$$

if  $\sum_{i=1}^N |d_i| > 1$ . Note that the proper choice of the parameter  $d_i$  may make  $D$  very close to 1, i.e. we obtain a line-like fractal, or very close to 2, i.e. we obtain a surface-like fractal which can be used to describe the rough contact interface. For more details see Panagiotopoulos et al. (1992) or Hu et al. (2000). An application is shown in Section 12.6.3.

# 5

---

## *Contact Boundary Value Problem and Weak Form*

For the formulation of the boundary value problem, we only have to discuss the additional terms due to contact in detail. The equations describing the behaviour of the bodies coming into contact do not change, and can be found in Chapter 2. For the treatment of different material behaviour like elasto-plasticity or visco-elasticity, see basic textbooks like Lubliner (1990), Malvern (1969), Maugin (1992) or for algorithms related to plasticity and visco-plasticity in combination with finite elements, see Simo and Hughes (1998).

Since the finite element method will be applied to solve the resulting nonlinear boundary value problems, weak formulations have to be developed for contact problems. The main concern of this chapter is the incorporation of the constraint equations formulated for frictionless contact in Section 3.1, and of the interface laws related to stick and sliding in the contact interface. However, one of the major problems in contact mechanics in the algorithmic treatment, which is associated with the numerical treatment of contact problems, is the non-differentiability of normal contact and friction terms. To overcome these difficulties, different formulations have been developed which are discussed in section 5.3.

### **5.1 FRICTIONLESS CONTACT IN LINEAR ELASTICITY**

As an introduction for contact, a solid with a rigid surface is considered in three-dimensional linear elasticity. Due to the contact constraints a variational inequality will appear instead of the standard variational equations known from classical solid mechanics, see also Section 1.1.



Due to the fact that the motion of the body is constrained from one side by the rigid surface, the problem is called a unilateral contact problem. The derivation of the associated variational formulation is the same as for the example in Section 1.1. One only has to use the relevant equations for the linear elastic solid in three dimensions which are stated in Chapter 2.

In the case of linear elasticity, the equilibrium equation (2.39) is given by

$$-\text{Div } \sigma = \bar{\mathbf{f}} \quad \text{in } B, \quad (5.1)$$

where  $\sigma$  is the stress tensor at a point  $\mathbf{X}$  in the interior of body  $B$ . In linear elasticity the stress can be computed from the linear strain field  $\varepsilon$ , see (2.27),

$$\varepsilon(\mathbf{u}) = \frac{1}{2} (\text{Grad } \mathbf{u} + \text{Grad}^T \mathbf{u}) \quad (5.2)$$

via the classical law of HOOKE, see (2.95);

$$\sigma(\mathbf{u}) = \mathbb{C} [\varepsilon(\mathbf{u})], \quad (5.3)$$

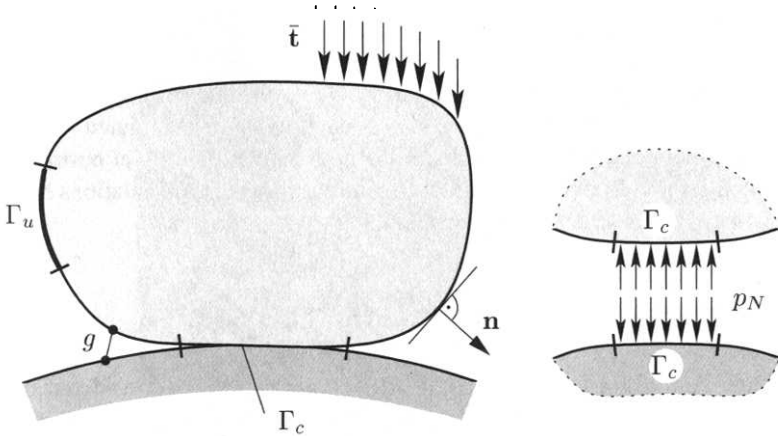
with the displacement field  $\mathbf{u}$  and the elasticity tensor  $\mathbb{C}$ . Let us assume that the following conditions hold on the boundary (see also Figure 5.1):

1. Displacement boundary conditions, also called DIRICHLET conditions:

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_u. \quad (5.4)$$

2. Traction boundary conditions, also called NEUMANN conditions:

$$\sigma \mathbf{n} = \bar{\mathbf{t}} \quad \text{on } \Gamma_\sigma, \quad (5.5)$$



**Fig. 5.1** : Unilateral contact of an elastic solid.

where  $\mathbf{n}$  is the outward normal of the surface of the solid and  $\hat{\mathbf{t}}$  denotes the applied traction on  $\Gamma_\sigma$ .

### 3. Contact conditions:

$$\begin{aligned} u_N - g &\leq 0, \\ p_N &\leq 0 \quad \text{on } \Gamma_c, \\ (u_N - g) p_N &= 0, \end{aligned} \quad (5.6)$$

with the normal component of the displacement field  $u_N = \mathbf{u} \cdot \mathbf{n}$  and the contact pressure  $p_N$  which is equivalent with the normal component of the traction vector  $p_N = \mathbf{t} \cdot \mathbf{n}$ .

This boundary value problem which yields the displacement field, includes the inequalities related to frictionless contact. It is called SIGNORINI'S problem, see Signorini (1933).

For a finite element solution of this system of equations, one needs the variational formulation of SIGNORINI'S problem. As usual, one can introduce a space  $\mathbf{V}$  of vector-valued, real functions defined on  $\mathcal{B}$ , e.g. see Kikuchi and Oden (1988). It is assumed that the test function  $\mathbf{v}$  (which in most of the engineering literature is called *virtual displacement*) fulfils the condition  $\mathbf{v} = \mathbf{0}$  on  $\Gamma_u$  and the contact condition  $v_N - g \leq 0 \forall \mathbf{v} \in \mathcal{V}$  with  $v_N = \mathbf{v} \cdot \mathbf{n}$ . Furthermore, let be  $\mathbf{u}$  the solution of SIGNORINI'S problem; then one can state the weak form of (5.1). This is equivalent to the virtual work produced by the stresses  $\boldsymbol{\sigma}$  and the virtual strains  $\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{v})$  due to the virtual displacements  $\mathbf{u} - \mathbf{v}$ , as well as the virtual work stemming from the external loads and body forces:

$$\begin{aligned} \int_B \boldsymbol{\sigma} \cdot \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{v}) dV &= \int_B \bar{\mathbf{f}} \cdot (\mathbf{u} - \mathbf{v}) dV + \int_{\Gamma_\sigma} \bar{\mathbf{t}} \cdot (\mathbf{u} - \mathbf{v}) d\Gamma \\ &+ \int_{\Gamma_c} p_N(\mathbf{u})(u_N - v_N) d\Gamma. \end{aligned} \quad (5.7)$$

Here  $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{u})$  is a function of the displacement defined via (5.3) and (5.2).  $\bar{\mathbf{f}}$  are the body forces and  $\bar{\mathbf{t}}$  denote the boundary tractions. Note that the term which includes  $\Gamma_u$  does not enter, since  $(\mathbf{u} - \mathbf{v})$  satisfies the boundary conditions, see (5.4). The last term in (5.7) can be reformulated with (5.6)<sub>3</sub> as

$$p_N(u_N - v_N) = p_N(u_N - v_N + g - g) = p_N(v_N - g) \geq 0, \quad (5.8)$$

where the greater equal sign results from (5.6)<sub>1,2</sub>; see also the discussion in Section 1.1. With this inequality the solution of the SIGNORINI problem defined by (5.1) to (5.6) has to fulfil

$$\int_B \boldsymbol{\sigma} \cdot \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{v}) dV \geq \int_B \bar{\mathbf{f}} \cdot (\mathbf{u} - \mathbf{v}) dV + \int_{\Gamma_\sigma} \bar{\mathbf{t}} \cdot (\mathbf{u} - \mathbf{v}) d\Gamma. \quad (5.9)$$

Thus a variational inequality, stemming from the contact constraint, characterizes the solution of SIGNORINI'S problem. This is different from the usual situation in solid mechanics, where the solutions have to fulfil variational equations. Due to the inequality constraint on the deformation field the contact problem is nonlinear even in the case of linear elasticity. Special algorithms have thus to be designed for problem (5.9).

The variational inequality can be recast in an abstract form, often used in theoretical and mathematical work regarding contact. In a more general situation, one can also formulate (5.9) for the contact between two deformable bodies. The analysis then yields the same structure as is inherent in (5.9); only the domain  $B = \cup_{\gamma=1}^2 B^\gamma$  is the union of both domains of the contacting bodies. Generally, one then defines the variational inequality by

$$a(\mathbf{u}, \mathbf{v} - \mathbf{u}) \geq f(\mathbf{v} - \mathbf{u}), \quad (5.10)$$

with

$$a(\mathbf{u}, \mathbf{v}) = \int_B \boldsymbol{\varepsilon}(\mathbf{u}) : \mathbb{C} : \boldsymbol{\varepsilon}(\mathbf{v}) dV, \quad (5.11)$$

$$f(\mathbf{v}) = \int_B \bar{\mathbf{b}} \cdot \mathbf{v} dV + \int_{\Gamma_\sigma} \bar{\mathbf{t}} \cdot \mathbf{v} d\Gamma, \quad (5.12)$$

where all quantities are defined by previous relations.

The problem is now stated by: find  $\mathbf{u} \in \mathbf{K}$  such that (5.10) is fulfilled for all  $\mathbf{v} \in \mathbf{K}$  with

$$\mathbf{K} = \{\mathbf{v} \in \mathbf{V} \mid (\mathbf{v}^2 - \bar{\mathbf{v}}^1) \cdot \bar{\mathbf{n}}^1 + g_0 \geq 0 \text{ on } \Gamma_c\}, \quad (5.13)$$

where  $\mathbf{V}$  represents the space of test functions.

The variational inequality (5.10) is stated here for frictionless contact. The problem is even more complicated when friction is present. In that case, not only are the inequality constraints in normal direction present, but there is also a special constitutive behaviour in the tangent direction at the contact interface. This is governed by sudden changes of the solution states such that the solution jumps from a state of stick (in which the tangential contact stresses follow as reactions from the stick conditions) to a state of sliding (in which the tangential stresses are computed from a constitutive equation, see Section 4.2). This special behaviour leads to even more mathematical difficulties when questions of existence and uniqueness of frictional contact problems are addressed. The mathematical structure of the variational inequality (5.10) is discussed in detail in, e.g., Duvaut and Lions (1976), Cocu (1984), Panagiotopoulos (1985), Rabier et al. (1986) or Kikuchi and Oden (1988). The latter reference also includes a mathematical analysis of the finite element method for contact problems. Special considerations concerning stability can be found in Klarbring (1988). Furthermore, examples for non-uniqueness or non-existence were discussed in Klarbring (1990) and Martins et al. (1994).

For dynamic contact problems, existence and uniqueness results can be found in Martins and Oden (1987), for example.

## 5.2 FRICTIONLESS CONTACT IN FINITE DEFORMATIONS PROBLEMS

All the equations needed to formulate the problem have already been discussed in Chapters 2–4. Here these basic equations are combined to obtain the boundary value problem for frictionless static contact. For this we formulate the local momentum equation for each body  $\mathcal{B}^\gamma$ , ( $\gamma = 1, 2$ ) in contact, see also (2.49):

$$\text{DIV } \mathbf{P}^\gamma + \bar{\mathbf{f}}^\gamma = \mathbf{0}, \quad (5.14)$$

where the inertia terms have been neglected.  $\mathbf{P}^\gamma$  denotes the first PIOLA–KIRCHHOFF stress tensor acting in the body  $\gamma$ , and  $\bar{\mathbf{f}}^\gamma = \rho_0^\gamma \bar{\mathbf{b}}^\gamma$  are the body forces. Next we formulate the boundary conditions for the deformation and the stress field

$$\varphi^\gamma = \bar{\varphi}^\gamma \quad \text{on } \Gamma_\varphi^\gamma, \quad (5.15)$$

$$\mathbf{t}^\gamma = \bar{\mathbf{t}}^\gamma \quad \text{on } \Gamma_\sigma^\gamma, \quad (5.16)$$

where  $\bar{\varphi}^\gamma$  and  $\bar{\mathbf{t}}^\gamma$  are described quantities. Furthermore, we have to account for the contact condition which is given by equation (3.6) when the bodies can come into contact. Together with the condition that no adhesion stresses can occur in the contact interface, we have from (4.2) the KUHN–TUCKER–KARUSH conditions for contact:

$$g_N \geq 0 \quad p_N \leq 0 \quad g_N p_N = 0 \quad \text{on } \Gamma_c. \quad (5.17)$$

Similar relations hold for the case in which constitutive equations are assumed in the contact interface  $\Gamma_c$ , and we can apply the relations defined in Section 4.1.2. In the case of frictional contact, the geometric relations derived in Section 3.2 and the constitutive laws given in Section 4.2 have to be used.

The constitutive equations for each body  $\mathcal{B}^\gamma$  have already been discussed in Section 2.4. For hyperelastic materials we have in general

$$\mathbf{P}^\gamma = \hat{\mathbf{P}}^\gamma(\mathbf{X}^\gamma, \mathbf{F}^\gamma, t), \quad (5.18)$$

for details see equations (2.74). Here we apply the equation as a model for nonlinear constitutive equations of the solid which is valid for finite elasticity. It leads in the current configuration to a nonlinear response function for the KIRCHHOFF stress  $\boldsymbol{\tau}$  in terms of the left CAUCHY–GREEN tensor  $\mathbf{b} = \mathbf{F}\mathbf{F}^T$ :  $\boldsymbol{\tau} = \mathbf{f}(\mathbf{b})$ . The KIRCHHOFF stress is related to the first PIOLA–KIRCHHOFF stress via  $\boldsymbol{\tau} = \mathbf{P}\mathbf{F}^T$ , with  $\mathbf{F}$  being the deformation gradient. The simplest example for hyperelasticity is the Neo-Hookian model, which can be applied, for example to rubber materials undergoing moderately large strains

$$\boldsymbol{\tau}^\gamma = \Lambda^\gamma (J^\alpha - 1) \mathbf{1} + \mu^\gamma (\mathbf{b}^\gamma - \mathbf{1}), \quad (5.19)$$

with the JACOBIAN of the deformation  $J^\gamma = \det \mathbf{F}^\gamma$ . The material parameters for the bodies  $\mathcal{B}^\gamma$  are the LAMÉ constants  $\Lambda^\gamma$  and  $\mu^\gamma$ .

Of course, it is possible to use more complicated constitutive relations which can also be of an inelastic nature. However, in such a case, no energy principle can be formulated.

For a numerical solution of the nonlinear boundary value problem summarized above, we will use the finite element method. Thus we need the weak form of the local field equations. Again, the weak formulation of contact problems leads to a variational inequality, see Duvaut and Lions (1976), which is stated in the next section.

Due to the fact that the constraint condition (5.17) is represented by an inequality, we also obtain a variational inequality which the solution of the contact problem has to fulfil. The general form can be written, as derived in Section 5.1 for the linear case, as

$$\sum_{\gamma=1}^2 \int_{B^\gamma} \boldsymbol{\tau}^\gamma \cdot \text{grad}(\boldsymbol{\eta}^\gamma - \boldsymbol{\varphi}^\gamma) dV \geq \sum_{\gamma=1}^2 \int_{B^\gamma} \bar{\mathbf{f}}^\gamma \cdot (\boldsymbol{\eta}^\gamma - \boldsymbol{\varphi}^\gamma) dV - \int_{\Gamma_{\sigma^\gamma}} \bar{\mathbf{t}}^\gamma \cdot (\boldsymbol{\eta}^\gamma - \boldsymbol{\varphi}^\gamma) dA. \quad (5.20)$$

Here the integration is performed with respect to the domain  $B^\gamma$  occupied by the body  $B^\gamma$  in the reference configuration. The stress tensor and the gradient operator "grad" are evaluated with respect to the current coordinates. Furthermore,  $\bar{\mathbf{f}}^\gamma$  denotes the body force of body  $B^\gamma$  and  $\bar{\mathbf{t}}^\gamma$  is the surface traction applied on the boundary of  $B^\gamma$ . The KIRCHHOFF stress  $\boldsymbol{\tau}^\gamma$  is defined in the previous section.

We now have to find the deformation  $(\boldsymbol{\varphi}^1, \boldsymbol{\varphi}^2) \in \mathbf{K}$  such that (5.20) is fulfilled for all  $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \in \mathbf{K}$  with

$$\mathbf{K} = \{ (\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \in \mathbf{V} \mid [\boldsymbol{\eta}^2 - \hat{\boldsymbol{\eta}}^1(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{n}}_1 \geq 0 \}. \quad (5.21)$$

This formulation holds for arbitrary constitutive equations, including inelastic effects.

If the problem can be described by hyperelastic material law, one can formulate the frictionless contact problem as follows:

$$\begin{aligned} \Pi &= \sum_{\gamma=1}^2 \Pi^\gamma \longrightarrow \text{MIN} \\ &\text{subject to } g_N \geq 0 \quad \text{on } \Gamma_c, \end{aligned} \quad (5.22)$$

which defines an optimization problem with inequality constraints. Here the energy function  $\Pi^\gamma$  for one body is given by (see also Section 2.3.3)

$$\Pi^\gamma = \int_{B^\gamma} W^\gamma(\mathbf{C}) dV - \int_{B^\gamma} \bar{\mathbf{f}}^\gamma \cdot \boldsymbol{\varphi}^\gamma dV - \int_{\Gamma_{\sigma^\gamma}} \bar{\mathbf{t}}^\gamma \cdot \boldsymbol{\varphi}^\gamma dA, \quad (5.23)$$

where  $W^\gamma(\mathbf{C})$  describes the strain energy function of the body  $B^\gamma$ , which is defined in Section 2.4.1 in detail.

In the case of finite elasticity, the existence of the solution of (5.22) can be proved, e.g. see Ciarlet (1988) or Curnier et al. (1992).

*REMARK 5.1: Mathematical analysis can be found for some cases of the unilateral contact undergoing large deflections in Kikuchi and Oden (1988) or Ciarlet (1988), for example. The linear problem has already been stated in Section 5.1.*

### 5.3 TREATMENT OF CONTACT CONSTRAINTS

In this section we shall discuss several different formulations that can be applied to incorporate the contact constraints into the variational formulation. Here frictionless as well as frictional contact formulations are derived.

Different possibilities exist for the numerical solution of these problems. Among them are the so-called *active set strategies*, which are applied in combination with LAGRANGE multipliers. Furthermore penalty techniques, e.g. see the text books of Bertsekas (1984) or Luenberger (1984), can be applied to solve contact problems. All these methods are well known in optimization theory. Other solution schemes are based on mathematical programming, e.g. see Conry and Seireg (1971) or Klarbring (1986), who applied this method to frictional contact problems.

Most standard finite element codes which are able to handle contact problems use either the penalty or the LAGRANGE multiplier method; for an overview and the mathematical framework, see Kikuchi and Oden (1988). Each of the methods has its own advantages and disadvantages, which will be discussed in detail in the following. The methods are designed to fulfil the constraint equations in the normal direction in the contact interface. For the tangential part we need in general constitutive relations when stick/slip motion occurs; associated techniques will be discussed in Chapter 9.

Here we concentrate in general on different possibilities to formulate the contact conditions. For a more simple representation we assume that the contact interface is known. This will often be the case later when an active set strategy is employed to solve the inequality (5.20).

Once the contact interface is known we can write the weak form as an equality. This means that we know the active set of constraints within an incremental solution step. Hence, equation (5.20) can be written as

$$\sum_{\gamma=1}^2 \left\{ \int_{B^\gamma} \boldsymbol{\tau}^\gamma \cdot \text{grad } \boldsymbol{\eta}^\gamma dV - \int_{B^\gamma} \bar{\mathbf{f}}^\gamma \cdot \boldsymbol{\eta}^\gamma dV - \int_{\Gamma_{\sigma^\gamma}} \bar{\mathbf{t}}^\gamma \cdot \boldsymbol{\eta}^\gamma dA \right\} + C_c = 0, \quad (5.24)$$

where  $C_c$  are contact contributions associated with the active constraint set.  $\boldsymbol{\eta}^\gamma \in V$  is the so-called test function or virtual displacement, which is zero at the boundary  $\Gamma_\varphi^\gamma$  where the deformations are prescribed.

In the case of hyperelastic materials, the starting point to derive equation (5.24) is the minimization of the total energy of the two bodies in contact (see Section 2.3.3)

$$\sum_{\gamma=1}^2 \left\{ \int_{B^\gamma} W^\gamma(\mathbf{C}) dV - \int_{B^\gamma} \bar{\mathbf{f}}^\gamma \cdot \boldsymbol{\phi}^\gamma dV - \int_{\Gamma_{\sigma^\gamma}} \bar{\mathbf{t}}^\gamma \cdot \boldsymbol{\phi}^\gamma dA \right\} + \Pi_c \implies MIN \quad (5.25)$$

where  $W^\gamma(\mathbf{C})$  is the strain energy related to body  $B^\gamma$ , which is discussed in detail in Section 2.4.1.  $\boldsymbol{\phi}^\gamma$  denotes the deformation of both bodies. The contributions due to the contact constraints are enclosed in  $\Pi_c$ . Note that this formulation is only valid for contact problems which do not include frictional sliding, since the friction process is dissipative and hence the solution becomes path-dependent.

For two bodies in contact we obtain the weak form or the energy related to the interface by assuming that contact is active at the surface  $\Gamma_c$ . Several different variants for the formulation of  $\Pi_c$  and  $C_c$  are discussed below:

1. The LAGRANGE multiplier method.
2. The penalty method.
3. The method of direct elimination of the geometrical contact constraints.
4. The formulation of constitutive equations in the contact interface.
5. The NITSCHKE method, which enforces geometrical constraints in a weak sense.
6. The perturbed LAGRANGE formulation which combines (1) and (2) in a mixed form.
7. The barrier method.
8. The augmented LAGRANGE method.
9. The cross-constraint method which combines the penalty and barrier methods.

This large variety of formulations also reflects the large number of different algorithms which have so far been developed to solve contact problems; see also Section 9.3.

Let us note that thermo-mechanical coupling, which is described in Chapter 10, can be formulated without particular problem using the above-mentioned formulations of the constraint terms. This is related to the fact that, independent from the method used, one has to compute the contact pressure and the distance between the mean planes of the rough contact surfaces from the normal constitutive law. Such evaluations are not influenced by the strategy adopted to solve the normal contact. However, not all strategies described below will turn out to be efficient for thermo-mechanical contact problems (see Chapter 10).

### 5.3.1 Lagrange multiplier method

A classical method is the use of LAGRANGE multipliers to add constraints to a weak form. We then formulate for the contact contribution  $\Pi_c$  as

$$\Pi_c^{LM} = \int_{\Gamma_c} (\lambda_N g_N + \lambda_T \cdot \mathbf{g}_T) dA, \quad (5.26)$$

where  $\lambda_N$  and  $\lambda_T$  are the LAGRANGE multiplier.  $g_N$  and  $\mathbf{g}_T$  are the normal and tangential gap functions. The variation of  $\Pi_c$  then leads to the constraint formulation

$$C_c^{LM} = \int_{\Gamma_c} (\lambda_N \delta g_N + \lambda_T \cdot \delta \mathbf{g}_T) dA + \int_{\Gamma_c} (\delta \lambda_N g_N + \delta \lambda_T \cdot \mathbf{g}_T) dA. \quad (5.27)$$

The first integral is associated with the virtual work of the LAGRANGE multipliers along the variation of the gap functions in normal and tangential directions. The second integral describes the enforcement of the constraints. Note that the LAGRANGE multiplier  $\lambda_N$  can be identified as the contact pressure  $p_N$ .  $\delta g_N$  is the variation of the normal gap, which is discussed in detail in Section 3.3. The terms  $\lambda_T \cdot \delta \mathbf{g}_T$  and  $\delta \lambda_T \cdot \mathbf{g}_T$  are associated with the tangential stick. A slip motion due to friction needs further consideration. In the case of pure stick the relative tangential slip  $\mathbf{g}_T$  is zero, which yields a constraint equation from which  $\lambda_T$  follows as a reaction. In the case of sliding, a tangential stress vector  $\mathbf{t}_T$  is determined from the constitutive law for frictional slip (see Section 4.2) and thus we should instead write  $\lambda_T \cdot \delta \mathbf{g}_T \rightarrow \mathbf{t}_T \cdot \delta \mathbf{g}_T$ , leading to

$$C_c^{slip} = \int_{\Gamma_c} (\lambda_N \delta g_N + \mathbf{t}_T \cdot \delta \mathbf{g}_T) dA + \int_{\Gamma_c} \delta \lambda_N g_N dA. \quad (5.28)$$

Again, the variation of the tangential slip can be found in Section 3.3. The LAGRANGE multiplier formulation is also the basis for the so-called *mortar* method used for connection of different non-matching meshes in the domain decomposition approaches for parallel computing.

*REMARK 5.2:* When only stick occurs in the contact interface, then we do not have to distinguish between the normal and tangential directions in the contact interface. Thus the constraint condition is given directly in terms of the deformation at the slave point; see the minimum distance problem (3.2)

$$\mathbf{x}^2 - \mathbf{x}^1(\bar{\xi}) = \mathbf{x}^2 - \bar{\mathbf{x}}^1 = \mathbf{0}. \quad (5.29)$$

With this, we obtain a simple expression for the contact contribution:

$$C_c^{LM} = \int_{\Gamma_c} \lambda \cdot (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) dA + \int_{\Gamma_c} \delta \lambda \cdot [\mathbf{x}^2 - \mathbf{x}^1(\bar{\xi})] dA. \quad (5.30)$$

Note that the tangential component  $\bar{\mathbf{x}}_{,\alpha}^1 \delta \xi^\alpha$ , which occurs when taking the variation of  $\bar{\mathbf{x}}^1$ , can be neglected in (5.30), since the point  $\mathbf{x}^2$  sticks at position  $\bar{\xi}$ , hence the convective coordinates  $\xi^\alpha$  do not change.

### 5.3.2 Penalty method

In this formulation a penalty term due to the constraint condition (3.7) is added to  $\Pi$  in (5.25) as follows:

$$\Pi_c^P = \frac{1}{2} \int_{\Gamma_c} (\epsilon_N (g_N^-)^2 + \epsilon_T \mathbf{g}_T \cdot \mathbf{g}_T) dA, \quad \epsilon_N, \epsilon_T > 0. \quad (5.31)$$

$\epsilon_N$  and  $\epsilon_T$  represent the penalty parameters. The penalty term  $\Pi_c^P$  is only added for active constraints which are defined by the penetration function  $g_N^-$ , see (3.7), and



has to be formulated for normal and tangential contacts, the latter in the case of stick. The variation of (5.31) yields

$$C_c^P = \int_{\Gamma_c} (\epsilon_N g_N^- \delta g_N^- + \epsilon_T \mathbf{g}_T \cdot \delta \mathbf{g}_T) dA, \quad \epsilon_N, \epsilon_T > 0. \quad (5.32)$$

It can be shown, e.g. see Luenberger (1984), that the solution of the LAGRANGE multiplier method is recovered from this formulation for  $\epsilon_N \rightarrow \infty$  and  $\epsilon_T \rightarrow \infty$ ; however, large numbers for  $\epsilon_N$  and  $\epsilon_T$  will lead to an ill-conditioned numerical problem. As in the LAGRANGE multiplier method, we have to distinguish between pure stick in the contact interface which yields (5.32), and the slip condition which leads to

$$C_c^{slip} = \int_{\Gamma_c} (\epsilon_N g_N^- \delta g_N^- + \mathbf{t}_T \cdot \delta \mathbf{g}_T) dA, \quad \epsilon_N > 0. \quad (5.33)$$

In the latter equation, one of the frictional laws from Section 4.2 has to be applied.

**REMARK 5.3:** When only stick occurs in the contact interface, then one does not need to distinguish between the normal and tangential directions and hence can proceed as in REMARK 5.2. In such a case we can choose equal penalty parameters for all directions  $\epsilon = \epsilon_N = \epsilon_T$ . With this and the constraint equation (5.29), the penalty term yields a simple expression for the contact contribution

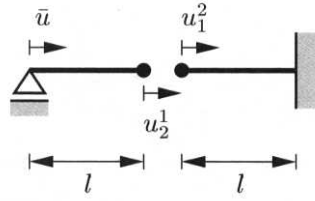
$$C_c^{stick} = \int_{\Gamma_c} \epsilon (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) dA. \quad (5.34)$$

Again, the tangential component of the variation  $\bar{\mathbf{x}}_{,\alpha}^1 \delta \xi^\alpha$  can be neglected (see Section 3.3.2), since the point sticks at position  $\bar{\boldsymbol{\xi}}$ , and hence the convective coordinates  $\xi^\alpha$  do not change. For application and for comparison of this formulation to the standard approach, see Section 8.2.

**REMARK 5.4:** Due to the fact that a high penalty parameter leads to an ill-conditioned problem, one is restricted in the choice of the penalty parameter for a given problem. The choice of the penalty parameter affects the solution in different ways. First the constraint equation is only fulfilled approximately. This of course also results in a deformation field  $\boldsymbol{\varphi}_\epsilon$ , which differs from the exact displacement field  $\boldsymbol{\varphi}$ . One has to show that

$$\|\boldsymbol{\varphi} - \boldsymbol{\varphi}_\epsilon\| \longrightarrow 0$$

for  $\epsilon \rightarrow \infty$  (for details, see e.g. Kikuchi and Oden (1988)). However, in practice a value for  $\epsilon$  has to be chosen such that ill-conditioning of the overall system of equations is avoided. In that case we have to live with the approximate solution  $\boldsymbol{\varphi}_\epsilon$ . In the case of a structure loaded only by surface tractions and body forces (see Figure 5.1) this affects only the displacement field since the loads are transferred through the contact surface by equilibrium, and hence the stress field is not disturbed too much when compared to the solution with an exact enforcement of the constraint equations. This situation changes when the structure is loaded by displacement boundary conditions (see Figure 5.2). In this case, the reactions and stresses depend upon the prescribed displacements. As an example, we consider the simple truss structure consisting of two bars of equal length  $l$  and equal axial stiffness  $EA$ . The system is fixed at the right



**Fig. 5.2** System loaded by prescribed displacement  $\bar{u}$ .

side and loaded by a prescribed displacement  $\bar{u}$  at the left side (see Figure 5.2). The initial gap between both bars is assumed to be zero. The contact constraint in this case is given by  $(u_2^1 - u_1^2) \leq 0$ . The formulation of the problem using the penalty method leads to the matrix system

$$\begin{bmatrix} \frac{EA}{l} + \epsilon & -\epsilon \\ -\epsilon & \frac{EA}{l} + \epsilon \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_1^2 \end{Bmatrix} = \begin{Bmatrix} \frac{EA}{l} \bar{u} \\ 0 \end{Bmatrix}, \quad (5.35)$$

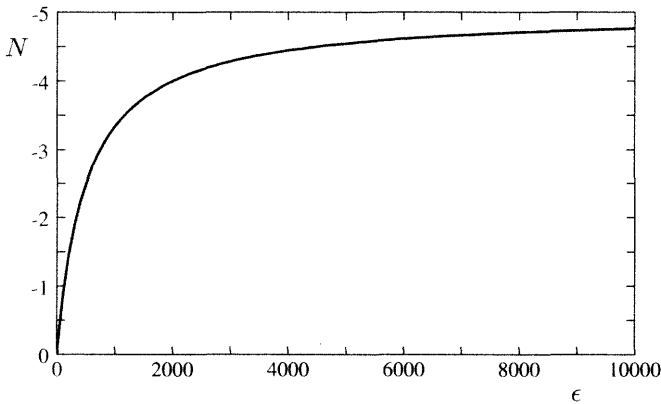
with the solution

$$\begin{Bmatrix} u_2^1 \\ u_1^2 \end{Bmatrix} = \frac{\bar{u}}{2 + \frac{EA}{l\epsilon}} \begin{Bmatrix} 1 + \frac{EA}{l\epsilon} \\ 1 \end{Bmatrix}. \quad (5.36)$$

The normal force in the left bar is then given by

$$N = \frac{EA}{l} (u_2^1 - \bar{u}) = \frac{EA}{l} \left( \frac{1 + \frac{EA}{l\epsilon}}{2 + \frac{EA}{l\epsilon}} - 1 \right) \bar{u}. \quad (5.37)$$

The limiting cases for  $\epsilon \rightarrow 0$  and  $\epsilon \rightarrow \infty$  can be deduced from (5.37), which yields  $\epsilon \rightarrow 0 : N \rightarrow 0$  and  $\epsilon \rightarrow \infty : N \rightarrow -5$ , which is equivalent to the exact enforcement of the constraint.



**Fig. 5.3** Solution dependency of normal force on the penalty parameter.

**Table 5.1** Deviation of normal force from exact solution

$\epsilon / EA$	$N$	Deviation %
1	-3.333	33.33
10	-4.762	4.76
100	-4.975	0.50
1000	-4.998	0.04

The normal force  $N$  is plotted in Figure 5.3 for different values of the penalty parameter  $\epsilon$  when we use for the data  $EA = 1000$ ,  $l = 1$  and  $\bar{u} = 0.01$ . Observe that low penalty parameters lead to a large deviation of the normal force from that obtained by the exact enforcement of the gap condition. This effect is also shown in Table 5.1.

Note that even in this simple example, one needs a penalty number which is at least 100 times  $EA$  to have a good approximation of the normal force in the bars. Due to this observation, it is clear that one has to adjust the penalty parameter in finite element approximations of contact problems. With refined meshes and hence better finite element approximations of the solution field, the error due to a too small penalty parameter also has to be reduced. This leads to a choice of the penalty parameter as a function of the mesh size.

### 5.3.3 Direct constraint elimination

The constraint equations (3.6), (3.13) or (3.14) can be enforced directly in the case of contact. This leads to a coupling in the virtual work expression (5.24) or in the total energy (5.25). The number of unknowns thus reduces. However, an efficient enforcement of the constraint depends heavily upon the discretization. Basically, one could formulate the inequality constraint (3.6) as an equality constraint

$$g_N = (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \mathbf{n}^1 = 0,$$

which yields

$$\mathbf{x}^2 \cdot \mathbf{n}^1 = \bar{\mathbf{x}}^1 \cdot \mathbf{n}^1 \longrightarrow (\mathbf{X}^2 + \mathbf{u}^2) \cdot \mathbf{n}^1 = (\bar{\mathbf{X}}^1 + \bar{\mathbf{u}}^1) \cdot \mathbf{n}^1 \quad (5.38)$$

for the current coordinates and the displacements. This local elimination works well for node-to-node contact elements, but not for arbitrary discretizations. In such a case, the point of departure is the LAGRANGE multiplier method. Using

$$\int_{\Gamma_c} \delta \lambda_N g_N d\Gamma = \int_{\Gamma_c} \delta \lambda_N (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \mathbf{n}^1 d\Gamma$$

one can project the constraints

$$\int_{\Gamma_c} \delta \lambda_N \mathbf{x}^2 \cdot \mathbf{n}^1 d\Gamma = \int_{\Gamma_c} \delta \lambda_N \bar{\mathbf{x}}^1 \cdot \mathbf{n}^1 d\Gamma, \quad (5.39)$$

and apply this equation to eliminate the unknowns on one side of the contact interface  $\Gamma_c$ . Here the choice of the interpolation for the LAGRANGE multiplier is essential for a stable method. Such methods are discussed in more detail in Section 7.4.2. In total, the direct elimination of the constraints reduces the problem size by all degrees of freedoms which are associated with the unknowns on one side of the contact interface  $\Gamma_c$ . Note that the elimination can be performed so that the positive definite structure of (5.25) is not destroyed.

### 5.3.4 Constitutive equation in the interface

In this case the constitutive equations which have been discussed in Sections 4.1 and 4.2 will be applied for the determination of  $p_N$  and  $\mathbf{t}_T$ . Hence, we do not add a constraint equation as in the case of the LAGRANGE multiplier or penalty method. The contact term to be used in the functional (5.24) when the constraint is active is then given by

$$C_c = \int_{\Gamma_c} (p_N \delta g_N + \mathbf{t}_T \cdot \delta \mathbf{g}_T) dA. \quad (5.40)$$

One can easily see that the introduction of the constitutive equation for the normal pressure (4.12) can be interpreted as a nonlinear penalty functional for the normal contact. The standard penalty method can be recovered from relation (4.12) using  $n = 1$ . However, such a choice would be artificial, since the usual range of the constitutive parameter  $n$ , stemming from experiments, is in the range  $2 \leq n \leq 3.33$ .

*REMARK 5.5: Due to the fact that the constitutive laws which have to be applied for  $p_N$  in (5.40) represent very stiff nonlinear springs in the contact interface, the use of this formulation within the finite element method often leads to ill-conditioned systems of equations. Hence, methods like the augmented LAGRANGE approach (see Section 5.3.8) are needed to avoid this numerical problem.*

### 5.3.5 Nitsche method

Another formulation which can be applied to enforce the contact constraints was derived in Nitsche (1970) and applied within domain decomposition methods with non-matching grids in Becker and Hansbo (1999). It is based on a different concept in which, instead of the LAGRANGE multipliers, the stress vector in the contact interface is computed from the stress field of the bodies. This leads to another set of boundary terms, which are stated here for the frictionless contact of two bodies:

$$\Pi_c^N = - \int_{\Gamma_c} \frac{1}{2} (p_N^1 + p_N^2) g_N dA + \frac{1}{2} \int_{\Gamma_c} \epsilon_N [g_N]^2 dA, \quad (5.41)$$

where the superscript of the contact pressure  $p_N^\gamma$  is associated with the body  $B^\gamma$ . Note that in this case, a contribution from both bodies enters the formulation in a mean sense. The last term in (5.41) represents the standard penalty term of Section 5.3.2. It

is included only to avoid ill-conditioning of the global equation system resulting from this formulation. However, since this formulation enforces the constraint exactly, the penalty term is not active, and hence the solution, contrary to the penalty method, does not depend on the penalty parameter  $\epsilon_N$ , which will be shown in Section 5.4 by means of an example.

The contact stresses  $p_N^\gamma$  in (5.41) are defined in terms of the displacement field, which for linear elasticity with CAUCHY'S theorem (2.40) leads to

$$p_N^\gamma = \mathbf{t}^\gamma \cdot \bar{\mathbf{n}}^\gamma = \bar{\mathbf{n}}^1 \cdot \boldsymbol{\sigma}(\mathbf{u}^\gamma) \mathbf{n}^\gamma = \bar{\mathbf{n}}^\gamma \cdot \mathbb{C}^\gamma[\boldsymbol{\varepsilon}(\mathbf{u}^\gamma)] \mathbf{n}^\gamma = \mathbf{N}_c^\gamma \cdot \nabla^s \mathbf{u}^\gamma, \quad (5.42)$$

where the last term was introduced to shorten the notation (the structure of the symmetrical displacement gradient  $\nabla^s \mathbf{u}$  is defined in (2.62), and the structure of  $\mathbf{N}_c^\gamma$  will be defined in Section 7.4.3). The variation of (5.41) can be computed, which yields

$$C_c^N = - \int_{\Gamma_c} \frac{1}{2} (\delta p_N^1 + \delta p_N^2) g_N dA - \int_{\Gamma_c} \frac{1}{2} (p_N^1 + p_N^2) \delta g_N dA + \int_{\Gamma_c} \epsilon_N g_N \delta g_N dA \quad (5.43)$$

with

$$\delta p_N^\gamma = \bar{\mathbf{n}}^\gamma \cdot \mathbb{C}^\gamma[\boldsymbol{\varepsilon}(\boldsymbol{\eta}^\gamma)] \mathbf{n}^\gamma = \mathbf{N}_c^\gamma \cdot \nabla^s \boldsymbol{\eta}^\gamma, \quad (5.44)$$

where  $\boldsymbol{\eta}^\gamma$  is the variation related to body  $\mathcal{B}^\gamma$ . Combining (5.42) to (5.44), one finally obtains

$$\begin{aligned} C_c^N &= - \int_{\Gamma_c} \frac{1}{2} (\mathbf{N}_c^1 \cdot \nabla^s \boldsymbol{\eta}^1 + \mathbf{N}_c^2 \cdot \nabla^s \boldsymbol{\eta}^2) g_N dA \\ &\quad - \int_{\Gamma_c} \frac{1}{2} (\mathbf{N}_c^1 \cdot \nabla^s \mathbf{u}^1 + \mathbf{N}_c^2 \cdot \nabla^s \mathbf{u}^2) \delta g_N dA \\ &\quad + \int_{\Gamma_c} \epsilon_N g_N^- \delta g_N^- dA, \end{aligned} \quad (5.45)$$

which shows that the NITSCHKE method yields a formulation which only depends upon the primary displacement variables.

A similar formulation can also be stated for the stick case, leading with (5.30) to

$$\Pi_c^N = - \int_{\Gamma_c} \frac{1}{2} (\mathbf{t}^1 + \mathbf{t}^2) \cdot (\mathbf{x}^2 - \bar{\mathbf{x}}^1) dA + \int_{\Gamma_c} \epsilon (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot (\mathbf{x}^2 - \bar{\mathbf{x}}^1) dA. \quad (5.46)$$

Again, the stress vector (traction vector)  $\mathbf{t}^\gamma$  is computed via CAUCHY'S theorem:

$$\mathbf{t}^\gamma = \boldsymbol{\sigma}(\mathbf{u}^\gamma) \mathbf{n}^\gamma = \mathbb{C}^\gamma[\boldsymbol{\varepsilon}(\mathbf{u}^\gamma)] \mathbf{n}^\gamma. \quad (5.47)$$

The variation yields

$$\begin{aligned} C_c^N &= - \int_{\Gamma_c} \frac{1}{2} (\delta \mathbf{t}^1 + \delta \mathbf{t}^2) \cdot (\mathbf{x}^2 - \bar{\mathbf{x}}^1) dA - \int_{\Gamma_c} \frac{1}{2} (\mathbf{t}^1 + \mathbf{t}^2) \cdot (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) dA \\ &\quad + \int_{\Gamma_c} \epsilon (\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) dA, \end{aligned} \quad (5.48)$$

with the variation of the traction vector

$$\delta \mathbf{t}^\gamma = \boldsymbol{\sigma}(\boldsymbol{\eta}^\gamma) \mathbf{n}^\gamma = \mathbb{C}^\gamma [\boldsymbol{\varepsilon}(\boldsymbol{\eta}^\gamma)] \mathbf{n}^\gamma. \quad (5.49)$$

In contrast to the LAGRANGE multiplier method, one does not need to introduce additional variables, since the stresses have to be inserted into (5.46) or (5.48) using (5.47) and (5.49).

In the nonlinear case the NITSCHKE method becomes more complex, since the variations of the tractions depend upon the type of constitutive equations used to model the solid. They are thus more difficult to compute than the variations of the LAGRANGE multipliers in Section 5.3.1.

### 5.3.6 Perturbed Lagrange formulation

This special formulation can be used to combine both penalty and LAGRANGE multiplier methods in a mixed formulation, e.g. see Oden (1981) or Simo et al. (1985). In this case, the following functional can be formulated for two solids coming into contact:

$$\Pi^{PL} = \sum_{\alpha=1}^2 \Pi^\alpha + \Pi_c^{PL}, \quad (5.50)$$

where  $\sum_\alpha \Pi^\alpha$  defines the total energy of the two bodies and  $\Pi_c^{PL}$  is the energy related to the contact interface. In detail, the last term in (5.50) is given by

$$\Pi_c^{PL} = \int_{\Gamma_c} \left[ \lambda_N g_N - \frac{1}{2\epsilon_N} \lambda_N^2 + \boldsymbol{\lambda}_T \cdot \mathbf{g}_T - \frac{1}{\epsilon_T} \boldsymbol{\lambda}_T \cdot \boldsymbol{\lambda}_T \right] d\Gamma. \quad (5.51)$$

Here the LAGRANGE multiplier term is regularized by the second term in the integral, which can be viewed as the complementary energy due to the LAGRANGE multiplier. The variation leads to

$$\begin{aligned} C_c^{PL} = \int_{\Gamma_c} & \left[ \lambda_N \delta g_N + \delta \lambda_N \left( g_N - \frac{1}{\epsilon_N} \lambda_N \right) \right. \\ & \left. + \boldsymbol{\lambda}_T \cdot \delta \mathbf{g}_T + \delta \boldsymbol{\lambda}_T \cdot \left( \mathbf{g}_T - \frac{1}{\epsilon_T} \boldsymbol{\lambda}_T \right) \right] d\Gamma. \end{aligned} \quad (5.52)$$

The first and third terms are again associated with the LAGRANGE multiplier formulation, whereas the second and fourth terms yield the “constitutive laws”:  $\lambda_N = \epsilon_N g_N$  and  $\boldsymbol{\lambda}_T = \epsilon_T \mathbf{g}_T$  if evaluated locally. If we insert this result for  $\lambda_N$  in the first term of (5.52), we obtain the standard penalty formulation (5.33). Letting  $\epsilon_N \rightarrow \infty$  yields the classical LAGRANGE multiplier method. However, this formulation is only valid for the frictionless and the stick cases. In the case of sliding, we have to use an incremental constitutive equation like COULOMB’S law which cannot be stated in the form of a complementary energy. In this case one has to use either the LAGRANGE multiplier or the penalty method.

We note that equation (5.52) can also be a starting point for special mixed formulations (e.g. in finite element formulations when different interpolation functions are used for the LAGRANGE multiplier and the displacement field; this is described in more detail in Chapter 7).

### 5.3.7 Barrier method

Another technique for problems with inequality constraints is the so-called barrier method. It adds a constraint functional of the type

$$\Pi_c^B = \epsilon_N \int_{\Gamma_c} b(g_{N+}) d\Gamma \quad (5.53)$$

to equation (5.50), which is always active for all possible contact nodes. By  $g_{N+}$  we denote that in this method there are no values of  $g_N$  allowed which violate (3.6). The value  $\epsilon_N > 0$  is the so-called barrier parameter, which has to be chosen in an appropriate way or changed within the algorithmic treatment of the contact problem. The barrier function  $b$  can be chosen as

$$b(g_N) = -\frac{1}{g_N} \quad \text{or} \quad (5.54)$$

$$b(g_N) = -\ln[\min\{1, -g_N\}]. \quad (5.55)$$

The second function is not differentiable because of the expression  $\min\{1, -g_{N+}\}$ . However, one can show that the differentiable function  $b(g_{N+}) = -\ln[-g_{N+}]$  also leads to convergence. Due to the construction of the constraint functional, the solution always has to stay in the feasible region, which means that no penetration is allowed in any intermediate iteration step. To ensure this special safeguard algorithms are needed, e.g. see Bazaraa et al. (1993). The variation of (5.53) yields, for the function (5.54),

$$C_c^B = \int_{\Gamma_c} \frac{\epsilon_N}{g_N^2} \delta g_N d\Gamma. \quad (5.56)$$

The advantage of this method is that all constraints are always active, and no on-and-off switch has to be applied to distinguish between active and passive constraints, as in the LAGRANGE multiplier or penalty methods. A drawback, however, is that one has to find a feasible starting point which fulfils all constraints. Furthermore, ill-conditioning as in the penalty method can occur. The barrier method is, due to these drawbacks, not used very much in computational contact mechanics. However, there is some recent success in using barrier methods in combination with augmented LAGRANGE techniques, see Kloosterman et al. (2001). In this work the inequality constraints  $g_N \geq 0$  are enforced based on the term

$$\Pi_c^B = -\epsilon_N \int_{\Gamma_c} \bar{\lambda} \ln \left( 1 + \frac{g_{N+}}{\epsilon_N} \right) d\Gamma, \quad (5.57)$$

where  $\bar{\lambda} \geq 0$  is the fixed LAGRANGE multiplier and  $\epsilon_N > 0$  the barrier parameter.

### 5.3.8 Augmented Lagrange methods

Another method to regularize the non-differentiable normal contact and friction terms is provided by the augmented LAGRANGE formulation. This technique has been considered extensively within the context of incompressibility constraints in, for example, Glowinski and Le Tallec (1984), and was also applied to contact problems for frictionless contact in Wriggers et al. (1985) and Kikuchi and Oden (1988). Recently, this approach has been extended successfully to large displacement contact problems including friction, see Alart and Curnier (1991) or Laursen and Simo (1993a). The main idea is to combine either the penalty method or the constitutive interface laws with LAGRANGE multiplier methods. The augmented LAGRANGE formulation yields a  $C^1$ -differentiable saddle point functional which is described in detail in Pietrzak and Curnier (1999). In that paper the following augmented LAGRANGE functional is introduced for normal contact:

$$\Pi_N^{AM} = \begin{cases} \int_{\Gamma_c} (\lambda_N g_N + \frac{\epsilon_N}{2} g_N^2) d\Gamma & \text{for } \hat{\lambda}_N \leq 0, \\ \int_{\Gamma_c} -\frac{1}{2\epsilon_N} |\lambda_N|^2 d\Gamma & \text{for } \hat{\lambda}_N > 0, \end{cases} \quad (5.58)$$

with  $\hat{\lambda}_N = \lambda_N + \epsilon_N g_N$ . The structure of this functional is such that it holds not only for  $\hat{\lambda}_N \leq 0$ , but also for  $\hat{\lambda}_N > 0$ , where the latter case means that the gap is open. The variation of (5.58) yields

$$C_N^{AM} = \begin{cases} \int_{\Gamma_c} (\hat{\lambda}_N \delta g_N + \delta \lambda_N g_N) d\Gamma & \text{for } \hat{\lambda}_N \leq 0, \\ \int_{\Gamma_c} -\frac{1}{\epsilon_N} \lambda_N \delta \lambda_N d\Gamma & \text{for } \hat{\lambda}_N > 0. \end{cases} \quad (5.59)$$

In the same way, a similar formulation can be formulated for the classical COULOMB law. For this purpose we introduce the increment of the relative tangential movement by  $\check{\mathbf{g}}_T = \dot{\xi}^\alpha \mathbf{a}_\alpha^1 dt$  and the augmented LAGRANGE multiplier  $\hat{\lambda}_T = \lambda_T + \epsilon_T \check{\mathbf{g}}_T$ . Using these definitions we can write the following functional, see Pietrzak and Curnier (1999), for the state of contact ( $\hat{\lambda}_N \leq 0$ ):

$$\Pi_T^{AM} = \begin{cases} \int_{\Gamma_c} (\lambda_T \cdot \check{\mathbf{g}}_T + \frac{\epsilon_T}{2} \check{\mathbf{g}}_T \cdot \check{\mathbf{g}}_T) d\Gamma & \text{for } \|\hat{\lambda}_T\| \leq \mu \hat{p}_N, \\ \int_{\Gamma_c} -\frac{1}{2\epsilon_T} \left[ \|\lambda_T\|^2 - 2\mu \hat{p}_N \|\hat{\lambda}_T\| + (\mu \hat{p}_N)^2 \right] d\Gamma & \text{for } \|\hat{\lambda}_T\| > \mu \hat{p}_N, \end{cases} \quad (5.60)$$

where  $\mu$  is the friction coefficient and  $\hat{p}_N$  is the augmented normal contact pressure. For the state of no contact ( $\hat{\lambda}_N > 0$ ), we have analogous to (5.58) the following



functional:

$$\Pi_T^{AM} = \int_{\Gamma_c} -\frac{1}{2\epsilon_T} \|\lambda_T\|^2 d\Gamma \quad \forall \hat{\lambda}_T. \quad (5.61)$$

Thus the functional also holds in the case of gap opening. This results from a prolongation of COULOMB'S cone for positive values of the augmented normal contact pressure. Note that the functional is continuous at  $\hat{\lambda}_N = 0$  and  $\|\hat{\lambda}_T\| = \mu \hat{p}_N$ , which also ensures that the combined functional (5.25) is globally continuous. The variation of (5.60) and (5.61) for a closed gap ( $\hat{\lambda}_N \leq 0$ ) yields

$$C_T^{AM} = \begin{cases} \int_{\Gamma_c} (\lambda_T \cdot \delta \tilde{\mathbf{g}}_T + \delta \lambda_T \cdot \tilde{\mathbf{g}}_T) d\Gamma & , \quad \|\hat{\lambda}_T\| \leq \mu \hat{p}_N, \\ \int_{\Gamma_c} \left( \frac{\mu \hat{p}_N}{\|\hat{\lambda}_T\|} \hat{\lambda}_T \cdot \delta \tilde{\mathbf{g}}_T - \frac{1}{\epsilon_T} \left[ \lambda_T - \frac{\mu \hat{p}_N \hat{\lambda}_T}{\|\hat{\lambda}_T\|} \right] \cdot \delta \lambda_T \right) d\Gamma & , \quad \|\hat{\lambda}_T\| > \mu \hat{p}_N, \end{cases} \quad (5.62)$$

and for the open gap ( $\hat{\lambda}_N > 0$ )

$$C_T^{AM} = \int_{\Gamma_c} -\frac{1}{\epsilon_T} \lambda_T \cdot \lambda_T d\Gamma, \quad \forall \hat{\lambda}_T. \quad (5.63)$$

A major problem associated with the numerical treatment of the penalty method and the contact interface laws is the ill-conditioning which arises when the penalty parameters  $\epsilon_N, \epsilon_T$  or the stiffness due to constitutive interface laws are combined with the stiffness of the bodies within the finite element formulation. A standard method to overcome the problem of ill-conditioning is based on the augmented LAGRANGE technique, well known in optimization theory. A simplified variant of (5.58) and (5.60) is provided by a special assumption put on the LAGRANGE multipliers. This leads to a double loop algorithm in which the LAGRANGE multiplier  $\bar{\lambda}_N$  is held constant during an iteration loop to solve the weak form in the inner loop. Then within an outer loop the LAGRANGE multiplier is updated to a new value (see Section 9.3.5). This procedure is known as the USZAWA algorithm.

The formulation leading to the USZAWA algorithm can be derived from the above equations by keeping the LAGRANGE multipliers constant. Here we shall state the result only for the frictionless case, which leads to the following contact contribution in the weak form:

$$C_N^{UZ} = \int_{\Gamma_c} (\bar{\lambda}_N + \epsilon_N g_N) \delta g_N d\Gamma. \quad (5.64)$$

Since  $\bar{\lambda}_N$  is unknown, an update procedure for the LAGRANGE multiplier has to be constructed within an iteration loop. The simplest update is

$$\bar{\lambda}_{N_{new}} = \bar{\lambda}_{N_{old}} + \epsilon_N g_{N_{new}}, \quad (5.65)$$

which is only of first order accuracy. For other possibilities, see Bertsekas (1984), for example.

If constitutive equations are used in the contact interface, Wriggers and Zavarise (1993a) have developed a special update which fulfils the nonlinear interface law by the update of the LAGRANGE multiplier  $\bar{\lambda}_N$ . For this formulation the geometrical relation (4.9) has to be adopted. This leads to a reformulation of (5.64)

$$C_N^{UZO} = \int_{\Gamma_c} (\bar{\lambda}_N + \epsilon_N [g_N - \zeta + d(\bar{p}_N)]) \delta g_N d\Gamma. \quad (5.66)$$

Now the update formula is given by

$$\bar{p}_{N_{new}} = \bar{p}_{N_{old}} + \epsilon_N [g_{N_{new}} - \zeta + d(\bar{p}_{N_{old}})], \quad (5.67)$$

where  $\{\dots\}_{old}$  are the known quantities from the previous state. Due to the appearance of a nonlinear function, the update is related, but different, to the standard update procedure for the LAGRANGE multipliers, see equation (5.65).

### 5.3.9 Cross-constraint method

A relatively new method for the solution of contact problems which is a modification of standard strategies will be presented in this section. Here the idea is to satisfy the contact constraints by a nonlinear, smooth change of contact stiffness around the solution. Thus the adopted approach leads to an iterative method which does not depict numerical instabilities within the solution search process. This fact permits us to achieve a better convergence rate with respect to standard methods.

The cross-constraint method, whose basic philosophy is outlined in Zavarise and Wriggers (1996), adopts smooth analytical functions to represent the nonlinear behaviour in contact processes. The method can be used to solve contact problems with constraint equations (3.1), or those based on introducing constitutive laws for contact surfaces. In the latter case, the formulation permits us to deal with problems where high precision of the calculated approach is required. To achieve this, the cross-constraint method has to be enhanced by using as a limit function a stress-penetration relationship which is constructed on physical bases.

As in the previous sections, the mechanical contact problem is solved by a modification of the unconstrained potential  $\Pi(\varphi)$

$$\Pi(\varphi) + P_i [g(\varphi)] \rightarrow \text{extremum}, \quad (5.68)$$

where  $g(\varphi) \leq 0$  represents the constraint functions set. Standard penalty ( $P$ ), LAGRANGE multiplier ( $LM$ ), barrier method ( $B$ ) (see Section 5.3.7) and the new cross-constraint method ( $CC$ ) can be obtained by suitable particularizations of  $P_i [g(\varphi)]$ . The characteristics of the modification functions are summarized as follows:

$$P \equiv \begin{cases} P_P(0) = 0 \\ P'_P(0) = 0 \end{cases} \quad LM \equiv \begin{cases} P_L(0) = 0 \\ P'_L(0) = \text{unknown} \end{cases} \quad (5.69)$$

$$B \equiv \begin{cases} P_B(0) = \infty \\ P'_B(0) = \infty \end{cases} \quad CC \equiv \begin{cases} P_C(0) = \hat{P} \\ P'_C(0) = \hat{F} \end{cases}, \quad (5.70)$$

where  $P_i$  is the modification of the unconstrained potential  $\Pi$  and the derivative  $P'_i$  is related to the contact force. In the case of the penalty method, a parabolic function is added; the LAGRANGE multiplier method adds a linear function but introduces new variables, i.e. the contact forces, and the barrier method adds an hyperbolic function. Finally, the proposed method adds a function, smooth over the entire domain, which presents nonzero, finite values at the constraint limit. Both traditional penalty and barrier methods can be obtained as limit cases of the proposed technique. Exponential functions of the type

$$P_c = \frac{\hat{F}^2}{\hat{K}} \exp \left( \frac{\hat{F}}{\hat{K}} g_N \right) \quad (5.71)$$

have been successfully used in Zavarise and Wriggers (1996). All methods produce a shift of the minimum from the unconstrained point to a zone close to the constrained solution point. The proposed method does not take limit values at the constraint limit, but takes finite nonzero values, both for the potential and for the contact forces. Moreover, it does not introduce additional variables.

Values of the cross-constraint method in the satisfaction of constraint equations are between those for the penalty and barrier methods. The values of  $\hat{F}$  and  $\hat{K}$  have to be changed according to the characteristics of the problem to obtain a good approximation for  $g(\mathbf{x}) = 0$ . An iterative procedure can be designed to fulfil this condition up to a specified tolerance.

The cross-constraint function (5.71) is close to zero when the gap is open ( $g > 0$ ); this minimizes the disturbance of the solution. Furthermore, (5.71) also presents a smooth transition around the constraints limit. Numerical tests have shown that the employment of (5.71) yields a contact algorithm which performs well from the numerical point of view.

If we also want to take into account the physics of the problem, i.e. the micromechanics of the contact interface, we have to consider a relationship based on the microscopical roughness of the contacting surfaces, and determine a pressure versus mean-plane-distance law. This can be done using constitutive equations for contact, see Section 4.1.2.

The basic consideration to replace the analytical function (5.71) with a contact constitutive law is related to the fact that the shape of such a law poses all the characteristics required. The law we use here is based on the microscopical characterization of the contacting surfaces and a hypothesis of the plastic behaviour of contacting asperities. Suitable adaptation of the relationship to obtain a form useful for finite element computations is derived in Zavarise et al. (1992b). The requested contact law is given by

$$p_N = A S_{NC} e^{[S_{NE}(\xi - g_N)^2]} = f(A, g_N), \quad (5.72)$$

where  $S_{NC}$  and  $S_{NE}$  are two constants which contain a suitable combination of the statistical parameters that characterize the contacting surfaces.  $A$  is the contact area,  $\xi$  is the initial distance of the surfaces and  $g_N$  denotes the surface approach.

This relationship represents an implicit nonlinear dependence of exponential type on the contact force and the mean plane distance. The adoption of such a constitu-

tive law as a limit function of the cross-constraints method keeps the efficiency and stability of the method and adds also physical insight. This fact will be even more important when considering frictional or thermal contact behaviour. The law can be used within the range of physically reasonable values of the contact pressure. Numerical computations may require an extension in the range of high pressure, which can occur during the iterative solution process. In this case, the extension can be designed by a linear function with  $C^1$ -continuity.

The contact problem is then defined by

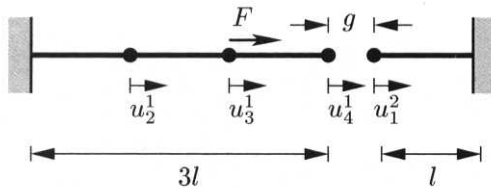
$$\delta\Pi(\varphi) + \int_{\Gamma_c} A S_{NC} e^{[S_{NE}(\xi - g_N)^2]} \delta g_N d\Gamma = 0, \quad (5.73)$$

which is solved by iterative (NEWTON-type) methods. Due to the continuity of the function around the constraint limit, typical numerical instabilities that may take place when the solution jumps between two states, one characterized by  $g_N > 0$ , and the other characterized by  $g_N = 0$ , are avoided. Hence the contact law represents a regularization of the non-smoothness of the standard contact formulation.

## 5.4 COMPARISON OF DIFFERENT METHODS

We shall illustrate the basic features of methods which lead to a linear system of equations for a given set of active constraints. These are the LAGRANGE multiplier, the penalty, the direct elimination and the NITSCHKE formulations, which are presented by means of a simple example. Furthermore, the different methods are compared with respect to the accuracy of the solutions.

Let us consider the two trusses depicted in Figure 5.4. The length of the left truss is  $3l$ , its axial stiffness  $EA/l$ . It is discretized by three truss elements with linear shape functions. The right truss has an axial stiffness of  $2EA/l$  and a length of  $l$ . Both trusses are clamped at the ends. The left truss is loaded by a point load  $F$ . When the load increases, the initial gap  $g$  between the two trusses closes at a certain value. This value can be computed from the equation system of the unconstrained system which already includes the boundary conditions at the left and right end of the bar



**Fig. 5.4** Truss structure with initial gap.

system ( $v_1^1 = v_2^2 = 0$ ),

$$\mathbf{K}_0 \mathbf{u} = \mathbf{f} \rightarrow \frac{EA}{l} \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \\ u_1^2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ F \\ 0 \\ 0 \end{Bmatrix}. \quad (5.74)$$

The solution can be expressed in closed form

$$\mathbf{u}^T = \frac{Fl}{EA} \{1, 2, 2, 0\}^T.$$

The force to close an assumed gap of  $g = 0.01 l$  follows from  $2 Fl / EA = 0.01 l$  as  $F = \frac{EA}{2} g$ . Once the gap is closed, we have to formulate the constraint problem by using one of the methods described in the previous sections.

For all methods used to enforce the contact constraint, we have to formulate the gap condition and its variation. Since the  $x$ -direction is also the normal direction, from the general formula (3.6) for the gap, we obtain

$$g_N = g - (u_4^1 - u_1^2) \quad (5.75)$$

and for the variation

$$\delta g_N = -\delta u_4^1 + \delta u_1^2. \quad (5.76)$$

In this special case we have only one point in the system (see Figure 5.4) where contact occurs. Hence for some methods to enforce the constraints we can solve two linear problems. One which was already defined in (5.74) is related to the unconstrained problem. The other is associated with the enforcement of the contact constraint condition. For the LAGRANGE multiplier formulation or the penalty method we can then set up a linear system of equations which directly gives the solution. This is not possible for the barrier or the augmented LAGRANGE methods, since in these approaches either a nonlinear constraint is present or an iteration has to be used. In these cases, a nonlinear system of equations can be formulated and solved iteratively.

**Lagrange multiplier method.** Using formulation (5.27) one has to add for the LAGRANGE multiplier approach the terms  $\lambda_N \delta g_N + \delta \lambda_N g_N$ , which with the explicit forms the gap function and its variation (5.75) and (5.76) yields the following system of equations  $\mathbf{K}_{LM} \mathbf{u}_{LM} = \mathbf{f}_{LM}$ . This can be written explicitly as

$$\begin{bmatrix} 2 \frac{EA}{l} & -\frac{EA}{l} & 0 & 0 & 0 \\ -\frac{EA}{l} & 2 \frac{EA}{l} & -\frac{EA}{l} & 0 & 0 \\ 0 & -\frac{EA}{l} & \frac{EA}{l} & 0 & -1 \\ 0 & 0 & 0 & 2 \frac{EA}{l} & 1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \\ u_1^2 \\ \lambda_N \end{Bmatrix} = \begin{Bmatrix} 0 \\ F \\ 0 \\ 0 \\ -g \end{Bmatrix}. \quad (5.77)$$

Here we used the fact that the displacements are zero in the first step of the contact algorithm. Since the problem is geometrically linear, in (5.77) we can use on the right hand side instead of the gap  $g_N$  as given in (5.75) just the initial gap  $g$ .

The solution of LAGRANGE multiplier formulation (5.77) can be stated in closed form as

$$\mathbf{u}_{LM} = \frac{2}{7EA} \begin{Bmatrix} (2EA g + 3Fl) / 2 \\ 2EA g + 3Fl \\ 3EA g + Fl \\ (-EA g + 2Fl) / 2 \\ EA (-EA g + 2Fl) / l \end{Bmatrix}. \quad (5.78)$$

where the last term represents the LAGRANGE multiplier or the reaction force in the contact interface.

**Penalty method.** In the case of the penalty method, the constraint is added by  $\epsilon_N g_N \delta g_N$ , which leads with (5.75) and (5.76) to the system of equations  $\mathbf{K}_P \mathbf{u}_P = \mathbf{f}_P$ , or explicitly to

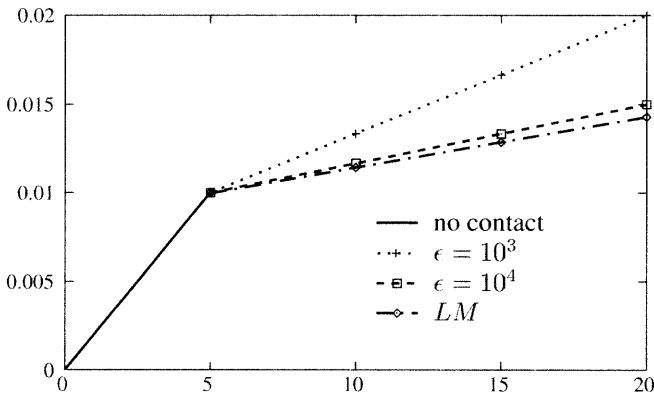
$$\begin{bmatrix} 2\frac{EA}{l} & -\frac{EA}{l} & 0 & 0 \\ -\frac{EA}{l} & 2\frac{EA}{l} & -\frac{EA}{l} & 0 \\ 0 & -\frac{EA}{l} & \frac{EA}{l} + \epsilon_N & -\epsilon_N \\ 0 & 0 & -\epsilon_N & 2\frac{EA}{l} + \epsilon_N \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \\ u_2^2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ F \\ \epsilon_N g \\ -\epsilon_N g \end{Bmatrix}. \quad (5.79)$$

Again, it is possible to find a closed form solution for the linear system of equations. For the variable  $u_4^1$  we obtain, for example,

$$u_4^1 = \frac{2(3EA\epsilon_N g + 2EAF l + \epsilon_N Fl^2)}{2(EA)^2 + 7\epsilon_N EA l}. \quad (5.80)$$

As discussed in Section 5.3.2,  $\epsilon_N \rightarrow \infty$  the penalty solution approaches for the solution for  $u_4^1$  in (5.78) of the LAGRANGE multiplier method, which can be easily verified using (5.80).

The displacement  $u_4^1$  is plotted in Figure 5.5 for different values of  $F$  when we select  $EA = 1000$ ,  $l = 1$  and  $g = 0.01$  for the data. One can see in Figure 5.5 that



**Fig. 5.5** Comparison of LAGRANGE and penalty methods.

the stiffness of the system changes once the gap closes for  $F = 5$ . The solution using the penalty method approaches the LAGRANGE Multiplier (LM) solution for large values of  $\epsilon_N$ , which can be observed from the solutions for  $\epsilon_N = 10^3$  and  $10^4$ .

It is also clear that the penalty solution is only an approximation of the correct enforcement of the constraint condition obtained by the LAGRANGE multiplier approach. It has been shown above that the penalty method yields in the limit  $\epsilon_N \rightarrow \infty$  the exact solution. However, in real computational applications it is not possible to use very large penalty parameters, since these lead to an ill-conditioning of the system of equations (5.79), hence the UZAWA scheme related to the augmented LAGRANGE multiplier method can be applied to improve the solution for a given penalty parameter  $\epsilon_N$ . We will not formulate the iterative scheme for the truss structure here, since this is done in Section 9.3.5.

**Direct elimination method.** This method was described in (5.3.3) and is based on the fact that the constraint (5.75) is enforced directly. For this purpose, we construct a projection matrix  $\mathbf{P}$  which reduces the displacement variables appearing in the constraint (5.75) by one

$$\bar{\mathbf{u}} = \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \\ u_1^2 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ -g \end{Bmatrix} = \mathbf{P} \mathbf{u} - \mathbf{g}. \quad (5.81)$$

This projection, without  $\mathbf{g}$ , is also valid for the variations. Hence the matrix system (5.74) of the unconstraint problem can be pre- and post-multiplied by  $\mathbf{P}$ , leading to the reduced system

$$\mathbf{K}_E \bar{\mathbf{u}} = \mathbf{P} \mathbf{K}_0 \mathbf{P}^T \bar{\mathbf{u}} = \mathbf{P} \mathbf{f} + \frac{2EA}{l} \mathbf{g}, \quad (5.82)$$

or in explicit form,

$$\frac{EA}{l} \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \end{Bmatrix} = \begin{Bmatrix} 0 \\ F \\ \frac{2EA}{l} g \end{Bmatrix}. \quad (5.83)$$

Note that the symmetric structure of the stiffness matrix is not destroyed. The solution of this equation system yields

$$\bar{\mathbf{u}} = \frac{2}{7EA} \begin{Bmatrix} (2EA g + 3Fl) / 2 \\ 2EA g + 3Fl \\ 3EA g + Fl \end{Bmatrix}, \quad (5.84)$$

which is equivalent to (5.78). The displacement  $u_1^2$  which was eliminated can be computed from the constraint equation (5.75)

$$u_1^2 = u_4^1 - g = \frac{2}{7EA} (3EA g + Fl) - g = \frac{2}{7EA} \left( -\frac{1}{2}EA g + Fl \right),$$

which is the exact value also given in  $\mathbf{u}_{LM}$ . The contact force corresponds to the normal force in the elements adjacent to the gap (see also next section). More generally, one can compute  $\lambda_N$  from (5.77) once the displacements are known. By decomposing this equation system in

$$\begin{aligned} \mathbf{K}_0 \mathbf{u} + \mathbf{C} \lambda_N &= \mathbf{f} \quad \text{with } \mathbf{C}^T = \langle 0, 0, -1, 1 \rangle \\ \mathbf{C}^T \mathbf{u} &= -g, \end{aligned}$$

one can solve the first equation for  $\lambda_N$  once  $\mathbf{v}$  is known. By pre-multiplying with  $\mathbf{C}^T$  this leads to

$$\mathbf{C}^T \mathbf{C} \lambda_N = \mathbf{C}^T \mathbf{f} - \mathbf{C}^T \mathbf{K}_0 \mathbf{u}, \quad (5.85)$$

where  $\mathbf{u}$  is the same as in (5.74). The multiplication with  $\mathbf{C}^T$  is necessary since  $\mathbf{C}^{-1}$  does not exist. For this example, we have  $\mathbf{C}^T \mathbf{C} = 2$ ,  $\mathbf{C}^T \mathbf{f} = 0$ , and hence

$$\lambda_N = -\frac{1}{2} \langle 0, 0, -1, 1 \rangle \frac{EA}{l} \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \\ u_1^2 \end{Bmatrix} = \frac{4}{7} F - \frac{2}{7} \frac{EA g}{l}, \quad (5.86)$$

which is also the result stated in (5.78) for the LAGRANGE multiplier.

**Nitsche method.** When the NITSCHKE method described in Section 5.3.5 is applied to the problem defined in Figure 5.4, one has to formulate the constraint conditions via equation (5.43). In this equation the contact pressure  $p_N^\gamma$  has to be inserted, which is computed from (5.44). For the problem at hand, the contact pressure is represented by the normal force. Hence one has to compute the normal force in the elements adjacent to the gap

$$N_3^1 = \frac{EA}{l} (u_4^1 - u_3^1) \quad \text{and} \quad N_3^2 = \frac{2EA}{l} (0 - u_1^2). \quad (5.87)$$

With this result and the explicit expressions for the gap (5.75) and its variation (5.75), equation (5.43) can be formulated as

$$\begin{aligned} C_c^N &= -\frac{1}{2} \frac{EA}{l} (u_4^1 - u_3^1 - 2u_1^2) (\delta u_4^1 - \delta u_1^2) \\ &\quad - \frac{1}{2} \frac{EA}{l} (\delta u_4^1 - \delta u_3^1 - 2\delta u_1^2) (u_4^1 - u_1^2 - g) \\ &\quad + \epsilon_N (u_4^1 - u_1^2 - g) (\delta u_4^1 - \delta u_1^2). \end{aligned} \quad (5.88)$$

This leads to the system of equations  $\mathbf{K}_N \mathbf{u}_N = \mathbf{f}_N$ , or explicitly to

$$\begin{bmatrix} 2\frac{EA}{l} & -\frac{EA}{l} & 0 & 0 \\ -\frac{EA}{l} & 2\frac{EA}{l} & -\frac{EA}{2l} & -\frac{EA}{2l} \\ 0 & -\frac{EA}{2l} & \epsilon_N & \frac{3EA}{2l} - \epsilon_N \\ 0 & -\frac{EA}{2l} & \frac{3EA}{2l} - \epsilon_N & \epsilon_N \end{bmatrix} \begin{Bmatrix} u_2^1 \\ u_3^1 \\ u_4^1 \\ u_1^2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ F + \frac{EA g}{2l} \\ \epsilon_N g - \frac{EA g}{2l} \\ \frac{EA g}{l} - \epsilon_N g \end{Bmatrix}, \quad (5.89)$$



which has the solution

$$\mathbf{u}_N = \frac{2}{7EA} \begin{Bmatrix} (2EA g + 3Fl)/2 \\ 2EA g + 3Fl \\ 3EA g + Fl \\ (-EA g + 2Fl)/2 \end{Bmatrix}. \quad (5.90)$$

Note that in this solution, the penalty parameter  $\epsilon_N$  does not occur, which means that (5.90) is independent of  $\epsilon_N$ . This is due to the fact that the first two constraint equations in (5.88) lead to an exact enforcement of the contact constraint (5.75). Hence, the penalty stabilization is not needed here. However, for large equation systems it is necessary to add the stabilization, since zeros occur in the diagonals of the stiffness matrix, see (5.89) for  $\epsilon_N = 0$ . The contact stress can now be computed from the normal force (5.87)

$$N_c = p_N A = N_3^1 = \frac{EA}{l} (u_4^1 - u_3^1) = \frac{4}{7} F - \frac{2}{7} \frac{EA g}{l}, \quad (5.91)$$

which is exactly the same result as was obtained for the LAGRANGE multiplier in (5.78).

## 5.5 LINEARIZATION OF THE CONTACT CONTRIBUTIONS

For an iterative solution of the nonlinear equation system associated with the weak form (5.24) and the various contact terms described above, different methods can be applied. A fast and reliable method is the NEWTON scheme, which is available in several variants to improve the global convergence properties of the method. These formulations rely on a linearization of the weak form. For the solids the linearizations were already derived in Sections 6.2.2 and 6.2.4. Thus, we can concentrate here on the linearization of the contact contributions.

As has been shown above, these have a different structure. However, the main ingredients are the same, and we will write here a general form of the contact contribution to see which terms need a detailed derivation for the linearization. For this purpose we distinguish between the normal and tangential contact terms in (5.24). Based on this we can write

$$C_c^G = \int_{\Gamma_c} c_N(\lambda_N, g_N, \delta\lambda_N, \delta g_N) d\Gamma + \int_{\Gamma_c} c_T(\lambda_T, \mathbf{g}_T, \delta\lambda_T, \delta\mathbf{g}_T; p_N) d\Gamma, \quad (5.92)$$

where  $c_N$  and  $c_T$  are functions of variations and variables, e.g. in the case of frictionless contact within the LAGRANGE multiplier method (see Section 5.3.1),

$$c_N = \lambda_N \delta g_N + \delta\lambda_N g_N,$$

or for the penalty method (see Section 5.3.2),

$$c_N = \epsilon_N g_N \delta g_N.$$

Note that, due to the geometrical nonlinearity of the finite deformation problem, the variations  $\delta g_N$  and  $\delta \mathbf{g}_T$  also contribute to the linearization, often resulting in very complex terms, see below (this is of course not the case when geometrically linear problems are formulated). On the contrary, the linearization of the LAGRANGE parameters  $\lambda_N$  and  $\lambda_T$  is zero. However, the variation of the tangential contributions is more involved, since one has to distinguish between the stick and the slip case in frictional contact. This is denoted by the dependence on the contact pressure  $p_N$  which occurs in the friction law.

### 5.5.1 Normal contact

The linearization of  $c_N$  can be computed from

$$\frac{\partial c_N}{\partial \mathbf{u}} \Delta \mathbf{u} = \frac{\partial c_N}{\partial \lambda_N} \Delta \lambda_N + \frac{\partial c_N}{\partial g_N} \Delta g_N + \frac{\partial c_N}{\partial \delta g_N} \Delta \delta g_N, \quad (5.93)$$

where the differentiation with respect to  $\delta \lambda_N$  has been neglected, since this term is zero. For the LAGRANGE multiplier method we now have

$$\frac{\partial c_N}{\partial \lambda_N} = \delta g_N, \quad \frac{\partial c_N}{\partial g_N} = \delta \lambda_N \quad \text{and} \quad \frac{\partial c_N}{\partial \delta g_N} = \lambda_N.$$

In the same way, the penalty method yields

$$\frac{\partial c_N}{\partial \lambda_N} = 0, \quad \frac{\partial c_N}{\partial g_N} = \epsilon_N \delta g_N \quad \text{and} \quad \frac{\partial c_N}{\partial \delta g_N} = \epsilon_N g_N.$$

Now we have to compute linearizations  $\Delta g_N$  and  $\Delta \delta g_N$  in terms of the displacement fields of both bodies in contact. The linearization of  $\Delta g_N$  has the same structure as the variation of  $g_N$ . Hence with (3.29), by exchanging  $\boldsymbol{\eta}^\alpha$  by  $\Delta \mathbf{u}^\alpha$ , we obtain

$$\Delta g_N = [\Delta \mathbf{u}^2 - \Delta \mathbf{u}^1(\bar{\xi}^1, \bar{\xi}^2)] \cdot \mathbf{n}^1(\bar{\xi}^1, \bar{\xi}^2). \quad (5.94)$$

The linearization of the gap in the normal direction  $\Delta(\delta g_N)$  has to be computed from the full variation (3.28), since terms which are zero in (3.28) (see Section 3.3) can also contribute to the tangent. We start with (3.6) from

$$\mathbf{x}^2 - \bar{\mathbf{x}}^1 = \mathbf{x}^2 - \mathbf{x}^1(\bar{\xi}^1, \bar{\xi}^2) = g_N \bar{\mathbf{n}}^1, \quad (5.95)$$

and obtain the variation in a way which can be used to compute  $\Delta \delta g_N$ :

$$\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1 - \bar{\mathbf{x}}_{,\alpha}^1 \delta \xi^\alpha = \delta g_N \bar{\mathbf{n}}^1 + g_N \delta \bar{\mathbf{n}}^1. \quad (5.96)$$

This yields  $\delta g_N$ , see (3.29), when multiplied by  $\bar{\mathbf{n}}^1$ . Since  $\boldsymbol{\eta}^1$  is a function of  $\boldsymbol{\xi}$  which again depends on the displacement field, a lot of terms arise from this equation. To shorten the notation we denote values related to the projection point  $\bar{\boldsymbol{\xi}}$  by just a bar, however we have to bear in mind that these terms depend on the deformation dependent surface coordinates  $\boldsymbol{\xi}$ . Now from (5.96) we obtain the linearization

$$\begin{aligned} & -[\bar{\boldsymbol{\eta}}_{,\alpha}^1 \Delta \xi^\alpha + \Delta \bar{\mathbf{u}}_{,\alpha}^1 \delta \xi^\alpha + \bar{\mathbf{x}}_{,\alpha\beta}^1 \Delta \xi^\beta \delta \xi^\alpha + \bar{\mathbf{x}}_{,\alpha}^1 \Delta \delta \xi^\alpha] \\ & = \Delta(\delta g_N) \bar{\mathbf{n}}^1 + \delta g_N \Delta \bar{\mathbf{n}}^1 + \Delta g_N \delta \bar{\mathbf{n}}^1 + g_N \Delta \delta \bar{\mathbf{n}}^1, \end{aligned} \quad (5.97)$$

where  $\Delta \mathbf{u}^\gamma$  are the increments of  $\mathbf{x}^\gamma = \mathbf{X}^\gamma + \mathbf{u}^\gamma$  and  $\Delta \xi^\alpha$  denote the increments of  $\xi^\alpha$ . In this equation the quantities  $\Delta \bar{\xi}^\alpha$ ,  $\delta \xi^\alpha$ ,  $\Delta \delta \xi^\alpha$ ,  $\Delta \bar{\mathbf{n}}^1$ ,  $\delta \bar{\mathbf{n}}^1$  and  $\Delta \delta \bar{\mathbf{n}}^1$  are still unknown. First we observe that variation and linearization  $\delta \bar{\xi}^\alpha$  and  $\Delta \xi^\alpha$  have the same structure, and hence the linearization  $\Delta \xi^\alpha$  can be computed analogous to (3.21) by interchanging the variations  $\boldsymbol{\eta}^\gamma$ , related to body  $B^\gamma$ , by the increments  $\Delta \mathbf{u}^\gamma$ :

$$\Delta \bar{\xi}^\beta = \bar{H}_{\alpha\beta}^{-1} [ (\Delta \mathbf{u}^2 - \Delta \bar{\mathbf{u}}^1) \cdot \bar{\mathbf{a}}_\alpha + g_N \bar{\mathbf{n}}^1 \cdot \Delta \bar{\mathbf{u}}_{,\alpha}^1 ] . \quad (5.98)$$

The tensor  $\bar{H}_{\alpha\beta}$  is exactly the same as given in equation (3.19). By multiplication with  $\bar{\mathbf{n}}^1$ , (5.97) can be solved for  $\Delta \delta g_N$ :

$$\Delta(\delta g_N) = -(\bar{\boldsymbol{\eta}}_{,\alpha}^1 \Delta \xi^\alpha + \Delta \bar{\mathbf{u}}_{,\alpha}^1 \delta \xi^\alpha + \bar{\mathbf{x}}_{,\alpha\beta}^1 \Delta \xi^\beta \delta \xi^\alpha) - g_N \bar{\mathbf{n}}^1 \cdot \Delta \delta \bar{\mathbf{n}}^1 . \quad (5.99)$$

Note that several terms disappear due to the identities  $\bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{n}}^1 = 0$  and  $\bar{\mathbf{n}}^1 \cdot \bar{\mathbf{x}}_\alpha^1 = 0$ . Thus the only term unknown is  $\bar{\mathbf{n}}^1 \cdot \Delta \delta \bar{\mathbf{n}}^1$ , which can be rewritten as  $\Delta(\bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{n}}^1) = \Delta \bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{n}}^1 + \bar{\mathbf{n}}^1 \cdot \Delta \delta \bar{\mathbf{n}}^1 = 0$ . From this identity, the last term in (5.99) follows as

$$-g_N \bar{\mathbf{n}}^1 \cdot \Delta \delta \bar{\mathbf{n}}^1 = g_N \Delta \bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{n}}^1 .$$

Now we have to compute the linearization and variation of the normal at the projection point  $\bar{\xi}$ . From the orthogonality condition  $\bar{\mathbf{n}}^1 \cdot \bar{\mathbf{a}}_\alpha^1 = 0$ , where  $\bar{\mathbf{a}}_\alpha^1 = \mathbf{x}_{,\alpha}^1$  is the tangent vector at the projection point, we compute

$$\delta \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{a}}_\alpha^1 = -\bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{a}}_\alpha^1 . \quad (5.100)$$

Since  $\bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{n}}^1 = 0$  we can solve (5.100) for the variation of the normal vector

$$\delta \bar{\mathbf{n}}^1 = -[\bar{\mathbf{a}}^1 \cdot \bar{\mathbf{n}}^1] \delta \bar{\mathbf{a}}_\alpha^1 = -(\bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{a}}_\alpha^1) \bar{\mathbf{a}}^1 \cdot \bar{\mathbf{a}}_\alpha^1 = -(\bar{\mathbf{n}}^1 \cdot \delta \bar{\mathbf{a}}_\alpha^1) \bar{a}^{\alpha\beta} \bar{\mathbf{a}}_\beta^1 , \quad (5.101)$$

and in the same way derive the linearization of the normal vector

$$\Delta \bar{\mathbf{n}}^1 = -[\bar{\mathbf{a}}^1 \cdot \bar{\mathbf{n}}^1] \Delta \bar{\mathbf{a}}_\alpha^1 = -(\bar{\mathbf{n}}^1 \cdot \Delta \bar{\mathbf{a}}_\alpha^1) \bar{\mathbf{a}}^1 \cdot \bar{\mathbf{a}}_\alpha^1 = -(\bar{\mathbf{n}}^1 \cdot \Delta \bar{\mathbf{a}}_\alpha^1) \bar{a}^{\alpha\beta} \bar{\mathbf{a}}_\beta^1 . \quad (5.102)$$

The contravariant base vector has been expressed by the covariant tangent vector as  $\bar{\mathbf{a}}^1 \cdot \bar{\mathbf{a}}_\alpha^1 = \bar{a}^{\alpha\beta} \bar{\mathbf{a}}_\beta^1$ . The variation and linearization of the tangent vector are given by

$$\delta \bar{\mathbf{a}}_\alpha^1 = \delta \bar{\mathbf{u}}_{,\alpha}^1 + \bar{\mathbf{x}}_{,\alpha\beta}^1 \delta \xi^\beta \quad \text{and} \quad \Delta \bar{\mathbf{a}}_\alpha^1 = \Delta \bar{\mathbf{u}}_{,\alpha}^1 + \bar{\mathbf{x}}_{,\alpha\beta}^1 \Delta \xi^\beta . \quad (5.103)$$

Once the linearization of the base vector  $\bar{\mathbf{a}}_\alpha^1$  is expressed by the incremental displacements the linearization of the normal vector is completed. What remains is to insert these results into (5.99) to derive the linearization of the variation of the gap vector. The final result is then given after some algebraic manipulations:

$$\begin{aligned} \Delta \delta g_N = & -(\bar{\boldsymbol{\eta}}_{,\alpha}^1 \Delta \xi^\alpha + \Delta \bar{\mathbf{u}}_{,\alpha}^1 \delta \xi^\alpha + \bar{\mathbf{x}}_{,\alpha\beta}^1 \Delta \xi^\beta \delta \xi^\alpha) \cdot \bar{\mathbf{n}}^1 \\ & + g_N \bar{\mathbf{n}}^1 \cdot (\bar{\boldsymbol{\eta}}_{,\alpha}^1 + \bar{\mathbf{x}}_{,\alpha\beta}^1 \delta \xi^\beta) \bar{a}^{\alpha\gamma} (\Delta \bar{\mathbf{u}}_{,\gamma}^1 + \bar{\mathbf{x}}_{,\gamma\theta}^1 \delta \xi^\theta) \cdot \bar{\mathbf{n}}^1 . \end{aligned} \quad (5.104)$$

Note that  $\Delta \delta g_N$  is symmetric with respect to variation and linearization. Hence within the finite element discretization this leads to a symmetric contribution to the

contact tangent stiffness matrix, see Chapter 8. Equation (5.104) is valid for general three-dimensional contact problems. For two-dimensional application it reduces significantly. In that case, the metric tensor  $\bar{a}_{\alpha\beta}$  is simply one number,  $\bar{a}_{11}$ , which is associated with the length of the tangent vector. Hence we define  $\bar{a}_{11} = \bar{l}^2$ , and with  $\xi^1 = \xi$  obtain

$$\begin{aligned} \Delta\delta g_N = & -(\bar{\eta}_{,\xi}^1 \Delta\xi + \Delta\bar{\mathbf{u}}_{,\xi}^1 \delta\xi + \bar{\mathbf{x}}_{,\xi\xi}^1 \Delta\xi \delta\xi) \cdot \bar{\mathbf{n}}^1 \\ & + \frac{g_N}{\bar{l}^2} (\bar{\eta}_{,\xi}^1 + \bar{\mathbf{x}}_{,\xi\xi}^1 \delta\xi) \cdot [\bar{\mathbf{n}}^1 \otimes \bar{\mathbf{n}}^1] (\Delta\bar{\mathbf{u}}_{,\xi}^1 + \bar{\mathbf{x}}_{,\xi\xi}^1 \delta\xi), \end{aligned} \quad (5.105)$$

a symmetric form in variation  $\boldsymbol{\eta}$  and linearization  $\Delta\mathbf{u}$ .

**REMARK 5.6**

1. In case that the factor  $g_N / \bar{l}^2$  is very small, one could be tempted to neglect the second term in (5.104) or (5.105) which makes the formulation much simpler. It is clear that then the property of quadratic convergence in NEWTON'S method is lost. However, the convergence rate can still be superlinear, and thus be sufficient for many applications.
2. Some discretization techniques are derived on the basis of flat element geometries. In such a case, equation (5.104) reduces considerably since all terms which include second derivatives are zero. With this assumption equation (5.104) collapses to

$$\Delta\delta g_N = -(\bar{\eta}_{,\alpha}^1 \Delta\xi^\alpha + \Delta\bar{\mathbf{u}}_{,\alpha}^1 \delta\xi^\alpha) \cdot \bar{\mathbf{n}}^1 + g_N \bar{a}^{\alpha\gamma} (\bar{\mathbf{n}}^1 \cdot \bar{\eta}_{,\alpha}^1) (\Delta\bar{\mathbf{u}}_{,\gamma}^1 \cdot \bar{\mathbf{n}}^1). \quad (5.106)$$

Note also that the tensor  $\bar{H}_{\alpha\beta}$  simplifies in this case, which is due to the fact that the second term in (3.21) disappears. Hence  $\bar{H}_{\alpha\beta}$  is equal to metric tensor  $\bar{a}_{\alpha\beta}$ , and thus we can use

$$\begin{aligned} \delta\xi^\beta &= \bar{a}^{\alpha\beta} [(\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot \bar{\mathbf{a}}_\alpha + g_N \bar{\mathbf{n}}^1 \cdot \bar{\boldsymbol{\eta}}_{,\alpha}^1] \\ \Delta\xi^\beta &= \bar{a}^{\alpha\beta} [(\Delta\mathbf{u}^2 - \Delta\bar{\mathbf{u}}^1) \cdot \bar{\mathbf{a}}_\alpha + g_N \bar{\mathbf{n}}^1 \cdot \Delta\bar{\mathbf{u}}_{,\alpha}^1], \end{aligned}$$

in (5.106) instead of (3.21) and (5.98).

3. Furthermore, the associated two-dimensional version of the linearization of the variation of the gap function (5.105) for flat surfaces has the structure

$$\Delta\delta g_N = -(\bar{\eta}_{,\xi}^1 \Delta\xi + \Delta\bar{\mathbf{u}}_{,\xi}^1 \delta\xi) \cdot \bar{\mathbf{n}}^1 + \frac{g_N}{\bar{l}^2} \bar{\eta}_{,\xi}^1 \cdot [\bar{\mathbf{n}}^1 \otimes \bar{\mathbf{n}}^1] \Delta\bar{\mathbf{u}}_{,\xi}^1. \quad (5.107)$$

Here  $\delta\xi$  and  $\Delta\xi$  are, according to (3.24), given by

$$\begin{aligned} \delta\xi &= \frac{1}{\bar{l}^2} [(\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot \bar{\mathbf{x}}_{,\xi}^1 + g_N \bar{\mathbf{n}}^1 \cdot \bar{\boldsymbol{\eta}}_{,\xi}^1] \\ \Delta\xi &= \frac{1}{\bar{l}^2} [(\Delta\mathbf{u}^2 - \Delta\bar{\mathbf{u}}^1) \cdot \bar{\mathbf{x}}_{,\xi}^1 + g_N \bar{\mathbf{n}}^1 \cdot \Delta\bar{\mathbf{u}}_{,\xi}^1]. \end{aligned}$$

### 5.5.2 Tangential contact

The linearization of the tangential contact contribution  $c_T$  in (5.92) follows from

$$\frac{\partial c_T}{\partial \mathbf{u}} \Delta\mathbf{u} = \frac{\partial c_T}{\partial \boldsymbol{\lambda}_T} \Delta\boldsymbol{\lambda}_T + \frac{\partial c_T}{\partial \mathbf{g}_T} \Delta\mathbf{g}_T + \frac{\partial c_T}{\partial \delta \mathbf{g}_T} \Delta\delta \mathbf{g}_T + \frac{\partial c_T}{\partial p_N} \Delta p_N. \quad (5.108)$$

However, here we have to distinguish between the stick and the slip case. We shall thus rewrite the general form of (5.92) for stick and frictional slip in such a way that we introduce the tangential stress vector  $\mathbf{t}_T$ . It can be either interpreted as a LAGRANGE multiplier in the case of stick, or as a stress vector stemming from a constitutive equation like COULOMB's friction law. With this we can write in general

$$c_T = \mathbf{t}_T \cdot \delta \mathbf{g}_T + \delta \mathbf{t}_T \cdot \mathbf{g}_T, \quad (5.109)$$

where the first term reflects the virtual work of the stress vector along the relative tangential displacement. This can be either due to a constitutive relation in the interface or due to a LAGRANGE multiplier. The second term appears only in the case of a LAGRANGE multiplier formulation of the stick part, since in frictional sliding  $\mathbf{g}_T \neq 0$ . This equation can be simplified even more by noting that  $\delta \mathbf{g}_T = \delta \xi^\alpha \bar{\mathbf{a}}_\alpha^1$ . By now introducing the component form of the stress vector,  $\mathbf{t}_T = t_{T\alpha} \bar{\mathbf{a}}^1{}^\alpha$ , the first term in (5.109) yields

$$\mathbf{t}_T \cdot \delta \mathbf{g}_T = t_{T\alpha} \delta \xi^\alpha. \quad (5.110)$$

With this the linearization of  $c_T$  in the stick case for a LAGRANGE multiplier formulation leads to

$$\frac{\partial c_T}{\partial \mathbf{u}} \Delta \mathbf{u} = \Delta \lambda_{T\alpha} \delta \xi^\alpha + \lambda_{T\alpha} \Delta \delta \xi^\alpha + \delta \lambda_{T\alpha} \Delta \xi^\alpha. \quad (5.111)$$

In the frictional sliding case we have

$$\frac{\partial c_T}{\partial \mathbf{u}} \Delta \mathbf{u} = \Delta t_{T\alpha} \delta \xi^\alpha + t_{T\alpha} \Delta \delta \xi^\alpha. \quad (5.112)$$

In these equations the following quantities have to be derived:  $\Delta t_{T\alpha}$ ,  $\Delta \xi^\alpha$  and  $\Delta \delta \xi^\alpha$ . Since the linearization of the tangential part of the stress vector depends heavily on the constitutive equation in the contact interface (see Section 4.2) we discuss here only the second and third terms. The linearization of the stress vector  $\Delta t_{T\alpha}$  can be computed once the update formula is known in terms of the displacement variables (e.g. see Section 9.5.2).

Equations analogous to (5.111) and (5.112) can also be derived for the penalty formulation. However, the relevant equations will be found in more detail in Chapter 8. This is due to the fact that in this formulation, only knowledge of  $\Delta \xi^\alpha$  and  $\Delta \delta \xi^\alpha$  is needed, apart from the penalty terms which are related to the stress vector.

The linearization of  $\xi^\alpha$  has already been given in (5.98), thus it remains to compute  $\Delta \delta \xi^\alpha$ . For this purpose we start again from (5.97). This equation is multiplied by  $\bar{\mathbf{a}}_\gamma^1$  which yields

$$\begin{aligned} -\bar{a}_{\alpha\gamma}^1 \Delta \delta \xi^\alpha &= [\bar{\eta}_{,\alpha}^1 \Delta \xi^\alpha + \Delta \bar{\mathbf{u}}_{,\alpha}^1 \delta \xi^\alpha + \bar{\mathbf{x}}_{,\alpha\beta}^1 \Delta \xi^\beta \delta \xi^\alpha] \cdot \bar{\mathbf{a}}_\gamma^1 \\ &\quad + [\delta g_N \Delta \bar{\mathbf{n}}^1 + \Delta g_N \delta \bar{\mathbf{n}}^1 + g_N \Delta \delta \bar{\mathbf{n}}^1] \cdot \bar{\mathbf{a}}_\gamma^1, \end{aligned} \quad (5.113)$$

where terms which cancel out have been omitted. With

$$\begin{aligned} \Delta [\delta (\bar{\mathbf{a}}_\gamma^1 \cdot \bar{\mathbf{n}}^1)] &= \Delta [\delta \bar{\mathbf{a}}_\gamma^1 \cdot \bar{\mathbf{n}}^1 + \bar{\mathbf{a}}_\gamma^1 \cdot \delta \bar{\mathbf{n}}^1] \\ &= \Delta \delta \bar{\mathbf{a}}_\gamma^1 \cdot \bar{\mathbf{n}}^1 + \delta \bar{\mathbf{a}}_\gamma^1 \cdot \Delta \bar{\mathbf{n}}^1 + \Delta \bar{\mathbf{a}}_\gamma^1 \cdot \delta \bar{\mathbf{n}}^1 + \bar{\mathbf{a}}_\gamma^1 \cdot \Delta \delta \bar{\mathbf{n}}^1, \end{aligned}$$

one can rewrite the last term in (5.113) as

$$\bar{\mathbf{a}}_\gamma^1 \cdot \Delta \delta \bar{\mathbf{n}}^1 = -(\Delta \delta \bar{\mathbf{a}}_\gamma^1 \cdot \bar{\mathbf{n}}^1 + \delta \bar{\mathbf{a}}_\gamma^1 \cdot \Delta \bar{\mathbf{n}}^1 + \Delta \bar{\mathbf{a}}_\gamma^1 \cdot \delta \bar{\mathbf{n}}^1). \quad (5.114)$$

Finally, the evaluation of  $\Delta \delta \bar{\mathbf{a}}_\gamma^1$  is given by

$$\begin{aligned} \Delta \delta \bar{\mathbf{a}}_\gamma^1 &= \Delta(\bar{\boldsymbol{\eta}}_\gamma^1 + \bar{\mathbf{x}}_{,\gamma\alpha}^1 \delta \xi^\alpha) \\ &= \bar{\boldsymbol{\eta}}_{,\gamma\alpha}^1 \Delta \xi^\alpha + \Delta \bar{\mathbf{u}}_{,\gamma\alpha}^1 \delta \xi^\alpha + \bar{\mathbf{x}}_{,\gamma\alpha\beta}^1 \delta \xi^\alpha \Delta \xi^\beta + \bar{\mathbf{x}}_{,\gamma\alpha}^1 \Delta \delta \xi^\alpha. \end{aligned} \quad (5.115)$$

With this result all quantities in (5.113) are known. These can be combined for the final result which obtains a nice structure when the definition of the gap function  $g_N \bar{\mathbf{n}}^1 = \mathbf{x}^2 - \bar{\mathbf{x}}^1$  is used. Combining the fourth and the fifth terms on the right-hand side of (5.113) with the last two components in (5.114) yields

$$\begin{aligned} &\delta g_N \Delta \bar{\mathbf{n}}^1 + \Delta g_N \delta \bar{\mathbf{n}}^1 - g_N (\delta \bar{\mathbf{a}}_\gamma^1 \cdot \Delta \bar{\mathbf{n}}^1 + \Delta \bar{\mathbf{a}}_\gamma^1 \cdot \delta \bar{\mathbf{n}}^1) \\ &= -[\delta(g_N \bar{\mathbf{n}}^1) \cdot \Delta \bar{\mathbf{a}}_\gamma + \Delta(g_N \bar{\mathbf{n}}^1) \cdot \delta \bar{\mathbf{a}}_\gamma] \\ &= -[\delta(\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \Delta \bar{\mathbf{a}}_\gamma + \Delta(\mathbf{x}^2 - \bar{\mathbf{x}}^1) \cdot \delta \bar{\mathbf{a}}_\gamma] \\ &= -(\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot \Delta \bar{\mathbf{a}}_\gamma - (\Delta \mathbf{u}^2 - \Delta \bar{\mathbf{u}}^1) \cdot \delta \bar{\mathbf{a}}_\gamma + \bar{\mathbf{x}}_\alpha^1 \delta \xi^\alpha \cdot \Delta \bar{\mathbf{a}}_\gamma + \bar{\mathbf{x}}_\alpha^1 \Delta \xi^\alpha \cdot \delta \bar{\mathbf{a}}_\gamma \end{aligned} \quad (5.116)$$

All results obtained so far can be substituted into (5.113), which after some algebra and grouping of all terms with  $\Delta \delta \xi^\alpha$  leads on the left-hand side to an expression for the linearization of the variation of  $\xi^\alpha$ :

$$\begin{aligned} \Delta \delta \xi^\alpha &= [-\bar{\mathbf{a}}_\beta^1 \cdot (\delta \xi^\gamma \Delta \bar{\mathbf{u}}_{,\gamma}^1 + \bar{\boldsymbol{\eta}}_{,\gamma}^1 \Delta \xi^\gamma) - (\bar{\mathbf{a}}_\beta^1 \cdot \bar{\mathbf{a}}_{\gamma,\theta}^1 - g_N \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{a}}_{\beta,\gamma\theta}^1) \delta \xi^\gamma \Delta \xi^\theta \\ &\quad + g_N (\bar{\boldsymbol{\eta}}_{,\beta\gamma}^1 \Delta \xi^\gamma + \Delta \bar{\mathbf{u}}_{,\beta\gamma}^1 \delta \xi^\gamma) \cdot \bar{\mathbf{n}}^1 \\ &\quad - (\bar{\boldsymbol{\eta}}_{,\beta}^1 + \bar{\mathbf{a}}_{\beta,\gamma}^1 \delta \xi^\gamma) \cdot \bar{\mathbf{a}}_\theta^1 \Delta \xi^\theta - (\Delta \bar{\mathbf{u}}_{,\beta}^1 + \bar{\mathbf{a}}_{\beta,\gamma}^1 \Delta \xi^\gamma) \cdot \bar{\mathbf{a}}_\theta^1 \delta \xi^\theta \\ &\quad + (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot (\Delta \bar{\mathbf{u}}_{,\beta}^1 + \bar{\mathbf{a}}_{\beta,\gamma}^1 \Delta \xi^\gamma) \\ &\quad + (\Delta \mathbf{u}^2 - \Delta \bar{\mathbf{u}}^1) \cdot (\bar{\boldsymbol{\eta}}_{,\beta}^1 + \bar{\mathbf{a}}_{\beta,\gamma}^1 \delta \xi^\gamma)] \bar{H}^{\alpha\beta}. \end{aligned} \quad (5.117)$$

This expression is very complex since, additionally, the quantities  $\delta \xi^\alpha$  and  $\Delta \xi^\alpha$  also have to be inserted as functions of the incremental displacements  $\Delta \mathbf{u}^\alpha$  and the variations  $\boldsymbol{\eta}^\alpha$ , see (3.21) and (5.98). Furthermore, the tensor  $\bar{H}^{\alpha\beta}$  is the inverse of  $\bar{H}_{\alpha\beta}$  defined in (5.98).

**REMARK 5.7**

1. The term  $\Delta \delta \xi^\alpha$  simplifies for flat geometries which occur when linear interpolations are applied to discretize the contact terms. One derives from (5.117)

$$\begin{aligned} \Delta \delta \xi^\alpha &= [-\bar{\mathbf{a}}_\beta^1 \cdot (\delta \xi^\gamma \Delta \bar{\mathbf{u}}_{,\gamma}^1 + \bar{\boldsymbol{\eta}}_{,\gamma}^1 \Delta \xi^\gamma) - \bar{\boldsymbol{\eta}}_{,\beta}^1 \cdot \bar{\mathbf{a}}_\theta^1 \Delta \xi^\theta - \Delta \bar{\mathbf{u}}_{,\beta}^1 \cdot \bar{\mathbf{a}}_\theta^1 \delta \xi^\theta \\ &\quad + (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot \Delta \bar{\mathbf{u}}_{,\beta}^1 + (\Delta \mathbf{u}^2 - \Delta \bar{\mathbf{u}}^1) \cdot \bar{\boldsymbol{\eta}}_{,\beta}^1] \bar{a}^{\alpha\beta} \end{aligned} \quad (5.118)$$

which omits a lot of terms.

2. In the two-dimensional situation, most of the terms in (5.117) remain in the linearization of the variation of the gap function

$$\Delta \delta \xi = [-\bar{\mathbf{a}}^1 \cdot (\delta \xi \Delta \bar{\mathbf{u}}_{,\xi}^1 + \bar{\boldsymbol{\eta}}_{,\xi}^1 \Delta \xi) - (\bar{\mathbf{a}}^1 \cdot \bar{\mathbf{a}}_{,\xi}^1 - g_N \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{a}}_{,\xi\xi}^1) \delta \xi \Delta \xi]$$

$$\begin{aligned}
& +g_N (\bar{\eta}_{,\xi\xi}^1 \Delta\xi + \Delta\bar{u}_{,\xi\xi}^1 \delta\xi) \cdot \bar{n}^1 \\
& -(\bar{\eta}_{,\xi}^1 + \bar{a}_{,\xi}^1 \delta\xi) \cdot \bar{a}^1 \Delta\xi - (\Delta\bar{u}_{,\xi}^1 + \bar{a}_{,\xi}^1 \Delta\xi) \cdot \bar{a}^1 \delta\xi \\
& +(\eta^2 - \bar{\eta}^1) \cdot (\Delta\bar{u}_{,\xi}^1 + \bar{a}_{,\xi}^1 \Delta\xi) + (\Delta u^2 - \Delta\bar{u}^1) \cdot (\bar{\eta}_{,\xi}^1 + \bar{a}_{,\xi}^1 \delta\xi) \Big] \bar{H}^{11}.
\end{aligned} \tag{5.119}$$

3. However, in the two-dimensional flat case a lot of terms drop out. This leads to

$$\begin{aligned}
\Delta\delta\xi = & \left[ -2\bar{a}^1 \cdot (\delta\xi \Delta\bar{u}_{,\xi}^1 + \bar{\eta}_{,\xi}^1 \Delta\xi) \right. \\
& \left. +(\eta^2 - \bar{\eta}^1) \cdot \Delta\bar{u}_{,\xi}^1 + (\Delta u^2 - \Delta\bar{u}^1) \cdot \bar{\eta}_{,\xi}^1 \right] \frac{1}{l^2},
\end{aligned} \tag{5.120}$$

where the terms  $\Delta\xi$  and  $\delta\xi$  are computed as in (5.107).

### 5.5.3 Special case of stick

When the formulation discussed in Remarks 5.2 and 5.3 is used to enforce the constraint condition for stick, then we do not have to distinguish between the normal and tangential directions in the contact interface, and the constraint equation is given by (5.29):  $\mathbf{g} = \mathbf{x}^2 - \mathbf{x}^1(\bar{\xi}) = \mathbf{x}^2 - \bar{\mathbf{x}}^1 = \mathbf{0}$ . This leads to the LAGRANGE multiplier formulation for stick

$$\Pi_c^{LM} = \int_{\Gamma_c} \lambda \cdot \mathbf{g} dA. \tag{5.121}$$

Since the stick condition requires that point  $\mathbf{x}^2$  remains at the same position on the master surface, the convective coordinates does not change:  $\delta\xi^\alpha = \text{const.}$  Hence the variation of the stick condition (5.29) is simply given by

$$\delta\mathbf{g} = \eta^2 - \bar{\eta}^1, \tag{5.122}$$

where the variation of  $\bar{\xi}^\alpha$  can be neglected. From  $\Pi_c^{LM}$  we obtain the associated weak form

$$C_c^{LM} = \int_{\Gamma_c} (\delta\lambda \cdot \mathbf{g} + \lambda \cdot \delta\mathbf{g}) dA. \tag{5.123}$$

The linearization of this term is then given by

$$\Delta C_c^{LM} = \int_{\Gamma_c} (\delta\lambda \cdot \Delta\mathbf{g} + \Delta\lambda \cdot \delta\mathbf{g}) dA, \tag{5.124}$$

where  $\Delta\delta\mathbf{g} = \mathbf{0}$  and  $\Delta\mathbf{g}$  has the same structure as (5.122):

$$\Delta\mathbf{g} = \Delta\mathbf{u}^2 - \Delta\bar{\mathbf{u}}^1. \tag{5.125}$$

All other terms are zero, which is a considerable simplification regarding all complicated expressions in the last two sections.

In the same way, for the penalty formulation in Remark 5.3 from (5.34), we obtain

$$\Delta C_c^{\text{stick}} = \int_{\Gamma_c} \epsilon (\Delta\mathbf{u}^2 - \Delta\bar{\mathbf{u}}^1) \cdot (\eta^2 - \bar{\eta}^1) dA. \tag{5.126}$$

Also, this expression is very simple, and should be used to formulate stick conditions within the penalty method.

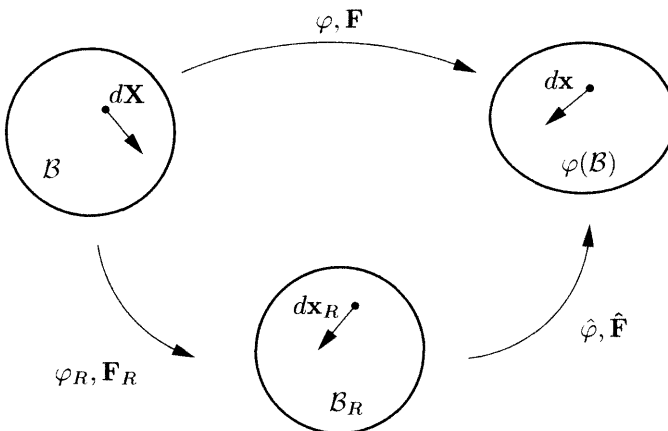
## 5.6 ROLLING CONTACT

Rolling contact has its technical application for the interaction between railroad wheels and train tracks and for the analysis of car or lorry tyres. In both cases one has to distinguish between stationary and non-stationary loading and response. Due to the technical importance, many contributions regarding rolling contact can be found in the literature. Railroad wheels in contact with the track were investigated in detail in Kalker (1990) using special algorithms and formulations based on linear elasticity. Finite element treatment based on the ALE-formulation for rolling contact problems can be found in e.g. Padovan and Zeid (1984), Oden and Lin (1986), Tallec and Rahier (1994), Nackenhorst (2000) or Hu and Wriggers (2002) for general applications.

Since the ALE-formulation introduces a new reference configuration, all equations developed so far for contact problems have to be reformulated.

### 5.6.1 Special reference frames for rolling contact

In the case of rolling contact, it can be useful not to apply the LAGRANGE description as given before, but to use a special rotating reference frame. From the continuum point of view, we distinguish between three different configurations of the body, see Figure 5.6.  $\mathcal{B}$  denotes the initial configuration of the body. With  $\varphi(\mathcal{B})$  we describe the deformed configuration. Finally,  $\varphi_R(\mathcal{B})$  is the rotating reference configuration



**Fig. 5.6** Reference frames to describe the deformation.



which rotates with a given angular velocity  $\omega_R$ . Thus, for the deformation map we have

$$\varphi = \hat{\varphi} \otimes \varphi_R \quad (5.127)$$

where  $\hat{\varphi}$  is the deformation relative to the rotation frame. For the line elements – needed later to define the strain measures – we have

$$\begin{aligned} d\mathbf{x} &= \mathbf{F} d\mathbf{X} = \hat{\mathbf{F}} d\mathbf{x}_R \\ d\mathbf{x}_R &= \mathbf{F}_R d\mathbf{X} \end{aligned} \quad (5.128)$$

with  $\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}$ ,  $\hat{\mathbf{F}} = \frac{\partial \mathbf{x}}{\partial \mathbf{x}_R}$ ,  $\mathbf{F}_R = \frac{\partial \mathbf{x}_R}{\partial \mathbf{X}}$ . Hence, from this a multiplicative decomposition of the deformation gradients related to the different configurations follows:

$$\boxed{d\mathbf{x} = \hat{\mathbf{F}} \mathbf{F}_R d\mathbf{X}}. \quad (5.129)$$

If the reference configuration  $\varphi_R(\mathcal{B})$  is given by a pure spinning rotation, then we can write

$$\mathbf{x}_R = \mathbf{R} \mathbf{X}, \quad (5.130)$$

with the orthogonal matrix  $\mathbf{R}$  describing the rotation. Thus

$$\boxed{d\mathbf{x} = \hat{\mathbf{F}} \mathbf{R} d\mathbf{X}}. \quad (5.131)$$

The velocity is defined by the material time derivative  $\mathbf{v} = \frac{d\varphi}{dt}$ . Using the special decomposition of the motion into the rotating frame  $\varphi_R$  and the deformation relative to this rotating reference configuration  $\hat{\varphi}$ , see (5.127), with  $\mathbf{x} = \varphi(\mathbf{X}, t) = \hat{\varphi}(\varphi_R(\mathbf{X}, t), t)$  by the chain rule, one obtains

$$\boxed{\mathbf{v} = \frac{d\varphi}{dt} = \frac{\partial \hat{\varphi}}{\partial t} + \frac{\partial \hat{\varphi}}{\partial \varphi_R} \frac{d\varphi_R}{dt}}. \quad (5.132)$$

This equation can be rewritten with (5.128), leading to

$$\mathbf{v} = \frac{d\varphi}{dt} = \frac{\partial \hat{\varphi}}{\partial t} + \hat{\mathbf{F}} \mathbf{v}_R, \quad (5.133)$$

where the velocity  $\mathbf{v}_R$  is defined by  $\mathbf{v}_R = \frac{d\varphi_R}{dt}$ . For computation of the acceleration, the chain rule has to be applied again to (5.132), which yields

$$\mathbf{a} = \frac{d^2 \varphi}{dt^2} = \frac{\partial^2 \hat{\varphi}}{\partial t^2} + 2 \frac{\partial^2 \hat{\varphi}}{\partial \varphi_R \partial t} \frac{d\varphi_R}{dt} + \frac{\partial^2 \hat{\varphi}}{\partial \varphi_R^2} \left[ \frac{d\varphi_R}{dt} \right]^2 + \frac{\partial \hat{\varphi}}{\partial \varphi_R} \frac{d^2 \varphi_R}{dt^2}. \quad (5.134)$$

This equation can be shortened using (5.128), which leads to

$$\boxed{\mathbf{a} = \frac{\partial^2 \hat{\varphi}}{\partial t^2} + 2 \dot{\hat{\mathbf{F}}} \mathbf{v}_R + \frac{\partial}{\partial \varphi_R} \hat{\mathbf{F}} \mathbf{v}_R^2 + \hat{\mathbf{F}} \mathbf{a}_R}, \quad (5.135)$$

where  $\mathbf{a}_R$  is defined by  $\mathbf{a}_R = \frac{d^2 \varphi_R}{dt^2}$ . In the case of a steady state spinning process,  $\hat{\varphi}$  is not explicitly time-dependent, which simplifies (5.132) and (5.134) considerably:

$$\begin{aligned} \mathbf{v} &= \frac{\partial \hat{\varphi}}{\partial \varphi_R} \frac{d\varphi_R}{dt} = \hat{\mathbf{F}} \mathbf{v}_R, \\ \mathbf{a} &= \frac{\partial^2 \hat{\varphi}}{\partial \varphi_R^2} \left[ \frac{d\varphi_R}{dt} \right]^2 + \frac{\partial \hat{\varphi}}{\partial \varphi_R} \frac{d^2 \varphi_R}{dt^2} = \frac{\partial}{\partial \varphi_R} \hat{\mathbf{F}} \mathbf{v}_R^2 + \hat{\mathbf{F}} \mathbf{a}_R. \end{aligned} \quad (5.136)$$

In this case, the reference frame has a constant angular velocity  $\boldsymbol{\omega}_R$  with respect to a given axis. This defines a rigid body motion of the reference frame. Hence, using (5.130) for the velocity  $\mathbf{v}_R$  in the reference frame, one derives

$$\mathbf{v}_R = \frac{\partial \mathbf{x}_R}{\partial t} + \boldsymbol{\Omega}_R \mathbf{x}_R, \quad (5.137)$$

with  $\mathbf{X} = \mathbf{R}^{-1} \mathbf{x}_R = \mathbf{R}^T \mathbf{x}_R$  and  $\dot{\mathbf{R}} \mathbf{R}^T = \boldsymbol{\Omega}_R$ . The associated acceleration  $\mathbf{a}_R$  is then computed from

$$\mathbf{a}_R = \frac{\partial^2 \mathbf{x}_R}{\partial t^2} + \dot{\boldsymbol{\Omega}}_R \mathbf{x}_R + \boldsymbol{\Omega}_R \left( \frac{\partial \mathbf{x}_R}{\partial t} + \boldsymbol{\Omega}_R \mathbf{x}_R \right). \quad (5.138)$$

To the skew symmetric tensor  $\boldsymbol{\Omega}_R$ , one can associate an axial vector such that

$$\mathbf{v}_R = \frac{\partial \mathbf{x}_R}{\partial t} + \boldsymbol{\omega}_R \times \mathbf{x}_R. \quad (5.139)$$

The first term disappears for a constant time-independent rotation. Combining (5.136) and (5.137) yields

$$\mathbf{v} = \hat{\mathbf{F}} \boldsymbol{\Omega}_R \mathbf{x}_R. \quad (5.140)$$

Furthermore, equations (5.136) and (5.138) then lead to the acceleration

$$\mathbf{a} = \left( \frac{\partial}{\partial \varphi_R} \hat{\mathbf{F}} \boldsymbol{\Omega}_R \mathbf{x}_R + \hat{\mathbf{F}} \boldsymbol{\Omega}_R \right) \boldsymbol{\Omega}_R \mathbf{x}_R. \quad (5.141)$$

## 5.6.2 Strain measures

With respect to Section 5.6.1, one has to investigate how the strain measures, developed in Section 2.1.2, are affected by the rotating reference configuration introduced. For this purpose, (5.128) is applied, which gives  $\mathbf{F} = \hat{\mathbf{F}} \mathbf{R}$ . This result can be used in (2.15) to compute the right CAUCHY GREEN strains,

$$\mathbf{C} = (\hat{\mathbf{F}} \mathbf{R})^T \hat{\mathbf{F}} \mathbf{R} = \mathbf{R}^T \hat{\mathbf{F}}^T \hat{\mathbf{F}} \mathbf{R} = \mathbf{R}^T \hat{\mathbf{C}} \mathbf{R}. \quad (5.142)$$

Thus, by the choice of the rotating reference configuration, the strains,  $\hat{\mathbf{C}}$ , producing stresses exclusively stem from the motion relative to the rotating frame. By inserting

the kinematical relations developed for the rotated reference frame, the GREEN-LAGRANGIAN strains follow from (5.142):

$$\mathbf{E} = \frac{1}{2} (\mathbf{R}^T \hat{\mathbf{C}} \mathbf{R} - \mathbf{1}) = \mathbf{R}^T \frac{1}{2} (\hat{\mathbf{C}} - \mathbf{1}) \mathbf{R} = \mathbf{R}^T \hat{\mathbf{E}} \mathbf{R}, \quad (5.143)$$

which have the same structure as (5.142).

### 5.6.3 Weak Form

For numerical simulations the weak form of a contact problem has to be formulated in such a way that leads to the most efficient solution scheme. In the case of rolling contact a formulation with respect to the rotating reference configuration  $\mathcal{B}_R$  is optimal. Due to that the weak form (2.59) is transformed to the rotating reference configuration

$$\begin{aligned} \int_{\mathcal{B}_R} \mathbf{F} \mathbf{S} \cdot \text{Grad } \boldsymbol{\eta} dV &= \int_{\mathcal{B}_R} \hat{\mathbf{F}} \mathbf{R} \mathbf{S} \cdot \widehat{\text{Grad}} \boldsymbol{\eta} \mathbf{R} dV_R \\ &\quad \int_{\mathcal{B}_R} \hat{\mathbf{F}} \hat{\mathbf{S}} \cdot \widehat{\text{Grad}} \boldsymbol{\eta} dV_R, \end{aligned} \quad (5.144)$$

where the rotated second PIOLA-KIRCHHOFF stress tensor  $\hat{\mathbf{S}}$  has been defined according to  $\hat{\mathbf{S}} = \mathbf{R} \mathbf{S} \mathbf{R}^T$ . Furthermore, the inertia term yields, for a constant spinning motion,

$$\begin{aligned} \int_{\mathcal{B}} \rho_0 \dot{\mathbf{v}} \cdot \boldsymbol{\eta} dV &= \int_{\mathcal{B}_R} \rho_R \boldsymbol{\Omega}_R^2 \mathbf{x}_R \cdot \boldsymbol{\eta} dV_R \\ &\quad - \int_{\mathcal{B}_R} \rho_R (\hat{\mathbf{F}} \boldsymbol{\Omega}_R \mathbf{x}_R) \cdot (\widehat{\text{Grad}} \boldsymbol{\eta} \boldsymbol{\Omega}_R \mathbf{x}_R) dV_R, \end{aligned} \quad (5.145)$$

where the first term denotes the body forces due to spinning. The second term is associated with the inertia forces due to constant spinning.

Using (5.144) and (5.145) in (2.59), we obtain the weak form with respect to the rotating frame

$$\begin{aligned} G(\hat{\boldsymbol{\varphi}}, \boldsymbol{\eta}) &= - \int_{\mathcal{B}_R} [\rho_R (\hat{\mathbf{F}} \boldsymbol{\Omega}_R \mathbf{x}_R) \cdot (\widehat{\text{Grad}} \boldsymbol{\eta} \boldsymbol{\Omega}_R \mathbf{x}_R) - \hat{\mathbf{F}} \hat{\mathbf{S}} \cdot \widehat{\text{Grad}} \boldsymbol{\eta}] dV_R \\ &\quad - \int_{\mathcal{B}_R} \rho_R (\bar{\mathbf{b}} - \boldsymbol{\Omega}_R^2 \mathbf{x}_R) \cdot \boldsymbol{\eta} dV_R - \int_{\partial \mathcal{B}_{\sigma_R}} \bar{\mathbf{t}} \cdot \boldsymbol{\eta} dA_R = 0. \end{aligned}$$

(5.146)

The linearization of the weak form, defined in quantities of the rotated frame, follows from (5.146). Since  $\mathbf{R}$  and  $\boldsymbol{\Omega}_R$  are constant, we obtain the same result as in (2.122). However, there is one extra term stemming from the inertia forces, the first term in (5.146). The linearization of this term is trivial since it is linear in the relevant

deformation  $\hat{\varphi}$ . In total, for the linearization at the known state  $\hat{\varphi}$  we derive:

$$\begin{aligned} DG(\hat{\varphi}, \eta) \cdot \Delta \mathbf{u} = & \int_{\mathcal{B}_R} \{ \widehat{\text{Grad}} \Delta \mathbf{u} \tilde{\mathbf{S}} \cdot \widehat{\text{Grad}} \eta + \delta \tilde{\mathbf{E}} \cdot \tilde{\mathbf{C}} [\Delta \tilde{\mathbf{E}}] \} dV_R \\ & - \int_{\mathcal{B}_R} \rho_R (\widehat{\text{Grad}} \Delta \mathbf{u} \Omega_R \mathbf{x}_R) \cdot (\widehat{\text{Grad}} \eta \Omega_R \mathbf{x}_R) dV_R. \end{aligned} \quad (5.147)$$

### 5.6.4 Constitutive equation

Due to the introduction of the rotating reference frame, a special treatment is needed for constitutive equations which depend upon history variables, like viscoelastic or elastoplastic models. The difference between a purely elastic and an inelastic constitutive law in a rotating reference configuration is shown in the next two sections.

**Elastic response function.** The constitutive equation which represents the hyperelastic response of the solid (2.74) can be transformed into the rotated frame which is applied for rolling contact. With  $\hat{\mathbf{S}} = \mathbf{R} \mathbf{S} \mathbf{R}^T$ ,  $\mathbf{C} = \mathbf{R}^T \hat{\mathbf{C}} \mathbf{R}$  and  $\mathbf{C}^{-1} = \mathbf{R}^T \hat{\mathbf{C}}^{-1} \mathbf{R}$ , one can write (2.74) directly in terms of the stress  $\hat{\mathbf{S}}$  and the strain  $\hat{\mathbf{C}}$  as follows:

$$\hat{\mathbf{S}} = \frac{\Lambda}{2} (J^2 - 1) \hat{\mathbf{C}}^{-1} + \mu (\mathbf{1}, -\hat{\mathbf{C}}^{-1}) \quad (5.148)$$

since the rotation tensor multiplies the tensors on both sides of (2.74) in the same way, and thus cancels out. Hence, no special treatment for the elastic constitutive equation is needed in the case of rolling contact. This is also true for the incremental constitutive tensor. One only has to exchange  $\mathbf{C}$  by  $\hat{\mathbf{C}}$  in (2.89) to compute the associated incremental constitutive tensor in  $\mathcal{B}_R$ .

**Viscoelastic response function.** Car tyres, which are often investigated using rolling contact formulations, are made of rubber which not only responds to deformations like a hyperelastic material, but also shows inelastic behaviour which can be described by a viscoelastic material model. Here again, one has to investigate whether a description using the rotated frame does change the constitutive equation or not (see above). Using a standard viscoelastic model, see Christensen (1980), we obtain for the second PIOLA-KIRCHHOFF stresses,

$$\mathbf{S}(t) = \bar{\mathbf{S}}^e[\mathbf{E}(t)] + \nu \int_{-\infty}^t e^{-\frac{t-s}{\tau}} \frac{\partial \mathbf{E}}{\partial s} ds. \quad (5.149)$$

Here  $t$  denotes the time,  $\mathbf{E}$  is the GREEN-LAGRANGIAN strain tensor,  $\bar{\mathbf{S}}^e$  is the hyperelastic response function (see last section) and  $\nu$  and  $\tau$  are constitutive parameters. All quantities in (5.149) refer to the initial configuration  $\mathcal{B}$ . Transformation with respect to the rotating reference configuration with  $\hat{\mathbf{S}} = \mathbf{R} \mathbf{S} \mathbf{R}^T$  and  $\mathbf{E} = \mathbf{R}^T \hat{\mathbf{E}} \mathbf{R}$

yields

$$\hat{\mathbf{S}}(t) = \bar{\mathbf{S}}^e[\hat{\mathbf{E}}(t)] + \nu \int_{-\infty}^t e^{-\frac{t-s}{\tau}} \frac{d}{ds} [\mathbf{R}(t) \mathbf{R}^T(s) \hat{\mathbf{E}}(s) \mathbf{R}(s) \mathbf{R}^T(t)] ds. \quad (5.150)$$

In this model, the rotation  $\mathbf{R}$  depends upon the complete motion, and hence its history has to be known during the whole simulation. However, the rotation only appears as  $\mathbf{R}(t) \mathbf{R}^T(s)$ , which means that only the relative rotation between the two times  $t$  and  $s$  is involved. Since we assume steady state rotation, the relative rotation is explicitly given by  $e^{\boldsymbol{\Omega}(t-s)}$ . For a more detailed treatment of viscoelastic constitutive equations for rolling motions, see Tallec and Rahier (1994) or Govindjee and Mihalic (1998).

### 5.6.5 Contact kinematics

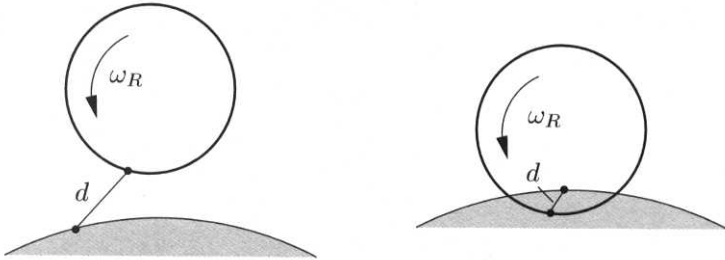
Relations for contact kinematics have to be derived based on the rotating reference frame. This leads, especially for tangential contact, to different formulations, when compared to the results presented in Chapter 3.

**Normal contact.** To formulate the non-penetration condition for rolling contact, we define a minimum distance problem between the rolling object and the rigid surface, see Figure 5.7. The reference configuration which is used to define the contact conditions can be either the rotating frame or the initial configuration. In cases when we have to distinguish both, we use the reference to  $\mathcal{B}$  or  $\mathcal{B}_R$ .

For a mathematical description of the problem it is useful to introduce convective coordinates  $\boldsymbol{\xi} = (\xi^1, \xi^2)$  on the surface with which the rolling object is in contact, from now on called the *master* surface. This leads to the definition of the master surface, described by the position vector  $\mathbf{X}_0 = \mathbf{X}_0(\xi^1, \xi^2)$ . Now for every point  $x$  on the deformed boundary of the rolling object we can define the minimum distance problem:

$$d = \text{MIN} \|\mathbf{x} - \mathbf{X}_0(\xi^1, \xi^2)\|. \quad (5.151)$$

The solution of this problem provides the pair of convective coordinates,  $\bar{\boldsymbol{\xi}}$ , which denotes the point on the master surface closest to point  $\mathbf{x}$ .  $\bar{\boldsymbol{\xi}}$  follows from the condition,



**Fig. 5.7** (a) Non-penetration, (b) penetration.

see also (3.3),

$$\frac{d}{d\xi^\alpha} d = \frac{1}{\|\mathbf{x} - \mathbf{X}_0(\xi^1, \xi^2)\|} [\mathbf{x} - \mathbf{X}_0(\xi^1, \xi^2)] \cdot \mathbf{X}_{0,\alpha}(\xi^1, \xi^2) = 0, \quad (5.152)$$

which is the closest point projection of point  $\mathbf{x}$  onto the master surface. Since  $\bar{\mathbf{X}}_{0,\alpha} = \mathbf{X}_{0,\alpha}(\bar{\xi})$  are the tangent vectors to the convective coordinates of the master surface at the solution point,  $\bar{\xi}$ , the vector  $\mathbf{x} - \bar{\mathbf{X}}_0(\bar{\xi})$  points in a direction normal to the master surface. Hence it can be used to define the non-penetration condition.

With

$$g_N = [\mathbf{x} - \bar{\mathbf{X}}_0(\bar{\xi})] \cdot \bar{\mathbf{N}}_0(\bar{\xi}) \quad (5.153)$$

we define the gap function where

$$\begin{aligned} \bar{\mathbf{N}}_0(\bar{\xi}) &= \frac{\mathbf{x} - \bar{\mathbf{X}}_0(\bar{\xi})}{\|\mathbf{x} - \bar{\mathbf{X}}_0(\bar{\xi})\|} \quad \text{or} \\ &= \frac{\bar{\mathbf{X}}_{0,1}(\bar{\xi}) \times \bar{\mathbf{X}}_{0,2}(\bar{\xi})}{\|\bar{\mathbf{X}}_{0,1}(\bar{\xi}) \times \bar{\mathbf{X}}_{0,2}(\bar{\xi})\|}. \end{aligned} \quad (5.154)$$

Both definitions can be used in (5.153), though the first is not well behaved for  $g_N \rightarrow 0$ . Hence the second condition should be applied in (5.153). Function  $g_N$  then describes the state at the interface as follows:

$$\begin{aligned} g_N > 0 & \quad \text{gap opening,} \\ g_N = 0 & \quad \text{perfect contact,} \\ g_N < 0 & \quad \text{penetration.} \end{aligned}$$

Thus contact is formulated by the inequality constraint

$$g_N \geq 0. \quad (5.155)$$

In case of a flat master surface which is often the case when rolling contact is considered we can simplify the representation of the master surface by cartesian coordinates. By defining the base vectors of the master surface as  $\mathbf{E}_1$  and  $\mathbf{E}_2$  we obtain for the normal  $\mathbf{N}_0 = \mathbf{E}_3$ . In this case the closest point projection (5.152) yields

$$[\mathbf{x} - \bar{\mathbf{X}}_0(X_1, X_2)] \cdot \mathbf{E}_3 = 0 \quad (5.156)$$

with the solution point  $(\bar{X}_1, \bar{X}_2)$ . Furthermore the gap is given by

$$g_N = [\mathbf{x} - \bar{\mathbf{X}}_0(\bar{X}_1, \bar{X}_2)] \cdot \mathbf{E}_3. \quad (5.157)$$

**Tangential contact.** The kinematical relations for the tangential motion in the contact area have lead to the definition of the relative tangential velocity. This quantity can be obtained in the case of rolling contact by the derivative of condition (5.152) with respect to time. This yields

$$(\mathbf{v} - \dot{\bar{\mathbf{X}}}_0) \cdot \bar{\mathbf{X}}_{0,\alpha} + (\mathbf{x} - \bar{\mathbf{X}}_0) \cdot \dot{\bar{\mathbf{X}}}_{0,\alpha} - \bar{\mathbf{X}}_{0,\beta} \dot{\bar{\xi}}^\beta \cdot \bar{\mathbf{X}}_{0,\alpha} = 0. \quad (5.158)$$

Here  $\mathbf{v}$  is the velocity associated with point  $x$ . With the metric  $\bar{A}_{\alpha\beta} = \bar{\mathbf{X}}_{0,\alpha} \cdot \bar{\mathbf{X}}_{0,\beta}$  the components of the relative gap velocity in the tangential direction are

$$\bar{A}_{\alpha\beta} \dot{\xi}^\beta = (\mathbf{v} - \bar{\mathbf{V}}_0) \cdot \bar{\mathbf{X}}_{0,\alpha} + (\mathbf{x} - \bar{\mathbf{X}}_0) \cdot \dot{\bar{\mathbf{X}}}_{0,\alpha}, \quad (5.159)$$

where we have set  $\bar{\mathbf{V}}_0 = \dot{\bar{\mathbf{X}}}_0$ .

In the case of a flat master surface, equation (5.159) simplifies with  $\dot{\bar{\mathbf{X}}}_{0,\alpha} = 0$ ,  $\bar{\mathbf{X}}_{0,\alpha} = \mathbf{E}_\alpha$  and  $\bar{A}_{\alpha\beta} = \delta_{\alpha\beta}$  to

$$\dot{X}_\alpha = (\mathbf{v} - \bar{\mathbf{V}}_0) \cdot \mathbf{E}_\alpha, \quad (5.160)$$

where  $\dot{X}_\alpha$  denotes the change in time of the projection point  $(\bar{X}_1, \bar{X}_2)$  on the master surface, and thus is the relative tangential velocity in the contact area. With the projection tensor  $\mathbf{P}_\perp = [\mathbf{E}_\alpha \otimes \mathbf{E}_\alpha]$  we can reformulate (5.160) as

$$\mathbf{v}_\perp = \mathbf{P}_\perp (\mathbf{v} - \mathbf{V}_0), \quad (5.161)$$

with  $\mathbf{v}_\perp = \dot{X}_\alpha \mathbf{E}_\alpha$  being the tangential relative velocity vector in  $\Gamma_c$ .

With these relations we can now formulate the tangential contact conditions. The first is the non-slip or stick condition

$$\mathbf{v}_\perp = 0, \quad (5.162)$$

and hence imposes a constraint on the relative tangential motion. It means that locally, the rotating object is rolling and not sliding on the surface.

If the tangential forces exceed a certain limit in  $\Gamma_c$ , then slip occurs. In that case, the associated relative tangential velocity follows from a constitutive relation. Classically Coulomb's law is applied to determine the slip velocity, however more complicated constitutive equations can also be used to model the frictional behaviour in the contact interface. These constitutive equations will be considered in the next section.

Generally slip as well as stick can occur in the contact area. Thus we can subdivide the contact area  $\Gamma_c$  into  $\Gamma_c^{slip} \cup \Gamma_c^{stick} = \Gamma_c$ . Within the numerical method applied to solve the rolling contact problem, the stick and the slip area have to be computed.

In some cases it might be necessary in the numerical implementation of rolling contact, e.g. see Nackenhorst (1993), to enforce the stick condition in a weak sense. Following Nackenhorst (1993) one can use a least square fit

$$\int_{\Gamma_c^{stick}} \mathbf{v}_\perp^2 d\Gamma \rightarrow MIN \implies \int_{\Gamma_c^{stick}} \mathbf{v}_\perp \cdot \delta \mathbf{v}_\perp d\Gamma = 0, \quad (5.163)$$

which leads with (5.160) and (5.162) to

$$\int_{\Gamma_c^{stick}} \mathbf{P}_\perp \delta \mathbf{v} \cdot \mathbf{P}_\perp (\mathbf{v} - \bar{\mathbf{V}}_0) d\Gamma = 0, \quad (5.164)$$

and hence yields a system of equations to determine the tangential components of  $\mathbf{v}$  in the stick interface  $\Gamma_c^{stick}$  which fulfil the stick condition (5.162) in the weak sense

$$\int_{\Gamma_c^{stick}} \delta \mathbf{v} \cdot \mathbf{P}_\perp \mathbf{v} d\Gamma = \int_{\Gamma_c^{stick}} \delta \mathbf{v} \cdot \mathbf{P}_\perp \bar{\mathbf{V}}_0 d\Gamma. \quad (5.165)$$

Note that if slip occurs in a steady state computation of rolling contact, then the dissipation due to the frictional forces in the slip zone has to be compensated for by a moment around the spinning axis of the rolling body to preserve stationary motion.

**Definition of creepage.** For further reference we also define another kinematical quantity which measures the creepage in the contact interface. Its definition is given by

$$\mathbf{s} = \frac{\dot{\mathbf{X}}_0 - \omega_R \times \mathbf{R}}{\|\dot{\mathbf{X}}_0\|}. \quad (5.166)$$

The creepage vector  $\mathbf{s}$  can be decomposed into a part which is related to the flattening of the rolling body

$$\mathbf{s}_F = \frac{\omega_R \times (\mathbf{r} - \mathbf{R})}{\|\dot{\mathbf{X}}_0\|}, \quad (5.167)$$

and a partial slip due to rolling in the contact area

$$\mathbf{s}_S = \frac{\dot{\mathbf{X}}_0 - \omega_R \times \mathbf{r}}{\|\dot{\mathbf{X}}_0\|}, \quad (5.168)$$

so that we have  $\mathbf{s} = \mathbf{s}_F + \mathbf{s}_S$ .



**This page intentionally left blank**

# 6

## *Discretization of the Continuum*

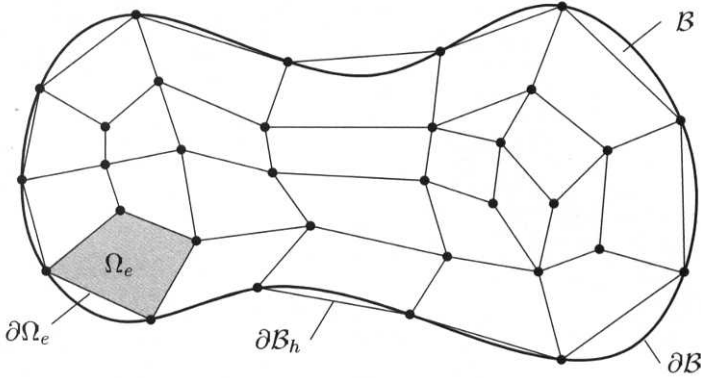
Discretization of the domain contributions of bodies in contact is not the objective of this work. For a detailed treatment with respect to the finite element implementations of boundary-value problems regarding large deformations, see Oden (1972), Crisfield (1991), Crisfield (1997), Zienkiewicz and Taylor (2000a), Bathe (1996) or Wriggers (2001) and the references therein. However, in this chapter, the discretization of continua undergoing large strains using isoparametric elements is discussed briefly for completeness.

Within the finite element method we have different approximations. These are geometrical approximations of the domain  $\mathcal{B}$  on which the boundary value problem is defined. Furthermore, the associated fields, deformations or stresses have to be approximated. Also, the integrals are not evaluated exactly, since, as they are evaluated for the weak form, they have to be computed via numerical integration procedures. Collectively, these approximations are sources for errors inherent in the finite element method. The estimation of these errors is the contents of Chapter 12.

In this section a description of the interpolations, which are the basis for a treatment using isoparametric elements, is given. Within this framework, we assume that the domain  $\mathcal{B}$  is discretized by  $n_e$  finite elements, which leads to its geometrical approximation  $\mathcal{B}^h$ :

$$\overline{\mathcal{B}} \approx \mathcal{B}^h = \bigcup_{e=1}^{n_e} \Omega_e. \quad (6.1)$$

The configuration of one element is  $\Omega_e \subset \mathcal{B}^h$ , as shown in Figure 6.1 for a two-dimensional case.  $\partial\mathcal{B}^h$  denotes the boundary of the discretization  $\mathcal{B}^h$ , which is in general also an approximation of the function describing the real boundary  $\partial\mathcal{B}$ .



**Fig. 6.1** Discretization of body  $B$ .

## 6.1 ISOPARAMETRIC CONCEPT

The finite element method requires that the field variables be approximated by a finite element  $\Omega_e$ . We write the displacement field  $\mathbf{u}(\mathbf{X})$  as

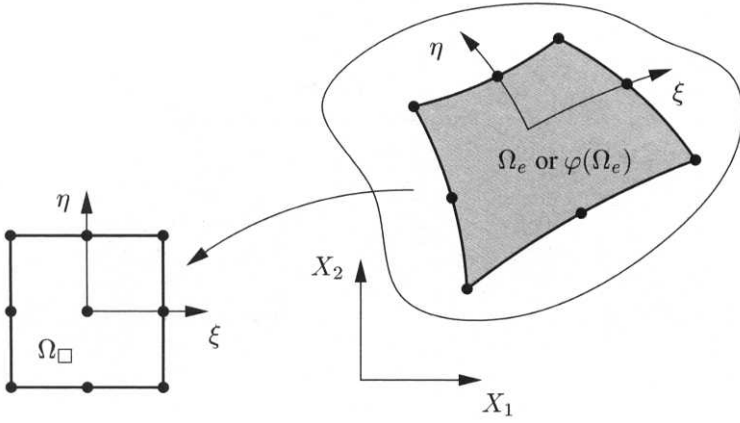
$$\mathbf{u}_{\text{exact}}(\mathbf{X}) \approx \mathbf{u}^h(\mathbf{X}) = \sum_{I=1}^N N_I(\mathbf{X}) \mathbf{u}_I, \quad (6.2)$$

where  $\mathbf{X}$  is the position vector in  $\Omega_e$ ,  $N_I(\mathbf{X})$  are interpolation or basis functions which are defined on  $\Omega_e$ , and  $\mathbf{u}_I$  denote the unknown nodal variables. Here,  $\mathbf{u}_I = \{u_1, u_2, u_3\}_I^T$  are the nodal variables of the displacement field.

During the development of finite element methods, many possibilities for interpolation of the unknown functions within an element have been exploited. Due to its general applicability, especially when arbitrary geometries have to be discretized, the isoparametric concept is widely used. On the isoparametric approach, we approximate geometry and field variables by the same interpolation functions, see Figure 6.2:

$$\mathbf{X}_e^h = \sum_{I=1}^n N_I(\xi) \mathbf{X}_I, \quad \text{and} \quad \mathbf{u}_e^h = \sum_{I=1}^n N_I(\xi) \mathbf{u}_I. \quad (6.3)$$

The interpolation functions in equations (6.3) of the element in  $B^h$  have been represented by interpolation functions  $N_I(\xi)$  defined on the reference element  $\Omega_{\square}$ , see Figure 6.2. Thus, for every element  $\Omega_e$ , there exists a transformation (6.3)<sub>1</sub> which relates the coordinates  $\mathbf{X}_e = \mathbf{X}_e(\xi)$  to the coordinates  $\xi$  of the reference element  $\Omega_{\square}$ . Hence all computations are performed with respect to the reference configuration. Only in very special cases do the initial and current configurations



**Fig. 6.2** Isoparametric mapping.

of a finite element coincide. However, this transformation is numerically easy to handle, and allows transformation of the reference element to arbitrary geometries. This feature leads to the fact that, in implementing of the method, there is almost no difference in the formulation of finite elements with respect to the current or initial configurations.

Figure 6.3 depicts the two possibilities to describe deformation in continuum mechanics using the isoparametric concept. It can be seen easily that Figure 6.3 is a discrete version of Figure 2.1, where additionally we have now introduced the reference configuration  $\Omega_\square$ . The kinematical relations within one element are

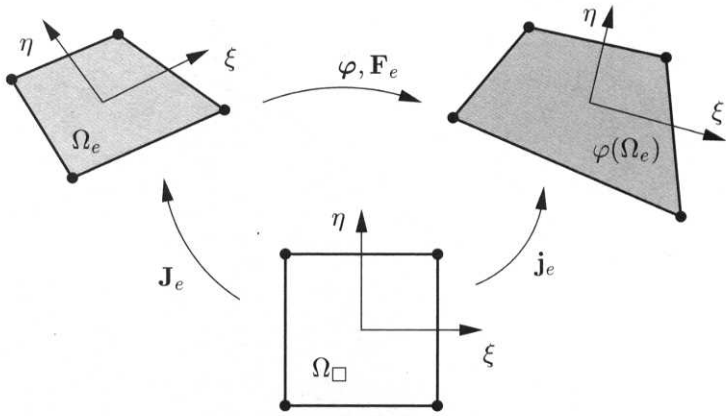
$$\mathbf{F}_e = \mathbf{j}_e \mathbf{J}_e^{-1} \quad \text{and} \quad J_e = \det \mathbf{F}_e = \frac{\det \mathbf{j}_e}{\det \mathbf{J}_e}, \quad (6.4)$$

which show that the deformation gradient is uniquely defined by the isoparametric mapping of  $\Omega_\square$  onto  $\Omega_e$  in the initial configuration, or onto  $\varphi(\Omega_e)$  in the current configuration. In these equations, the gradients  $\mathbf{j}_e$  and  $\mathbf{J}_e$  are defined as follows:

$$\begin{aligned} \mathbf{j}_e &= \text{Grad}_\xi \mathbf{x} = \frac{\partial \mathbf{x}}{\partial \xi} = \sum_{I=1}^n N_{I,\xi}(\xi) \mathbf{x}_I \otimes \mathbf{E}_\xi, \\ \mathbf{J}_e &= \text{Grad}_\xi \mathbf{X} = \frac{\partial \mathbf{X}}{\partial \xi} = \sum_{I=1}^n N_{I,\xi}(\xi) \mathbf{X}_I \otimes \mathbf{E}_\xi. \end{aligned} \quad (6.5)$$

Since the derivatives  $N_{I,\xi}$  are scalar quantities, we can move them in front of the base vectors  $\mathbf{E}_\xi$ . This yields

$$\mathbf{j}_e = \sum_{I=1}^n \mathbf{x}_I \otimes N_{I,\xi}(\xi) \mathbf{E}_\xi = \sum_{I=1}^n \mathbf{x}_I \otimes \nabla_\xi N_I,$$



**Fig. 6.3** Isoparametric description of deformations.

$$\mathbf{J}_e = \sum_{I=1}^n \mathbf{X}_I \otimes N_{I,\xi}(\boldsymbol{\xi}) \mathbf{E}_\xi = \sum_{I=1}^n \mathbf{X}_I \otimes \nabla_\xi N_I. \quad (6.6)$$

$\nabla_\xi N_I$  is the gradient of the scalar function  $N_I$  with respect to the coordinates  $\boldsymbol{\xi}$ .

With this, it is simple to compute gradients with respect to the initial or current configurations. For a vector field this reads as  $\mathbf{u}^h$ ,

$$\begin{aligned} \text{Grad } \mathbf{u}^h &= \sum_{I=1}^n \mathbf{u}_I \otimes \nabla_X N_I, \\ \text{grad } \mathbf{u}^h &= \sum_{I=1}^n \mathbf{u}_I \otimes \nabla_x N_I. \end{aligned} \quad (6.7)$$

Analogous to the transformation of the derivatives between different configurations, see (2.24), we obtain

$$\nabla_\xi N_I = \mathbf{J}_e^T \nabla_X N_I \quad \text{and} \quad \nabla_\xi N_I = \mathbf{j}_e^T \nabla_x N_I, \quad (6.8)$$

or the inverse relations

$$\nabla_X N_I = \mathbf{J}_e^{-T} \nabla_\xi N_I, \quad \text{and} \quad \nabla_x N_I = \mathbf{j}_e^{-T} \nabla_\xi N_I, \quad (6.9)$$

such that the gradient in (6.7) is completely defined in quantities which are defined in the reference configuration  $\Omega_\square$  as

$$\begin{aligned} \text{Grad } \mathbf{u} &= \sum_{I=1}^n \mathbf{u}_I \otimes \mathbf{J}_e^{-T} \nabla_\xi N_I, \\ \text{grad } \mathbf{u} &= \sum_{I=1}^n \mathbf{u}_I \otimes \mathbf{j}_e^{-T} \nabla_\xi N_I. \end{aligned} \quad (6.10)$$

The only difference in the formulation of both gradients in (6.10) lies in the exchange of the gradients  $\mathbf{j}_e$  and  $\mathbf{J}_e$ , and therefore this approach is advantageous, especially for large deformation finite element formulations.

### 6.1.1 Isoparametric interpolation functions

Within the different possibilities to construct interpolation functions for isoparametric elements, we here follow the concept of the LAGRANGE interpolation, e.g. see Zienkiewicz and Taylor (1989). For a LAGRANGE polynomial of power  $n-1$ , in the one-dimensional case we obtain

$$N_I(\xi) = \prod_{\substack{J=1 \\ J \neq I}}^n \frac{(\xi_J - \xi)}{(\xi_J - \xi_I)}. \quad (6.11)$$

For two- or three-dimensional interpolations, we choose a product formulation

$$N_J(\xi, \eta) = N_I(\xi) N_K(\eta) \quad \text{or} \quad N_J(\xi, \eta, \zeta) = N_I(\xi) N_K(\eta) N_L(\zeta), \quad (6.12)$$

with  $J = 1, \dots, n^{dim}$  and  $I, K, L = 1, \dots, n$  ( $dim$  is the spatial dimension of the problem). The interpolation or shape functions are defined in the local coordinate system  $\xi = \{\xi, \eta, \zeta\}$ .

In the next section we specify the isoparametric shape functions for one-, two- and three-dimensional problems.

### 6.1.2 One-dimensional shape functions

Here we briefly discuss one-dimensional shape functions which are  $C^0$ -continuous. These can be found in many places in the literature, and thus only the final equations which are needed in subsequent sections will be given.

In one-dimensional problems we only have one component of our field variables, thus we can write with (6.3)

$$X_e^h = \sum_{I=1}^n N_I(\xi) X_I, \quad u_e^h = \sum_{I=1}^n N_I(\xi) u_I, \quad (6.13)$$

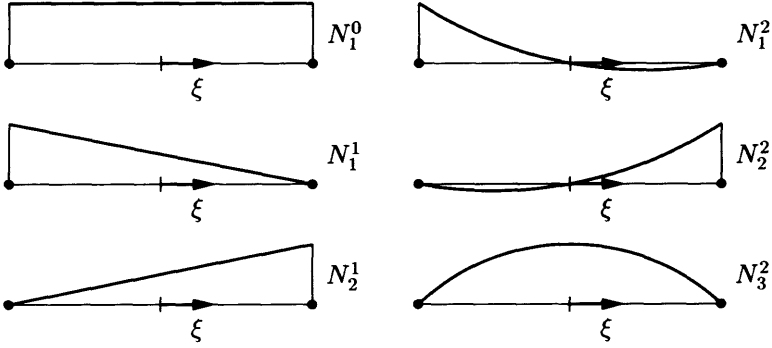
for coordinate  $X$  and the associated field variable  $u$ .  $n$  is the number of shape functions, and  $\xi \in [-1, 1]$  is the coordinate in the reference configuration, see Figure 6.4. The shape functions  $N_I(\xi)$  follow from (6.11) and are different with respect to the choice of the polynomial order. We have

- Constant shape functions

$$N_1(\xi) = 1. \quad (6.14)$$

- Linear shape functions

$$N_1(\xi) = \frac{1}{2}(1 - \xi), \quad N_2(\xi) = \frac{1}{2}(1 + \xi). \quad (6.15)$$



**Fig. 6.4** One-dimensional shape functions.

• Quadratic shape functions

$$N_1(\xi) = \frac{1}{2}\xi(\xi - 1), \quad N_3(\xi) = (1 - \xi^2), \quad N_2(\xi) = \frac{1}{2}\xi(1 + \xi). \quad (6.16)$$

The isoparametric mapping of function  $u$  onto the reference element is obtained with equation (6.13).

**Computation of derivatives.** For the computation of strains, its variations or linearizations, we need the derivatives of the field variable  $u$ . Within the isoparametric concept we have to use the chain rule

$$\frac{\partial u_e^h}{\partial X} = \sum_{I=1}^n \frac{\partial N_I(\xi)}{\partial X} u_I, \quad (6.17)$$

for the derivative of  $u$

$$\frac{\partial u_e^h}{\partial X} = \frac{\partial u_e^h}{\partial \xi} \frac{\partial \xi}{\partial X} = \left( \sum_{I=1}^n \frac{\partial N_I(\xi)}{\partial \xi} u_I \right) \frac{\partial \xi}{\partial X}. \quad (6.18)$$

The derivative  $\frac{\partial \xi}{\partial X}$  is computed with the interpolation for the geometry (6.13),

$$\frac{\partial \xi}{\partial X} = \left( \frac{\partial X}{\partial \xi} \right)^{-1} = \left( \sum_{I=1}^n \frac{\partial N_I(\xi)}{\partial \xi} X_I \right)^{-1} = J_e(\xi)^{-1}, \quad (6.19)$$

where we have made use of the abbreviation  $\frac{\partial X_e^h}{\partial \xi} = J_e$ .

In the special case of linear shape functions (6.15), with the length of an element ( $L_e = X_2 - X_1$ ) we obtain

$$\sum_{I=1}^n \frac{\partial N_I(\xi)}{\partial \xi} X_I = \frac{1}{2}(X_2 - X_1) = \frac{1}{2}L_e, \quad (6.20)$$

and inserting this into (6.18), the explicit formula

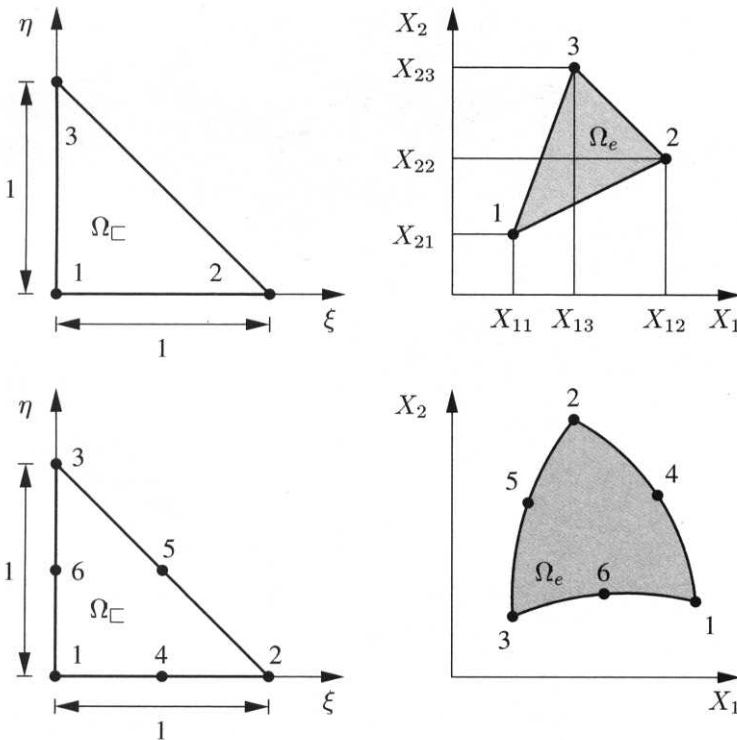
$$\frac{\partial u_e^h}{\partial X} = \frac{u_2 - u_1}{L_e}, \quad (6.21)$$

which yields a constant value.

### 6.1.3 Two-dimensional shape functions

In the two-dimensional case, quadrilateral and triangular finite elements have to be distinguished. Here we discuss  $C^0$ -continuous shape functions which are linear as well as quadratic.

First triangular elements are considered. The simplest element with linear shape functions consists of three nodes; an element with quadratic interpolation needs six nodes to define the fields and geometry within an element. In Figure 6.5 the triangular element is depicted for the quadratic interpolation. For a linear element only the vertices 1 to 3 are necessary to define the interpolation. The element is shown in



**Fig. 6.5** Three- and six node triangular element.



Figure 6.5 in its reference configuration  $\Omega_{\square}$ , denoted by the  $\xi$ - $\eta$  coordinates, and in its physical space, denoted by the  $X_1$ - $X_2$  coordinate system.

The shape functions for the linear case are defined by

$$N_1 = 1 - \xi - \eta, \quad N_2 = \xi, \quad N_3 = \eta. \quad (6.22)$$

Here all partial derivatives with respect to  $\xi$  and  $\eta$  are constant.

The shape functions for the quadratic element are

$$\begin{aligned} N_1 &= \lambda (2\lambda - 1), & N_4 &= 4\xi\lambda, \\ N_2 &= \xi (2\xi - 1), & N_5 &= 4\xi\eta, \\ N_3 &= \eta (2\eta - 1), & N_6 &= 4\eta\lambda, \end{aligned} \quad (6.23)$$

with the abbreviation  $\lambda = 1 - \xi - \eta$ .

Next the shape functions for quadrilateral elements are defined. The simplest quadrilateral has four nodes. The associated interpolation for geometry and field variables is bilinear. The product formula (6.12), together with the shape functions (6.15), yields

$$N_I(\xi, \eta) = \frac{1}{2} (1 + \xi_I \xi) \frac{1}{2} (1 + \eta_I \eta), \quad (6.24)$$

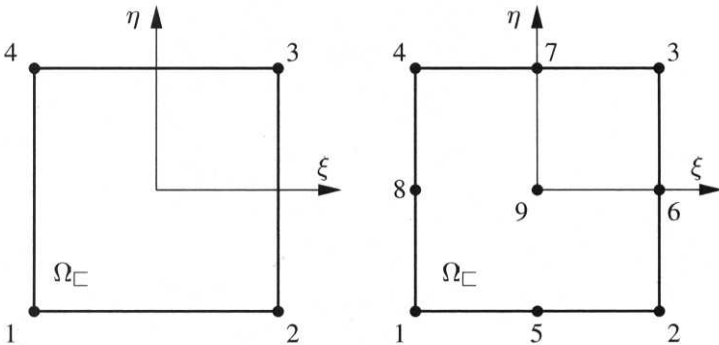
where the coordinates  $\xi_I$  and  $\eta_I$  are associated with the vertices (see Figure 6.6 on the reference element  $\Omega_{\square}$ )

$$\xi_1 = (-1, -1) \quad \xi_2 = (1, -1) \quad \xi_3 = (1, 1) \quad \xi_4 = (-1, 1). \quad (6.25)$$

The shape functions for the quadratic nine-node element again follow from the product formula (6.12) using the quadratic interpolation (6.16). For the nodes (see Figure 6.6), we obtain

- Vertices ( $I = 1, 2, 3, 4$ ):

$$N_I(\xi, \eta) = \frac{1}{4} (\xi^2 + \xi_I \xi) (\eta^2 + \eta_I \eta), \quad (6.26)$$



**Fig. 6.6** Isoparametric quadrilateral elements.

- Mid nodes ( $I = 5, 6, 7, 8$ ):

$$N_I(\xi, \eta) = \frac{1}{2} \xi_I^2 (\xi^2 + \xi_I \xi) (1 - \eta^2) + \frac{1}{2} \eta_I^2 (\eta^2 + \eta_I \eta) (1 - \xi^2). \quad (6.27)$$

- Central node ( $I = 9$ ):

$$N_9(\xi, \eta) = (1 - \xi^2)(1 - \eta^2). \quad (6.28)$$

It should be noted that this is not the only possibility to define these nine shape functions. Often a hierarchical formulation is used, e.g. see Zienkiewicz and Taylor (1989) or Bathe (1982).

The derivatives of the shape functions defined in the reference coordinates with respect to the coordinates in the physical space follow within the isoparametric concept by the chain rule

$$\frac{\partial \mathbf{u}_e^h}{\partial X_\alpha} = \sum_{I=1}^n \frac{\partial N_I(\xi, \eta)}{\partial X_\alpha} \mathbf{u}_I, \quad (\alpha = 1, 2). \quad (6.29)$$

Here the partial derivative of  $N_I$  with respect to  $X_\alpha$  is computed according to (6.9),

$$\nabla_X N_I = \begin{Bmatrix} N_{I,1} \\ N_{I,2} \end{Bmatrix} = \mathbf{J}_e^{-T} \begin{Bmatrix} N_{I,\xi} \\ N_{I,\eta} \end{Bmatrix}, \quad (6.30)$$

with the JACOBI matrix  $\mathbf{J}_e$  of an  $\Omega_e$  element for the transformation between the reference and initial configuration,

$$\mathbf{J}_e = \sum_{I=1}^n \mathbf{X}_I \otimes \nabla_\xi N_I = \sum_{I=1}^n \begin{Bmatrix} X_{1,I} \\ X_{2,I} \end{Bmatrix} \begin{Bmatrix} N_{I,\xi} \\ N_{I,\eta} \end{Bmatrix}^T = \begin{bmatrix} X_{1,\xi} & X_{1,\eta} \\ X_{2,\xi} & X_{2,\eta} \end{bmatrix},$$

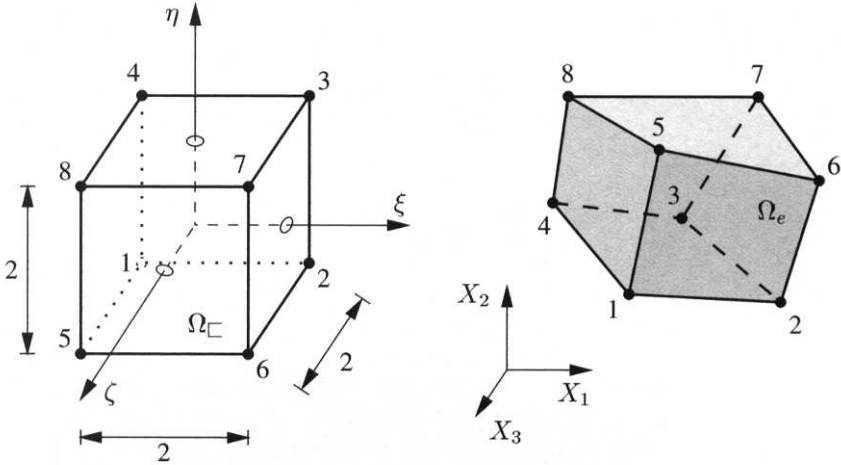
$$\text{with} \quad X_{\alpha,\beta} = \sum_{I=1}^n N_{I,\beta} X_{\alpha I}. \quad (6.31)$$

This leads to an explicit form which allows us to compute in (6.29) the derivatives with respect to  $\mathbf{X}$ :

$$\begin{Bmatrix} N_{I,1} \\ N_{I,2} \end{Bmatrix} = \frac{1}{\det \mathbf{J}_e} \begin{bmatrix} X_{2,\eta} & -X_{2,\xi} \\ -X_{1,\eta} & X_{1,\xi} \end{bmatrix} \begin{Bmatrix} N_{I,\xi} \\ N_{I,\eta} \end{Bmatrix}. \quad (6.32)$$

### 6.1.4 Three-dimensional shape functions

Finite elements for three-dimensional problems are either brick or tetrahedron elements. Also, isoparametric interpolations are advantageous here when arbitrary geometries have to be discretized. Besides bricks and tetrahedrons, more elements are of course possible, e.g. prismatic elements, which will not be discussed here. For general shape functions, see Dhatt and Touzot (1985). For the three-dimensional



**Fig. 6.7** Isoparametric 8 node brick element.

brick element, shown in Figure 6.7, we have the shape functions

$$N_I = \frac{1}{2} (1 + \xi_I \xi) \frac{1}{2} (1 + \eta_I \eta) \frac{1}{2} (1 + \phi_I \phi), \quad (6.33)$$

which follow from the product formula (6.12) with (6.15). Figure 6.7 depicts the associated element in its reference configuration,  $\Omega_\square$ , and the initial configuration,  $\Omega_e$ . Quadratic elements can be designed with (6.12) and (6.16). This yields an interpolation with 27 nodes per element. However, we will not give the explicit representation here, which can be found in Zienkiewicz and Taylor (1989) or Dhatt and Touzot (1985), for example

Shape functions for the tetrahedron elements can be developed analogous to the two-dimensional case. We obtain

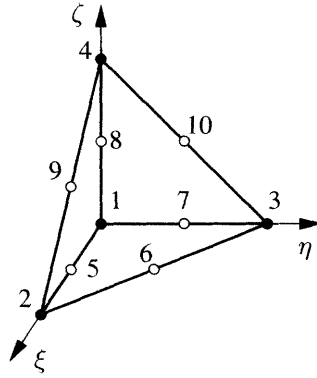
- 4-node tetrahedron (linear interpolation)

$$N_1 = 1 - \xi - \eta - \zeta, \quad N_2 = \xi, \quad N_3 = \eta, \quad N_4 = \zeta. \quad (6.34)$$

- 10-nodes tetrahedron (quadratic interpolation)

$$\begin{aligned} N_1 &= \lambda(2\lambda - 1), & N_6 &= 4\xi\eta, \\ N_2 &= \xi(2\xi - 1), & N_7 &= 4\eta\lambda, \\ N_3 &= \eta(2\eta - 1), & N_8 &= 4\zeta\lambda, \\ N_4 &= \zeta(2\zeta - 1), & N_9 &= 4\xi\zeta, \\ N_5 &= 4\xi\lambda, & N_{10} &= 4\eta\zeta, \end{aligned} \quad (6.35)$$

with  $\lambda = 1 - \xi - \eta - \zeta$ .



**Fig. 6.8** Isoparametric tetrahedra elements, local node numbers.

The local node numbers associated with these shape functions are depicted in Figure 6.8.

The derivatives of the shape functions with respect to the coordinates of the initial or current configuration can be computed using (6.8). For the derivatives with respect to the coordinates of the initial configuration, we have

$$\nabla_X N_I = \begin{Bmatrix} N_{I,1} \\ N_{I,2} \\ N_{I,3} \end{Bmatrix} = \mathbf{J}_e^{-T} \begin{Bmatrix} N_{I,\xi} \\ N_{I,\eta} \\ N_{I,\zeta} \end{Bmatrix}. \quad (6.36)$$

The JACOBI matrix  $\mathbf{J}_e$  of element  $\Omega_e$ , which is needed in this derivation, is given by (6.6) from

$$\mathbf{J}_e = \sum_{I=1}^n \mathbf{X}_I \otimes \nabla_\xi N_I = \begin{bmatrix} X_{1,\xi} & X_{1,\eta} & X_{1,\zeta} \\ X_{2,\xi} & X_{2,\eta} & X_{2,\zeta} \\ X_{3,\xi} & X_{3,\eta} & X_{3,\zeta} \end{bmatrix}. \quad (6.37)$$

Within this formula, the components of  $\mathbf{J}_e$  are computed from

$$X_{m,k} = \sum_{I=1}^n N_{I,k} X_{mI},$$

where the partial derivative with respect to  $k$  stands for a derivative with respect to  $\xi$ ,  $\eta$  or  $\zeta$ .

## 6.2 DISCRETIZATION OF THE WEAK FORMS

In general we can now apply the one-, two- or three-dimensional shape functions to describe the interpolation for the geometry and the field variables within the weak

forms. In this chapter we do this in a brief form for equations (2.59) and (2.63). Furthermore, the linearizations of the weak forms are considered. Following Figure 6.1, a domain is subdivided into  $n_e$  finite elements. This discretization leads to an approximation of the geometry, which affects the representation of the boundary of the domain under consideration.

An interpolation as described in (6.3) is chosen for each finite element  $\Omega_e$ , which approximates the displacement field  $\mathbf{u}$  and the geometry. The integrals of the weak form can then be written with the isoparametric interpolation as

$$\int_B (\dots) dV \approx \int_{B_h} (\dots) dV = \bigcup_{e=1}^{n_e} \int_{\Omega_e} (\dots) d\Omega = \bigcup_{e=1}^{n_e} \int_{\Omega_\square} (\dots) d\square. \quad (6.38)$$

The operator  $\cup$  is introduced instead of a sum sign to denote the assembly process which has to be performed to obtain the set of nonlinear algebraic equations following from (6.38). The polynomial shape functions of the isoparametric interpolation ensures fulfillment of the inter-element continuity conditions, as well as fulfillment of the boundary conditions within the global system of equations. Since the assembly process is standard and well known, it is not described in detail here (see Bathe (1982), Zienkiewicz and Taylor (1989), Knothe and Wessels (1991), or Gross et al. (1999), for example).

### 6.2.1 FE formulation of the weak form with regard to the initial configuration

The approximation of the weak form (2.59) requires discretization of the virtual internal work  $\int_B \mathbf{S} \cdot \delta \mathbf{E} dV$ , of the inertia terms  $\int_B \rho_0 \dot{\mathbf{v}} \cdot \boldsymbol{\eta} dV$  and of the volume- and surface loads  $\int_B \rho_0 \bar{\mathbf{b}} \cdot \boldsymbol{\eta} dV + \int_\Gamma \bar{\mathbf{t}} \cdot \boldsymbol{\eta} dA$ . For the virtual internal work, we need the variation of the GREEN-LAGRANGIAN strain tensors within the element  $\Omega_e$ , see (6.38). With (2.58) and (6.7), one obtains

$$\delta \mathbf{E}^h = \frac{1}{2} \sum_{I=1}^n \left[ \mathbf{F}_e^T (\boldsymbol{\eta}_I \otimes \nabla_X N_I) + (\nabla_X N_I \otimes \boldsymbol{\eta}_I) \mathbf{F}_e \right], \quad (6.39)$$

where the same interpolation was used for the deformation  $\boldsymbol{\varphi}$  and the variation  $\boldsymbol{\eta}$ . In this equation a finite element approximation of the deformation gradient (2.6) has to be applied, which can be written with (6.7) within the element  $\Omega_e$  as

$$\mathbf{F}_e = \sum_{K=1}^n (\mathbf{x}_K \otimes \nabla_X N_K). \quad (6.40)$$

For the derivation of the matrix formulation needed within the computer implementation of finite elements, index notation is necessary. For (6.39) this yields

$$\delta E_{AB}^h = \frac{1}{2} \sum_{I=1}^n [F_{kA} N_{I,B} + N_{I,A} F_{kB}] \eta_{kI} \quad (6.41)$$

with components of the deformation gradients  $F_{kB} = \sum_{J=1}^n x_{kJ} N_{J,B}$ .

Within the matrix formulation we can consider the symmetry of the GREEN-LAGRANGIAN strain tensor and its variation. Thus, it is possible to introduce only six components of nine components for the three-dimensional strain tensor instead

$$\delta \mathbf{E} = \begin{Bmatrix} \delta E_{11} \\ \delta E_{22} \\ \delta E_{33} \\ 2 \delta E_{12} \\ 2 \delta E_{23} \\ 2 \delta E_{13} \end{Bmatrix} = \sum_{I=1}^n \mathbf{B}_{L I} \boldsymbol{\eta}_I, \quad (6.42)$$

which can be approximated as a sum over the element nodes  $I$  with the matrices

$$\mathbf{B}_{L I} = \begin{bmatrix} F_{11} N_{I,1} & F_{21} N_{I,1} & F_{31} N_{I,1} \\ F_{12} N_{I,2} & F_{22} N_{I,2} & F_{32} N_{I,2} \\ F_{13} N_{I,3} & F_{23} N_{I,3} & F_{33} N_{I,3} \\ F_{11} N_{I,2} + F_{12} N_{I,1} & F_{21} N_{I,2} + F_{22} N_{I,1} & F_{31} N_{I,2} + F_{32} N_{I,1} \\ F_{12} N_{I,3} + F_{13} N_{I,2} & F_{22} N_{I,3} + F_{23} N_{I,2} & F_{32} N_{I,3} + F_{33} N_{I,2} \\ F_{11} N_{I,3} + F_{13} N_{I,1} & F_{21} N_{I,3} + F_{23} N_{I,1} & F_{31} N_{I,3} + F_{33} N_{I,1} \end{bmatrix}. \quad (6.43)$$

The index  $L$  in (6.42) depicts that the matrix  $\mathbf{B}_{L I}$  is linear in the displacements, since we have  $\mathbf{F}^h = \mathbf{1} + \text{Grad } \mathbf{u}^h$ .

The stresses follow from the constitutive equation, which will be specified in the associated sections. However, note that the stresses have to be computed pointwise within the element, and result for instance in finite elasticity from a pure function evaluation of the response function. Since also the second PIOLA-KIRCHHOFF stress tensor is symmetric, we only need its six independent components which yields the vector  $\mathbf{S} = \{S_{11}, S_{22}, S_{33}, S_{12}, S_{23}, S_{13}\}^T$ . With these preliminaries, the virtual internal work can be written as

$$\begin{aligned} \int_B \delta \mathbf{E}^h \cdot \mathbf{S}^h dV &= \bigcup_{e=1}^{n_e} \int_{\Omega_e} \delta \mathbf{E}^T \mathbf{S} d\Omega \\ &= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \boldsymbol{\eta}_I^T \int_{\Omega_e} \mathbf{B}_{L I}^T \mathbf{S} d\Omega \\ &= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \boldsymbol{\eta}_I^T \int_{\Omega_\square} \mathbf{B}_{L I}^T \mathbf{S} \det \mathbf{J}_e d\Omega. \end{aligned} \quad (6.44)$$

The last term in (6.45) already reflects the evaluation of the integrals with respect to the configuration of the isoparametric reference element. To shorten the notation, we introduce the vector

$$\mathbf{R}_I(\mathbf{u}_e) = \int_{\Omega_e} \mathbf{B}_{L I}^T \mathbf{S} d\Omega, \quad (6.45)$$

and reformulate the virtual internal work

$$\int_B \delta \mathbf{E}^h \cdot \mathbf{S}^h dV = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \boldsymbol{\eta}_I^T \mathbf{R}_I(\mathbf{u}_e) = \boldsymbol{\eta}^T \mathbf{R}(\mathbf{u}). \quad (6.46)$$

In this equation  $\boldsymbol{\eta}$  is the test function of virtual displacement and  $\mathbf{R}(\mathbf{u})$  is the stress divergence term, also often called the residual force vector, which results from the assembly of all finite elements to the complete structure. Note, that  $\delta \mathbf{E}^h$  is linear with respect to the displacement field, whereas the stress tensor  $\mathbf{S}^h$  can still depend in an arbitrary nonlinear form upon the displacements.

The inertia term, defined by  $\int_B \rho_0 \dot{\mathbf{v}} \cdot \boldsymbol{\eta} dV$ , in the weak form (2.59) is computed with interpolation of the velocity, using standard shape functions  $N_K$  for the spatial discretization,

$$\mathbf{v}(\mathbf{X}, t) = \sum_{K=1}^n N_K(\boldsymbol{\xi}) \mathbf{v}_K(t).$$

The acceleration is given by derivation of the nodal values  $\mathbf{v}_K(t)$ , since the shape functions  $N_K$  depend only upon the spatial coordinates,

$$\dot{\mathbf{v}}(\mathbf{X}, t) = \sum_{K=1}^n N_K(\boldsymbol{\xi}) \dot{\mathbf{v}}_K. \quad (6.47)$$

Inserting this result in the associated inertia term in (2.59), and applying the same discretization as in (6.39) for the variations  $\boldsymbol{\eta}$ , leads to

$$\begin{aligned} \int_B \rho_0 \boldsymbol{\eta} \cdot \dot{\mathbf{v}} dV &= \bigcup_{e=1}^{n_e} \int_{\Omega_e} \rho_0 \boldsymbol{\eta}^T \dot{\mathbf{v}} dV \\ &= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \boldsymbol{\eta}_I^T \int_{\Omega_e} N_I \rho_0 N_K d\Omega \dot{\mathbf{v}}_K. \end{aligned}$$

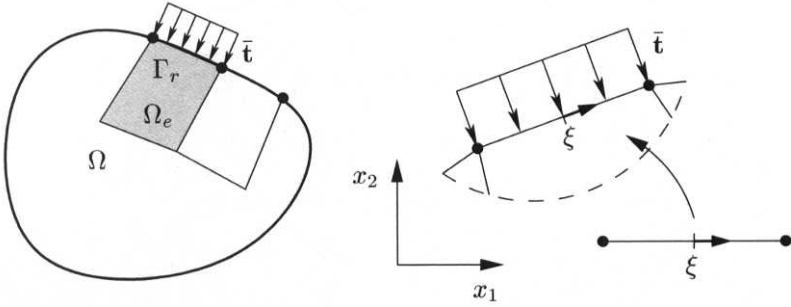
By introducing the unit matrix  $\mathbf{I}$  and application to the nodal velocities  $\dot{\mathbf{v}}_K = \mathbf{I} \dot{\mathbf{v}}_K$ , we obtain the mass matrix for a nodal pair  $I$  and  $K$  of an element  $\Omega_e$

$$\mathbf{M}_{IK} = \int_{\Omega_e} N_I \rho_0 N_K d\Omega \mathbf{I}, \quad (6.48)$$

and hence the inertia term for the global system as

$$\int_B \rho_0 \boldsymbol{\eta} \cdot \dot{\mathbf{v}} dV = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \boldsymbol{\eta}_I^T \mathbf{M}_{IK} \dot{\mathbf{v}}_K = \boldsymbol{\eta}^T \mathbf{M} \dot{\mathbf{v}}, \quad (6.49)$$

where  $\mathbf{M}$  is the mass matrix and  $\dot{\mathbf{v}}$  the acceleration vector after assembly of the global structure.



**Fig. 6.9** Discretization of surface loads.

The loading terms are determined in an analogous way. After inserting the finite element approximations for the test function  $\eta$ , it follows

$$\begin{aligned} \int_B \rho_0 \eta \cdot \bar{\mathbf{b}} dV + \int_{\Gamma_\sigma} \eta \cdot \bar{\mathbf{t}} dA &= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \eta_I^T \int_{\Omega_e} \rho_0 \bar{\mathbf{b}} N_I d\Omega \\ &+ \bigcup_{r=1}^{n_r} \sum_{I=1}^m \eta_I^T \int_{\Gamma_r} N_I \bar{\mathbf{t}} d\Gamma, \end{aligned}$$

where  $n_r$  are the number of loaded element boundaries and  $\Gamma_l$  is the element surface of an element which is subjected to a surface load defined by the stress vector  $\bar{\mathbf{t}}$ , see Figure 6.9. Observe that for the interpolation function of the surface loads, we can use a function which is reduced by one dimension. Thus the surface loads in Figure 6.9, which depicts a two-dimensional body, need as an approximation for the test function along the boundary a one-dimensional function, defined by  $m$  surface nodes (in Figure 6.9 we have  $m = 2$  nodes). Also, here we can simplify by matrix notation, and with

$$\mathbf{P}_I = \int_{\Omega_e} N_I \rho \bar{\mathbf{b}} d\Omega \quad \text{and} \quad \mathbf{P}_I^\sigma = \int_{\Gamma_r} N_I \bar{\mathbf{t}} d\Gamma \quad (6.50)$$

obtain the load vectors

$$\int_B \rho \eta \cdot \bar{\mathbf{b}} dV + \int_{\Gamma_\sigma} \eta \cdot \bar{\mathbf{t}} dA = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \eta_I^T \mathbf{P}_I + \bigcup_{r=1}^{n_r} \sum_{I=1}^m \eta_I^T \mathbf{P}_I^\sigma = \eta^T \mathbf{P}. \quad (6.51)$$

The vector  $\mathbf{P}$  contains all information with regard to the loads acting on the structure.

The matrix notation in (6.46), (6.49) and (6.51) yields, for the weak form (2.59),

$$\eta^T [\mathbf{M} \dot{\mathbf{v}} + \mathbf{R}(\mathbf{u}) - \mathbf{P}] = 0. \quad (6.52)$$



Due to the fact that the test function  $\eta$  is arbitrary, this leads to a nonlinear system of ordinary differential equations:

$$\mathbf{M} \dot{\mathbf{v}} + \mathbf{R}(\mathbf{u}) - \mathbf{P} = \mathbf{0} \quad \forall \mathbf{u} \in \mathbb{R}^N. \quad (6.53)$$

In (6.53) all quantities are evaluated with respect to the initial configuration.  $N$  is the total number of degrees of freedoms contained in the unknown displacement vector  $\mathbf{u}$ .  $\dot{\mathbf{v}}$  is the acceleration vector and  $\mathbf{M}$  denotes the mass matrix.

In case that the inertia forces are zero ( $\mathbf{M} \dot{\mathbf{v}} = \mathbf{0}$ ), from the system of ordinary differential equations we obtain a nonlinear algebraic system of equations which has to be solved by an iterative procedure. In general we apply NEWTON's method, and hence need the linearization of (6.53), which is discussed in the next section.

## 6.2.2 Linearization of the weak form in the initial configuration

For an efficient solution of the nonlinear algebraic equation systems (6.53), NEWTON's method is applied which requires the linearization of (6.53). We derive the linearization in the following by assuming that the inertia terms can be neglected. The linearization can be obtained by a direct discretization of the continuous formulation (2.122)

$$DG(\bar{\varphi}, \eta) \cdot \Delta \mathbf{u} = \int_B \{ \text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} \cdot \text{Grad } \eta + \delta \bar{\mathbf{E}} \cdot \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}] \} dV. \quad (6.54)$$

For the first term with

$$\begin{aligned} \text{Grad } \Delta \mathbf{u}^h &= \sum_{K=1}^n \Delta \mathbf{u}_K \otimes \nabla_X N_K, \\ \text{Grad } \eta &= \sum_{I=1}^n \eta_I \otimes \nabla_X N_I \end{aligned} \quad (6.55)$$

we obtain directly the discretization

$$\int_B \text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} \cdot \text{Grad } \eta dV = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \int_{\Omega_e} (\Delta \mathbf{u}_K \otimes \nabla_X N_K) \bar{\mathbf{S}} \cdot (\eta_I \otimes \nabla_X N_I) d\Omega,$$

which yields, with the rules for the dyadic and scalar products and with  $\Delta \mathbf{u}_K \cdot \eta_I = \eta_I^T \Delta \mathbf{u}_K = \eta_I^T \mathbf{I} \Delta \mathbf{u}_K$ ,

$$\int_B \text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} \cdot \text{Grad } \eta dV = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \eta_I^T \int_{\Omega_e} \bar{G}_{IK} \mathbf{I} d\Omega \Delta \mathbf{u}_K, \quad (6.56)$$

where the abbreviation

$$\bar{G}_{IK} = (\nabla_X N_I)^T \bar{\mathbf{S}} \nabla_X N_K \quad (6.57)$$

has been used. The matrix form of the scalar product (6.57) can be derived if the gradients are described as vectors. This leads to

$$\bar{G}_{IK} = [N_{I,1} \quad N_{I,2} \quad N_{I,3}] \begin{bmatrix} \bar{S}_{11} & \bar{S}_{12} & \bar{S}_{13} \\ \bar{S}_{21} & \bar{S}_{22} & \bar{S}_{23} \\ \bar{S}_{31} & \bar{S}_{32} & \bar{S}_{33} \end{bmatrix} \begin{Bmatrix} N_{K,1} \\ N_{K,2} \\ N_{K,3} \end{Bmatrix}. \quad (6.58)$$

Relation (6.56) is independent from the constitutive equation, since only the stress at configuration  $\bar{\varphi}$  has to be considered. Hence the matrix which defined by (6.56) is often called the INITIAL STRESS MATRIX.

The second term in (2.122)

$$\int_B \delta \bar{\mathbf{E}} \cdot \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}] dV$$

depends upon the incremental constitutive tensor  $\bar{\mathbf{C}}$  which has to be evaluated at configuration  $\bar{\varphi}$ , and thus is directly connected to the constitutive equation. For elastic materials this tensor has been given in Section 2.4.2 (e.g. see (2.89)). For elasto-plastic or other constitutive equations, the associated matrix formulation can be found in (2.92). Since  $\Delta \bar{\mathbf{E}}$  has the same structure as  $\delta \bar{\mathbf{E}}$ , with (6.39) we can write

$$\Delta \mathbf{E}^h = \frac{1}{2} \sum_{I=1}^n \left[ \mathbf{F}_e^T (\Delta \mathbf{u}_I \otimes \nabla_X N_I) + (\nabla_X N_I \otimes \Delta \mathbf{u}_I) \mathbf{F}_e \right]. \quad (6.59)$$

From this relation, the matrix formulation follows with (6.43)

$$\Delta \mathbf{E} = \sum_{I=1}^n \mathbf{B}_{L I} \Delta \mathbf{u}_I. \quad (6.60)$$

Introduction of this relation, together with the incremental constitutive tensor  $\bar{\mathbf{D}}$ , yields

$$\int_B \delta \bar{\mathbf{E}} \cdot \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}] dV = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \eta_I^T \int_{\Omega_e} \bar{\mathbf{B}}_{L I}^T \bar{\mathbf{D}} \bar{\mathbf{B}}_{L K} d\Omega \Delta \mathbf{u}_K. \quad (6.61)$$

Thus we can summarize, and obtain the discretization

$$\int_B \{ \text{Grad } \Delta \mathbf{u} \bar{\mathbf{S}} \cdot \text{Grad } \boldsymbol{\eta} + \delta \bar{\mathbf{E}} \cdot \bar{\mathbf{C}} [\Delta \bar{\mathbf{E}}] \} dV = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \eta_I^T \bar{\mathbf{K}}_{T I K} \Delta \mathbf{u}_K. \quad (6.62)$$

Here matrix  $\bar{\mathbf{K}}_{I K}$  denotes the "tangent matrix" because it represents the tangent to the deformation at  $\bar{\varphi}$ :

$$\bar{\mathbf{K}}_{T I K} = \int_{\Omega_e} \left[ (\nabla_X N_I)^T \bar{\mathbf{S}} \nabla_X N_K + \bar{\mathbf{B}}_{L I}^T \bar{\mathbf{D}} \bar{\mathbf{B}}_{L K} \right] d\Omega. \quad (6.63)$$

It is stated for the nodal combination  $I, K$  within a finite element  $\Omega_e$ . In this notation the submatrix  $\bar{K}_{TIK}$  has the size  $n_{dof} \times n_{dof}$ , where  $n_{dof}$  is the number of degrees of freedom for one node within the finite element (in three-dimensional problems in continuum mechanics we have three degrees of freedom for each point, hence  $n_{dof} = 3$ ). Indices  $I$  and  $K$  are nodes of an element, and thus directly associated with the discretization. For example, for a ten node tetrahedron we have  $n = 10$ , hence the total size of the tangent matrix  $\bar{K}_{T_e}$  for one element is  $(n \cdot n_{dof}) \times (n \cdot n_{dof}) = 30 \times 30$ .

**Two-dimensional element for St. Venant material.** In this section we derive as an example the matrix formulation for a two-dimensional element with respect to the initial configuration. As a constitutive relation, St. Venant material is considered which relates the GREEN-LAGRANGIAN strains and the second PIOLA-KIRCHHOFF stresses by a linear relation. Furthermore, plain strain is assumed.

For computation of the stress divergence term resulting from the weak form, we have to specify (6.45). Thus for a displacement formulation we need the stresses as a function of the strains and the strains field in terms of the displacement gradients. The ST. VENANT constitutive equation is given by

$$\mathbf{S} = (\Lambda \mathbf{1} \otimes \mathbf{1} + 2\mu \mathbf{I}) [\mathbf{E}]. \quad (6.64)$$

This form can be easily converted into a matrix formulation. For the two-dimensional case, we obtain

$$\mathbf{S} = \mathbf{D} \mathbf{E} = \begin{Bmatrix} S_{11} \\ S_{22} \\ S_{12} \end{Bmatrix} = \begin{bmatrix} \Lambda + 2\mu & \Lambda & 0 \\ \Lambda & \Lambda + 2\mu & 0 \\ 0 & 0 & \mu \end{bmatrix} \begin{Bmatrix} E_{11} \\ E_{22} \\ 2E_{12} \end{Bmatrix}. \quad (6.65)$$

Now it remains to compute the GREEN-LAGRANGIAN strain tensor.

The components of the matrix form of the GREEN-LAGRANGIAN strain tensor, see (6.65), follow from (2.16) with (6.40) in the case of two dimensions for a finite element  $\Omega_e$ :

$$\mathbf{E} = \frac{1}{2} (\mathbf{F}_e^T \mathbf{F}_e - \mathbf{I}) \quad \text{with} \quad \mathbf{F}_e = \sum_{K=1}^n \begin{bmatrix} x_{1K} N_{K,1} & x_{1K} N_{K,2} \\ x_{2K} N_{K,1} & x_{2K} N_{K,2} \end{bmatrix}. \quad (6.66)$$

Here the nodal coordinates  $x_{\alpha K} = X_{\alpha K} + u_{\alpha K}$  belong to the current configuration  $\bar{\varphi}$ .

The approximation of the virtual strains  $\delta \mathbf{E}$  with (6.42) yields the matrix  $\mathbf{B}_{L I}$ , which in the two-dimensional case has the representation

$$\mathbf{B}_{L I} = \begin{bmatrix} F_{11} N_{I,1} & F_{21} N_{I,1} \\ F_{12} N_{I,2} & F_{22} N_{I,2} \\ F_{11} N_{I,2} + F_{12} N_{I,1} & F_{21} N_{I,2} + F_{22} N_{I,1} \end{bmatrix}. \quad (6.67)$$

We can also express the virtual strain with  $\mathbf{F} = \mathbf{1} + \text{Grad} \mathbf{u}$  different from (6.43) by

$$\delta \mathbf{E} = \sum_{I=1}^4 [\mathbf{B}_{0 I} + \mathbf{B}_{V I}(\mathbf{u})] \boldsymbol{\eta}_I. \quad (6.68)$$

In that case, the matrices  $\mathbf{B}_{0I}$  and  $\mathbf{B}_{VI}$  have the explicit form

$$\mathbf{B}_{0I} = \begin{bmatrix} N_{I,1} & 0 \\ 0 & N_{I,2} \\ N_{I,2} & N_{I,1} \end{bmatrix} \quad (6.69)$$

and

$$\mathbf{B}_{VI} = \begin{bmatrix} u_{1,1} N_{I,1} & u_{2,1} N_{I,1} \\ u_{1,2} N_{I,2} & u_{2,2} N_{I,2} \\ u_{1,1} N_{I,2} + u_{1,2} N_{I,1} & u_{2,2} N_{I,1} + u_{2,1} N_{I,2} \end{bmatrix}. \quad (6.70)$$

The derivative  $u_{\alpha,\beta}$  can be computed for each quadrature point analogous to the components of the deformation gradient in (6.41). We obtain  $u_{\alpha,\beta} = \sum_{K=1}^4 N_{K,\beta} u_{\alpha K}$ , where the indices  $\alpha$  and  $\beta$  take values of 1 and 2. Note that the nonlinear part defined by matrix  $\mathbf{B}_{VI}$  disappears for  $\mathbf{u} = \text{const.}$

The stress divergence term (6.45) is then obtained with (6.68) for an element  $\Omega_e$

$$\mathbf{R}_I(\mathbf{u}_e) = \int_{\Omega_e} (\mathbf{B}_{0I} + \mathbf{B}_{VI})^T \mathbf{S} d\Omega. \quad (6.71)$$

The load vector can be computed using equation (6.50), but we do not want to specify it here in detail.

Linearization of (6.71) at  $\bar{\varphi}$  leads to the tangential stiffness matrix of a finite element. It is given with (6.68) analogous to (6.63)

$$\bar{\mathbf{K}}_{TIK} = \int_{\Omega_e} [(\mathbf{B}_{0I} + \bar{\mathbf{B}}_{VI})^T \mathbf{D}(\mathbf{B}_{0K} + \bar{\mathbf{B}}_{VK}) + \bar{\mathbf{G}}_{IK} \mathbf{I}] d\Omega. \quad (6.72)$$

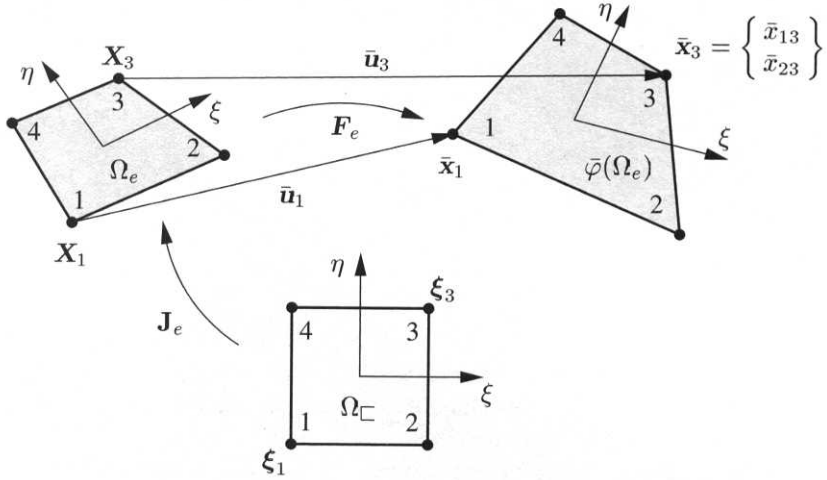
Note that all quantities with a bar have to be evaluated at  $\bar{\varphi}$ . Due to the special ST. VENANT model for elasticity, the incremental constitutive tensor  $\mathbf{D}$ , defined in (6.65), is constant. A more compact notation of the tangent matrix follows with (6.67):

$$\bar{\mathbf{K}}_{TIK} = \int_{\Omega_e} [\bar{\mathbf{B}}_{LI}^T \mathbf{D} \bar{\mathbf{B}}_{LK} + \bar{\mathbf{G}}_{IK} \mathbf{I}] d\Omega. \quad (6.73)$$

The term  $\bar{\mathbf{G}}_{IK}$  is given for the two-dimensional problem by the product

$$\bar{\mathbf{G}}_{IK} = [N_{I,1} \quad N_{I,2}] \begin{bmatrix} \bar{S}_{11} & \bar{S}_{12} \\ \bar{S}_{21} & \bar{S}_{22} \end{bmatrix} \begin{Bmatrix} N_{K,1} \\ N_{K,2} \end{Bmatrix}. \quad (6.74)$$

Both equations (6.72) and (6.74) have to be evaluated at  $\bar{\varphi}$  at which the linearization takes place. The stresses in (6.74) are computed via the constitutive relation (6.65), based on the strains (6.66). The integrals in (6.71) and (6.72) have to be computed using a numerical quadrature formula. In this case, it is efficient to refer to the reference element, see (6.45) and Figure 6.10. This yields, with (6.45), a  $n_{dof} \times 1 =$



**Fig. 6.10** Isoparametric transformation and deformed state at  $\bar{\varphi}$ .

$2 \times 1$  vector for the stress divergence term which is associated with node  $I$ :

$$\begin{aligned} \mathbf{R}_I(\mathbf{u}_e) &= \int_{\Omega_e} (\mathbf{B}_{0I} + \mathbf{B}_{VI})^T \mathbf{S} d\Omega \\ &\approx \sum_{p=1}^{n_p} W_p [\mathbf{B}_{0I}(\xi_p, \eta_p) + \mathbf{B}_{VI}(\xi_p, \eta_p)]^T \mathbf{S}(\xi_p, \eta_p) \det \mathbf{J}_e(\xi_p, \eta_p). \end{aligned} \quad (6.75)$$

For evaluation of the integrals a GAUSS quadrature, with  $n_p = 2 \times 2 = 4$  points, is sufficient if the element is a four node element with bilinear shape functions. The coordinates of the GAUSS points  $\xi_p, \eta_p$  and the associated weights  $W_p$  are given in Table A.2 in the appendix. For computation of the stresses at a GAUSS point  $\mathbf{S}(\xi_p, \eta_p)$ , we need the deformation gradient, see (6.66), which has to be evaluated at  $(\xi_p, \eta_p)$

$$\mathbf{F}_e(\xi_p, \eta_p) = \sum_{K=1}^n \begin{bmatrix} x_{1K} N_{K,1}(\xi_p, \eta_p) & x_{1K} N_{K,2}(\xi_p, \eta_p) \\ x_{2K} N_{K,1}(\xi_p, \eta_p) & x_{2K} N_{K,2}(\xi_p, \eta_p) \end{bmatrix}. \quad (6.76)$$

The stress at a GAUSS quadrature point then follows with the strains (6.66) from (6.65). Note that in equation (6.76), the summation (Index  $K$ ) has to include all nodes, since all interpolation functions have a contribution to the deformation at one GAUSS quadrature point within the element.

In an analogous way, the numerical integration of the integrals concerning the tangent matrix (6.72) has to be performed.

The submatrices for the indices  $I$  and  $K$  are  $2 \times 2$  matrices. Based upon the result (6.73), their explicit form using the GAUSS quadrature follows:

$$\begin{aligned} \bar{\mathbf{K}}_{T_{IK}} &= \int_{\Omega_e} [\bar{\mathbf{B}}_{L I}^T \mathbf{D} \bar{\mathbf{B}}_{L K} + \bar{\mathbf{G}}_{IK} \mathbf{I}] d\Omega \\ &\approx \sum_{p=1}^{n_p} W_p \left[ \bar{\mathbf{B}}_{L I}^T(\xi_p, \eta_p) \mathbf{D} \bar{\mathbf{B}}_{L K}(\xi_p, \eta_p) + \bar{\mathbf{G}}_{IK}(\xi_p, \eta_p) \mathbf{I} \right] \det \mathbf{J}_e(\xi_p, \eta_p). \end{aligned} \quad (6.77)$$

The stress divergence vectors,  $\mathbf{R}_I$ , and the submatrices of the tangent matrix,  $\bar{\mathbf{K}}_{T_{IK}}$ , have to be ordered in the following way to obtain the stress divergent vector and the tangent matrix for the element  $\Omega_e$ :

$$\mathbf{R}_e = \left\{ \begin{matrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_3 \\ \mathbf{R}_4 \end{matrix} \right\}_{8 \times 1} \quad \bar{\mathbf{K}}_{T_e} = \begin{bmatrix} \bar{\mathbf{K}}_{T_{11}} & \bar{\mathbf{K}}_{T_{12}} & \bar{\mathbf{K}}_{T_{13}} & \bar{\mathbf{K}}_{T_{14}} \\ & \bar{\mathbf{K}}_{T_{22}} & \bar{\mathbf{K}}_{T_{23}} & \bar{\mathbf{K}}_{T_{24}} \\ & & \bar{\mathbf{K}}_{T_{33}} & \bar{\mathbf{K}}_{T_{34}} \\ \text{symm.} & & & \bar{\mathbf{K}}_{T_{44}} \end{bmatrix}_{8 \times 8}. \quad (6.78)$$

The total size follows from the number of nodes, 4, and the number of degrees of freedom, 2.

#### REMARK 6.1

- The order of the nonlinearity of (6.71) depends only upon the constitutive equation. For the ST. VENANT material used here, (6.71) is a cubic polynomial in  $\mathbf{u}$ . However, this no longer holds once a constitutive equation like the NEO-HOOKIAN material, see (2.74), is used.
- If stresses have to be computed in a design analysis then a transformation of the second PIOLA-KIRCHHOFF stresses to the CAUCHY stresses must be performed on the basis of (2.50).
- In the case of the linear theory of small strains and displacements, the terms  $\bar{\mathbf{G}}_{IK}$  and  $\bar{\mathbf{B}}_{V I}$ ,  $\bar{\mathbf{B}}_{V K}$  disappear in (6.71) and (6.72), thus the resulting equations are linear in  $\mathbf{u}$  and represent the equations of the classical linear theory of elasticity. In the two-dimensional form of (2.95), we then obtain the stiffness matrix

$$\mathbf{K}_{IK} = \int_{\Omega_e} \mathbf{B}_{0 I}^T \mathbf{D}_0 \mathbf{B}_{0 K} d\Omega.$$

### 6.2.3 FE formulation of the weak form in the current configuration

The derivation of the matrix formulation for the weak form with respect to the current configuration follows analogous to the derivation of equation (2.59), but we use as a basic equation (2.61). Within the integrals the *push forward* of the variation of the GREEN-LAGRANGIAN strain tensor  $\delta \mathbf{E} = \nabla^S \boldsymbol{\eta}$  is needed, see (2.62). Hence (2.62) has to be approximated. With equations (6.7)<sub>2</sub>, this leads to

$$\nabla^S \boldsymbol{\eta}^h = \frac{1}{2} \sum_{I=1}^n [(\boldsymbol{\eta}_I \otimes \nabla_x N_I) + (\nabla_x N_I \otimes \boldsymbol{\eta}_I)]. \quad (6.79)$$

As in the last section, it is advantageous to switch to index notation to derive the matrix formulation. We obtain

$$(\nabla^S \boldsymbol{\eta})_{im}^h = \frac{1}{2} \sum_{I=1}^n [\eta_{iI} N_{I,k} + N_{I,i} \eta_{kI}] , \quad (6.80)$$

where now  $N_{I,m} = \partial N_I / \partial x_m$  is the partial derivative of the shape functions with respect to the spatial coordinates  $x_m$ . These derivatives can be computed using (6.9)<sub>2</sub>

$$N_{I,k} = \{j_e^{-1}\}_{1k} N_{I,\xi} + \{j_e^{-1}\}_{2k} N_{I,\eta} + \{j_e^{-1}\}_{3k} N_{I,\zeta} , \quad (6.81)$$

where  $\{j_e^{-1}\}_{ik}$  are the associated components of the inverse of the JACOBI matrix  $\mathbf{j}_e$ . Equation (6.80) yields the components of  $\nabla^S \boldsymbol{\eta}^h$ . Due to symmetry, the components can be assembled in the vector  $(\nabla^S \boldsymbol{\eta}^h)^T = [\eta_{1,1}, \eta_{2,2}, \eta_{3,3}, (\eta_{1,2} + \eta_{2,1}), (\eta_{2,3} + \eta_{3,2}), (\eta_{1,3} + \eta_{3,1})]$ . With this the approximation of the spatial gradient is given by

$$\nabla^S \boldsymbol{\eta}^h = \sum_{I=1}^n \begin{bmatrix} N_{I,1} & 0 & 0 \\ 0 & N_{I,2} & 0 \\ 0 & 0 & N_{I,3} \\ N_{I,2} & N_{I,1} & 0 \\ 0 & N_{I,3} & N_{I,2} \\ N_{I,3} & 0 & N_{I,1} \end{bmatrix} \begin{Bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{Bmatrix}_I = \sum_{I=1}^n \mathbf{B}_{0I} \boldsymbol{\eta}_I . \quad (6.82)$$

Note that matrix  $\mathbf{B}_{0I}$  does not contain any displacements, which is indicated by the index "0".

**REMARK 6.2** In contrast to matrix  $\mathbf{B}_{LI}$ , matrix  $\mathbf{B}_{0I}$  has a sparse structure. Half of its entries are zero. It is thus easily concluded that the associated zero components can be neglected in any multiplication of  $\mathbf{B}_{0I}$  with vectors or matrices. Thus, the finite element formulation with regard to the current configuration is much more efficient on the element level.

Furthermore, we note that the structure of  $\mathbf{B}_{0I}$  is exactly the same as the  $\mathbf{B}$ -matrix of linear theory, see Zienkiewicz and Taylor (1989). The only difference is that in the linear theory, all derivatives are with respect to the coordinates  $\mathbf{X}$  of the initial configuration, while here all derivatives have to be computed with respect to the coordinates  $\mathbf{x}$  of the current configuration according to (6.80) and (6.81).

With these preliminary remarks and the introduction of a vector  $\boldsymbol{\sigma}$  which contains the independent components,  $\boldsymbol{\sigma} = \{\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{23}, \sigma_{13}\}^T$ , of the CAUCHY stress tensor, the internal virtual work in (2.61) can be written as

$$\begin{aligned} \int_{\varphi(B)} \nabla^S \boldsymbol{\eta}^h \cdot \boldsymbol{\sigma}^h dv &= \bigcup_{e=1}^{n_e} \int_{\varphi(\Omega_e)} (\nabla^S \boldsymbol{\eta}^h)^T \boldsymbol{\sigma}^h d\omega \\ &= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \boldsymbol{\eta}_I^T \int_{\varphi(\Omega_e)} \mathbf{B}_{0I}^T \boldsymbol{\sigma} d\omega \end{aligned} \quad (6.83)$$

$$= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \eta_I^T \int_{\Omega_\square} \mathbf{B}_{0I}^T \boldsymbol{\sigma} \det \mathbf{J}_e d\Omega.$$

The last form in (6.84) already contains the reference to the isoparametric base element  $\Omega_\square$ . A comparison with the associated relation in (6.45) shows that both formulations distinguish each other by the  $\mathbf{B}$ -Matrix, the determinant of the isoparametric mapping (6.5) and, of course, the stress tensor. By introducing

$$\mathbf{r}_I(\mathbf{u}_e) = \int_{\varphi(\Omega_e)} \mathbf{B}_{0I}^T \boldsymbol{\sigma} d\omega, \quad (6.84)$$

we can shorten the notation, and for the virtual internal work obtain

$$\int_{\varphi(B)} \nabla^S \boldsymbol{\eta}^h \cdot \boldsymbol{\sigma}^h dv = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \eta_I^T \mathbf{r}_I(\mathbf{u}_e) = \boldsymbol{\eta}^T \mathbf{r}(\mathbf{u}). \quad (6.85)$$

With the transformation for the volume elements  $dv = J dV$  and the relation between the CAUCHY stress tensor and the KIRCHHOFF stress tensor, see (2.51), which yields  $\boldsymbol{\tau} = J \boldsymbol{\sigma}$ , we can transform the integral representing the virtual internal work in (6.84) to the reference configuration:

$$\int_{\varphi(B)} \nabla^S \boldsymbol{\eta}^h \cdot \boldsymbol{\sigma}^h dv = \int_B \nabla^S \boldsymbol{\eta}^h \cdot \boldsymbol{\tau}^h dV. \quad (6.86)$$

Discretization with finite elements leads to

$$\begin{aligned} \int_B \nabla^S \boldsymbol{\eta}^h \cdot \boldsymbol{\tau}^h dV &= \bigcup_{e=1}^{n_e} \int_{\Omega_e} (\nabla^S \boldsymbol{\eta}^h)^T \boldsymbol{\tau}^h d\Omega, \\ &= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \eta_I^T \int_{\Omega_e} \mathbf{B}_{0I}^T \boldsymbol{\tau} d\Omega, \\ &= \bigcup_{e=1}^{n_e} \sum_{I=1}^n \eta_I^T \int_{\Omega_\square} \mathbf{B}_{0I}^T \boldsymbol{\tau} \det \mathbf{J}_e d\Omega. \end{aligned} \quad (6.87)$$

Hence, in this case the residual vector denoting the stress divergence term is defined by

$$\mathbf{r}_I(\mathbf{u}_e) = \int_{\Omega_\square} \mathbf{B}_{0I}^T \boldsymbol{\tau} d\Omega. \quad (6.88)$$

The total internal work follows from (6.85).

The approximation of the inertia terms is performed according to (6.49). In the same way, (6.51), the load terms are formulated. Thus we can summarize the finite



element discretization of the weak form with respect to the current configuration (2.61):

$$\boldsymbol{\eta}^T [\mathbf{M} \dot{\mathbf{v}} + \mathbf{r}(\mathbf{u}) - \mathbf{P}] = 0, \quad (6.89)$$

which for arbitrary values of the test function  $\boldsymbol{\eta}$  yields the nonlinear ordinary differential system

$$\mathbf{M} \dot{\mathbf{v}} + \mathbf{r}(\mathbf{u}) - \mathbf{P} = \mathbf{0}. \quad (6.90)$$

For static problems this system reduces to a nonlinear algebraic system of equations for the unknown nodal displacements  $\mathbf{u}$ :

$$\mathbf{g}(\mathbf{u}) = \mathbf{r}(\mathbf{u}) - \mathbf{P} = \mathbf{0}. \quad (6.91)$$

The vector representing the stress divergence term  $\mathbf{r}(\mathbf{u})$  can be computed in equations (6.90) or (6.91) either by (6.84) or (6.88). Both formulations are equivalent. Note that the relation (6.85) looks like the formulation in the linear theory, only the quantities  $\delta \mathbf{e}$  and  $\boldsymbol{\sigma}$  are evaluated with respect to the current configuration.

## 6.2.4 Linearization of the weak form in the current configuration

In the last section we derived two weak forms, equations (6.85) and (6.88), which differ only in the region of integration,  $\varphi(B^h)$  or  $B^h$ . The linearization of these forms is described in Section 2.5.3, thus we only have to apply the discretization to these results.

Linearization of the weak form (6.85) follows from equation (2.127) as

$$Dg(\bar{\varphi}, \boldsymbol{\eta}) \cdot \Delta \mathbf{u} = \int_{\bar{\varphi}(B)} \{ \overline{\text{grad}} \Delta \mathbf{u} \cdot \bar{\boldsymbol{\sigma}} \cdot \overline{\text{grad}} \boldsymbol{\eta} + \bar{\nabla}^S \boldsymbol{\eta} \cdot \bar{\hat{\mathbf{c}}} [\bar{\nabla}^S \Delta \mathbf{u}] \} dv. \quad (6.92)$$

The first term has exactly the same form as the associated term in the formulation with respect to the initial configuration. Hence the discretization is the same, and can be directly adopted from the discretization in the initial configuration, see (6.56). Only the derivatives are now with respect to the coordinates  $\bar{x}_i$  of the current configuration  $\varphi(\bar{B})$ . With the discretization of the gradient

$$\begin{aligned} \overline{\text{grad}} \Delta \mathbf{u}^h &= \sum_{K=1}^n \Delta \mathbf{u}_K \otimes \bar{\nabla}_x N_K, \\ \overline{\text{grad}} \boldsymbol{\eta}^h &= \sum_{I=1}^n \boldsymbol{\eta}_I \otimes \bar{\nabla}_x N_I, \end{aligned} \quad (6.93)$$

we obtain the first part of the integrals

$$\int_{\bar{\varphi}(B)} \overline{\text{grad}} \Delta \mathbf{u} \cdot \bar{\boldsymbol{\sigma}} \cdot \overline{\text{grad}} \boldsymbol{\eta} dv = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \boldsymbol{\eta}_I^T \int_{\bar{\varphi}(\Omega_e)} \bar{g}_{IK} \mathbf{I} d\Omega \Delta \mathbf{u}_K. \quad (6.94)$$

Within this term, the abbreviation

$$\bar{g}_{IK} = (\bar{\nabla}_x N_I)^T \bar{\sigma} \bar{\nabla}_x N_K \quad (6.95)$$

has been used. The matrix from of the scalar product follows, as in (6.58), as

$$\bar{g}_{IK} = [\bar{N}_{I,1} \quad \bar{N}_{I,2} \quad \bar{N}_{I,3}] \begin{bmatrix} \bar{\sigma}_{11} & \bar{\sigma}_{12} & \bar{\sigma}_{13} \\ \bar{\sigma}_{21} & \bar{\sigma}_{22} & \bar{\sigma}_{23} \\ \bar{\sigma}_{31} & \bar{\sigma}_{32} & \bar{\sigma}_{33} \end{bmatrix} \begin{Bmatrix} \bar{N}_{K,1} \\ \bar{N}_{K,2} \\ \bar{N}_{K,3} \end{Bmatrix}. \quad (6.96)$$

This equation is independent from the constitutive equation, as is (6.56), since only the stresses of the configuration  $\bar{\varphi}$  enter the integral.

The second term in (2.122)

$$\int_{\bar{\varphi}(\mathcal{B})} \bar{\nabla}^S \eta \cdot \bar{\mathfrak{c}} [\bar{\nabla}^S \Delta \mathbf{u}] dv$$

depends upon the incremental constitutive tensor  $\bar{\mathfrak{c}}$ , evaluated at the current configuration  $\bar{\varphi}$ , and thus directly from the constitutive equation (e.g. see Section 2.4.2, equation (2.90)). Using the same arguments as for linearization with respect to the initial configuration and (6.82), we obtain

$$\int_{\bar{\varphi}(\mathcal{B})} \bar{\nabla}^S \eta \cdot \bar{\mathfrak{c}} [\bar{\nabla}^S \Delta \mathbf{u}] dv = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \eta_I^T \int_{\bar{\varphi}(\Omega_e)} \bar{\mathbf{B}}_{0I}^T \bar{\mathbf{D}}^M \bar{\mathbf{B}}_{0K} d\Omega \Delta \mathbf{u}_K, \quad (6.97)$$

where all quantities in the integrals have to be evaluated at  $\bar{\varphi}$ . In summary, we find the discretization

$$\int_{\bar{\varphi}(\mathcal{B})} \{ \overline{\text{grad}} \Delta \mathbf{u} \cdot \overline{\text{grad}} \eta + \bar{\nabla}^S \eta \cdot \bar{\mathfrak{c}} [\bar{\nabla}^S \Delta \mathbf{u}] \} dv = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \eta_I^T \bar{\mathbf{K}}_{TIK}^M \Delta \mathbf{u}_K, \quad (6.98)$$

where matrix  $\bar{\mathbf{K}}_{TIK}^M$  is the tangent matrix with respect to the current configuration,

$$\bar{\mathbf{K}}_{TIK}^M = \int_{\bar{\varphi}(\Omega_e)} \left[ (\bar{\nabla}_x N_I)^T \bar{\sigma} \bar{\nabla}_x N_K + \bar{\mathbf{B}}_{0I}^T \bar{\mathbf{D}}^M \bar{\mathbf{B}}_{0K} \right] d\omega. \quad (6.99)$$

It is defined for the combination of nodes  $I, K$  within the element  $\Omega_e$ , see also Section 6.2.2. The discretization of the weak form (6.88) in the current configuration follows in an analogous way, we just state here the final result:

$$\int_{\mathcal{B}} \{ \overline{\text{grad}} \Delta \mathbf{u} \cdot \bar{\tau} \cdot \overline{\text{grad}} \eta + \bar{\nabla}^S \eta \cdot \bar{\mathfrak{c}} [\bar{\nabla}^S \Delta \mathbf{u}] \} dv = \bigcup_{e=1}^{n_e} \sum_{I=1}^n \sum_{K=1}^n \eta_I^T \bar{\mathbf{K}}_{TIK}^{MR} \Delta \mathbf{u}_K, \quad (6.100)$$

where matrix  $\bar{\mathbf{K}}_{T_{IK}}^{MR}$  is the tangent matrix with respect to the current configuration:

$$\bar{\mathbf{K}}_{T_{IK}}^{MR} = \int_{\Omega_e} \left[ (\bar{\nabla}_x N_I)^T \bar{\tau} \bar{\nabla}_x N_K + \bar{\mathbf{B}}_{0I}^T \bar{\mathbf{D}}^{MR} \bar{\mathbf{B}}_{0K} \right] d\Omega. \quad (6.101)$$

The matrix form  $\bar{\mathbf{D}}^{MR}$  of the incremental constitutive tensor  $\bar{\mathfrak{c}}$  can be found for a NEO-HOOKE material, e.g. in (2.92). The associated form for  $\bar{\mathbf{D}}^M$  results from the transformation with the JACOBI determinant  $J$ , as given in (2.126).

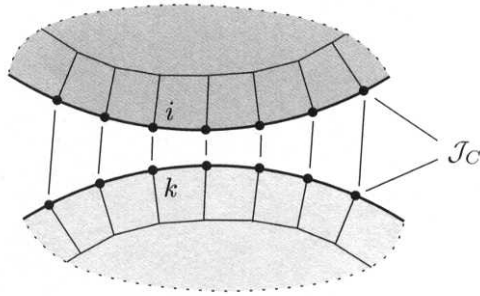
# 7

---

## *Discretization, Small Deformation Contact*

In the first applications of finite elements to contact problems of two deformable bodies only small changes in the geometry were assumed so that the geometrically linear theory could be applied. In that case it is possible to incorporate the contact constraints on a purely nodal basis, e.g. see Francavilla and Zienkiewicz (1975). Later, contact elements were also developed, which resulted from a degenerated solid element, e.g. see Stadter and Weiss (1979) or the textbook of Kikuchi and Oden (1988). A mathematical study of these classes of elements which also accounts for the correct integration rules can be found in Oden (1981) and Kikuchi and Oden (1988). All of the above-mentioned elements need a discretization in which the element nodes match each other at the contact interface. For the general case of nodes being arbitrarily distributed along the possible contact interface between two bodies, which can occur when automatic meshing is used for two different bodies, Simo et al. (1985) developed a segment approach to discretize the contact interface for the two-dimensional case. Also, first attempts have been made to use the *hp*-version of finite elements for the discretization of contact problems, see Paczelt et al. (1999). Recently, such discretizations gained more attention due to automatic meshing tools and domain decomposition methods for large problems. We refer to mathematical literature like Belgacem et al. (1997) and Belgacem et al. (1999) for the development of so-called mortar methods, and to Rebel et al. (2000) or McDevitt and Laursen (2000) for an engineering treatment of such methods.

In general, in small deformation contact, we do not need search algorithms since no large slip can occur on the contact surface. Hence when the first discretization takes into account the contact interface as depicted in Figure 7.1, then the element sizes coincide at the interface, and one can define contact elements which have nodes



**Fig. 7.1** Contact discretization for small deformations.

on the surface of both bodies. These contain pairs  $\{i, k\} \in \mathcal{J}_C$  which might possibly come into contact. Note also that in frictional contact problems, the mesh topology does not change during the loading process, since the slip is infinitesimally small. Thus in the case of friction, it is also possible to develop contact interface elements on a nodal basis, which can be simply added in the assembly procedure to the general stiffness matrix like standard finite elements.

In the following we discuss the case of two deformable bodies being in contact. The special case of the classical SIGNORINI problem, where a deformable body is in contact with a rigid obstacle, follows directly from the equations developed by assuming one body is rigid, and by defining the normal vector on the rigid obstacle.

## 7.1 GENERAL APPROACH FOR CONTACT DISCRETIZATION

There are different possibilities to formulate (and hence discretize) the contact constraint. This was discussed in Section 5.3, leading to the main formulations of LAGRANGE multipliers and the penalty approach. Discretization in the framework of these two methods will be discussed in this section.

In general, for penalty and LAGRANGE multiplier formulations, different discretizations of  $\Gamma_c$  are possible which depend on the problem, on the discretization of the bodies which come into contact, and on the type of constitutive interface law.

### 7.1.1 Lagrange multiplier method

The LAGRANGE multiplier method (5.27) is a mixed method. This means that the field of LAGRANGE multipliers  $\lambda_N$  and the displacement fields  $\mathbf{u}^\alpha$  of both contacting bodies, which define the gap function  $g_N$ , as well as its variations like  $\delta g_N$  (here shown for the frictionless case) have to be discretized. This yields for the residual

$$\int_{\Gamma_c} \lambda_N \delta g_N d\Gamma \longrightarrow \sum_{i=1}^{n_c} \int_{\Gamma_i^h} \lambda_N^h \delta g_N^h d\Gamma, \quad (7.1)$$

and for the constraint equation

$$\int_{\Gamma_c} \delta \lambda_N g_N d\Gamma = 0 \longrightarrow \sum_{i=1}^{n_c} \int_{\Gamma_i^h} \delta \lambda_N^h g_N^h d\Gamma = 0. \quad (7.2)$$

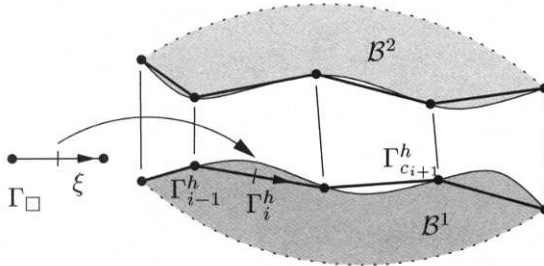
The interpolations for  $\lambda_N^h$  and  $\delta g_N^h$  are defined on  $\Gamma_{ci}^h$  by

$$\lambda_N^h = \sum_K M_K(\xi) \lambda_{NK} \quad \text{and} \quad \delta g_N^h = \sum_I N_I(\xi) \delta g_{NI}. \quad (7.3)$$

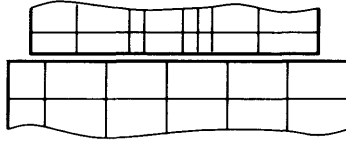
$\xi$  is a local, in general, convective coordinate which defines the shape functions  $N_I$  and  $M_K$  on the reference element; see Section 6.1.1 and Figure 7.2.  $\lambda_{NK}$  and  $\delta g_{NI}$  are the nodal values of the LAGRANGE multiplier and gap functions, respectively.  $n_c$  denotes the number of active contact constraints  $\mathcal{J}_A \in \mathcal{J}_C$  which have to be determined by the contact algorithm. For evaluation of the integrals in (7.1), it is not always clear on which side of the interface (body  $B^1$  or  $B^2$ ) this integration has to be carried out. Thus, one has to choose one of the surfaces of the bodies in the contact interface as the reference or *master* surface, and then perform the integration in a way that is consistent with the discretization. In Figure 7.2 the boundary  $\partial B^1$  is chosen as the master surface. The figure describes this discretization for the two-dimensional case when linear interpolations are used. In this case, the surface of the bodies coming into contact is also discretized by linear shape functions. It can also be seen that the sum in (7.1) has to be applied to add up all contributions of the contact elements associated with  $\Gamma_i^h$ . Using (7.3) in (7.1), one obtains

$$\int_{\Gamma_c} \lambda_N \delta g_N d\Gamma \longrightarrow \sum_{i=1}^{n_c} \sum_I \sum_K \delta g_{NI} \int_{(\xi)} N_I(\xi) M_K(\xi) \det \mathbf{J}_\Gamma(\xi) d\xi \lambda_{NK}, \quad (7.4)$$

where  $\det \mathbf{J}_\Gamma(\xi)$  is the transformation of a surface element in  $\Gamma_i^h$  to the reference element  $\Gamma_\square$ . By comparing this result to (6.48), we see that the integral in (7.4) has the structure of a mass matrix. This is especially true when the same interpolation functions are used for gap and LAGRANGE multipliers.



**Fig. 7.2** Contact discretization, isoparametric formulation.



**Fig. 7.3** Non-matching meshes at the contact interface.

Note that the interpolations have to be chosen in such a way that they fulfil the BABUSKA–BREZZI BB-condition for this mixed formulation (e.g. see Kikuchi and Oden (1988), Belgacem et al. (1999) or El-Abbasi and Bathe (2001)). This is usually no problem if the nodes of the contacting bodies assume the same position at the interface, see Figure 7.2.

However, if the bodies coming into contact are discretized by using different finite element meshes, then the nodes no longer match each other, see Figure 7.3. In this case, one has to carefully investigate the discretization in the light of the BB-condition. An early suggestion for a discretization for non-matching grids can be found in Simo et al. (1985). Recently, new methods, so-called mortar methods, were designed for domain decomposition in which unstructured grids are connected within a parallel finite element solution. These methods are well understood mathematically (see Bernadi et al. (1994), Tallec and Sassi (1995) or Wohlmuth (2000a)). It is obvious that one can also apply such a strategy to finite element contact problems when the nodes in the contact interface do not coincide, as can be seen in, for example, Figure 8.1. Work along these lines related to contact problems can be found in the mathematical literature in Belgacem et al. (1999) or Krause and Wohlmuth (2001) and in the engineering literature in Rebel et al. (2000) or McDevitt and Laursen (2000). Another approach for non-matching grids in the contact interface is provided by the NITSCHKE method, which only works with the primary displacement variables as described in Becker and Hansbo (1999). The mortar and NITSCHKE methods are discussed in Section 7.4.

Based on the LAGRANGE multiplier method, the following general matrix formulation can be derived. The discretization of a body  $B^\gamma$  has already been discussed in detail in Chapter 6. We recall these results for the linear elastic case of both bodies, which leads to the discrete form of the potential energy (5.25),

$$\Pi(\mathbf{u}) = \frac{1}{2} \mathbf{u}^T \mathbf{K} \mathbf{u} - \mathbf{u}^T \bar{\mathbf{f}}, \quad (7.5)$$

where the displacement vector  $\mathbf{u}$  includes the nodal displacements of both bodies. In the same way, the stiffness matrix  $\mathbf{K}$  is associated with both bodies, and  $\bar{\mathbf{f}}$  contains body forces and surface tractions of both bodies. In more detail, we can write

$$\Pi(\mathbf{u}) = \frac{1}{2} \langle \mathbf{u}^1, \mathbf{u}^2 \rangle \begin{bmatrix} \mathbf{K}^1 & \mathbf{0} \\ \mathbf{0} & \mathbf{K}^2 \end{bmatrix} \begin{Bmatrix} \mathbf{u}^1 \\ \mathbf{u}^2 \end{Bmatrix} - \langle \mathbf{u}^1, \mathbf{u}^2 \rangle \begin{Bmatrix} \bar{\mathbf{f}}^1 \\ \bar{\mathbf{f}}^2 \end{Bmatrix}. \quad (7.6)$$

Observe that the bodies are not yet coupled. This occurs due to the additional terms which are derived from (7.1) and (7.2). To obtain a matrix form for these terms we

introduce for the gap function in each contact element  $\Gamma_i^h$

$$g_N^h = \widehat{C}_i^T \mathbf{u}, \quad (7.7)$$

where  $\widehat{C}_i$  depends upon the choice of discretization. By integration, using (7.4), with (7.7) we obtain the vector  $C_i$ . Now the contributions can be assembled into a matrix which then contains all  $n_c$  constraints. With

$$C = [C_1 | C_2 | \dots | C_{n_c}] \quad (7.8)$$

we obtain the discrete form of (5.25), together with (5.26):

$$\Pi^{LM}(\mathbf{u}, \Lambda) = \frac{1}{2} \mathbf{u}^T \mathbf{K} \mathbf{u} - \mathbf{u}^T \bar{\mathbf{f}} + \Lambda^T C^T \mathbf{u}. \quad (7.9)$$

$\Lambda$  is the vector of all LAGRANGE multipliers associated with the  $n_c$  contact constraints. Variation of  $\Pi^{LM}(\mathbf{u}, \Lambda)$  with respect to displacements and LAGRANGE multipliers yields with  $\eta = \delta \mathbf{u}$  two equations

$$\begin{aligned} \eta^T [\mathbf{K} \mathbf{u} - \bar{\mathbf{f}} + C \Lambda] &= 0, \\ \delta \Lambda^T [C^T \mathbf{u}] &= 0. \end{aligned} \quad (7.10)$$

By the fundamental theorem of variations, these equations can be arranged in matrix form

$$\begin{bmatrix} \mathbf{K} & C \\ C^T & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \Lambda \end{Bmatrix} = \begin{Bmatrix} \bar{\mathbf{f}} \\ \mathbf{0} \end{Bmatrix}. \quad (7.11)$$

Note that this is a linear system of equations for the given number of  $n_c$  contact constraints. In a real contact analysis this number is not known. Thus, we then have a nonlinear problem in which, in addition to the displacement field and the LAGRANGE multipliers, the correct contact zone also has to be determined. Algorithms for this will be stated in Chapter 9.

In the following sections we shall discuss details for the different discretization which are only associated with the contact terms in (7.11).

### 7.1.2 Penalty method

Contrary to the LAGRANGE multiplier method, the penalty method only needs discretization of the displacement variables. The contribution of contact constraints to the weak form leads, with (3.7) and (5.32), to

$$\int_{\Gamma_c} \epsilon_N g_N^- \delta g_N^- d\Gamma \longrightarrow \int_{\Gamma_c^h} \epsilon_N g_N^h \delta g_N^h d\Gamma, \quad (7.12)$$

with the interpolation for the gap function and its variation

$$g_N^h = \sum_I N_I(\xi) g_{NI} \quad \text{and} \quad \delta g_N^h = \sum_I N_I(\xi) \delta g_{NI}. \quad (7.13)$$



But here also one has to be careful when choosing the interpolation for continuous contact. Since the penalty method is equivalent to a mixed method the BB-condition plays the same role as for the LAGRANGE multiplier approach (e.g. see Oden (1981) for a detailed discussion of this matter). Hence, when applying the penalty method one has to take special care to choose the correct discretization in the case of non-matching grids, see Figure 7.3.

The matrix form for the penalty method follows from (5.25), together with (5.31). Again we have to perform the integration in (7.12), which analogous to the definition (7.8) leads to

$$\Pi^P(\mathbf{u}) = \frac{1}{2} \mathbf{u}^T \mathbf{K} \mathbf{u} - \mathbf{u}^T \bar{\mathbf{f}} + \frac{\epsilon_N}{2} \mathbf{u}^T \mathbf{C} \mathbf{C}^T \mathbf{u}. \quad (7.14)$$

Variation of  $\Pi^P(\mathbf{u})$  with  $\boldsymbol{\eta} = \delta \mathbf{u}$  yields

$$\boldsymbol{\eta}^T \left[ \mathbf{K} \mathbf{u} - \bar{\mathbf{f}} + \epsilon_N \mathbf{C} \mathbf{C}^T \mathbf{u} \right] = 0, \quad (7.15)$$

which leads to the matrix form

$$\left[ \mathbf{K} + \mathbf{K}^P \right] \mathbf{u} = \bar{\mathbf{f}}, \quad \text{with } \mathbf{K}^P = \epsilon_N \mathbf{C} \mathbf{C}^T. \quad (7.16)$$

Again, this linear system of equations is given for the fixed number of  $n_c$  contact constraints. Algorithms to solve this form will be stated in Chapter 9.

In the following we discuss discretizations related to the methods mentioned above, and to the formulation using constitutive equations for the contact interface. Furthermore, to simplify notation, we will drop the superscript  $^h$  which denotes the approximation using finite elements.

## 7.2 NODE-TO-NODE CONTACT ELEMENT

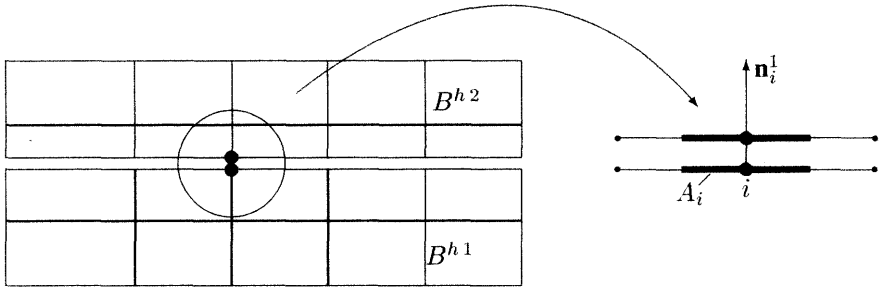
The simplest formulation for contact is a discretization which establishes constraint equations and contact interface constitutive equations on a purely nodal basis. Such a formulation will be called a *node-to-node contact*. For this discretization the frictionless as well as the frictional contact formulation is developed below.

### 7.2.1 Frictionless contact

A node-to-node contact can only be applied to geometrically linear problems, since large relative tangential movement of the nodes is not allowed in the contact area. The constraint equation for contact can then be formulated directly for each nodal pair, denoted in Figure 7.4. The geometrical contact constraint condition for the normal contact was stated in (3.11). It reads in this case for one node pair  $i$

$$g_{Ni} = (\mathbf{u}_i^2 - \mathbf{u}_i^1) \cdot \mathbf{n}_i^1 + g_i \geq 0. \quad (7.17)$$

Here  $\mathbf{u}_i^a$  are the displacement vectors of bodies  $B^a$  of the nodal pair associated with  $i$ .  $g_i$  is the initial gap between both nodes.



**Fig. 7.4** Node-to-node contact element.

For small strains the change in the normal is neglected within the computations. Since the initial gap is independent of the displacement field, the variation of (7.17) is given by

$$\delta g_{N i} = (\boldsymbol{\eta}_i^2 - \boldsymbol{\eta}_i^1) \cdot \mathbf{n}_i^1. \quad (7.18)$$

Due to its simplicity, and assuming that the contact constraint is active for  $n_c$  nodes, we can express the integral (7.1) for the contact contributions in the weak form by a sum over all active contact nodes.

**Lagrange multiplier formulation.** For the LAGRANGE multiplier method this leads to

$$\int_{\Gamma_c} \lambda_N \delta g_N d\Gamma \longrightarrow \sum_{i=1}^{n_c} \lambda_{N i} \delta g_{N i} A_i = \sum_{i=1}^{n_c} \lambda_{N i} (\boldsymbol{\eta}_i^2 - \boldsymbol{\eta}_i^1) \cdot \mathbf{n}_i^1 A_i, \quad (7.19)$$

and the weak form of the constraint equation

$$\int_{\Gamma_c} \delta \lambda_N g_N d\Gamma \longrightarrow \sum_{i=1}^{n_c} \delta \lambda_{N i} g_{N i} A_i = \sum_{i=1}^{n_c} \delta \lambda_{N i} [(\mathbf{u}_i^2 - \mathbf{u}_i^1) \cdot \mathbf{n}_i^1 + g_i] A_i = 0, \quad (7.20)$$

where  $n_c$  are the active contact nodes in  $\Gamma_c^h$ . The test functions  $\boldsymbol{\eta}_i^\alpha$  and the normal vector  $\mathbf{n}_i^1$  are defined for the node  $i$  as depicted in Figure 7.4. The product of LAGRANGE multiplier  $\lambda_{N i}$  and area  $A_i$  related to node  $i$  is the contact nodal force, and  $\lambda_{N i}$  is equivalent to the contact pressure  $p_{N i}$  associated with node  $i$ . The matrix form for this discretization can be obtained by introducing two vectors, one for the variations  $\{\boldsymbol{\eta}_i^\alpha, \delta \lambda_i\}$ , and one for the increments of the variables  $\{\Delta \mathbf{u}_i^\alpha, \Delta \lambda_i\}$  at

node  $i$ :

$$\hat{\eta}_i = \left\{ \begin{matrix} \eta_i \\ \delta \lambda_i \end{matrix} \right\}, \quad \Delta \hat{\mathbf{u}}_i = \left\{ \begin{matrix} \Delta \mathbf{u}_i \\ \Delta \lambda_i \end{matrix} \right\} \quad \text{with} \quad \eta_i = \left\{ \begin{matrix} \eta_i^2 \\ \eta_i^1 \end{matrix} \right\} \quad \text{and} \quad \Delta \mathbf{u}_i = \left\{ \begin{matrix} \Delta \mathbf{u}_i^2 \\ \Delta \mathbf{u}_i^1 \end{matrix} \right\}. \quad (7.21)$$

Inserting this into (7.19) and (7.20) yields the contact residuum

$$\int_{\Gamma_c} (\delta \lambda_N g_N + \lambda_N \delta g_N) d\Gamma \approx \sum_{i=1}^{n_c} \hat{\eta}_i^T \mathbf{G}_i^{cL} \quad \text{with} \quad \mathbf{G}_i^{cL} = \left\{ \begin{matrix} \lambda_{Ni} \mathbf{C}_i \\ \mathbf{u}_i^T \mathbf{C}_i \end{matrix} \right\} A_i \quad (7.22)$$

and

$$\mathbf{C}_i = \left\{ \begin{matrix} \mathbf{n}_i^1 \\ -\mathbf{n}_i^1 \end{matrix} \right\} \quad \text{and} \quad \mathbf{u}_i = \left\{ \begin{matrix} \mathbf{u}_i^2 \\ \mathbf{u}_i^1 \end{matrix} \right\}. \quad (7.23)$$

Furthermore from the linearization of (7.22) we obtain the expression

$$\hat{\eta}_i^T \mathbf{K}_i^{cL} \Delta \hat{\mathbf{u}}_i \quad (7.24)$$

with the contact stiffness matrix for one node  $i$ ,

$$\mathbf{K}_i^{cL} = \begin{bmatrix} \mathbf{0} & \mathbf{C}_i \\ \mathbf{C}_i^T & 0 \end{bmatrix} A_i. \quad (7.25)$$

**Penalty formulation.** In the case of the penalty method, we have to discretize equation (7.19) for  $\lambda_N = \epsilon_N g_N$ , leading with (7.21) and (7.23) to

$$\int_{\Gamma_c} \epsilon_N g_N \delta g_N d\Gamma \longrightarrow \sum_{i=1}^{n_c} \epsilon_N g_{Ni} \delta g_{Ni} A_i = \sum_{i=1}^{n_c} \eta_i^T \epsilon_N g_{Ni} \mathbf{C}_i A_i. \quad (7.26)$$

Often the area  $A_i$  which is associated with the contact point  $i$  (see Figure 7.4) is neglected (or “hidden” in the penalty parameter  $\epsilon_N$ ) in the node-to-node contact formulation. This means that the contact stress  $p_{Ni} = \epsilon_N g_{Ni}$  becomes a contact (nodal) force  $f_{Ni} = \epsilon_N A_i g_{Ni} = \bar{\epsilon}_N g_{Ni}$ . An evaluation of a contact interface law like (4.11) is not possible with this simplifying discretization.

The associated matrix formulation for (7.26) results in the geometrically linear case for node  $i$  to the definition of the contact residual  $G_i^c = \eta_i^T \mathbf{G}_i^{cP}$  and its associated tangent matrix  $\mathbf{K}_i^{cP}$ . The explicit form can be stated by using the notation defined in equation (7.23):

$$\mathbf{G}_i^{cP} = \epsilon_N g_{Ni} A_i \mathbf{C}_i, \quad \mathbf{K}_i^{cP} = \epsilon_N A_i \mathbf{C}_i \mathbf{C}_i^T. \quad (7.27)$$

LAGRANGE multiplier and penalty formulation, as given here, are valid for one-, two- and three-dimensional formulations. One only has to adapt the definition of the normal  $\mathbf{n}_i^1$ , the variations  $\eta_i^\alpha$  and the displacements  $\mathbf{u}_i^\alpha$ , at the contacting node  $i$  in relation to the spatial dimension of the problem. As examples we state here the penalty residuals and tangents for the one- and two-dimensional cases.

- **One-dimensional case.** If contact between one-dimensional structures is considered, then the normal vector  $\mathbf{n}_i^1$  is given by the scalar  $n_x = 1$ . Hence the matrix  $\mathbf{C}_i$  is simply  $\mathbf{C}_i^T = \langle 1, -1 \rangle$ , which yields the residual and tangent

$$\mathbf{G}_i^{cP} = \epsilon_N g_{Ni} A_i \begin{Bmatrix} 1 \\ -1 \end{Bmatrix} \quad \text{and} \quad \mathbf{K}_i^{cP} = \epsilon_N A_i \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (7.28)$$

This result can be compared with the formulation used in Section 5.4 to evaluate different methods to formulate the contact constraints.

- **Two-dimensional case.** For two-dimensional problems the normal vector has two components  $(\mathbf{n}_i^1)^T = \langle n_x, n_y \rangle_i$ . Thus the matrix  $\mathbf{C}_i$ , and hence also the residual matrix, now have four components. Explicitly, one obtains

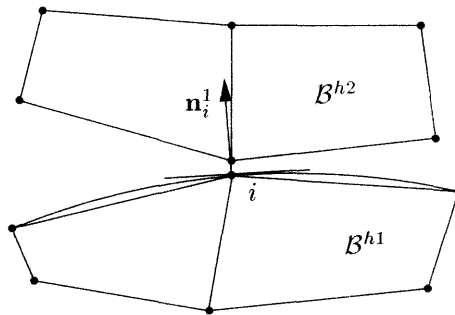
$$\mathbf{G}_i^{cP} = \epsilon_N g_{Ni} A_i \begin{Bmatrix} n_x \\ n_y \\ -n_x \\ -n_y \end{Bmatrix}_i \quad (7.29)$$

and

$$\mathbf{K}_i^{cP} = \epsilon_N A_i \begin{bmatrix} n_x^2 & n_x n_y & -n_x^2 & -n_x n_y \\ n_x n_y & n_y^2 & -n_x n_y & -n_y^2 \\ -n_x^2 & -n_x n_y & n_x^2 & n_x n_y \\ -n_x n_y & -n_y^2 & n_x n_y & n_y^2 \end{bmatrix}_i. \quad (7.30)$$

We note that it is not always obvious how to define the normal vector for a given interpolation of the contact boundary. This is especially true when a linear interpolation, as shown in Figure 7.5, is used. Here it can be seen that there is a jump in the normal at point  $i$ . Thus, the definition of the normal at point  $i$  is not unique. A way out is given by using a normal  $\mathbf{n}_i^1$ , which is obtained from a smooth interpolation of the discretized contact surface by, for example, BÉZIER or spline functions, as depicted in Figure 7.5.

The problem of defining the normal at nodal points can also be circumvented by introducing an isoparametric interpolation for the contact surface, see next section.



**Fig. 7.5** Definition of the contact normal.

## 7.2.2 Contact with friction

In the case of frictional contact, we have to distinguish between stick and slip motion. The theoretical background can be found in Section 4.2. Here only the classical law of COULOMB is applied. The relative tangential movement between the two contacting nodes is given with respect to (3.23) by

$$\mathbf{g}_{Ti} = (\mathbf{e}_{T\alpha}^1 \otimes \mathbf{e}_{T\alpha}^1) (\mathbf{u}_i^2 - \mathbf{u}_i^1), \quad (7.31)$$

where  $\mathbf{e}_{T\alpha}^1$  is the unit tangent vector at node  $i$  in the direction  $\alpha = 1, \dots, 2$ . The tangent vectors have to be constructed such that they form an orthogonal frame with the normal vector  $\mathbf{n}_i^1$ , hence  $\mathbf{e}_{T\alpha}^1 \mathbf{e}_{T\beta}^1 = \delta_{\alpha\beta}$ . In the same way, we define the relative tangential velocity

$$\dot{\mathbf{g}}_{Ti} = (\mathbf{e}_{T\alpha}^1 \otimes \mathbf{e}_{T\alpha}^1) (\dot{\mathbf{u}}_i^2 - \dot{\mathbf{u}}_i^1). \quad (7.32)$$

First we consider the stick case. This will be done here using the penalty method, see Section 5.3.2. Again, the integral in (5.32) is discretized by a sum over the nodes in contact

$$\begin{aligned} \int_{\Gamma_c} \epsilon_T \mathbf{g}_T \cdot \delta \mathbf{g}_T dA &\longrightarrow \sum_{i=1}^{n_c} \epsilon_T \delta \mathbf{g}_{Ti}^T \mathbf{g}_{Ti} A_i \\ &= \sum_{i=1}^{n_c} \epsilon_T [(\mathbf{u}_i^2 - \mathbf{u}_i^1)^T \mathbf{e}_{T\alpha}^1] [(\boldsymbol{\eta}_i^2 - \boldsymbol{\eta}_i^1)^T \mathbf{e}_{T\alpha}^1] A_i. \end{aligned} \quad (7.33)$$

By introducing the matrix

$$\mathbf{T}_{\alpha i} = \begin{Bmatrix} \mathbf{e}_{T\alpha}^1 \\ -\mathbf{e}_{T\alpha}^1 \end{Bmatrix}, \quad (7.34)$$

we can state the matrix formulation for (7.33). This results in the geometrically linear case for node  $i$  to the definition of the contact residual for tangential stick

$$\int_{\Gamma_c} \epsilon_T \mathbf{g}_T \cdot \delta \mathbf{g}_T dA = \sum_{i=1}^{n_c} \boldsymbol{\eta}_i^T \mathbf{G}_{Ti}^{cst}. \quad (7.35)$$

The explicit form can be stated by using the notation defined in equation (7.23):

$$\mathbf{G}_{Ti}^{cst} = \epsilon_T A_i \sum_{\alpha=1}^2 g_{T\alpha i} \mathbf{T}_{\alpha i}, \quad (7.36)$$

with  $g_{T\alpha i} = \mathbf{T}_{\alpha i}^T \mathbf{u}_i$ . The associated tangent matrix follows from the linearization of (7.36), which only depends in  $g_{T\alpha i}$  in the geometrical linear case upon the displacement field. With  $\Delta g_{T\alpha i} = \mathbf{T}_{\alpha i}^T \Delta \mathbf{u}_i$  we obtain

$$\mathbf{K}_{Ti}^{cst} = \epsilon_T A_i \sum_{\alpha=1}^2 \mathbf{T}_{\alpha i} \mathbf{T}_{\alpha i}^T. \quad (7.37)$$

In the case of slip, the virtual work expression for the contact contribution is derived using an implicit backward EULER integration of (4.27), see Sections 4.2.4 and 9.5.2. Using these results we obtain

$$\int_{\Gamma_c} \mathbf{t}_{T\,n+1} \cdot \delta \mathbf{g}_T dA = \int_{\Gamma_c} \mu p_{N\,n+1} \mathbf{n}_{T\,n+1}^{tr} \cdot \delta \mathbf{g}_T dA = \sum_{i=1}^{n_c} \boldsymbol{\eta}_i^T \mathbf{G}_{T\,i}^{c\,slip}. \quad (7.38)$$

In this equation the residual slip vector  $\mathbf{G}_{T\,i}^{c\,slip}$  for contact node  $i$  is given by

$$\mathbf{G}_{T\,i}^{c\,slip} = \mu p_{N\,n+1} A_i \sum_{\alpha=1}^2 n_{T\,\alpha\,i}^{tr} \mathbf{T}_{\alpha\,i} \quad \text{with} \quad n_{T\,\alpha\,i}^{tr} = (\mathbf{n}_{T\,n+1}^{tr})^T \mathbf{e}_{T\,\alpha}^1. \quad (7.39)$$

The linearization of (7.38) is now constructed from two terms, since  $\mathbf{n}_{T\,n+1}^{tr}$  as well as  $p_{N\,n+1}$  depend upon the displacement field. The linearization of  $\mathbf{n}_{T\,n+1}^{tr}$  yields for a node  $i$

$$\Delta \mathbf{n}_{T\,n+1}^{tr} = \Delta \frac{\mathbf{t}_{T\,n+1}^{tr}}{\|\mathbf{t}_{T\,n+1}^{tr}\|} = \frac{1}{\|\mathbf{t}_{T\,n+1}^{tr}\|} [1 - \mathbf{n}_{T\,n+1}^{tr} (\mathbf{n}_{T\,n+1}^{tr})^T] \Delta \mathbf{t}_{T\,n+1}^{tr} \quad (7.40)$$

where the linearization of the tangential nodal force with (7.32) yields

$$\Delta \mathbf{t}_{T\,n+1}^{tr} = \epsilon_T \Delta \mathbf{g}_{T\,n+1} = \epsilon_T \sum_{\alpha=1}^2 \left( \mathbf{T}_{\alpha\,i}^T \Delta \mathbf{u}_i \right) \mathbf{e}_{T\,\alpha}^1. \quad (7.41)$$

Combining the last two equations, the contribution to the linearized weak form can be computed with (7.39) as

$$\delta \mathbf{g}_{T\,i}^T \Delta \mathbf{t}_{T\,n+1}^{tr} = \epsilon_T \sum_{\alpha=1}^2 \sum_{\beta=1}^2 \boldsymbol{\eta}_i^T \left[ (\delta_{\alpha\beta} - n_{T\,\alpha}^{tr} n_{T\,\beta}^{tr}) \mathbf{T}_{\alpha\,i} \mathbf{T}_{\beta\,i}^T \right] \Delta \mathbf{u}_i. \quad (7.42)$$

For node  $i$  the linearization of the normal contact pressure  $p_{N\,n+1}$  follows with the matrices introduced for the frictionless case in (7.26),

$$\Delta p_{N\,n+1} = \epsilon_N \Delta g_{N\,i} = \epsilon_N \mathbf{C}_i^T \Delta \mathbf{u}_i. \quad (7.43)$$

This adds

$$\delta \mathbf{g}_{T\,i}^T (\mu \Delta p_{N\,n+1} \mathbf{n}_{T\,n+1}) = \mu \epsilon_N \sum_{\alpha=1}^2 \boldsymbol{\eta}_i^T n_{T\,\alpha}^{tr} \mathbf{T}_{\alpha\,i} \mathbf{C}_i^T \Delta \mathbf{u}_i \quad (7.44)$$

to the linearized weak form. Combining (7.42) and (7.44) yields the tangent matrix for node  $i$  which is in the slip state

$$\mathbf{K}_{T\,i}^{c\,slip} = \epsilon_T \sum_{\alpha=1}^2 \left\{ \sum_{\beta=1}^2 \left[ (\delta_{\alpha\beta} - n_{T\,\alpha}^{tr} n_{T\,\beta}^{tr}) \mathbf{T}_{\alpha\,i} \mathbf{T}_{\beta\,i}^T \right] + \mu \epsilon_N n_{T\,\alpha}^{tr} \mathbf{T}_{\alpha\,i} \mathbf{C}_i^T \right\}. \quad (7.45)$$

Due to the second term, the tangent matrix is non-symmetric. This is because the COULOMB law of friction can be viewed as a non-associative constitutive equation, and hence cannot be represented by a positive strain energy function.

Note that the first term in (7.45) is zero in the two-dimensional case. This stems from the fact that the quantity  $n_{T\alpha}^{tr}$  can then be represented by a signum function, since  $t_{Tn+1}^{tr} / |t_{Tn+1}^{tr}| = \text{Sign}(t_{T\alpha}^{tr})$ , which leads to  $(\delta_{\alpha\beta} - n_{T\alpha}^{tr} n_{T\beta}^{tr}) = 1 - \text{Sign}(t_{Tn+1}^{tr})^2 = 0$ . Thus only the non-symmetric part of (7.45) remains in the two-dimensional case:

$$\mathbf{K}_{Ti}^{c\ slip} = \mu \epsilon_N \text{Sign}(t_{Tn+1}^{tr}) \mathbf{T}_i \mathbf{C}_i^T \quad \text{with} \quad \mathbf{T}_i = \begin{Bmatrix} \mathbf{e}_{Ti}^1 \\ -\mathbf{e}_{Ti}^1 \end{Bmatrix}, \quad (7.46)$$

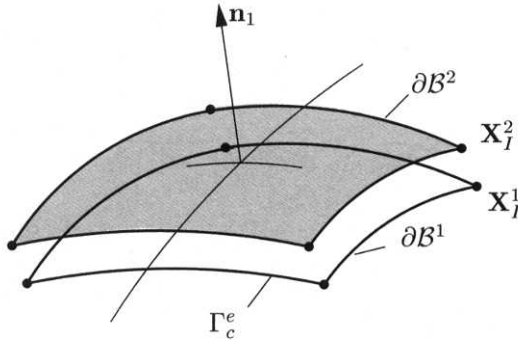
where  $\mathbf{e}_{Ti}^1$  is the unit tangent vector.

### 7.3 ISOPARAMETRIC DISCRETIZATION OF THE CONTACT CONTRIBUTION

Another possibility to discretize the contact surface is given by a direct interpolation of the surface using the isoparametric formulation. Also, this contact element does not allow large relative tangential movement in the contact area, and thus is only valid for geometrically linear applications, as discussed below. Such an element connects the surfaces  $\partial B^1$  and  $\partial B^2$  at the contact interface, as shown in Figure 7.6.

The gap function  $g_N$ , stated in (3.11) for geometrically linear analysis, is discretized by an isoparametric interpolation, also leading to a well defined contact pressure. For the gap function we obtain its variation and linearization

$$\begin{aligned} g_N &= \sum_I N_I(\xi, \eta) g_{NI}, \\ \delta g_N &= \sum_I N_I(\xi, \eta) (\eta_I^2 - \eta_I^1) \cdot \mathbf{n}^1, \end{aligned} \quad (7.47)$$



**Fig. 7.6** Isoparametric contact element for three-dimensional problems.

$$\Delta g_N = \sum_I N_I(\xi, \eta) (\Delta \mathbf{u}_I^2 - \Delta \mathbf{u}_I^1) \cdot \mathbf{n}^1,$$

where  $I$  denotes the nodal points of the isoparametric surface elements.

The contact normal used in (7.48) can be constructed directly from the interpolation. This is due to the fact that the tangent vectors to the surface can be obtained directly from the convective description of the surface by the isoparametric interpolation, see Figure 7.7 and (B.2) in Appendix B. With this we compute the normal vector to the discretized surface of body  $\mathcal{B}^1$ :

$$\mathbf{n}^1 = \frac{\mathbf{G}_1^1 \times \mathbf{G}_2^1}{\|\mathbf{G}_1^1 \times \mathbf{G}_2^1\|}. \quad (7.48)$$

The tangent vectors are obtained by the partial derivative of the position vector to the initial configuration of the contact surface,  $\mathbf{G}_\alpha^1 = \mathbf{X}_{,\alpha}^1$ , with respect to the coordinates  $\xi$  and  $\eta$ . Note that all vectors are now defined on  $\partial \mathcal{B}^1$ , which is denoted by the superscript 1. Thus, for the contact normal we have

$$\mathbf{n}^1 = \frac{\mathbf{X}_{,\xi}^1 \times \mathbf{X}_{,\eta}^1}{\|\mathbf{X}_{,\xi}^1 \times \mathbf{X}_{,\eta}^1\|}. \quad (7.49)$$

As we apply an isoparametric interpolation, we use the same shape functions for displacements and coordinates. The position vector  $\mathbf{X}^1$  is approximated by

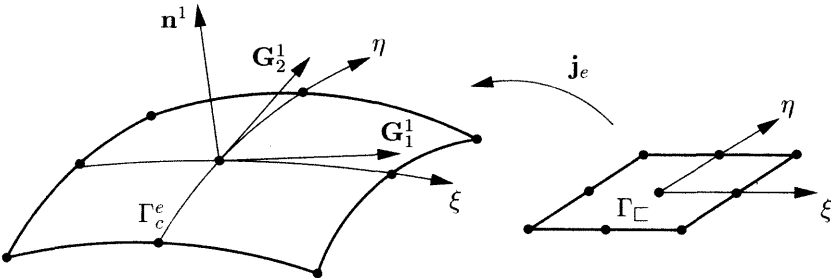
$$\mathbf{X}^1 = \sum_{I=1}^m N_I(\xi, \eta) \mathbf{X}_I^1, \quad (7.50)$$

which leads to the derivatives for its components  $X_i$ ,

$$X_{i,\alpha}^1 = \sum_{I=1}^m N_I(\xi, \eta)_{,\alpha} X_{iI}^1. \quad (7.51)$$

With this we can explicitly evaluate the cross product in (7.49), which yields

$$\mathbf{N}^1(\xi, \eta) = \mathbf{X}_{,\xi}^1 \times \mathbf{X}_{,\eta}^1 = \begin{Bmatrix} X_{2,\xi}^1 X_{3,\eta}^1 - X_{3,\xi}^1 X_{2,\eta}^1 \\ X_{3,\xi}^1 X_{1,\eta}^1 - X_{1,\xi}^1 X_{3,\eta}^1 \\ X_{1,\xi}^1 X_{2,\eta}^1 - X_{2,\xi}^1 X_{1,\eta}^1 \end{Bmatrix}. \quad (7.52)$$



**Fig. 7.7** Computation of the normal vector.



The matrix formulation for (7.48) can be derived by introducing the vectors

$$\eta_{cI} = \begin{Bmatrix} \eta_I^2 \\ \eta_I^1 \end{Bmatrix}, \quad \Delta \mathbf{u}_{cI} = \begin{Bmatrix} \Delta \mathbf{u}_I^2 \\ \Delta \mathbf{u}_I^1 \end{Bmatrix} \quad \text{and} \quad \hat{\mathbf{N}} = \begin{Bmatrix} \mathbf{N}^1 \\ -\mathbf{N}^1 \end{Bmatrix}. \quad (7.53)$$

This leads to the variation and linearization of the gap function:

$$\begin{aligned} \delta g_N &= \sum_I \eta_{cI}^T [N_I(\xi, \eta) \hat{\mathbf{N}}(\xi, \eta) / \|\mathbf{N}^1\|], \\ \Delta g_N &= \left[ \sum_I N_I(\xi, \eta) \hat{\mathbf{N}}(\xi, \eta)^T / \|\mathbf{N}^1\| \right] \Delta \mathbf{u}_{cI}. \end{aligned} \quad (7.54)$$

**Lagrange multiplier formulation.** Using LAGRANGE multipliers we have, in addition to discretization (7.54), to introduce a discretization for the LAGRANGE multiplier and its variation:

$$\lambda_N = \sum_K M_K(\xi, \eta) \lambda_{NK} \quad \text{and} \quad \delta \lambda_N = \sum_K M_K(\xi, \eta) \delta \lambda_{NK}. \quad (7.55)$$

Note that interpolation for the LAGRANGE multiplier has to be chosen in relation to the interpolation of the displacement field (7.48) such that the BB-condition is fulfilled, e.g. see Kikuchi and Oden (1988). Here for the moment we use the general formulation with different interpolations for  $g_N$  and  $\lambda_N$ , without specifying the shape functions  $N_I$  and  $M_K$  explicitly. With (7.52), for the contact contribution of the weak form of equilibrium we can write

$$\begin{aligned} \int_{\Gamma_c^e} \lambda_N \delta g_N d\Gamma &\approx \sum_{I=1}^n \eta_{cI}^T \mathbf{G}_{cI}^u \\ \text{with} \quad \mathbf{G}_{cI}^u &= \int_{-1}^{+1} \int_{-1}^{+1} \lambda_N(\xi, \eta) N_I(\xi, \eta) \hat{\mathbf{N}}(\xi, \eta) d\xi d\eta. \end{aligned} \quad (7.56)$$

The area element  $d\Gamma$  in the contact surface  $\Gamma_c^e$  can be computed using the reference configuration  $\Gamma_\square$ , see Figure 7.7. The area element  $d\Gamma$  is given by  $d\Gamma = \|\mathbf{X}_{,\xi}^1 \times \mathbf{X}_{,\eta}^1\| d\xi d\eta = \|\mathbf{N}^1\| d\xi d\eta$ . Note that  $\|\mathbf{N}^1\|$  appears in the denominator of (7.54), and as a factor in the area element in the reference configuration, hence this term is cancelled in the weak form (7.56). Furthermore, for the weak form of the constraint we obtain

$$\begin{aligned} \int_{\Gamma_c^e} \delta \lambda_N g_N d\Gamma &\approx \sum_{K=1}^m \delta \lambda_{NK} G_{cK}^\lambda = 0 \\ \text{with} \quad G_{cK}^\lambda &= \int_{-1}^{+1} \int_{-1}^{+1} M_K(\xi, \eta) g_N(\xi, \eta) \|\mathbf{N}^1\| d\xi d\eta. \end{aligned} \quad (7.57)$$

In matrix form, (7.56) is represented by

$$\int_{\Gamma_e^c} \lambda_N \delta g_N d\Gamma \approx \langle \boldsymbol{\eta}_{c1}^T, \dots, \boldsymbol{\eta}_{cn}^T \rangle \begin{Bmatrix} \mathbf{G}_{c1}^u \\ \vdots \\ \mathbf{G}_{cn}^u \end{Bmatrix}, \quad (7.58)$$

and (7.57) yields

$$\int_{\Gamma_e^c} \delta \lambda_N g_N d\Gamma \approx \langle \delta \lambda_{N1}, \dots, \delta \lambda_{Nm} \rangle \begin{Bmatrix} G_{c1}^\lambda \\ \vdots \\ G_{cm}^\lambda \end{Bmatrix}. \quad (7.59)$$

The linearization of (7.56) and (7.57) leads to

$$\int_{\Gamma_e^c} \Delta \lambda_N \delta g_N d\Gamma \approx \sum_{I=1}^n \sum_{K=1}^m \boldsymbol{\eta}_{cI}^T \mathbf{C}_{IK} \Delta \lambda_{NK} \quad (7.60)$$

$$\text{with} \quad \mathbf{C}_{IK} = \int_{-1}^{+1} \int_{-1}^{+1} M_K(\xi, \eta) N_I(\xi, \eta) \hat{\mathbf{N}}(\xi, \eta) d\xi d\eta,$$

and

$$\int_{\Gamma_e^c} \delta \lambda_N \Delta g_N d\Gamma \approx \sum_{I=1}^n \sum_{K=1}^m \delta \lambda_{NK} \mathbf{C}_{KI} \Delta \mathbf{u}_I, \quad (7.61)$$

with  $\mathbf{C}_{KI} = \mathbf{C}_{IK}^T$ . Hence, the complete matrix form for the linearization is represented for one contact element by

$$\langle \boldsymbol{\eta}_{c1}^T, \dots, \boldsymbol{\eta}_{cn}^T, \delta \lambda_{N1}, \dots, \delta \lambda_{Nm} \rangle \mathbf{K}_e^{LM} \begin{Bmatrix} \Delta \mathbf{u}_1 \\ \vdots \\ \Delta \mathbf{u}_n \\ \Delta \lambda_{N1} \\ \vdots \\ \Delta \lambda_{Nm} \end{Bmatrix}, \quad (7.62)$$

with

$$\mathbf{K}_e^{LM} = \begin{bmatrix} & & & \mathbf{C}_{11} & \dots & \mathbf{C}_{1m} \\ & \mathbf{0} & & \vdots & \vdots & \vdots \\ \mathbf{C}_{11}^T & \dots & \mathbf{C}_{n1}^T & \mathbf{C}_{n1} & \dots & \mathbf{C}_{nm} \\ \vdots & \vdots & \vdots & & & \\ \mathbf{C}_{1m}^T & \dots & \mathbf{C}_{nm}^T & & \mathbf{0} & \end{bmatrix}, \quad (7.63)$$

where the matrix elements  $\mathbf{C}_{IK}$  are defined in (7.60). In this expression, the indices for the displacement variables run over the number of nodes used to specify the

isoparametric interpolation, e.g.  $n = 4$  for bilinear and  $n = 9$  for biquadratic shape functions. The same holds for the LAGRANGE multipliers, where now  $m$  represents the number of nodes used for the interpolation.

**Penalty formulation.** The weak form for the penalty method is given in equation (7.12). With (7.48) we can represent the weak form contribution of one element by further using (7.52), and introduce the finite element discretization directly to obtain for (7.12) a compact form

$$\int_{\Gamma_c} \epsilon_N g_N \delta g_N d\Gamma \approx \bigcup_{c=1}^{n_c} \sum_{I=1}^n \eta_{cI}^T \mathbf{G}_I^c, \quad (7.64)$$

with

$$\mathbf{G}_I^c = \int_{\Gamma_\square} \epsilon_N g_N(\xi, \eta) N_I(\xi, \eta) \hat{\mathbf{N}} d\xi d\eta, \quad (7.65)$$

where  $\bigcup_{c=1}^{n_c}$  denotes the assembly of  $n_c$  active contact elements and  $\Gamma_\square$  is the contact surface of the reference element, see Figure 7.7. In (7.64) the contact pressure  $p_N$  is given by  $p_N = \epsilon_N g_N(\xi, \eta)$  which is defined for every point of the contact element domain  $\Gamma_c^e$ .

The linearization of contact residual (7.64) yields, with (7.54), the result

$$\begin{aligned} \int_{\Gamma_c^e} \epsilon_N \delta g_N \Delta g_N d\Gamma &\approx \sum_I \sum_K \eta_{cI}^T \mathbf{K}_{JK}^P \Delta \mathbf{u}_{cK} \\ \text{with } \mathbf{K}_{IK}^P &= \int_{-1}^{+1} \int_{-1}^{+1} \epsilon_N N_I(\xi, \eta) N_K(\xi, \eta) \hat{\mathbf{N}} \hat{\mathbf{N}}^T / \|\mathbf{N}^1\| d\xi d\eta. \end{aligned} \quad (7.66)$$

Both, residual vector (7.64) and tangent matrix (7.66)<sub>2</sub> have to be evaluated using numerical integration. For the proper choice of the numerical integration rule, see Oden (1981), who has discussed this topic in the context of perturbed LAGRANGE formulations. To avoid locking and pressure oscillations, a reduced GAUSS integration or an integration based on SIMPSON'S rule works well. This means that, for instance, a one-point GAUSS integration should be used for a bilinear interpolation. In the case of the application of SIMPSON'S rule, the integral has to be evaluated at the four nodal points of the bilinear element.

The isoparametric discretization leads to a contact element which can be applied in the context with four, nine node quadrilaterals or triangular elements for three-dimensional continuum problems. Due to the smooth discretization, a good approximation of the contact pressure is obtained.

### 7.3.1 Examples for isoparametric contact elements

To give a more detailed description of this type of contact discretization, we consider some formulations in more detail. These are mixed formulations based on the so-called Q1/P0 approximation, and several two-dimensional formulations for the

LAGRANGE, penalty and perturbed LAGRANGE methods.

**Example 1:** As a first example we derive the explicit matrix formulation for a contact element with bilinear interpolation for the displacement field and constant approximation of the contact pressure in the case of frictionless contact. This is equivalent to the well known Q1/P0 element used in solid mechanics for incompressible problems. To make the formulation more flexible, we start from the perturbed LAGRANGE form discussed in Section 5.3.6. In that case, one has to approximate equation (5.52), which involves LAGRANGE multipliers and a regularization term. The weak form of the contact contributions is then

$$\delta C_c^P = \int_{\Gamma_c} \left[ \lambda_N \delta g_N + \delta \lambda_N \left( g_N - \frac{1}{\epsilon_N} \lambda_N \right) \right] d\Gamma. \quad (7.67)$$

Using the bilinear shape function, given explicitly in (6.24), we can discretize (7.67). Furthermore, the LAGRANGE multiplier  $\lambda_N$  is approximated by a constant value in the element represented here by  $\bar{\lambda}_N$ . This interpolation leads, with (7.48), (7.52) and (7.56), to the weak form contribution to equilibrium of a single contact element:

$$\int_{\Gamma_c^e} \lambda_N \delta g_N d\Gamma \approx \bar{\lambda}_N \sum_I \hat{\eta}_I^T \int_{-1}^{+1} \int_{-1}^{+1} N_I(\xi, \eta) \hat{\mathbf{N}} d\xi d\eta. \quad (7.68)$$

Furthermore, for the weak form of the constraint equation we obtain

$$\begin{aligned} \int_{\Gamma_c^e} \delta \lambda_N \left( g_N - \frac{1}{\epsilon_N} \lambda_N \right) d\Gamma \\ \approx \delta \bar{\lambda}_N \left[ \int_{-1}^{+1} \int_{-1}^{+1} g_N(\xi, \eta) \|\mathbf{N}^1\| d\xi d\eta - \frac{1}{\epsilon_N} \bar{\lambda}_N A_e \right] = 0, \end{aligned} \quad (7.69)$$

where the  $A_e$  represents the area of the contact element. This equation represents a regularization (constitutive equation) for the impenetrability constraint condition (3.6). It has to be fulfilled in an average sense over the finite element  $\Gamma_c^e$ . In case the second term is zero, we recover the classical LAGRANGE multiplier method. We can also obtain a penalty formulation by eliminating the LAGRANGE multiplier using (7.69). Solving this equation for the LAGRANGE multiplier  $\bar{\lambda}_N$  yields

$$\bar{\lambda}_N = \frac{\epsilon_N}{A_e} \int_{-1}^{+1} \int_{-1}^{+1} g_N(\xi, \eta) \|\mathbf{N}^1\| d\xi d\eta. \quad (7.70)$$

This result can be inserted into (7.68), which leads to

$$\int_{\Gamma_c^e} \lambda_N \delta g_N d\Gamma \approx \frac{\epsilon_N}{A_e} \sum_I \hat{\eta}_I^T \int_{-1}^{+1} \int_{-1}^{+1} g_N \|\mathbf{N}^1\| d\xi d\eta \mathbf{C}_I \quad (7.71)$$

with

$$C_I = \int_{-1}^{+1} \int_{-1}^{+1} N_I \hat{N} d\xi d\eta. \quad (7.72)$$

Note, however, that in this special formulation a double integration is necessary where  $g_N$ ,  $N^1$ ,  $N_I$  and  $\hat{N}$  are functions of  $\xi$  and  $\eta$ . With this the LAGRANGE multiplier is eliminated from the formulation, which is now a penalty formulation with a special interpolation (Q1/P0: linear for the displacement field and constant for the contact pressure  $\lambda_N$ ). The check for penetration has to be performed here for the complete contact element by evaluation of the integral condition

$$\bar{g}_N < 0 \implies \text{contact} \quad \text{with} \quad \bar{g}_N = \int_{-1}^{+1} \int_{-1}^{+1} g_N(\xi, \eta) \|N^1\| d\xi d\eta. \quad (7.73)$$

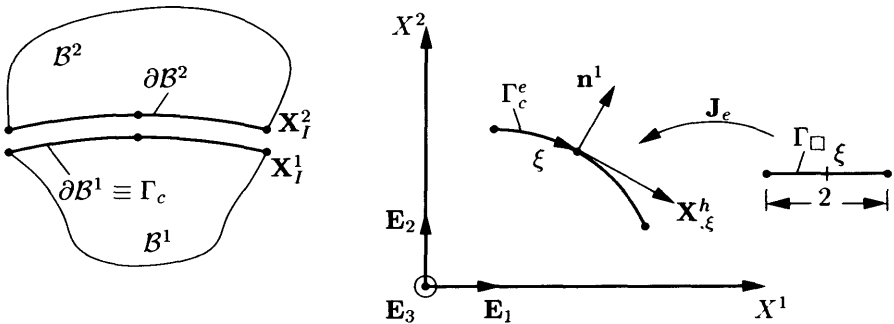
The linearization of (7.71) with respect to the unknown displacements only affects the first integral, which contains  $g_N$ . Since  $\Delta g_N$  has the same structure as  $\delta g_N$ , we obtain with (7.72)

$$\int_{\Gamma_c^e} \Delta \lambda_N \delta g_N d\Gamma \approx \frac{\epsilon_N}{A_e} \sum_I \sum_K \hat{\eta}_I^T C_I C_K^T \Delta u_K. \quad (7.74)$$

**Example 2:** As a second example, we formulate the contact element for two-dimensional problems. It then reduces to a one-dimensional element, see Figure 7.8. As in (7.48), for the gap function and its variation we obtain

$$g_N = \sum_I N_I(\xi) g_{NI} \quad \text{and} \quad \delta g_N = \sum_I N_I(\xi) (\eta_I^2 - \eta_I^1) \cdot n^1, \quad (7.75)$$

where now the isoparametric shape functions can be chosen according to Section 6.1.2. The contact normal can be constructed directly from the interpolation by the cross product of the tangent vector and the vector  $E_3$  perpendicular to the plane, see



**Fig. 7.8** Isoparametric contact element for two-dimensional problems.

Figure 7.8. With this we compute the normal vector to the discretized surface of body  $\mathcal{B}^1$  from

$$\mathbf{n}^1 = \frac{\mathbf{E}_3 \times \mathbf{X}_{,\xi}^1}{\|\mathbf{X}_{,\xi}^1\|}, \quad (7.76)$$

where the tangent vector is computed as in (7.49). Since the line element  $dS$  in the contact line can be computed using the reference configuration  $\Gamma_{\square}$  of the isoparametric formulation, we have  $dS = \|\mathbf{X}_{,\xi}^1\| d\xi$ . Now for the different contact formulations with (7.76), we can state the weak form contribution of the contact stresses for one element:

1. LAGRANGE multiplier method:

$$\int_{\Gamma_c^e} \lambda_N \delta g_N d\Gamma = \int_{-1}^{+1} \lambda_N(\xi) (\boldsymbol{\eta}^2 - \boldsymbol{\eta}^1) \cdot (\mathbf{E}_3 \times \mathbf{X}_{,\xi}^1) d\xi \quad (7.77)$$

$$\int_{\Gamma_c^e} \delta \lambda_N g_N d\Gamma = \int_{-1}^{+1} \delta \lambda_N(\xi) g_N(\xi) \|\mathbf{X}_{,\xi}^1\| d\xi. \quad (7.78)$$

2. Penalty method:

$$\int_{\Gamma_c^e} \epsilon_N g_N \delta g_N d\Gamma = \int_{-1}^{+1} \epsilon_N g_N(\xi) (\boldsymbol{\eta}^2 - \boldsymbol{\eta}^1) \cdot (\mathbf{E}_3 \times \mathbf{X}_{,\xi}^1) d\xi. \quad (7.79)$$

3. Perturbed LAGRANGE formulation (see also (7.67)):

$$\begin{aligned} \int_{\Gamma_c^e} \lambda_N \delta g_N d\Gamma &= \int_{-1}^{+1} \lambda_N(\xi) (\boldsymbol{\eta}^2 - \boldsymbol{\eta}^1) \cdot (\mathbf{E}_3 \times \mathbf{X}_{,\xi}^1) d\xi \quad (7.80) \\ \int_{\Gamma_c^e} \delta \lambda_N \left[ g_N - \frac{1}{\epsilon} \lambda_N \right] d\Gamma &= \int_{-1}^{+1} \delta \lambda_N(\xi) \left[ g_N(\xi) - \frac{1}{\epsilon} \lambda_N(\xi) \right] \|\mathbf{X}_{,\xi}^1\| d\xi. \end{aligned}$$

Based on (7.77), (7.79) or (7.80), the finite element discretization is obtained. We here restrict the matrix formulation to the penalty method and to the perturbed LAGRANGE formulation.

(a) **Penalty method.** With the isoparametric interpolation

$$\mathbf{X}^1 = \sum_{A=1}^m N_A(\xi) \mathbf{X}_A^1 \quad (7.81)$$

and the resulting derivatives for the components of  $\mathbf{X}^1$

$$X_{i,\xi}^1 = \sum_{A=1}^m N_A(\xi)_{, \xi} X_{iA}^1, \quad (7.82)$$

the cross product in (7.76) yields

$$\mathbf{N}^1(\xi) = \mathbf{E}_3 \times \mathbf{X}_{,\xi}^1 = \begin{Bmatrix} -X_{2,\xi}^1 \\ X_{1,\xi}^1 \end{Bmatrix}. \quad (7.83)$$

This leads to the discretized form of (7.79)

$$\int_{\Gamma_c^\varepsilon} \epsilon_N g_N \delta g_N d\Gamma \approx \int_{-1}^{+1} \epsilon_N g_N(\xi) \sum_I N_I(\xi) (\eta_I^2 - \eta_I^1) \cdot \mathbf{N}^1(\xi) d\xi. \quad (7.84)$$

The linearization of contact residual (7.84) yields, with (7.66),

$$\int_{\Gamma_c^\varepsilon} \epsilon_N \delta g_N \Delta g_N d\Gamma \approx \int_{-1}^{+1} \epsilon_N \sum_I N_I(\eta_I^2 - \eta_I^1) \cdot \mathbf{N}^1 \sum_K N_K (\Delta \mathbf{u}_I^2 - \Delta \mathbf{u}_I^1) \cdot \frac{\mathbf{N}^1}{\|\mathbf{X}_{,\xi}^1\|} d\xi. \quad (7.85)$$

The matrix formulation for (7.84) and (7.85) are then based on definitions (7.53), in which (7.83) has to be used. For the contact residual, this leads to

$$\int_{\Gamma_c^\varepsilon} \epsilon_N g_N \delta g_N d\Gamma \approx \bigcup_{c=1}^{n_c} \sum_{I=1}^m \eta_{cI}^T \mathbf{G}_I^c, \text{ with } \mathbf{G}_I^c = \int_{-1}^{+1} \epsilon_N g_N(\xi) N_I(\xi) \hat{\mathbf{N}} d\xi, \quad (7.86)$$

where the integration has to be performed on the reference element, see Figure 7.8.

For the tangent stiffness we obtain with (7.66) the matrix form

$$\int_{\Gamma_c^\varepsilon} \epsilon_N \Delta g_N \delta g_N d\Gamma \approx \bigcup_{c=1}^{n_c} \sum_{I=1}^m \sum_{K=1}^m \eta_{cI}^T \mathbf{C}_{IK} \Delta \mathbf{u}_{cK}, \quad (7.87)$$

with

$$\mathbf{C}_{IK} = \int_{-1}^{+1} \epsilon_N N_I(\xi) N_K(\xi) \hat{\mathbf{N}} \hat{\mathbf{N}}^T \frac{1}{\|\mathbf{X}_{,\xi}^1\|} d\xi. \quad (7.88)$$

For proper choice of the numerical integration rule, we refer to the remark made for the three-dimensional contact formulation. It can easily be seen that the contact residual and tangent changes when using different integration rules. However, since only polynomials appear in (7.86), an exact integration is also possible for this integral. For linear shape functions the polynomial order is 2, thus a two-point GAUSS rule is sufficient. A one-point GAUSS rule then leads to an under integration. For a quadratic

interpolation the maximal polynomial order is 6. In this case four GAUSS points have to be used for an exact integration, but also a three-point rule is sufficient.

**(b) Perturbed Lagrange formulation.** For this formulation the matrix form for a Q1/P0 interpolation will be developed. This yields matrices which can be evaluated in closed form, and thus are used to demonstrate in detail the matrix formulation of the finite element discretization. Note also that in this case the contact interface is approximated by a piecewise linear interpolation. We start from the weak formulation (7.80) for one contact element. By using the interpolations for the displacement field, the coordinates and the LAGRANGE multiplier are

$$\mathbf{u}^\alpha = \sum_{I=1}^2 N_I(\xi) \mathbf{u}_I^\alpha = \sum_{I=1}^2 \frac{1}{2} (1 + \xi \xi_I) \mathbf{u}_I^\alpha, \quad (7.89)$$

$$\mathbf{X}^1 = \sum_{I=1}^2 N_I(\xi) \mathbf{X}_I^1 = \sum_{I=1}^2 \frac{1}{2} (1 + \xi \xi_I) \mathbf{X}_I^1, \quad (7.90)$$

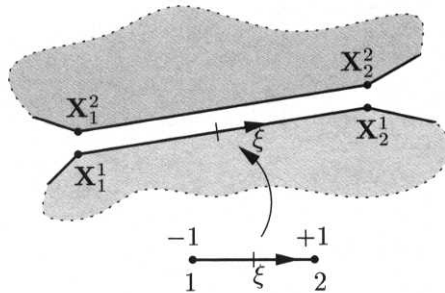
$$\lambda_N = \bar{\lambda}_N, \quad (7.91)$$

with  $\xi_1 = -1$  and  $\xi_2 = +1$ . Using this interpolation, the normal vector  $\mathbf{N}^1$  can be expressed explicitly

$$\mathbf{N}^1 = \begin{Bmatrix} -X_{2,\xi}^1 \\ X_{1,\xi}^1 \end{Bmatrix} = \frac{1}{2} \begin{Bmatrix} X_{21}^1 - X_{22}^1 \\ X_{12}^1 - X_{11}^1 \end{Bmatrix}, \quad (7.92)$$

with the components  $(\mathbf{X}_1^1)^T = \{X_{11}^1, X_{12}^1\}$  and  $(\mathbf{X}_2^1)^T = \{X_{21}^1, X_{22}^1\}$  of the position vectors  $\mathbf{X}_I^1$  of the nodal points of the master surface, see also Figure 7.9. Note that the vector  $\mathbf{N}^1$  is constant, which is in accordance with the straight geometry of the contact element. Now we can write, with (7.91) and (7.92) for (7.80),

$$\int_{\Gamma_c^e} \lambda_N \delta g_N d\Gamma = \bar{\lambda}_N \int_{-1}^{+1} \sum_{I=1}^2 N_I(\eta_I^2 - \eta_I^1) \cdot \mathbf{N}^1 d\xi. \quad (7.93)$$



**Fig. 7.9** Two-node contact element with Q1/P0 interpolation.



Since  $\eta_I^\alpha$  and  $\mathbf{N}^1$  are constant, we can integrate (7.93) analytically, and obtain with

$$\int_{-1}^{+1} \frac{1}{2} (1 + \xi_I \xi) d\xi = 1$$

the matrix form

$$\int_{\Gamma_c} \lambda_N \delta g_N d\Gamma = \bar{\lambda}_N \hat{\eta}^T \mathbf{G}^{PL} \quad \text{with} \quad \mathbf{C}^{PL} = \begin{Bmatrix} -\mathbf{N}^1 \\ -\mathbf{N}^1 \\ \mathbf{N}^1 \\ \mathbf{N}^1 \end{Bmatrix}, \quad (7.94)$$

where the vector  $\hat{\eta}$  which contains the variations  $\eta_I^\alpha$  has been introduced as

$$\hat{\eta}^T = \{ \eta_{11}^1, \eta_{12}^1, \eta_{21}^1, \eta_{22}^1, \eta_{11}^2, \eta_{12}^2, \eta_{21}^2, \eta_{22}^2 \}^T.$$

Now the constraint equation (7.80)<sub>2</sub> can be evaluated. For the chosen interpolation this leads to an expression for the LAGRANGE multiplier

$$\bar{\lambda}_N = \frac{\epsilon_N}{A_e} \int_{-1}^{+1} \sum_{I=1}^2 N_I(\xi) g_{NI} \|\mathbf{N}^1\| d\xi. \quad (7.95)$$

The element area  $A_e$  is given in the two-dimensional case as the element length  $L_e$ . Furthermore, we obtain  $\|\mathbf{N}^1\| = \frac{1}{2} \sqrt{(X_{22} - X_{22})^2 + (X_{12} - X_{11})^2} = L_e / 2$ . With this, for (7.95) we arrive at a simple expression for the LAGRANGE multiplier:

$$\bar{\lambda}_N = \frac{\epsilon_N}{2} (g_{N1} + g_{N2}). \quad (7.96)$$

Hence the constant LAGRANGE multiplier depends upon the average gap  $\bar{g}_N = (g_{N1} + g_{N2}) / 2$ , which is also used in this element formulation to establish whether the gap is closed or not, leading to the contact condition

$$\bar{g}_N < 0 \implies \text{contact}. \quad (7.97)$$

The linearization of the perturbed LAGRANGE formulation can be computed from (7.80) with (7.95). We then have

$$\begin{aligned} DC_c \cdot \Delta \mathbf{u} &= \int_{\Gamma_c} \Delta \lambda_N \delta g_N d\Gamma = \Delta \bar{\lambda}_N \int_{\Gamma_c} \delta g_N d\Gamma \\ &= \frac{\epsilon_N}{L_e} \int_{-1}^{+1} \sum_{I=1}^2 N_I (\Delta u_I^2 - \Delta u_I^1) \cdot \mathbf{N}^1 d\xi \int_{-1}^{+1} \sum_{K=1}^2 N_K (\eta_K^2 - \eta_K^1) \cdot \mathbf{N}^1 d\xi, \end{aligned} \quad (7.98)$$

which with

$$\Delta \hat{\mathbf{u}}^T = \{ \Delta u_{11}^1, \Delta u_{12}^1, \Delta u_{21}^1, \Delta u_{22}^1, \Delta u_{11}^2, \Delta u_{12}^2, \Delta u_{21}^2, \Delta u_{22}^2 \}^T$$

and (7.94) yields the matrix form of the tangent matrix for one contact element

$$DC_c \cdot \Delta \mathbf{u} = \hat{\eta}^T \mathbf{K}_c^{PL} \Delta \hat{\mathbf{u}} \quad \text{with} \quad \mathbf{K}_c^{PL} = \frac{\epsilon_N}{L_e} \mathbf{C}^{PL} (\mathbf{C}^{PL})^T. \quad (7.99)$$

## 7.4 DISCRETIZATION FOR NON-MATCHING MESHES

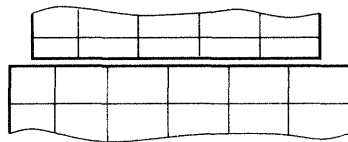
By using general mesh generators, unstructured meshes can be constructed in which nodes in the contact interface do not assume the same coordinates, see Figure 7.10. Then the formulations discussed above can no longer be applied, even in the geometrically linear case.

Methods for connecting finite element domains with non-matching grids (see Figure 7.10) are frequently used for parallel computations. Such formulations have different origins, and hence have also received special names. A commonly used approach is the mortar method. However, other methods like the NITSCHKE method exist. The formulations are designed in such a way that they fulfil the BB conditions, also called *inf-sup* conditions, and hence lead to a stable discretization. For a good literature overview and the underlying mathematical theory we recommend Wohlmuth (2000a). The basic difference between the mortar and the NITSCHKE methods lies in the fact that in a mortar discretization, one has to introduce LAGRANGE multipliers, whereas the formulation due to NITSCHKE is purely displacement-based, see Section 5.3.5. There are also approaches which were developed in the engineering literature, e.g. see Simo et al. (1985) or Papadopoulos and Taylor (1992). Here either a LAGRANGE or penalty formulation was applied.

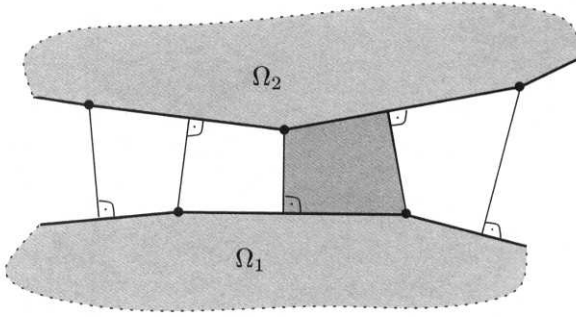
All methods will be discussed for the case of frictionless contact. Extensions to friction can be formulated analogous to Section 7.2.2. Treatments which include friction can also be found in Krause and Wohlmuth (2001) or McDevitt and Laursen (2000).

### 7.4.1 Discretization with contact segments

One idea to discretize the contact interface in the case of non-matching meshes is based on the introduction of so-called contact segments. The discretization of the contact interface by segments (see Figure 7.11) was introduced in Simo et al. (1985) for the geometrically linear case. Based on the definition of contact segments, it is possible to use different interpolations for the LAGRANGE multipliers and the gap function in a natural way. Hence, it is appropriate to employ the perturbed LAGRANGE formulation, see Section 5.3.6. This is true even in the case of a penalty formulation, since then this formulation gives some freedom in choosing the interpolation spaces for the displacements and the LAGRANGE multipliers.



**Fig. 7.10** Contact discretization for small deformations.



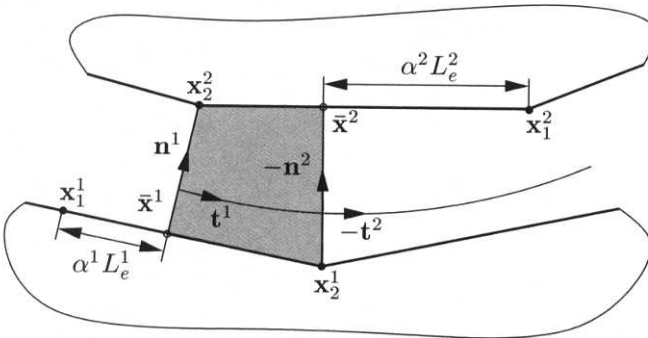
**Fig. 7.11** Contact segments for two-dimensional problems.

Following Simo et al. (1985) we define the contact segment by the geometry of the adjacent elements of the bodies in contact. Within the segments an intermediate surface  $\Gamma_c^s$  is introduced. A typical contact segment is shown in Figure 7.12. It is defined by two element edges with the nodes  $\mathbf{x}_2^1 - \mathbf{x}_1^1$  and  $\mathbf{x}_2^2 - \mathbf{x}_1^2$ . The superscripts refer to the body and the subscripts indicate the node number. The contact segment is now defined by a quadrilateral with four nodes, specified by the points  $\bar{\mathbf{x}}^1$ ,  $\mathbf{x}_2^1$ ,  $\bar{\mathbf{x}}^2$  and  $\mathbf{x}_2^2$ .  $\bar{\mathbf{x}}^1$  and  $\bar{\mathbf{x}}^2$  are the orthogonal projections of the nodes  $\mathbf{x}_2^2$  and  $\mathbf{x}_2^1$  onto the opposite element edge, see Figure 7.12. This projection is given by

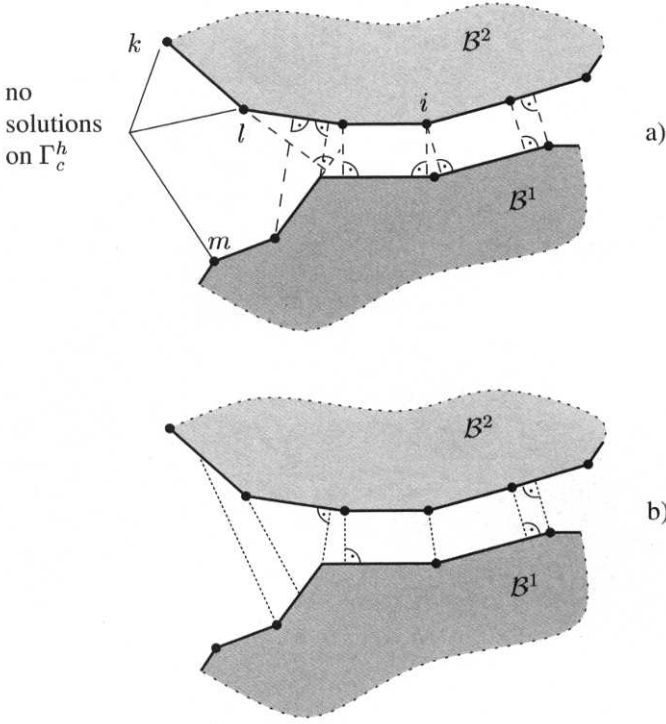
$$\bar{\mathbf{x}}^\gamma = (1 - \alpha^\gamma) \mathbf{x}_1^\gamma + \alpha^\gamma \mathbf{x}_2^\gamma, \quad (7.100)$$

where  $\gamma \in 1, 2$  refers to the body, and  $\alpha^\gamma$  follows from

$$\alpha^1 L_e^1 = (\mathbf{x}_2^2 - \mathbf{x}_1^1) \cdot \mathbf{t}^1 \quad \text{and} \quad \alpha^2 L_e^2 = (\mathbf{x}_2^1 - \mathbf{x}_1^2) \cdot \mathbf{t}^2, \quad (7.101)$$



**Fig. 7.12** Geometry of the contact segment.



**Fig. 7.13** Special cases for segment definitions.

where the tangent vectors are defined as

$$\mathbf{t}^\gamma = \frac{\mathbf{x}_2^\gamma - \mathbf{x}_1^\gamma}{\|\mathbf{x}_2^\gamma - \mathbf{x}_1^\gamma\|}. \quad (7.102)$$

Similar to the coordinates of the new nodes  $\bar{\mathbf{x}}^1$  and  $\bar{\mathbf{x}}^2$ , the displacement vector at these nodes can be stated:

$$\bar{\mathbf{u}}^\gamma = (1 - \alpha^\gamma) \mathbf{u}_1^\gamma + \alpha^\gamma \mathbf{u}_2^\gamma. \quad (7.103)$$

**REMARK 7.1:** Note that the projection used in Figure 7.12 does not work in all cases (see Figure 7.13a), where the projection is not unique for node  $i$ , and furthermore no solution exists for  $k, l, m$  within the segments defining the contact interface. In such situations, one has to redefine the segments. Possible definitions are stated in Figure 7.13b. However this definition leads to complex coding, since many special cases have to be checked, especially in three-dimensional situations where a lot of differently shaped segments can occur.

Within the segment, defined in Figure 7.12, the current coordinates, displacement fields and variations are interpolated with linear polynomials, and hence are given as

functions of the local coordinate  $\xi \in [0, 1]$

$$\begin{aligned} \mathbf{x}^1(\xi) &= (1 - \xi) \bar{\mathbf{x}}^1 + \xi \mathbf{x}_2^1, & \mathbf{x}^2(\xi) &= \xi \bar{\mathbf{x}}^2 + (1 - \xi) \mathbf{x}_2^2, \\ \mathbf{u}^1(\xi) &= (1 - \xi) \bar{\mathbf{u}}^1 + \xi \mathbf{u}_2^1, & \mathbf{u}^2(\xi) &= \xi \bar{\mathbf{u}}^2 + (1 - \xi) \mathbf{u}_2^2, \\ \boldsymbol{\eta}^1(\xi) &= (1 - \xi) \bar{\boldsymbol{\eta}}_1^1 + \xi \boldsymbol{\eta}_2^1, & \boldsymbol{\eta}^2(\xi) &= \xi \bar{\boldsymbol{\eta}}_1^2 + (1 - \xi) \boldsymbol{\eta}_2^2. \end{aligned} \quad (7.104)$$

Depending on the discretization used for the bodies  $\mathcal{B}^\gamma$ , higher order interpolations could also be applied here. With definitions (7.100) the gaps at the segment edges are obtained as

$$g_N^1 = (\mathbf{x}_2^2 - \bar{\mathbf{x}}^1) \cdot \mathbf{n}^1, \quad g_N^2 = (\mathbf{x}_2^1 - \bar{\mathbf{x}}^2) \cdot \mathbf{n}^2, \quad (7.105)$$

where the normal vector is given for the plane case by  $\mathbf{n}^\gamma = \mathbf{e}_3 \times \mathbf{t}^\gamma$ .

Now the interpolation for the gap function  $g_N(\xi)$  and its variation  $\delta g_N(\xi)$  can be defined according to (7.104)

$$g_N(\xi) = (1 - \xi) g_N^1 + \xi g_N^2, \quad \delta g_N(\xi) = (1 - \xi) \delta g_N^1 + \xi \delta g_N^2. \quad (7.106)$$

These interpolations are applied within a segment which is defined by the edge nodes  $\mathbf{x}_2^\gamma$  and the projections onto the other surface  $\bar{\mathbf{x}}^\gamma$ , see Figure 7.12. Now we can define an intermediate contact line in the segment which is  $C^1$ -continuous:

$$\boldsymbol{\gamma}_s^c(\xi) = \sum_{A=1}^2 H_A(\xi) \mathbf{y}^A + \bar{L}_s \sum_{A=1}^2 \bar{H}_A(\xi) \mathbf{t}^A, \quad (7.107)$$

where  $H_A$  and  $\bar{H}_A$  are the classical HERMITE polynomials

$$\begin{aligned} H_1(\xi) &= 1 - 3\xi^2 + 2\xi^3 & H_2(\xi) &= \xi^2(3 - 2\xi) \\ \bar{H}_1(\xi) &= \xi(1 - \xi)^2 & \bar{H}_2(\xi) &= \xi^2(1 - \xi), \end{aligned}$$

with  $\xi \in [0, 1]$ .  $\mathbf{t}^\gamma$  is already defined in (7.102) and  $\mathbf{y}^\gamma$  is the position vector of the beginning and end of the contact line  $\boldsymbol{\gamma}_s^c$  (see Figure 7.12):

$$\mathbf{y}^1 = (1 - \beta) \bar{\mathbf{x}}^1 + \beta \mathbf{x}_2^2, \quad \mathbf{y}^2 = (1 - \beta) \mathbf{x}_2^1 + \beta \bar{\mathbf{x}}^2. \quad (7.108)$$

The limiting choices of  $\beta = 0$  and  $\beta = 1$  correspond to selecting one of the contacting surfaces  $\partial\mathcal{B}^{\gamma^h}$  as an intermediate contact line. This does not mean, however, that the intermediate line is equivalent to the interpolation of one surface of the bodies, since it is still  $C^1$ -continuous, which is not true for interpolation (7.104). The length  $\bar{L}_s$  is defined as the distance between the end points  $\mathbf{y}^1$  and  $\mathbf{y}^2$ :  $\bar{L}_s = \|\mathbf{y}^2 - \mathbf{y}^1\|$ . The actual length of  $\boldsymbol{\gamma}_s^c$  is computed via

$$L_s = \int_0^1 \left\| \frac{d\boldsymbol{\gamma}_s^c}{d\xi} \right\| d\xi \quad (7.109)$$

which can be evaluated numerically using a quadrature rule.

The variational formulation is based on the perturbed LAGRANGE approach, see (7.80). The contact contributions take the form

$$\int_{\Gamma_c} \lambda_N \delta g_N d\Gamma = \sum_{s=1}^{n_{seg}} \int_{\gamma_s} \lambda_N \delta g_N d\gamma \quad (7.110)$$

$$\int_{\Gamma_c} (g_N - \frac{\lambda_N}{\epsilon_N}) \delta \lambda_N d\Gamma = \sum_{s=1}^{n_{seg}} \int_{\gamma_s} (g_N - \frac{\lambda_N}{\epsilon_N}) \delta \lambda_N d\gamma = 0. \quad (7.111)$$

The sum over the segments is carried out for all closed gaps. Depending on the interpolations, the latter equation can be solved for  $\lambda_N$  directly. With the interpolations (7.104), and assuming a constant contact pressure  $\bar{\lambda}_N = CONST$  within the segment, for the segment  $\gamma_s$  we obtain

$$\int_{\gamma_s} \lambda_N \delta g_N d\gamma = \bar{\lambda}_N \int_0^1 \delta g_N(\xi) \left\| \frac{d\gamma_s^c}{d\xi} \right\| d\xi \quad (7.112)$$

$$\int_{\gamma_s} (g_N - \frac{\lambda_N}{\epsilon_N}) \delta \lambda_N d\gamma = 0 \Rightarrow \bar{\lambda}_N = \frac{\epsilon_N}{L_s} \int_0^1 g(\xi) \left\| \frac{d\gamma_s^c}{d\xi} \right\| d\xi. \quad (7.113)$$

As has been shown in Simo et al. (1985), the evaluation of these integrals by the trapezoidal rule yields the simple formulas

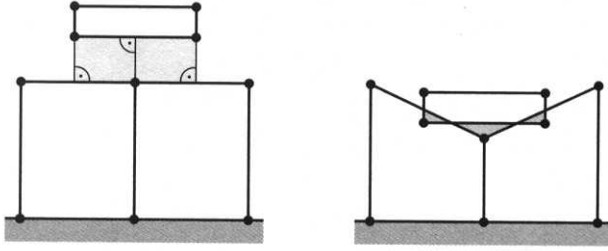
$$\bar{\lambda}_N \int_0^1 \delta g_N(\xi) \left\| \frac{d\gamma_s^c}{d\xi} \right\| d\xi \approx \frac{\bar{\lambda}_N}{2} (\delta g_N^1 + \delta g_N^2), \quad (7.114)$$

$$\bar{\lambda}_N \approx \frac{\epsilon_N}{2} (g_N^1 + g_N^2) = \epsilon_N \bar{g}_s, \quad (7.115)$$

where  $g_N^1$  and  $\delta g_N^1$  are the gaps defined in (7.105). Equation (7.115) has the interpretation that the constant LAGRANGE multiplier  $\bar{\lambda}_N$  is given in terms of the average gap  $\bar{g}_s$  in the segment. For a consistent formulation this average gap also has to be used for the contact check. A geometrical view of the enforcement of the contact constraints in the contact-segment approach is given in Figure 7.14, which depicts a fulfillment of the contact constraints in the middle of the left and right segments. This is equivalent to the fact that the constraint is fulfilled in an average sense, since the integral (7.113) is zero.

Finally, the evaluation of (7.114) can be combined with (7.115) to eliminate the LAGRANGE multiplier, and hence yield the contact contribution for the perturbed LAGRANGE formulation

$$\bar{\lambda}_N \int_0^1 \delta g_N(\xi) \left\| \frac{d\gamma_s^c}{d\xi} \right\| d\xi \approx \epsilon_N \bar{L}_s \bar{g}_s \delta \bar{g}_s, \quad (7.116)$$



**Fig. 7.14** Geometrical interpretation of the average gap.

with  $\delta \bar{g}_s = \frac{1}{2} (\delta g_N^1 + \delta g_N^2)$ . With (7.104) this equation can now be expressed in matrix form by introducing

$$\mathbf{x}_s^T = \langle \mathbf{x}_1^2, \mathbf{x}_2^2, \mathbf{x}_1^1, \mathbf{x}_2^1 \rangle \quad \text{and} \quad \boldsymbol{\eta}_s^T = \langle \boldsymbol{\eta}_1^2, \boldsymbol{\eta}_2^2, \boldsymbol{\eta}_1^1, \boldsymbol{\eta}_2^1 \rangle, \quad (7.117)$$

which leads to the contact residual within the perturbed LAGRANGE formulation with local elimination of the LAGRANGE multiplier

$$\epsilon_N \bar{g}_s \bar{L}_s \delta \bar{g}_s = \epsilon_N \bar{L}_s (\mathbf{x}_s^T \mathbf{C}_s) \boldsymbol{\eta}_s^T \mathbf{C}_s, \quad (7.118)$$

with

$$\mathbf{C}_s = \begin{Bmatrix} -\frac{1}{2} (1 - \alpha^2) \mathbf{n}^2 \\ \frac{1}{2} (\mathbf{n}^1 - \alpha^2 \mathbf{n}^2) \\ -\frac{1}{2} (1 - \alpha^1) \mathbf{n}^1 \\ \frac{1}{2} (\mathbf{n}^2 - \alpha^1 \mathbf{n}^1) \end{Bmatrix}. \quad (7.119)$$

In the case of established contact, the segment also contributes to the tangent stiffness. This contribution follows from the linearization of (7.118) with respect to the displacement variables. Since this term depends only upon the displacements within the product  $\mathbf{x}_s^T \mathbf{C}_s = (\mathbf{X}_s + \mathbf{u}_s)^T \mathbf{C}_s$ , for the tangent stiffness we obtain

$$\mathbf{K}_s = \epsilon_N \bar{L}_s \mathbf{C}_s \mathbf{C}_s^T, \quad (7.120)$$

which completes the discretization for contact segments. For more details, see Simo et al. (1985), and for its nonlinear extension see Papadopoulos and Taylor (1992).

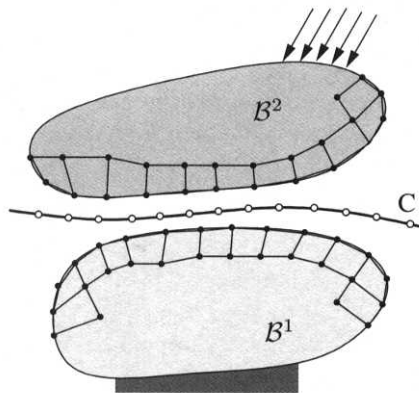
## 7.4.2 Mortar method

The mortar method is a special technique to enforce contact constraints in the discretized system for non-matching meshes. The method is based on a LAGRANGE multiplier formulation in which special interpolation functions are used to discretize the LAGRANGE multiplier in the contact interface. Two methods have been established during the last years within the mortar approach. One leads to a non-conforming method which is based on direct enforcements of the constraints, and hence is equivalent to the direct constraint elimination discussed in Section 5.3.3. This method

is described in Bernadi et al. (1994) and Wohlmuth (2000a). It leads to a positive definite system of equations, see also the example in Section 5.4. The other method is related to the weak enforcement of the constraints by applying the LAGRANGE multiplier method. The general idea is outlined in Section 5.3.1, however in the mortar approach different interpolations for the LAGRANGE multipliers are introduced. In general, one can use as ansatz space linear, quadratic or even higher order interpolation functions. However, due to the weak formulation of the constraints, mathematical conditions like the *inf-sup* or BABUSKA–BREZZI condition have to be fulfilled in order to achieve a stable discretization scheme; for details, see El-Abbasi and Bathe (2001) for a numerical and Wohlmuth (2000a) for a theoretical approach. In the following, we discuss two aspects related to discretization schemes based on the mortar approach. These are the introduction of the reference surfaces and the choice of the interpolation spaces.

Several techniques can be followed to define the contact surface. One is based on the use of an intermediate contact surface as the reference surface and to define the LAGRANGE multipliers on this surface. This intermediate contact surface  $C$  defines the *mortar* side in the interface, see Figure 7.15. Early formulations can be found in Simo et al. (1985), as described in Section 7.4.1. Lately mortar discretizations, based on the intermediate surface, have been developed in McDevitt and Laursen (2000) or Rebel et al. (2000).

Another choice is made in the mathematical literature, e.g. see Wohlmuth (2000b) and Krause and Wohlmuth (2001), which is based on the assumption that the mortar side is one of the surfaces of the bodies in the contact interface which would, for example, in our notation be the master surface. In Wohlmuth (2000b) it was shown that such formulation with the appropriate interpolation functions for the LAGRANGE multipliers fulfils the BB condition. Furthermore, the LAGRANGE multiplier interpolation can be constructed in such a way that the locality of the support of the nodal basis functions is preserved. Hence this formulation leads from a mathematical



**Fig. 7.15** Definition of intermediate mortar surface.



viewpoint to a stable discretization, and yields a good approximation of the contact stresses. The mortar method will be described here for the frictionless case.

The basic formulation starts from the LAGRANGE multiplier method stated in Section 5.3.1. From equation (5.26) for the frictionless case one obtains the energy related to the contact interface

$$\Pi_c = \int_{\Gamma_c} \lambda_N g_N dA. \quad (7.121)$$

$\lambda_N$  is the LAGRANGE multiplier in the normal direction of the contact interface and  $g_N$  measures the gap distance.  $\Gamma_c$  denotes the contact interface. The variation of this form leads to

$$C_c = \delta \Pi_c = \int_{\Gamma_c} (\delta \lambda_N g_N + \lambda_N \delta g_N) dA. \quad (7.122)$$

The first term describes the fulfillment of the constraint condition, and the second term yields the contact pressure (LAGRANGE multiplier) due to the enforcement of the constraint. Now the first term can be used to construct the non-conforming mortar scheme.

**Non-conforming mortar method.** Within the discretization of the first term in (7.122) one has to interpolate the gap function  $g_N$  and the LAGRANGE multiplier  $\lambda_N$ . The gap function is defined by the displacement field of the contacting surfaces. Hence the interpolation functions have to be the same as the interpolations used to discretize the contacting solids. Thus, the only free choice for the interpolation in the contact interface can be made for the LAGRANGE multipliers. These have to be interpolated in such way that the BB condition is fulfilled.

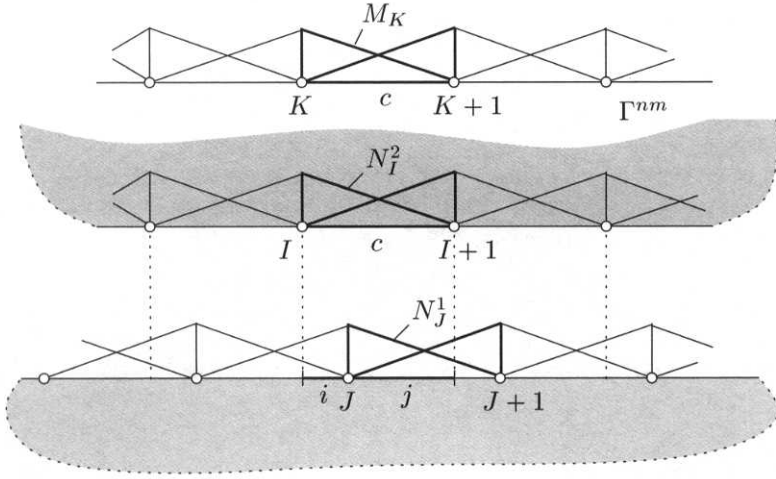
We start by introducing a LAGRANGE multiplier interpolation for the two-dimensional case

$$\lambda_N(\xi) = \sum_{K=1}^2 M_K(\xi) \lambda_{NK}, \quad (7.123)$$

with  $M_K$  being a linear function (see (6.15)), except at the ends of the contact area. At this end  $\lambda_N$  is chosen to be constant to achieve a stable discretization, e.g. see Wohlmuth (2000a). With respect to the non-mortar side, the constraint equation included in (7.122) can be written in a weak sense as

$$\sum_{c=1}^{n_c} \sum_I \sum_K \delta \lambda_{NK} \int_{\Gamma_c^{nm}} M_K N_I^2 d\Gamma \mathbf{u}_I^2 = \sum_{c=1}^{n_c} \sum_J \sum_K \delta \lambda_{NK} \int_{\Gamma_c^{nm}} M_K N_J^1 d\Gamma \mathbf{u}_J^1, \quad (7.124)$$

where  $n_c$  is the number of element sides in contact on the non-mortar side. On this side  $\Gamma_c^{nm}$  and also the shape functions  $N_I^2$  of body  $B^2$  are defined; see Figure 7.16 for a linear interpolation. The integration of the right-hand side of (7.124) also has to be performed on  $\Gamma_c^{nm}$ , however the shape functions  $N_J^1$  are defined on  $B^1$ . Hence the integration is more involved, since nodes  $I$  and  $J$  do not coincide. An



**Fig. 7.16** Shape functions for displacements and Lagrange multiplier in contact interface.

exact integration can be performed for the choice of linear shape functions using the segment definition (7.104) from Section 7.4.1. We observe that the integration along the non-mortar side yields a mass matrix; for the structure in the continuum case see (6.48). For a linear interpolation within a contact segment  $c$ , this leads to

$$\begin{aligned} \sum_I \sum_K \delta \lambda_{NK} \int_{-1}^1 M_K(\xi) N_I^2(\xi) d\xi \mathbf{u}_I^2 &= \langle \delta \lambda_{N1}, \delta \lambda_{N2} \rangle_c \frac{l}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}_c \left\{ \begin{matrix} u_1^2 \\ u_2^2 \end{matrix} \right\}_c \\ &= \delta \lambda_{Nc}^T \mathbf{M}_c^2 \mathbf{u}_c^2, \end{aligned} \quad (7.125)$$

where  $\mathbf{M}_c^2$  is the standard mass matrix of a bar. Assembly of all  $n_c$  terms on both sides of (7.124) gives

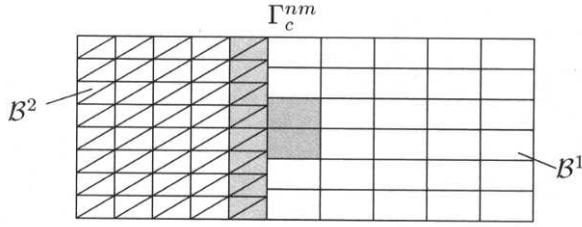
$$\delta \lambda_N^T (\mathbf{M}^2 \mathbf{u}^2 - \mathbf{M}^1 \mathbf{u}^1) = 0, \quad (7.126)$$

from which we can eliminate the displacements  $\mathbf{u}^2$  on the non-mortar side by

$$\mathbf{u}^2 = (\mathbf{M}^2)^{-1} \mathbf{M}^1 \mathbf{u}^1 \quad (7.127)$$

as discussed in Section 5.3.3.

In (7.126) we have computed matrix  $\mathbf{M}^1$  segment-wise, e.g.  $\mathbf{M}^1$  follows from an integration by dividing a segment  $c$  into  $i$  and  $j$  on the mortar side, see Figure 7.16. This integration can be performed exactly or by using a quadrature rule. For the linear interpolation with straight segments, a two-point GAUSS quadrature is sufficient, and yields an exact integration. For higher order isoparametric interpolations an exact



**Fig. 7.17** Support of the standard LAGRANGE multiplier spaces.

integration is more involved, since the contact surfaces can be curved, which leads to non-constant JACOBIANS.

Since  $\mathbf{M}^2$  is not a diagonal matrix the influence of one displacement  $u_j^1$  is coupled with all displacements  $\mathbf{u}^2$ . This situation is graphically depicted in Figure 7.17, where the large grey area on the non-mortar side shows the coupling due to  $\mathbf{M}^2$ , and the small grey zone on the mortar side shows the local distribution of  $u_j^1$  due to  $\mathbf{M}^1$ . The same is true when the LAGRANGE multipliers are kept within the formulation, as described in Section 5.3.1. Also, there the locality of the nodal basis function is lost.

**Mortar method using dual basis for Lagrange multipliers.** Following Wohlmuth (2000b) a dual LAGRANGE multiplier space is introduced. This yields the shape functions shown in Figure 7.18 for constant and linear interpolations. The condition for duality can be stated for a segment  $\Gamma_c^{nm}$  as follows:

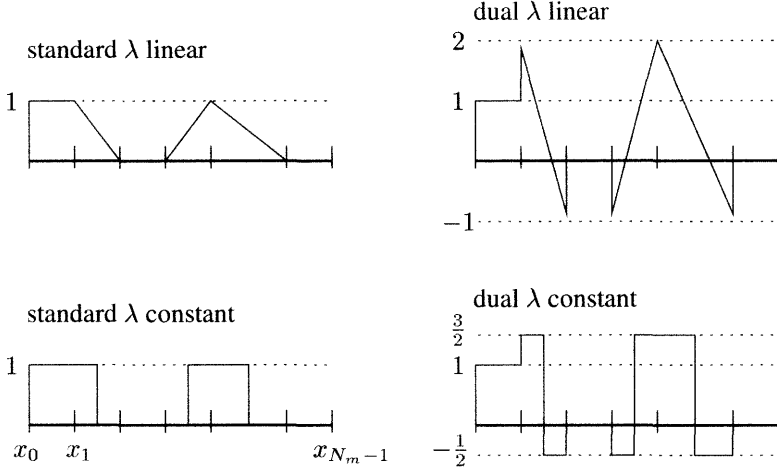
$$\int_{\Gamma_c^{nm}} N_I^2(\xi) \bar{M}_K(\xi) d\Gamma = \delta_{IK} c_I \int_{\Gamma_c^{nm}} [N_I^2(\xi)]^2 d\Gamma, \quad (7.128)$$

where  $N_I^2$  are the standard interpolation functions for the displacements on the non-mortar side, and  $\bar{M}_K$  is the dual basis used to interpolate the LAGRANGE multiplier. The interpolation functions which are depicted on the right side of Figure 7.18 are stated next in terms of local coordinates for the two-dimensional case; see also Wohlmuth (2000a). For a piecewise constant interpolation, we obtain

$$\begin{aligned} \bar{M}_1^0(\xi) &= \frac{3}{2} \text{ for } -1 \leq \xi \leq 0 \quad \text{and} \quad -\frac{1}{2} \text{ for } 0 < \xi \leq 1, \\ \bar{M}_2^0(\xi) &= -\frac{1}{2} \text{ for } -1 \leq \xi \leq 0 \quad \text{and} \quad \frac{3}{2} \text{ for } 0 < \xi \leq 1. \end{aligned} \quad (7.129)$$

Linear dual base functions are given by

$$\begin{aligned} \bar{M}_1^1(\xi) &= \frac{1}{2} (2|\xi - 1| - |\xi + 1|), \\ \bar{M}_2^1(\xi) &= \frac{1}{2} (2|\xi + 1| - |\xi - 1|); \end{aligned} \quad (7.130)$$



**Fig. 7.18** Constant and linear interpolation for the mortar method.

finally, quadratic dual basis functions are defined with (6.16) as

$$\begin{aligned}
 \bar{M}_1^2(\xi) &= \frac{1}{2} - \frac{3}{4} N_3(\xi) + N_1(\xi) = \frac{1}{4} (5\xi^2 - 2\xi - 1), \\
 \bar{M}_2^2(\xi) &= -1 + \frac{5}{2} N_3(\xi) = \frac{1}{2} (3 - 5\xi^2), \\
 \bar{M}_3^2(\xi) &= \frac{1}{2} - \frac{3}{4} N_3(\xi) + N_2(\xi) = \frac{1}{4} (5\xi^2 + 2\xi - 1).
 \end{aligned} \tag{7.131}$$

All these functions fulfil the orthogonality condition (7.128). These functions have to be inserted as interpolation functions for the LAGRANGE multipliers in (7.122). Note that no modification of these interpolation functions is needed at the end points, which has been shown in Wohlmuth (2000a). Due to this orthogonality property, using the shape functions (7.129) to (7.131), an assembled matrix form of the weak contact constraint equation is given by

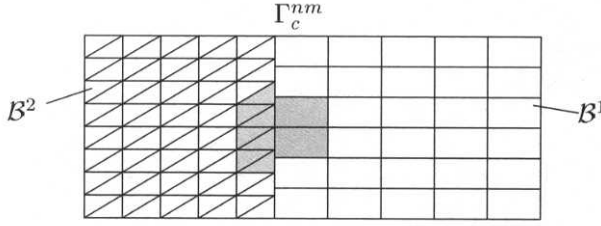
$$\delta \lambda_N^T (\mathbf{D}^2 \mathbf{u}^2 - \mathbf{M}^1 \mathbf{u}^1) = 0, \tag{7.132}$$

instead of (7.126). Hence the elimination (7.127) can now be expressed as

$$\mathbf{u}^2 = (\mathbf{D}^2)^{-1} \mathbf{M}^1 \mathbf{u}^1, \tag{7.133}$$

where  $\mathbf{D}^2$  is a diagonal matrix whose elements follow from (7.128). This leads to a contact interpolation with a local support; see Figure 7.19, which is computationally more efficient.

By denoting the displacements of nodes lying on the non-mortar side by  $\mathbf{u}_n$  and all others by  $\mathbf{u}_g$ , we can write the LAGRANGE multiplier equations according to (7.11)



**Fig. 7.19** Support of the dual LAGRANGE multiplier spaces.

as

$$\begin{bmatrix} \mathbf{K}_{gg} & \mathbf{K}_{gn} & \mathbf{C} \\ \mathbf{K}_{ng} & \mathbf{K}_{nn} & \mathbf{D} \\ \mathbf{C}^T & \mathbf{D}^T & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{u}_g \\ \mathbf{u}_n \\ \lambda \end{Bmatrix} = \begin{Bmatrix} \mathbf{f}_g \\ \mathbf{f}_n \\ \mathbf{0} \end{Bmatrix}, \quad (7.134)$$

where  $\mathbf{D}$  is the diagonal matrix introduced above, and  $\mathbf{C}$  corresponds to  $\mathbf{M}^1$  introduced in the last section. Due to the diagonal matrix  $\mathbf{D}$ , this equation has some features which can be exploited algorithmically. For example, if we know the displacements in an iteration step  $k$ , then we can directly compute the LAGRANGE multipliers from the second row of (7.134) by

$$\lambda^k = \mathbf{D}^{-1} (\mathbf{f}_n - \mathbf{K}_{ng} \mathbf{u}_g^k - \mathbf{K}_{nn} \mathbf{u}_n^k). \quad (7.135)$$

Since the inversion of the diagonal is trivial, an efficient two-step algorithm can be derived using the dual mortar method.

### 7.4.3 Nitsche method

Another method which can be applied for contact problems with non-matching meshes is the variational formulation due to NITSCHKE. The continuum formulation of this method is provided in Section 5.3.5. It leads to a weak form of the two bodies in contact, which is amended by three terms, see Becker and Hansbo (1999). These lead to a symmetrical global matrix system in the primary displacement variables. The variational formulation is based on (5.41) which, instead of the LAGRANGE multipliers, includes the contact pressures  $p_N^2$  stemming from the solids and a stabilizing second term which has the structure of a penalty term, see Section 5.3.2. It is given as

$$\Pi_c^N = - \int_{\Gamma_c} \frac{1}{2} (p_N^1 + p_N^2) g_N dA + \int_{\Gamma_c} \epsilon_N g_N^2 dA. \quad (7.136)$$

The variation of  $\Pi_c^N$  yields

$$C_c^N = - \int_{\Gamma_c} \frac{1}{2} [(\delta p_N^1 + \delta p_N^2) g_N + (p_N^1 + p_N^2) \delta g_N] dA + \int_{\Gamma_c} \epsilon_N g_N \delta g_N dA. \quad (7.137)$$

The discretization now starts from the definition of the contact pressures (5.44)

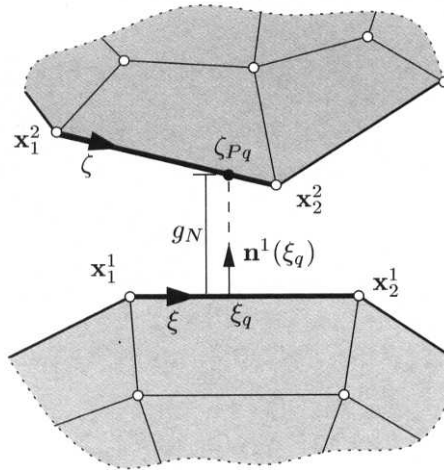
$$p_N^\gamma = \mathbf{n}^\gamma \cdot \mathbb{C}^\gamma[\boldsymbol{\varepsilon}(\mathbf{u}^\gamma)] \mathbf{n}^\gamma = \mathbf{N}^{\gamma T} \mathbf{E}^\gamma \sum_{I=1}^{n_e} \mathbf{B}_I^\gamma \mathbf{u}_I. \quad (7.138)$$

$\mathbf{E}^\gamma$  is the constitutive matrix of body  $\mathcal{B}^\gamma$  and  $\mathbf{B}_I^\gamma$  is defined in (6.82) for the current configuration, but can also be applied to geometrically linear problems when the current coordinate is replaced by the coordinate of the reference configuration. The vector which describes the projection of the stress field at the boundary in the normal direction  $\mathbf{N}^\gamma$  is defined by

$$\mathbf{N}^{\gamma T} = \langle n_1^2, n_2^2, n_3^2, 2n_1n_2, 2n_2n_3, 2n_1n_3 \rangle^\gamma. \quad (7.139)$$

It contains the components of the normal vector  $\mathbf{n}^\gamma$ . Equation (7.138) now has to be used within the variational formulation to enforce the constraints. Hence we have to insert this expression into (5.45). In this form an integration over the master surface is needed, which has to take into account the shape function of both sides on the contact interface.

We will restrict ourselves here to the two-dimensional case and bilinear interpolation functions for the finite elements at the contact interface, see Figure 7.20. The numerical integration will be performed using a two-point GAUSS rule for each master segment which is sufficient for linear shape functions. Hence one only has to evaluate the integral in (7.137) at two points, as shown in Figure 7.20 for one of the points  $\xi_q$ . For this integration the gap function  $g_N$  at each GAUSS point  $\xi_q$  has to be evaluated.  $g_N$  is obtained as the distance between the point  $\mathbf{x}^1(\xi_q) = (1 - \xi_q) \mathbf{x}_1^1 + \xi_q \mathbf{x}_2^1$  and the intersection of the normal at  $\xi_q$  with the boundary of the slave surface. This leads



**Fig. 7.20** Two-dimensional contact discretization for NITSCHÉ method.

to the condition

$$\mathbf{x}^1(\xi_q) + g_N \mathbf{n}^1(\xi_q) = (1 - \zeta_{Pq}) \mathbf{x}_1^2 + \zeta_{Pq} \mathbf{x}_2^2 = \mathbf{x}^2(\zeta_{Pq}). \quad (7.140)$$

The multiplication of this equation with the tangent vector related to the master surface  $\mathbf{e}_T^1 = \mathbf{x}_2^1 - \mathbf{x}_1^1$  yields an equation for the location of the intersection point in terms of the local coordinate  $\zeta_{Pq}$  related to the slave surface.

$$\zeta_{Pq} = \frac{[\mathbf{x}^1(\xi_q) - \mathbf{x}_1^2] \cdot \mathbf{e}_T^1}{(\mathbf{x}_2^2 - \mathbf{x}_1^2) \cdot \mathbf{e}_T^1}. \quad (7.141)$$

Note that the term with the gap function drops out. In the same way, one obtains the distance  $g_N$ ,

$$g_N = \frac{1}{\|\mathbf{n}^1\|} [\mathbf{x}^2(\zeta_{Pq}) - \mathbf{x}^1(\xi_q)] \cdot \mathbf{n}^1, \quad (7.142)$$

where, for generality, it is not assumed that  $\mathbf{n}^1$  is a unit vector. Hence  $\mathbf{n}^1$  can be defined simply by  $\mathbf{n}^1 = \mathbf{e}_3 \times \mathbf{e}_T^1$  in terms of the tangent  $\mathbf{e}_T^1$  to the master surface.

Now the contact pressures at the GAUSS point  $\xi_q$  on the master surface and its projection  $\zeta_{Pq}$  on the slave surface can be defined with (7.138) as

$$\begin{aligned} p_N^1 &= \mathbf{N}^{1T} \mathbf{E}^1 \sum_{I=1}^{n_e} \mathbf{B}_I^1(\xi_q) \mathbf{u}_I^1 = \mathbf{N}^{1T} \mathbf{E}^1 \bar{\mathbf{B}}_q^1 \mathbf{u}^1, \\ p_N^2 &= \mathbf{N}^{2T} \mathbf{E}^2 \sum_{I=1}^{n_e} \mathbf{B}_I^2(\zeta_{Pq}) \mathbf{u}_I^2 = \mathbf{N}^{2T} \mathbf{E}^2 \bar{\mathbf{B}}_{Pq}^2 \mathbf{u}^2. \end{aligned} \quad (7.143)$$

The same also follows for the variations of the contact pressures, which yields

$$\begin{aligned} \delta p_N^1 &= \mathbf{N}^{1T} \mathbf{E}^1 \sum_{I=1}^{n_e} \mathbf{B}_I^1(\xi_q) \eta_I^1 = \mathbf{N}^{1T} \mathbf{E}^1 \bar{\mathbf{B}}_q^1 \eta^1, \\ \delta p_N^2 &= \mathbf{N}^{2T} \mathbf{E}^2 \sum_{I=1}^{n_e} \mathbf{B}_I^2(\zeta_{Pq}) \eta_I^2 = \mathbf{N}^{2T} \mathbf{E}^2 \bar{\mathbf{B}}_{Pq}^2 \eta^2. \end{aligned} \quad (7.144)$$

This result can now be inserted into (7.137), which leads for a two-point GAUSS rule for one contact segment as defined in Figure 7.20 to

$$\begin{aligned} C_{ce}^N &= \sum_{q=1}^2 \left\{ -\frac{1}{2} \left[ \left( \eta^{1T} \bar{\mathbf{B}}_q^{1T} \mathbf{E}^1 \mathbf{N}^1 + \eta^{2T} \bar{\mathbf{B}}_{Pq}^{2T} \mathbf{E}^2 \mathbf{N}^2 \right) g_{Nq} \right. \right. \\ &\quad \left. \left. - \delta g_N \left( \mathbf{N}^{1T} \mathbf{E}^1 \bar{\mathbf{B}}_q^1 \mathbf{u}^1 + \mathbf{N}^{2T} \mathbf{E}^2 \bar{\mathbf{B}}_{Pq}^2 \mathbf{u}^2 \right) \right] \right. \\ &\quad \left. + \epsilon_N g_{Nq} \delta g_{Nq} \right\} J_q W_q, \end{aligned} \quad (7.145)$$

where  $J_q = \det \mathbf{J}_\Gamma(\xi_q)$  is the JACOBIAN of the surface transformation from the global coordinates to the local coordinates evaluated at the GAUSS point  $\xi_q$ .  $g_{Nq}$  is defined in (7.142), from which  $\delta g_{Nq}$  follows as

$$\delta g_{Nq} = \frac{1}{\|\mathbf{n}^1\|} \left[ (1 - \zeta_{Pq}) \boldsymbol{\eta}_1^2 + \zeta_{Pq} \boldsymbol{\eta}_2^2 - (1 - \xi_q) \boldsymbol{\eta}_1^1 - \xi_q \boldsymbol{\eta}_2^1 \right] \cdot \mathbf{n}^1. \quad (7.146)$$

The quantity  $W_q$  in (7.145) denotes the weighting factor for the GAUSS integration rule (which for two points is unity). Note that in the above equation, one has to change the notation according to the segment which is defined by the intersection of the normal at the GAUSS point and the slave surface, see Figure 7.20. Here the projection from the second GAUSS point may intersect with a different element of the slave surface.

With the introduction of a matrix form for  $g_{Nq}$  and  $\delta g_{Nq}$

$$g_{Nq} = \langle \mathbf{u}_1^2, \mathbf{u}_2^2, \mathbf{u}_1^1, \mathbf{u}_2^1 \rangle \begin{Bmatrix} (1 - \zeta_{Pq}) \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \\ \zeta_{Pq} \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \\ -(1 - \xi_q) \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \\ -\xi_q \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \end{Bmatrix} = \mathbf{u}^{12T} \mathbf{N}_q^{12},$$

$$\delta g_{Nq} = \langle \boldsymbol{\eta}_1^2, \boldsymbol{\eta}_2^2, \boldsymbol{\eta}_1^1, \boldsymbol{\eta}_2^1 \rangle \begin{Bmatrix} (1 - \zeta_{Pq}) \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \\ \zeta_{Pq} \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \\ -(1 - \xi_q) \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \\ -\xi_q \frac{\mathbf{n}^1}{\|\mathbf{n}^1\|} \end{Bmatrix} = \boldsymbol{\eta}^{12T} \mathbf{N}_q^{12}, \quad (7.147)$$

one can rewrite equation (7.145) as

$$\begin{aligned} C_{ce}^N &= -\frac{1}{2} \sum_{q=1}^2 \left[ \left( \boldsymbol{\eta}^{1T} \bar{\mathbf{B}}_q^{1T} \mathbf{E}^1 \mathbf{N}^1 + \boldsymbol{\eta}^{2T} \bar{\mathbf{B}}_{Pq}^{2T} \mathbf{E}^2 \mathbf{N}^2 \right) \mathbf{N}_q^{12T} \mathbf{u}^{12} \right. \\ &\quad \left. - \boldsymbol{\eta}^{12T} \mathbf{N}_q^{12} \left( \mathbf{N}^{1T} \mathbf{E}^1 \bar{\mathbf{B}}_q^1 \mathbf{u}^1 + \mathbf{N}^{2T} \mathbf{E}^2 \bar{\mathbf{B}}_{Pq}^2 \mathbf{u}^2 \right) \right] J_q W_q \\ &\quad + \sum_{q=1}^2 \epsilon_N \boldsymbol{\eta}^{12T} \mathbf{N}_q^{12} \mathbf{N}_q^{12T} \mathbf{u}^{12} J_q W_q. \end{aligned} \quad (7.148)$$

This equation can be recast in matrix form, leading to

$$C_{ce}^N = \sum_{q=1}^2 \langle \boldsymbol{\eta}^1, \boldsymbol{\eta}^2, \boldsymbol{\eta}^{12} \rangle \mathbf{R}_{cq}, \quad (7.149)$$

with the contact residual

$$\mathbf{R}_{cq} = \left\{ \begin{array}{l} -\frac{1}{2} \bar{\mathbf{B}}_q^{1T} \mathbf{E}^1 \mathbf{N}^1 g_{Nq} \\ -\frac{1}{2} \bar{\mathbf{B}}_{Pq}^{2T} \mathbf{E}^2 \mathbf{N}^2 g_{Nq} \\ \epsilon_N \mathbf{N}_q^{12} g_{Nq} - \frac{1}{2} \mathbf{N}_q^{12} \left( \mathbf{N}^{1T} \mathbf{E}^1 \bar{\mathbf{B}}_q^1 \mathbf{u}^1 + \mathbf{N}^{2T} \mathbf{E}^2 \bar{\mathbf{B}}_{Pq}^2 \mathbf{u}^2 \right) \end{array} \right\} J_q W_q. \quad (7.150)$$



The linearization of this linear residual term yields the contact tangent contribution for each GAUSS point:

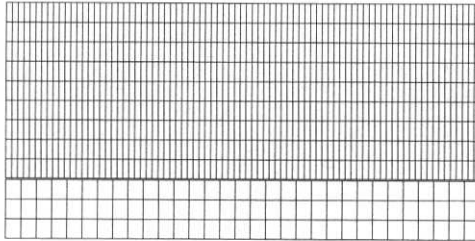
$$\mathbf{K}_{c q} = -\frac{1}{2} \begin{bmatrix} \mathbf{0} & \mathbf{0} & \bar{\mathbf{B}}_q^{1 T} \mathbf{E}^1 \mathbf{N}^1 \mathbf{N}_q^{12 T} \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{B}}_{P q}^{2 T} \mathbf{E}^2 \mathbf{N}^2 \mathbf{N}_q^{12 T} \\ \mathbf{N}_q^{12} \mathbf{N}^{1 T} \mathbf{E}^1 \bar{\mathbf{B}}_q^1 & \mathbf{N}_q^{12} \mathbf{N}^{2 T} \mathbf{E}^2 \bar{\mathbf{B}}_{P q}^2 & -2 \epsilon_N \mathbf{N}_q^{12} \mathbf{N}_q^{12 T} \end{bmatrix} J_q W_q. \quad (7.151)$$

Note that the stabilization due to the penalty method adds a term on the diagonal of the matrix.

With these matrices the contact formulation using the NITSCHKE method is completed. The method can also be applied to higher order interpolations. In that case, one has to iteratively find the intersection of the normal vector at the GAUSS points with the slave surface. Also, the normal vector is then no longer constant within the master segment. However, the main approach does not change, and the structure of the matrices which are obtained at a GAUSS point is the same as in (7.151).

A formulation for three-dimensional problems can be derived in an analogous way. The integration then has to be performed again with respect to the master surface. In that case, it is more complicated to compute the intersection of the normal vectors at the GAUSS points with the discretization of the slave surface. However, this is more a technical difficulty than a conceptual one.

To show the features of the NITSCHKE method, we consider a simple example with two elastic blocks. The dimension of the blocks are, respectively  $2 \times 0.25$  and  $2 \times 0.75$  *units* (see Figure 7.21); the meshes of the two blocks do not match in the interface, as depicted in Figure 7.21. The foundation and the block have been divided, respectively, into  $31 \times 3$  elements and  $81 \times 9$  elements. The upper and lower sides are clamped and a downward displacement of 0.02 *units* is imposed at the upper side. Material constants are:  $E = 10^8$  *units*,  $\nu = 0.0$ . A plane stress model is assumed. Results are collected in Table 7.1. These show that the NITSCHKE method is able to improve the penetration residual norm. This norm is introduced to monitor the accuracy of



**Fig. 7.21** Mesh of two elastic blocks with  $31 \times 3$  and  $81 \times 9$  elements.

Table 7.1 Residual norm for non-matching meshes with uniform pressure

# Gauss	Penalty $\varepsilon_N = 10^{10}$	NITSCHÉ-no-penalty	NITSCHÉ+Penalty
2	0.1980206E-03	0.2149006E-04	0.6970404E-06
5	0.1980199E-03	0.2628065E-05	0.1532591E-06
20	0.1980198E-03	0.4897944E-06	0.9072227E-08

the solutions by computing a

$$\|g_N\| = \sqrt{\frac{\sum_{i=1}^{n_c} g_{Ni}^2}{n_c}},$$

(7.152)

where  $n_c$  is the number of active contact points. As can be seen, the best result is obtained using a non-zero penalty term. The stress field is perfectly uniform, i.e. no disturbance of the stress field is generated at the interface for all the cases, with the exception of the case Nitsche-no-penalty with two GAUSS points. Here a small perturbation (1.03%) of the vertical stress field is produced, as depicted in Figure 7.22.

Next a non-uniform, linearly varying displacement field is applied to the problem shown in Figure 7.21. In this case, the global performance of the methods is still the same. However, a small disturbance of the stress field takes place in this case if the NITSCHÉ method is used with a zero penalty value, as depicted in Figure 7.23. It also has to be noticed that the reduction of the residual norm  $\|g_N\|$  with a non-zero penalty value is still remarkable (see Table 7.2). As a general rule, we can say that the NITSCHÉ method permits us to achieve a better residual of the penetrations.

A more general example has been developed to check the performance of the methods. For this purpose, two contacting blocks with curved contact surfaces are

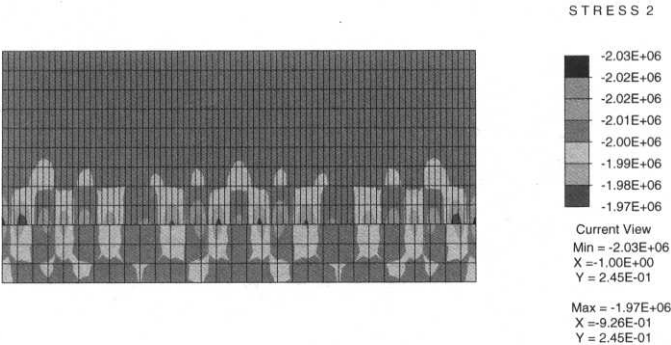
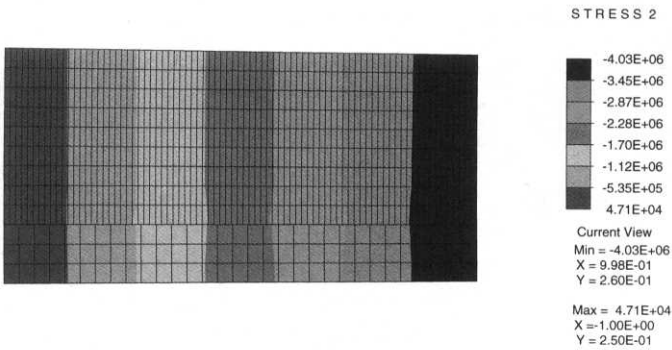


Fig. 7.22 Vertical stresses in both blocks for the NITSCHÉ-no-penalty case.

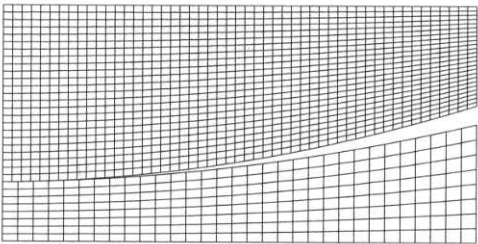


**Fig. 7.23** Mesh and vertical normal stress for example with non-uniform loading.

**Table 7.2** Residual norm for matching mesh with non-uniform loading

# Gauss	Penalty $\varepsilon_N = 10^{10}$	NITSCHKE-no-penalty	NITSCHKE+Penalty
2	0.2289195E-03	0.3857169E-04	0.1083247E-05
5	0.2289208E-03	0.5279852E-04	0.9166951E-06
20	0.2289215E-03	0.5503470E-04	0.9317194E-06

considered. The lower block has been discretized with  $31 \times 8$  elements, and the upper one with  $51 \times 24$  elements. Due to the parabolic shape of the surfaces, only the first node on the left is initially in contact, see Figure 7.24. The material data of the upper block is  $E = 10^7 \text{ units}$ ,  $\nu = 0.0$ , whereas the lower block assumes the same material parameters as given in the previous example. A downward uniform displacement of the top side is applied in three time steps such that the active contact zone varies with the steps.



**Fig. 7.24** Two blocks with curved surfaces.

Table 7.3 Residual norm for example with curved surfaces

STEP	Penalty $\varepsilon_N = 10^{10}$	NITSCHÉ-no-penalty	NITSCHÉ + Penalty
1	0.2055726E-03	no convergence	0.2064216E-04
2	0.3957804E-03	no convergence	0.2428796E-04
3	0.5856056E-03	no convergence	0.3814317E-04

Looking at the residual norms collected in Table 7.3, we again find a significant improvement in the solution achieved also in the previous examples. However, it should be noted that for this case the NITSCHÉ method does not converge without penalty stabilization. This observation was also made in Becker and Hansbo (1999) for an application of the NITSCHÉ method to non-matching grids in domain decomposition methods. The vertical stresses at the first load step are shown in Figure 7.25). It is also easy to see that the NITSCHÉ method produces a local disturbance of the

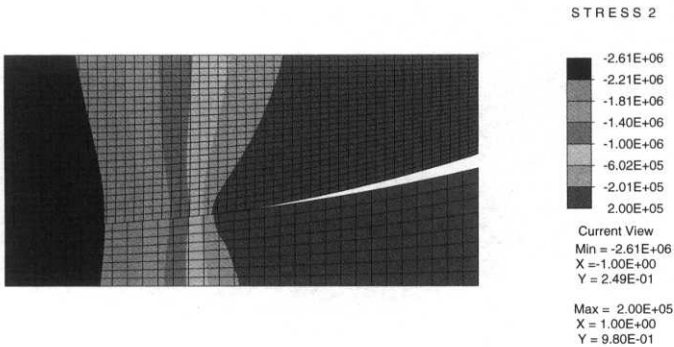


Fig. 7.25 Vertical stresses at first load step.

stress field in the contact zone. Its source can be the projection method which was used within the derivation of the method, e.g. see Figure 7.20.

**This page intentionally left blank**

# 8

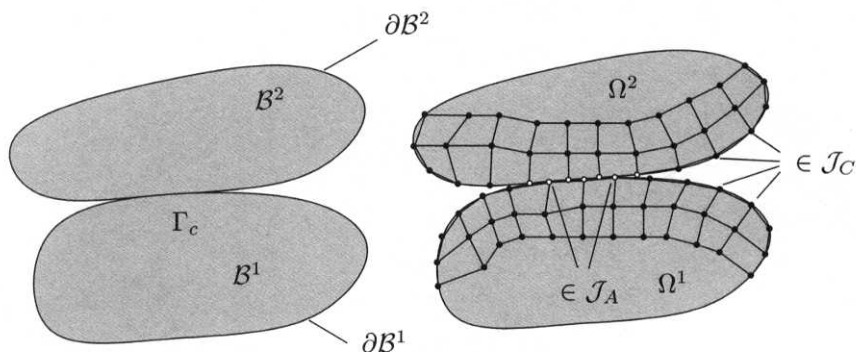
---

## *Discretization, Large Deformation Contact*

For the general case of contact including large deformations (see Figure 8.1) the active part of the contact interface  $\Gamma_c$ , where  $g_N \leq 0$ , has to be determined during the solution process. Currently, in finite deformations, most algorithms rely on discretizations which are based on nodes on the boundary of the bodies. Thus, the algorithm has to find the active contact constraints denoted by  $\mathcal{I}_A \in \mathcal{I}_C$ , where  $g_N = 0$  is fulfilled.  $\mathcal{I}_C$  are all possible contact nodes on the contact surface. In finite deformations this task requires an extensive contact search, which is discussed in more detail in Section 9.1.

The most frequently used discretization is the so-called node-to-segment approach. Here arbitrary sliding of a node over the entire contact area is allowed. Early implementations can be found in Hallquist (1979) or Hughes et al. (1977) and have been extended to more general cases, Hallquist et al. (1985), Bathe and Chaudhary (1985) and Wriggers et al. (1990). Now many finite element codes also include self-contact, e.g. see Hallquist et al. (1992). Also the idea of contact segments proposed by Simo et al. (1985), has been followed up and applied to problems involving large deformations, see Papadopoulos and Taylor (1992).

In this section we discuss discretization techniques for large deformation contact, where sliding of a contacting node or element over several elements of the other body is also allowed. To describe such a process properly, we introduce the master-slave concept in the current configuration in which a discrete slave point, denoted by the subscript  $s$  in the following, with coordinate  $\mathbf{x}_s^2$  comes into contact with a master segment, described by  $\mathbf{x}^1$ . The master segment will be a line in two-dimensional situations and a surface in three-dimensional contact problems. The line or surface describing the master segment is usually parameterized by convective coordinates



**Fig. 8.1** Discretization for large deformation contact.

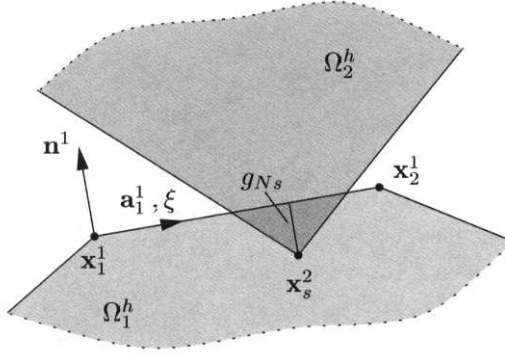
$\xi$ ; see Appendix B for the definition. This approach is well known from beam or shell theory. A parameterization of the master segment with convective coordinates,  $\mathbf{x}^1 = \mathbf{x}^1(\xi)$ , is also very close to the standard isoparametric formulation, see Chapter 6. However, when we discuss the  $C^1$ -continuous approximation of the contact surface, then the surface parameterization is not identically with the isoparametric interpolation of the underlying finite elements.

*REMARK 8.1: When the discretization of contact surfaces is considered, one has to distinguish between the contact of two deformable bodies or the contact of a deformable body with a rigid obstacle. At first glance, it seems that the latter case is simply a special case of the first problem, which is true. But due to the fact that the surface description of a rigid obstacle can be given once and for all by the correct geometrical model, this knowledge can be used within the discretization process. A formulation based on CAD-surfaces was developed in Hansson and Klarbring (1990), see also Heegaard and Curnier (1993), which includes applications to biomechanics. Williams and Pentland (1992) considered so-called superquadrics to specify the geometry of contacting objects, and Wriggers and Imhof (1993) formulated the contact problem with a rigid obstacle using splines.*

## 8.1 TWO-DIMENSIONAL NODE-TO-SEGMENT CONTACT DISCRETIZATION

The first formulation for finite sliding of a slave point over a master surface leads to a relatively simple element which is commonly included in nonlinear finite element codes. This discretization is called a node-to-segment (NTS) contact element, and is widely used in nonlinear finite element simulations of contact problems. Due to its importance, we consider this contact element in detail.

**Kinematics.** The kinematical relations can be directly computed using the equations stated in Chapter 3. We assume that the discrete slave point (s) with coordinate



**Fig. 8.2** Node-to-segment contact element.

$\mathbf{x}_s^2$  comes into contact with the master segment (1)–(2) defined by the nodal coordinates  $\mathbf{x}_1^1$  and  $\mathbf{x}_2^1$ , see Figure 8.2. With the interpolation for the master segment (contact surface of body  $\mathcal{B}^1$  based on the introduction of the surface coordinate  $\xi$  along the master surface

$$\hat{\mathbf{x}}^1(\xi) = \mathbf{x}_1^1 + (\mathbf{x}_2^1 - \mathbf{x}_1^1) \xi, \quad (8.1)$$

one can easily compute the tangent vector of the segment leading to

$$\bar{\mathbf{a}}_1^1 = \hat{\mathbf{x}}^1(\xi)_{,\xi} = (\mathbf{x}_2^1 - \mathbf{x}_1^1). \quad (8.2)$$

It is connected to a normalized base vector  $\mathbf{a}_1^1$  by  $\mathbf{a}_1^1 = \bar{\mathbf{a}}_1^1 / l$ , with  $l = \|\mathbf{x}_2^1 - \mathbf{x}_1^1\|$  being the current length of the master segment. With the unit tangent vector  $\mathbf{a}_1^1$  the unit normal to the segment (1)–(2) can be defined as  $\mathbf{n}^1 = \mathbf{e}_3 \times \mathbf{a}_1^1$ .

$\bar{\xi}$  and  $g_{Ns}$  are given by the solution of the minimal distance problem, i.e. by the projection of the slave node  $\mathbf{x}_s$  in (s) onto the master segment (1)–(2):

$$\bar{\xi} = \frac{1}{l} (\mathbf{x}_s^2 - \mathbf{x}_1^1) \cdot \mathbf{a}_1^1 \quad \text{and} \quad g_{Ns} = [\mathbf{x}_s^2 - (1 - \bar{\xi}) \mathbf{x}_1^1 - \bar{\xi} \mathbf{x}_2^1] \cdot \mathbf{n}^1. \quad (8.3)$$

From these equations and the local continuous formulation (3.29), we directly compute the variation of the gap function  $\delta g_N$  on the straight master segment (1)–(2)

$$\delta g_{Ns} = [\eta_s^2 - (1 - \bar{\xi}) \eta_1^1 - \bar{\xi} \eta_2^1] \cdot \mathbf{n}^1. \quad (8.4)$$

In the case of tangential sliding or stick, we have to distinguish between two different states, which are discussed below. For stick we introduce the relative tangential movement by

$$g_{Ts}^{st} = \int_{\xi_0}^{\bar{\xi}} l d\xi = (\bar{\xi} - \xi_0) l, \quad (8.5)$$

where  $\xi_0$  characterizes the stick point on the master segment and  $\bar{\xi}$  is the current projection defined in (8.3)<sub>1</sub>. Note that for an exact solution using the LAGRANGE



multiplier formulation  $g_{Ts}^{st}$  has to be zero, which means  $\bar{\xi} = \xi_0$ . In the case of the penalty method,  $g_{Ts}^{st}$  is the relative tangential motion which is controlled by the penalty stiffness, see Section 5.3.2. The variation of (8.5) yields the discrete form of the variation of the tangential gap

$$\delta g_{Ts}^{st} = l \delta \bar{\xi} + (\bar{\xi} - \xi_0) \delta l. \quad (8.6)$$

The variation of  $\bar{\xi}$  can be obtained using (3.31). For a straight segment this leads to

$$\begin{aligned} \bar{H}_{11} &= (a_{11} + g_{Ns} b_{11}) \implies \bar{H}_{11} = a_{11} = l^2, \\ \bar{R}_1 &= [\eta^2 - \hat{\eta}^1(\bar{\xi})] \cdot \bar{\mathbf{a}}_1^1 + g_{Ns} \bar{\mathbf{n}}^1 \cdot \hat{\eta}_{,\xi}^1(\bar{\xi}), \end{aligned} \quad (8.7)$$

which leads to

$$\delta \bar{\xi} = \frac{1}{l^2} \{ [\eta^2 - \hat{\eta}^1(\bar{\xi})] \cdot \bar{\mathbf{a}}_1^1 + g_{Ns} \bar{\mathbf{n}}^1 \cdot \hat{\eta}_{,\xi}^1(\bar{\xi}) \}. \quad (8.8)$$

Using these relations the variation of  $g_{Ts}^{st}$  for the NTS-element with  $\delta l = [\eta_2^1 - \eta_1^1] \cdot \mathbf{a}_1^1$  and  $\hat{\eta}^1(\xi) = \eta_2^1 + \xi(\eta_1^1 - \eta_2^1)$  is given by

$$\delta g_{Ts}^{st} = [\eta_s^2 - (1 - \bar{\xi}) \eta_1^1 - \bar{\xi} \eta_2^1] \cdot \mathbf{a}_1^1 + \frac{g_{Ns}}{l} [\eta_2^1 - \eta_1^1] \cdot \mathbf{n}^1 + \frac{g_{Ts}}{l} [\eta_2^1 - \eta_1^1] \cdot \mathbf{a}_1^1. \quad (8.9)$$

In the case of slip we no longer remain at the point  $\xi_0$ , and thus compute the relative tangential velocity from (3.24), and obtain

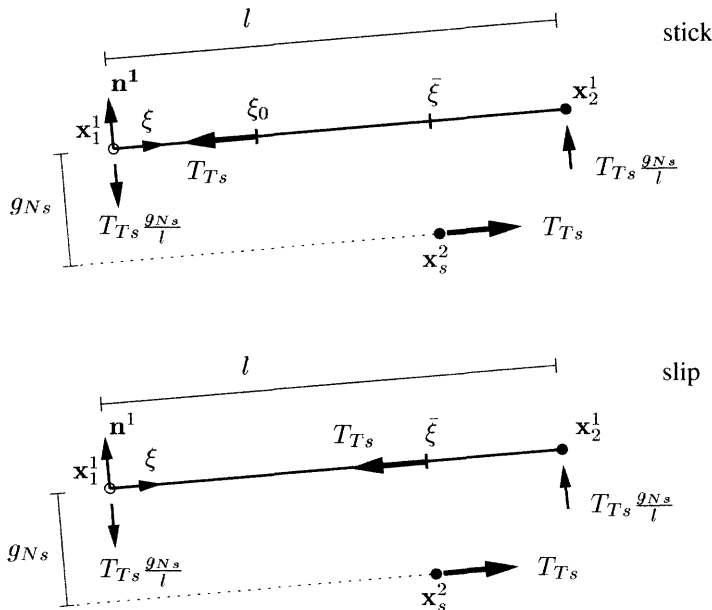
$$\delta g_{Ts}^{sl} = \delta \bar{\xi} l = [\eta_s^2 - (1 - \bar{\xi}) \eta_1^1 - \bar{\xi} \eta_2^1] \cdot \mathbf{a}_1^1 + \frac{g_{Ns}}{l} [\eta_2^1 - \eta_1^1] \cdot \mathbf{n}^1, \quad (8.10)$$

which differs from (8.9) by the last term. Equations (8.4), (8.9) and (8.10) characterize the main kinematical relations of the contact element in Figure 8.2.

**REMARK 8.2:** A similar approach which was based on pure geometrical considerations and hence started directly from the right side of (8.5) was applied in Zavarise and Taylor (1996) and Wriggers (1999); however, these authors use (8.9) instead of (8.10) for slip. It can be shown that this leads to an incorrect distribution of the tangential force in the contact interface, and hence the differentiation made in (8.9) and (8.10) is necessary. By using the definition  $g_{Ts} = (\bar{\xi} - \xi_0) l$  in (8.10), after some algebra one can obtain

$$\delta g_{Ts}^{st} = \delta \bar{\xi} l = [\eta_s^2 - (1 - \xi_0) \eta_1^1 - \xi_0 \eta_2^1] \cdot \mathbf{a}_1^1 + \frac{g_{Ns}}{l} [\eta_2^1 - \eta_1^1] \cdot \mathbf{n}^1, \quad (8.11)$$

which means that the tangential force acts on the master segment at  $\xi_0$  and not at  $\bar{\xi}$ . This is correct for the stick case, since  $\mathbf{x}_s^2$  is not allowed to move in the tangential direction. For the LAGRANGE multiplier method the projection we obtain is  $\bar{\xi} = \xi_0$ ; only in the case of the penalty approach does the tangential gap depicted here occur. However, for slip the tangential force has to be acting at the position given by the current position  $\bar{\xi}$ . Both facts are illustrated in Figure 8.3 for the node-to-segment element defined in Figure 8.2. The figures show the distribution of the tangential forces in the node-to-segment contact element for the stick and slip cases according to (8.11) and (8.10). Note that some forces occur in the normal direction.



**Fig. 8.3** Distribution of tangential contact forces for slip and stick.

This stems from the fact that the couple  $(T_{Ts} g_{Ns})$  due to the gap distance  $g_{Ns}$  has to be equilibrated.

It should be observed that the difference in using (8.9) or (8.10) is not significant as long as the time or load steps in an incremental solution procedure are small.

**Contact residual.** In what follows we compute the contribution of the node-to-segment element to the weak form (5.40). The basic formulation for this discretization is analogous to the node-to-node element. Thus we assume that we know the normal force  $P_{Ns} = p_{Ns} A_s$  and the tangential force  $T_{Ts} = t_{Ts} A_s$  at the discrete contact point (s) of the contact element under consideration, where  $A_s$  denotes the area of the contact element. Both forces,  $P_{Ns}$  and  $T_{Ts}$ , can be obtained from the constitutive relations discussed in Sections 4.1 and 4.2. This leads to

$$\int_{\Gamma_c} (p_N \delta g_N + t_T \delta g_T) d\Gamma \approx \sum_{s=1}^{n_c} (P_{Ns} \delta g_{Ns} + T_{Ts} \delta g_{Ts}), \quad (8.12)$$

where we need to distinguish between stick and slip in the second term. In practice we compute the normal force  $P_{Ns}$  either from equation (4.12) or from the penalty update  $P_{Ns} = \epsilon g_{Ns}$  multiplied by the area of the contact element. For the tangential force  $T_{Ts}$  we have to perform an algorithmic update, as we shall discuss in Section 9.5.2.

Thus, the contributions of one contact element take the form

$$\delta g_{Ns} P_{Ns} + \delta g_{Ts} T_{Ts} \quad (8.13)$$

for the discrete contact point (s) with the mechanical relative (LIE-type) variations analogous to (8.4) and (8.9). This equation can now be cast into a matrix formulation. For the normal part (8.13)<sub>1</sub> we set for the variation (8.4) of the penetration

$$\delta g_{N_s} = \eta_s^T N_s. \quad (8.14)$$

With the same notation we can express the variation (8.9) and (8.10) of the tangential gap for stick

$$\delta g_{T_s}^{st} = \eta_s^T \left( T_s + \frac{g_{N_s}}{l} N_{0s} + \frac{g_{T_s}}{l} T_{0s} \right) = \eta_s^T T_s^{st}, \quad (8.15)$$

and for slip

$$\delta g_{T_s}^{sl} = \eta_s^T \left( T_s + \frac{g_{N_s}}{l} N_{0s} \right) = \eta_s^T T_s^{sl}. \quad (8.16)$$

In these equations the following vectors have been used:

$$\eta_s = (\eta_s^2 \quad \eta_s^1 \quad \eta_s^1)^T, \quad (8.17)$$

$$N_s = \left\{ \begin{array}{c} \mathbf{n}^1 \\ -(1 - \bar{\xi}) \mathbf{n}^1 \\ -\bar{\xi} \mathbf{n}^1 \end{array} \right\}_s, \quad N_{0s} = \left\{ \begin{array}{c} \mathbf{0} \\ -\mathbf{n}^1 \\ \mathbf{n}^1 \end{array} \right\}_s, \quad (8.18)$$

and

$$T_s = \left\{ \begin{array}{c} \mathbf{a}_1^1 \\ -(1 - \bar{\xi}) \mathbf{a}_1^1 \\ -\bar{\xi} \mathbf{a}_1^1 \end{array} \right\}_s, \quad T_{0s} = \left\{ \begin{array}{c} \mathbf{0} \\ -\mathbf{a}_1^1 \\ \mathbf{a}_1^1 \end{array} \right\}_s. \quad (8.19)$$

Thus, the virtual mechanical work of the contact element can be written in the matrix formulation  $\eta_s^T G_s^c$  with the contact element residual

$$G_s^c = P_{N_s} N_s + T_{T_s} \tilde{T}_s, \quad (8.20)$$

where  $\tilde{T}_s$  stands either for stick or slip, see (8.15) or (8.16).

A pure displacement formulation of the contact problem is possible by expressing  $P_{N_s}$  either through (4.12) or (4.10), or by the penalty relation  $P_{N_s} = \epsilon_N g_{N_s}$ . This is in contrast to the LAGRANGE multiplier technique, where  $P_{N_s} = \lambda_{N_s}$ . But we observe that this discretization can be applied to both methods.

*REMARK 8.3: In the case of the USZAWA algorithm (see Section 5.3.8) within the augmented LAGRANGE method, we have to replace  $P_{N_s}$  in (8.20) in case of frictionless contact for the standard update technique by*

$$P_{N_s}^{new} = \bar{P}_{N_s}^{old} + \epsilon_N g_{N_s}^{new}, \quad (8.21)$$

and when using a constitutive model according to (5.67) by

$$P_{N_s}^{new} = \bar{P}_{N_s}^{old} + \epsilon_N \{ g_{N_s}^{new} - [\zeta - d(P_{N_s}^{old})] \}, \quad (8.22)$$

where  $g_{N_s}$  is given by (8.4).

**Linearization normal contact.** Often a NEWTON–RAPHSON iteration is used to solve the global set of equations. In that case linearization of (8.20) is needed to achieve quadratic convergence near the solution point. The linearization of the contact contribution will be discussed here for the case of the LAGRANGE multiplier and penalty methods. These linearizations were derived first in Wriggers and Simo (1985). Within the LAGRANGE multiplier approach, the starting equation for the NTS-element is for frictionless contact

$$P_{Ns} \delta g_{Ns} + \delta P_{Ns} g_{Ns}, \quad (8.23)$$

where  $P_{Ns}$  stands for the LAGRANGE multiplier. Linearization of this expression yields

$$\Delta P_{Ns} \delta g_{Ns} + P_{Ns} \Delta \delta g_{Ns} + \delta P_{Ns} \Delta g_{Ns}. \quad (8.24)$$

Since  $\delta g_{Ns}$  is given in (8.4) and  $\Delta g_{Ns}$  has the same structure as  $\delta g_{Ns}$ ,

$$\Delta g_{Ns} = [\Delta \mathbf{u}_s^2 - (1 - \bar{\xi}) \Delta \mathbf{u}_1^1 - \bar{\xi} \Delta \mathbf{u}_2^1] \cdot \mathbf{n}^1, \quad (8.25)$$

we only have to compute  $\Delta \delta g_{Ns}$ . The terms which depend upon the displacements in (8.4) are  $\bar{\xi}$  and  $\mathbf{n}^1$ , and we have to compute the linearizations of these quantities. The linearization of  $\bar{\xi}$  is equivalent to  $\delta \bar{\xi}$ , just the variations  $\boldsymbol{\eta}$  have to be exchanged by the displacement increments  $\Delta \mathbf{u}$ :

$$\Delta \bar{\xi} = \frac{1}{l} [\Delta \mathbf{u}_s^2 - (1 - \bar{\xi}) \Delta \mathbf{u}_1^1 - \bar{\xi} \Delta \mathbf{u}_2^1] \cdot \mathbf{a}_1^1 + \frac{g_{Ns}}{l^2} [\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1] \cdot \mathbf{n}^1. \quad (8.26)$$

To obtain the linearization of  $\mathbf{n}^1$  we start from the following expression:

$$\mathbf{n}^1 = \mathbf{e}_3 \times \mathbf{a}_1^1 \implies \Delta \mathbf{n}^1 = \mathbf{e}_3 \times \Delta \mathbf{a}_1^1, \quad (8.27)$$

where  $\mathbf{e}_3$  is the unit base vector perpendicular to the two-dimensional plane. The linearization of the unit tangent vector follows from (8.2) by a straightforward computation:

$$\begin{aligned} \Delta \mathbf{a}_1^1 &= \Delta \left[ \frac{1}{l} (\mathbf{x}_2^1 - \mathbf{x}_1^1) \right] = \frac{1}{l} (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1) - \frac{1}{l^2} (\mathbf{x}_2^1 - \mathbf{x}_1^1) [\mathbf{a}_1^1 \cdot (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1)] \\ &= \frac{1}{l} [\mathbf{1} - \mathbf{a}_1^1 \otimes \mathbf{a}_1^1] (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1), \end{aligned} \quad (8.28)$$

where the term  $[\mathbf{1} - \mathbf{a}_1^1 \otimes \mathbf{a}_1^1]$  can be rewritten with the definition of the unit tensor  $\mathbf{1} = \mathbf{a}_1^1 \otimes \mathbf{a}_1^1 + \mathbf{n}^1 \otimes \mathbf{n}^1$ . Hence

$$\Delta \mathbf{a}_1^1 = \frac{1}{l} [\mathbf{n}^1 \otimes \mathbf{n}^1] (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1), \quad (8.29)$$

and with (8.27), we finally obtain

$$\Delta \mathbf{n}^1 = -\frac{1}{l} [\mathbf{a}_1^1 \otimes \mathbf{n}^1] (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1). \quad (8.30)$$

These results can now be used to derive the expression for  $\Delta\delta g_{Ns}$

$$\begin{aligned}\Delta\delta g_{Ns} &= -\Delta\bar{\xi}(\eta_2^1 - \eta_1^1) \cdot \mathbf{n}^1 + [\eta_s^2 - (1 - \bar{\xi})\eta_1^1 - \bar{\xi}\eta_2^1] \cdot \Delta\mathbf{n}^1 \\ &= -\frac{1}{l} [\Delta\mathbf{u}_s^2 - (1 - \bar{\xi})\Delta\mathbf{u}_1^1 - \bar{\xi}\Delta\mathbf{u}_2^1] \cdot \mathbf{a}_1^1 (\eta_2^1 - \eta_1^1) \cdot \mathbf{n}^1 \\ &\quad - \frac{g_{Ns}}{l^2} (\eta_2^1 - \eta_1^1) \cdot \mathbf{n}^1 (\Delta\mathbf{u}_2^1 - \Delta\mathbf{u}_1^1) \cdot \mathbf{n}^1 \\ &\quad - \frac{1}{l} [\eta_s^2 - (1 - \bar{\xi})\eta_1^1 - \bar{\xi}\eta_2^1] \cdot \mathbf{a}_1^1 (\Delta\mathbf{u}_2^1 - \Delta\mathbf{u}_1^1) \cdot \mathbf{n}^1,\end{aligned}\quad (8.31)$$

which is symmetric with respect to variation and linearization. Using the matrices in (8.18), (8.19) and introducing  $\Delta\mathbf{u}_s^T = \{\Delta\mathbf{u}_s^2, \Delta\mathbf{u}_1^1, \Delta\mathbf{u}_2^1\}$ , we can state the matrix form of  $\Delta\delta g_{Ns}$

$$\Delta\delta g_{Ns} = \eta_s^T \mathbf{K}_{\Delta\delta} \Delta\mathbf{u}_s, \quad (8.32)$$

with

$$\mathbf{K}_{\Delta\delta} = -\frac{1}{l} [\mathbf{N}_{0s} \mathbf{T}_s^T + \mathbf{T}_s \mathbf{N}_{0s}^T + \frac{g_{Ns}}{l} \mathbf{N}_{0s} \mathbf{N}_{0s}^T]. \quad (8.33)$$

Using this result, we can establish the matrix form of (8.24) for the LAGRANGE multiplier method

$$\langle \eta_s^T, \delta P_{Ns} \rangle \mathbf{K}_s^{LM} \begin{Bmatrix} \Delta\mathbf{u}_s \\ \Delta P_{Ns} \end{Bmatrix} \quad \text{with} \quad \mathbf{K}_s^{LM} = \begin{bmatrix} P_{Ns} \mathbf{K}_{\Delta\delta} & \mathbf{N}_s \\ \mathbf{N}_s^T & \mathbf{0} \end{bmatrix}. \quad (8.34)$$

In the case of the penalty method, equation (8.24) reduces with  $P_{Ns} = \epsilon_N g_{Ns}$  to

$$\epsilon_N \Delta g_{Ns} \delta g_{Ns} + \epsilon_N g_{Ns} \Delta \delta g_{Ns}, \quad (8.35)$$

which yields the matrix form

$$\mathbf{K}_{Ns}^p = \epsilon_N \left[ \mathbf{N}_s \mathbf{N}_s^T - \frac{g_{Ns}}{l} \left( \mathbf{N}_{0s} \mathbf{T}_s^T + \mathbf{T}_s \mathbf{N}_{0s}^T + \frac{g_{Ns}}{l} \mathbf{N}_{0s} \mathbf{N}_{0s}^T \right) \right]. \quad (8.36)$$

All matrices in (8.33) to (8.36) are defined in (8.18) and (8.19).

#### REMARK 8.4

1. Note that in a geometrically linear case, all terms which are multiplied by  $g_{Ns}$  vanish. This gives the simple matrix  $\mathbf{K}_{Ns}^{LMp} = \epsilon_N \mathbf{N}_s \mathbf{N}_s^T$ .
2. When using a LAGRANGE multiplier method for the linear case the term  $P_{Ns} \mathbf{K}_{\Delta\delta}$  disappears in (8.34), which also leads to a simple matrix structure.
3. The terms multiplied by  $G_{Ns}$  may be omitted in (8.36) in nonlinear computations in the first stages of a NEWTON iteration. This is because quadratic convergence only occurs in the standard algorithm once the contact area is already known correctly, which results from the fact of non-differentiability of the gap function. Hence one does not need the full tangent in the first iterations, especially since the terms containing  $g_{Ns}$  can become quite large during the first iteration steps and might even lead to divergence for large incremental steps.

**Linearization for stick.** In the same way, the linearizations are obtained for the tangential stick and slip parts. The point of departure in the case of stick is the same as

for frictionless contact, only the variables are changed. For the LAGRANGE multiplier formulation this leads with  $\lambda_{Ts} = T_{Ts}$  to

$$\Delta T_{Ts} \delta g_{Ts}^{st} + T_{Ts} \Delta \delta g_{Ts}^{st} + \delta T_{Ts} \Delta g_{Ts}^{st}. \quad (8.37)$$

Again, the first and last terms contain known quantities where, according to (8.9), the linearization of the relative tangential gap is given by

$$\Delta g_{Ts}^{st} = [\Delta \mathbf{u}_s^2 - (1 - \xi_0) \Delta \mathbf{u}_1^1 - \xi_0 \Delta \mathbf{u}_2^1] \cdot \mathbf{a}_1^1 + \frac{g_{Ns}}{l} [\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1] \cdot \mathbf{n}^1. \quad (8.38)$$

The linearization of the variation  $\delta g_{Ts}^{st}$  is now obtained after some algebra:

$$\begin{aligned} \Delta \delta g_{Ts}^{st} &= \frac{1}{l} [\eta_s^2 - (1 - \xi_0) \eta_1^1 - \xi_0 \eta_2^1] \cdot [\mathbf{n}^1 \otimes \mathbf{n}^1] (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1) \\ &+ \frac{1}{l} [\Delta \mathbf{u}_s^2 - (1 - \xi_0) \Delta \mathbf{u}_1^1 - \xi_0 \Delta \mathbf{u}_2^1] \cdot [\mathbf{n}^1 \otimes \mathbf{n}^1] (\eta_2^1 - \eta_1^1) \\ &- \frac{g_{Ts}^{st}}{l^2} (\eta_2^1 - \eta_1^1) \cdot [\mathbf{n}^1 \otimes \mathbf{n}^1] (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1) \\ &- \frac{g_{Ns}}{l^2} (\eta_2^1 - \eta_1^1) \cdot [\mathbf{n}^1 \otimes \mathbf{a}_1^1 + \mathbf{a}_1^1 \otimes \mathbf{n}^1] (\Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1), \end{aligned} \quad (8.39)$$

or in matrix form

$$\Delta \delta g_{Ts}^{st} = \eta_s^T \mathbf{K}_{\Delta \delta}^{st} \mathbf{u}_s, \quad (8.40)$$

with

$$\mathbf{K}_{\Delta \delta}^{st} = \frac{1}{l} (N_{\xi 0} N_{0s}^T + N_{0s} N_{\xi 0}^T) - \frac{g_{Ts}^{st}}{l^2} N_{0s} N_{0s}^T - \frac{g_{Ns}}{l^2} (N_{0s} \mathbf{T}_{0s}^T + \mathbf{T}_{0s} N_{0s}^T). \quad (8.41)$$

Here the matrix  $N_{\xi 0}$  has the same structure as  $N_s$ , only the surface coordinate  $\bar{\xi}$  in (8.18) is exchanged by  $\xi_0$ .

The matrix form of (8.37) for the LAGRANGE multiplier method is computed analogous to (8.34), and leads to

$$\langle \eta_s^T, \delta P_{Ts} \rangle \mathbf{K}_{Ts}^{LM} \begin{Bmatrix} \Delta \mathbf{u}_s \\ \Delta P_{Ts} \end{Bmatrix} \quad \text{with} \quad \mathbf{K}_{Ts}^{LM} = \begin{bmatrix} T_{Ts} \mathbf{K}_{\Delta \delta}^{st} & \mathbf{T}_s^{st} \\ (\mathbf{T}_s^{st})^T & \mathbf{0} \end{bmatrix}. \quad (8.42)$$

In the case of the penalty method, equation (8.37) reduces with  $P_{Ts} = \epsilon_T g_{Ts}^{st}$  to

$$\epsilon_T \Delta g_{Ts}^{st} \delta g_{Ts}^{st} + \epsilon_T g_{Ts}^{st} \Delta \delta g_{Ts}^{st}, \quad (8.43)$$

which yields the matrix form for pure stick:

$$\begin{aligned} \mathbf{K}_{Ts}^{st} &= \epsilon_T \left\{ [\mathbf{T}_s^{st} (\mathbf{T}_s^{st})^T + \frac{g_{Ts}^{st}}{l} [N_{\xi 0} N_{0s}^T + N_{0s} N_{\xi 0}^T \right. \\ &\quad \left. - \frac{g_{Ns}}{l} (\mathbf{T}_{0s} N_{0s}^T + N_{0s} \mathbf{T}_{0s}^T) - \frac{g_{Ts}^{st}}{l} N_{0s} N_{0s}^T] \right\}. \end{aligned} \quad (8.44)$$

Note that the tangent matrices for the LAGRANGE multiplier and the penalty method are symmetric.

In (8.44) all terms containing  $g_{N_s}$  and  $g_{T_s}$  disappear in a geometrically linear situation, which yields  $\mathbf{K}_{T_s}^{st lin} = \epsilon_T \mathbf{T}_s \mathbf{T}_s^T$ . Within the LAGRANGE multiplier approach also, only  $\mathbf{T}_s$  is used in (8.42) for small deformations; furthermore, the tangent  $\mathbf{K}_{\Delta\delta}^{st}$  is zero.

**Linearization for slip.** The slip case starts from a constitutive equation for the tangential stress. Hence we have to linearize only the term  $T_{T_s} \delta g_{T_s}^{sl}$ , which yields

$$\Delta T_{T_s} \delta g_{T_s}^{sl} + T_{T_s} \Delta \delta g_{T_s}^{sl}, \quad (8.45)$$

where  $\Delta T_{T_s}$  may be computed according to Section 9.5.2 for a given tangential contact law of Section 4.2. Here we state the results for classical COULOMB'S law. For more complex constitutive models in the contact interface, see Section 9.5.2. According to (9.118), COULOMB'S law yields the tangential stress  $T_{T_{n+1}} = \mu P_{N_{n+1}} \text{sign } T_{T_{n+1}}^{tr}$  at time  $t_{n+1}$ . The linearization of  $T_{T_{n+1}}$  with respect to the displacement field is given for the penalty method by

$$\begin{aligned} \Delta T_{T_{n+1}} &= \mu \Delta P_{N_{n+1}} \text{sign } T_{T_{n+1}}^{tr} + \mu P_{N_{n+1}} \Delta (\text{sign } T_{T_{n+1}}^{tr}) \\ &= \mu \epsilon_N \Delta g_{N_{n+1}} \text{sign } T_{T_{n+1}}^{tr}, \end{aligned} \quad (8.46)$$

where the linearization of the last term in (9.118) disappears in a two-dimensional formulation, since the signum function "sign" is piecewise constant. Using (8.18) the matrix of the incremental tangential stress is given by

$$\Delta T_{T_{n+1}} = \mu \epsilon_N \text{sign } T_{T_{n+1}}^{tr} \mathbf{N}_{s_{n+1}}^T \Delta \mathbf{u}_s. \quad (8.47)$$

The linearization of  $\Delta \delta g_{T_s}^{sl}$  is performed analogous to (8.39). It will be stated here only in matrix form

$$\Delta \delta g_{T_s}^{sl} = \eta_s^T \mathbf{K}_{\Delta\delta}^{sl} \Delta \mathbf{u}_s, \quad (8.48)$$

with

$$\mathbf{K}_{\Delta\delta}^{sl} = \frac{1}{l} (N_s N_{0s}^T + N_{0s} N_s^T - T_{0s} \mathbf{T}_s^T) - \frac{g_{N_s}}{l^2} (N_{0s} \mathbf{T}_{0s}^T + 2 T_{0s} N_{0s}^T). \quad (8.49)$$

Using these results, the matrix form for the slip case can be stated for the contacting node (s) as

$$\mathbf{K}_{T_s}^{sl} = T_{T_s} \mathbf{K}_{\Delta\delta}^{sl} + \mu \epsilon_N \text{sign } T_{T_{n+1}}^{tr} \mathbf{T}_s^{sl} \mathbf{N}_s^T. \quad (8.50)$$

Note that this matrix is unsymmetric, which corresponds to the non-associativity of COULOMB'S frictional law.

## 8.2 ALTERNATIVE DISCRETIZATION FOR THE TWO-DIMENSIONAL NTS-CONTACT

In Remark 5.3 an alternative formulation was proposed which can be applied for frictional contact. It yields, for stick and slip, a matrix form which has fewer terms

and hence can be implemented in a more efficient way. The new formulation includes frictionless contact as a limiting case, which will be shown at the end of this section.

The main idea will be presented here for COULOMB'S law, which is commonly applied for frictional contact. Instead of using the return mapping procedure applied in the previous section, we base the formulation on the following observation:

1. The slave node lies within the friction cone in the case of stick. Mathematically this is described by the slip function

$$f_s = \|\mathbf{t}_T\| - \mu p_N < 0.$$

2. The slave node moves on the boundary of the friction cone when sliding, which leads to the equation

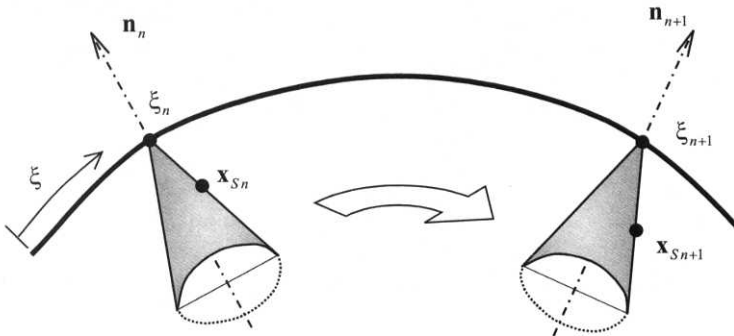
$$f_s = \|\mathbf{t}_T\| - \mu p_N = 0.$$

This feature is depicted in Figure 8.4, where the friction cone is drawn at time steps  $t_n$  and  $t_{n+1}$  for a sliding condition.

**Stick case.** The discretization is based on the NTS discussed in the last section. Following (5.34) the weak contribution for one slave node  $s$  is given for stick by

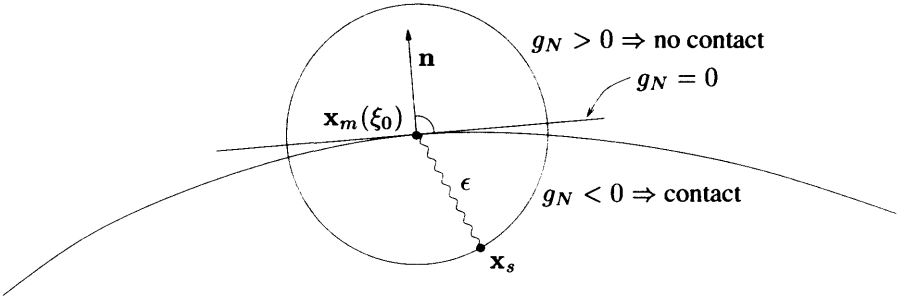
$$C_s^{st} = \mathbf{t} \cdot (\boldsymbol{\eta}_s^2 - \boldsymbol{\eta}_{m0}^1) = \mathbf{t} \cdot (\boldsymbol{\eta}_s^2 - (1 - \xi_0) \boldsymbol{\eta}_1^1 - \xi_0 \boldsymbol{\eta}_2^1), \quad (8.51)$$

where  $\mathbf{t} = \epsilon(\mathbf{x}_s^2 - \mathbf{x}_{m0}^1)$  is the total contact force.  $\mathbf{x}_s^2$  and  $\mathbf{x}_{m0}^1 = \mathbf{x}_m^1(\xi_0)$  are the current positions of the slave and master node, respectively, see Figure 8.5. We let  $\xi_0$  denote the position of the first contact of the slave node with the master surface or the last position after sliding at time  $t_n$ :  $\xi_0 = \xi_n$ . This position is fixed with respect to the surface coordinate, and thus the coordinate  $\xi$  does not change during the deformation as long as stick conditions are present. The master node  $\mathbf{x}_{m0}^1$  is related to the first



**Fig. 8.4** Moving friction cone.





**Fig. 8.5** Alternative form of node-to-segment contact element.

contact in this segment, before any sliding appears:  $\mathbf{x}_{m0}^1 = \mathbf{x}_m^1(\xi_n)$ . This position does not change on the master surface. Hence the force due to stick is given for the time step  $t_{n+1}$  by

$$\mathbf{t}_{n+1} = \epsilon (\mathbf{x}_{s\,n+1}^2 - \mathbf{x}_{m0\,n}^1) = \epsilon (\mathbf{x}_{s\,n+1}^2 - (1 - \xi_n) \mathbf{x}_{1\,n+1}^1 - \xi_n \mathbf{x}_{2\,n+1}^1). \quad (8.52)$$

$\epsilon = \epsilon_N = \epsilon_T$  is the penalty parameter used to enforce the constraint conditions for normal contact and stick. We note that the penalty parameter cannot be viewed as a constitutive parameter in this case, see also Section 4.2.4. This results from the fact that the same parameter is also chosen for the approach of the contact surfaces in the normal direction, and the constitutive laws for normal and tangential contact are usually not the same.

Inserting (8.52) into (8.51) leads to the virtual mechanical work of the NTS-contact element for stick. It can be written in the matrix formulation  $\boldsymbol{\eta}_s^T \mathbf{G}_s^{st}$  with the contact element residual

$$\mathbf{G}_s^{st} = \mathbf{B}_{s\,n} \mathbf{t}_{n+1} \quad \text{with} \quad \mathbf{B}_{s\,n} = \begin{Bmatrix} \mathbf{1} \\ -(1 - \xi_n) \mathbf{1} \\ -\xi_n \mathbf{1} \end{Bmatrix}, \quad (8.53)$$

where  $\boldsymbol{\eta}_s$  is defined in (8.17), and  $\mathbf{1}$  is a  $2 \times 2$  unit matrix. Since in stick conditions only the tangential force depends upon the current position vectors and the fixed coordinate  $\xi_n$ , the linearization of the discrete weak contribution  $\boldsymbol{\eta}_s^T \mathbf{G}_s^c$  of the slave node yields a simple expression for the tangent matrix:

$$\mathbf{K}_{Ts}^c = \epsilon \mathbf{B}_{s\,n} \mathbf{B}_{s\,n}^T. \quad (8.54)$$

Note the  $\mathbf{K}_{Ts}^c$  is symmetric<sup>1</sup>. When comparing this result with formulation (8.44) of the previous section, it is immediately clear that formulation (8.54) is much simpler. It is interesting to note that one can show the equivalence of formulation (8.54) and (8.44) together with (8.36). One only has to choose  $\epsilon_N = \epsilon_T = \epsilon$  in (8.36) and (8.44).

<sup>1</sup>In three dimensions the formulation is identical with (8.53) and (8.54), only  $\mathbf{1}$  is a  $3 \times 3$  matrix.

**Slip case.** In the case of frictional sliding, the COULOMB slip condition  $f_s$ , see (4.25), has to be evaluated. This is performed by computing the normal and tangential components of  $\mathbf{t}$  from the stick (trial) state

$$\mathbf{t}_{n+1}^{tr} = \epsilon \mathbf{g}_{sn} = \epsilon (\mathbf{x}_{sn+1}^2 - \mathbf{x}_m^1), \quad (8.55)$$

leading to

$$\mathbf{t}_{Tn+1}^{tr} = (\mathbf{1} - \mathbf{n}_{n+1}^1 \otimes \mathbf{n}_{n+1}^1) \mathbf{t}_{n+1}^{tr}, \quad (8.56)$$

$$p_{Nn+1} = \mathbf{t}_{n+1}^{tr} \cdot \mathbf{n}_{n+1}^1. \quad (8.57)$$

These values are now inserted in to the slip condition

$$f_s^{tr} = \|\mathbf{t}_{Tn+1}^{tr}\| - \mu p_{Nn+1} \leq 0. \quad (8.58)$$

In case this inequality is fulfilled, the slave node  $s$  remains in stick condition and hence within the slip cone related to  $\xi_n$  in Figure 8.4.

For  $f_s^{tr} > 0$  sliding occurs, and the unknown position  $\mathbf{x}_m(\xi_{n+1})$  has to be computed from the condition that  $\mathbf{x}_s$  now lies on the boundary of the slip cone. For a geometric description see Figure 8.6.

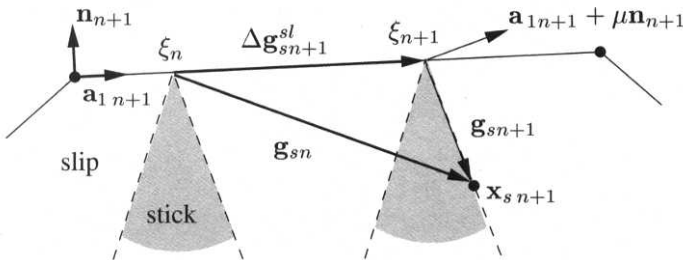
One can see for COULOMB'S law that  $\mathbf{x}_m(\xi_{n+1})$  results from the solution of

$$(\mathbf{x}_{sn+1}^2 - \mathbf{x}_m^1(\xi_{n+1})) \cdot [\zeta \mathbf{a}_{1n+1}^1 + \mu \mathbf{n}_{n+1}^1] = 0, \quad (8.59)$$

with  $\mathbf{x}_m^1(\xi_{n+1}) = (1 - \xi_{n+1}) \mathbf{x}_1^1 + \xi_{n+1} \mathbf{x}_2^1$ .  $\zeta = \text{sign}(\mathbf{t}_{Tn+1}^{tr})$  is the value ( $\zeta = \pm 1$ ) of the signum function depending on the slip direction.  $\zeta$  has to be introduced to acknowledge that the slave node lies on different boundaries of the cone for the two slip directions, which also changes the orthogonality condition (8.59). As the solution point  $\xi_{n+1}$  moves with the friction cone, this approach is called the *moving cone* algorithm.

Since (8.59) is linear in  $\xi$ , it can be solved directly, leading to

$$\xi_{n+1} = \frac{1}{l_{n+1}^2} (\mathbf{x}_{sn+1}^2 - \mathbf{x}_1^1) \cdot (\zeta \mathbf{a}_{1n+1}^1 + \mu \mathbf{n}_{n+1}^1), \quad (8.60)$$



**Fig. 8.6** Geometry in the case of friction for moving-cone NTS-element.

with  $l_{n+1}^2 = \mathbf{a}_{1\,n+1}^1 \cdot \mathbf{a}_{1\,n+1}^1$  and  $\mathbf{a}_{1\,n+1}^1 = \mathbf{x}_{2\,n+1}^1 - \mathbf{x}_{1\,n+1}^1$ . This result can be introduced into (8.55), which yields the total force due to frictional sliding (including the component for the normal pressure):

$$\mathbf{t}_{n+1} = \epsilon \left[ \mathbf{x}_{s\,n+1}^2 - (1 - \xi_{n+1}) \mathbf{x}_{1\,n+1}^1 - \xi_{n+1} \mathbf{x}_{2\,n+1}^1 \right]. \quad (8.61)$$

Inserting this result into (8.53) yields the discrete contribution to the weak formulation for a slave node in sliding contact. The weak form for frictional sliding is now

$$C_s^{sl} = \mathbf{t}_{n+1} \cdot \left[ \eta_s^2 - (1 - \xi_{n+1}) \eta_1^1 - \xi_{n+1} \eta_2^1 \right], \quad (8.62)$$

where the variation now has to be evaluated at the current point  $\xi_{n+1}$ . The matrix form of the contact residual can be stated according to (8.53) as

$$\mathbf{G}_s^{sl} = \mathbf{B}_{s\,n+1} \mathbf{t}_{n+1} \quad \text{with} \quad \mathbf{B}_{s\,n+1} = \left\{ \begin{array}{c} \mathbf{1} \\ -(1 - \xi_{n+1}) \mathbf{1} \\ -\xi_{n+1} \mathbf{1} \end{array} \right\}. \quad (8.63)$$

The differences to the stick case in (8.53) are that one has to use the force (8.61) and evaluate the matrix  $\mathbf{B}_s$  at the current position  $\xi_{n+1}$ .

The linearization of the residual involves linearization of the force (8.61) and linearization of  $\xi_{n+1}$  in the variation

$$\Delta C_s^{sl} = \Delta \mathbf{t}_{n+1} \cdot \left[ \eta_s^2 - (1 - \xi_{n+1}) \eta_1^1 - \xi_{n+1} \eta_2^1 \right] - \mathbf{t}_{n+1} \cdot (\eta_2^1 - \eta_1^1) \Delta \xi_{n+1}. \quad (8.64)$$

Since the surface coordinate depends upon the deformation, one obtains

$$\Delta \mathbf{t}_{n+1} = \epsilon \left[ \Delta \mathbf{u}_{s\,n+1}^2 - (1 - \xi_{n+1}) \Delta \mathbf{u}_{1\,n+1}^1 - \xi_{n+1} \Delta \mathbf{u}_{2\,n+1}^1 - \mathbf{a}_{1\,n+1}^1 \Delta \xi \right]. \quad (8.65)$$

Contrary to the classical formulation, one has to derive the linearization of  $\xi$ , which yields

$$\begin{aligned} \Delta \xi = \frac{1}{l^2} \left[ (\Delta \mathbf{u}_s^2 - (1 - \xi) \Delta \mathbf{u}_1^1 - \xi \Delta \mathbf{u}_2^1) \cdot (\zeta \mathbf{a}_1^1 + \mu \mathbf{n}^1) \right. \\ \left. + (\mathbf{x}_s^2 - (1 - \xi) \mathbf{x}_1^1 - \xi \mathbf{x}_2^1) \cdot (\zeta \Delta \mathbf{a}_1^1 + \mu \Delta \mathbf{n}^1) \right]. \end{aligned} \quad (8.66)$$

In this equation the subscript  $n + 1$  has been omitted for a more compact presentation; note that all variables and vectors have to be evaluated with respect to  $t_{n+1}$ . Furthermore, the definitions  $\Delta \mathbf{a}_1^1 = \Delta \mathbf{u}_2^1 - \Delta \mathbf{u}_1^1$  and  $\Delta \mathbf{n}^1 = \mathbf{e}_3 \times \Delta \mathbf{a}_1^1$  were used. To implement this formulation in a finite element code a matrix representation of equations (8.65) and (8.66) is needed. For this purpose the following matrices are introduced:

$$\mathbf{B}_C = \left\{ \begin{array}{c} \mathbf{0} \\ -\mathbf{1} \\ \mathbf{1} \end{array} \right\}, \quad (8.67)$$

and

$$\begin{aligned} \mathbf{M} &= \zeta \mathbf{1} + \mu \mathbf{E}_N \\ \mathbf{m}^1 &= \zeta \mathbf{a}_1^1 + \mu \mathbf{n}^1 \\ \mathbf{g}_s &= \mathbf{x}_s - (1 - \xi) \mathbf{x}_1 - \xi \mathbf{x}_2 \end{aligned} \quad (8.68)$$

with

$$\mathbf{E}_N = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (8.69)$$

where the matrix  $\mathbf{E}_N$  represents the cross product needed to compute the normal  $\mathbf{n}$ .

Introducing the last definitions in (8.65) finally yields a matrix expression for the incremental stress due to sliding:

$$\Delta \mathbf{t}_{n+1} = \epsilon \left[ \mathbf{B}_s^T - \frac{1}{l^2} \left\{ [\mathbf{a}_1^1 \otimes \mathbf{m}^1] \mathbf{B}_s^T - [\mathbf{a}_1^1 \otimes \mathbf{g}_s] \mathbf{M} \mathbf{B}_C^T \right\} \right]_{n+1} \Delta \mathbf{u}, \quad (8.70)$$

where  $\Delta \mathbf{u}$  is defined in the same manner as the variation  $\boldsymbol{\eta}$ . Note that all matrices have to be evaluated at time  $t_{n+1}$ . This result can now be introduced into the linearization of the weak form (8.64), which yields the non-symmetric matrix

$$\begin{aligned} \mathbf{K}_{Tsn+1}^{sl} = & \epsilon \left\{ \mathbf{B}_s \left[ \mathbf{1} - \frac{1}{l^2} \mathbf{a}_1^1 \mathbf{m}^{1T} \right] \mathbf{B}_s^T \right. \\ & \left. - \frac{1}{l^2} \left( \mathbf{B}_s [\mathbf{a}_1^1 \mathbf{g}_s^T \mathbf{M}] \mathbf{B}_C^T + \mathbf{B}_C [\mathbf{g}_s \mathbf{m}^{1T}] \mathbf{B}_s^T + \mathbf{B}_C [\mathbf{g}_s \mathbf{g}_s^T \mathbf{M}] \mathbf{B}_C^T \right) \right\}_{n+1}. \end{aligned} \quad (8.71)$$

Also, the matrix in (8.71) is far less complex than the matrix used in the standard NTS-element (see last section). All matrices  $[\cdot]$  in (8.71) are  $2 \times 2$  matrices while the  $\mathbf{B}$ -matrices  $\mathbf{B}_s$  and  $\mathbf{B}_C$  have size  $2 \times 6$ .

**Frictionless case.** By setting the friction coefficient in (8.59) to zero, from (8.60) as the limit case with  $\zeta = 1$ , one obtains

$$\xi_{n+1}^f = \frac{1}{l_{n+1}^2} (\mathbf{x}_{sn+1}^2 - \mathbf{x}_{1n+1}^1) \cdot \mathbf{a}_{1n+1}^1, \quad (8.72)$$

which is equivalent to (8.3). Inserting this result into (8.61) leads to the total force, which has only components in the normal direction,

$$\mathbf{t}_{n+1} = \epsilon [\mathbf{x}_{sn+1}^2 - (1 - \xi_{n+1}^f) \mathbf{x}_{1n+1}^1 - \xi_{n+1}^f \mathbf{x}_{2n+1}^1], \quad (8.73)$$

and to the weak form for frictionless sliding for the node  $s$

$$C_s^f = \mathbf{t}_{n+1} \cdot \left[ \boldsymbol{\eta}_s^2 - (1 - \xi_{n+1}^f) \boldsymbol{\eta}_1^1 - \xi_{n+1}^f \boldsymbol{\eta}_2^1 \right], \quad (8.74)$$

where the variation has to be evaluated at the current point  $\xi_{n+1}^f$ . The matrix form of the contact residual

$$\mathbf{G}_s^{sl} = \mathbf{B}_{sn+1} \mathbf{t}_{n+1} \quad \text{with} \quad \mathbf{B}_{sn+1} = \begin{Bmatrix} \mathbf{1} \\ -(1 - \xi_{n+1}^f) \mathbf{1} \\ -\xi_{n+1}^f \mathbf{1} \end{Bmatrix} \quad (8.75)$$

is equivalent to (8.63), hence the same formulation can be used; only the projection stemming from (8.59) is different.

The linearization of the residual follows from (8.74) with (8.73) as

$$\Delta C_s^{sl} = \Delta \mathbf{t}_{n+1} \cdot \left[ \boldsymbol{\eta}_s^2 - (1 - \xi_{n+1}^f) \boldsymbol{\eta}_1^1 - \xi_{n+1}^f \boldsymbol{\eta}_2^1 \right] - \mathbf{t}_{n+1} \cdot (\boldsymbol{\eta}_2^1 - \boldsymbol{\eta}_1^1) \Delta \xi_{n+1}^f. \quad (8.76)$$

Again, we can use the results (8.65) and (8.66) for  $\zeta = 1$  and  $\mu = 0$  to obtain the final tangent as

$$\begin{aligned} \mathbf{K}_{Ts}^f &= \epsilon \mathbf{B}_s \left[ \mathbf{I} - \frac{1}{l^2} \mathbf{a}_1^1 \mathbf{a}_1^{1T} \right] \mathbf{B}_s^T \\ &\quad - \frac{1}{l^2} \left[ \mathbf{B}_C \mathbf{g}_s \mathbf{g}_s^T \mathbf{B}_C^T + \mathbf{B}_s \mathbf{a}_1^1 \mathbf{g}_s^T \mathbf{B}_C^T + \mathbf{B}_C \mathbf{g}_s \mathbf{a}_1^{1T} \mathbf{B}_s^T \right]. \end{aligned} \quad (8.77)$$

All quantities have to be evaluated at time  $t_{n+1}$ . The matrix in (8.77) is not more complex than the tangent (8.36) which is used in the standard NTS-element (see the last section). Note, that the first linear term in (8.77) only depends on the normal direction, since in the two-dimensional case one can write for the unit tensor  $\mathbf{I} = \frac{1}{l^2} (\mathbf{a}_1^1 \otimes \mathbf{a}_1^1 + \mathbf{n}^1 \otimes \mathbf{n}^1)$ , and hence  $(\mathbf{I} - \frac{1}{l^2} \mathbf{a}_1^1 \mathbf{a}_1^{1T}) = \frac{1}{l^2} \mathbf{n}^1 \mathbf{n}^{1T}$ . The term  $l$  appears here in the denominator since  $\mathbf{a}_1^1$  and  $\mathbf{n}^1$  are not unit vectors. By using the tangent  $\mathbf{K}_{Ts}^f$ , no switch to the standard formulation has to be made in the frictionless case when the moving cone approach is applied. Equation (8.77) can also be obtained directly from (8.71) with  $\zeta = 1$  and  $\mu = 0$ . Finally, we note that the tangent in (8.77) is symmetric.

### 8.3 THREE-DIMENSIONAL CONTACT DISCRETIZATION

Discretizations for the contact of three-dimensional objects were first developed for explicit finite element codes like DYNA3D. The derivation of implicit formulations is more recent, especially the consistent linearization for use within NEWTON'S method. The matrix form for the frictionless three-dimensional case of node-to-surface discretizations was developed in Parisch (1989). A formulation for the case including friction was presented in Peric and Owen (1992), which relied on an approximation of the tangent matrix. A fully consistent linearization for a continuum-based approach to contact problems was derived in Laursen and Simo (1993b) and also in Parisch and Lübbing (1997). The formulation for friction was extended for large slip in Agelet de Saracibar (1997).

We should like to state the discrete form here based on the continuous formulation given in Section 5.3. The matrices are developed for three and four node master segments. In this section we denote by a superscript <sup>1</sup> the quantities belonging to the master surface of body  $B^1$ . Furthermore, quantities related to the slave node which comes into contact with the master surface are denoted by the superscript <sup>2</sup>, body  $B^2$ .

During finite deformations within the contact surface, a node of the contacting body (slave-node) eventually slides over several elements of the other body (master surface). We thus have to consider three different possible geometrical situations:

1. Slave-node contacts the surface of the master element (NTS).
2. Slave-node contacts an edge of the master element (NTE).
3. Slave-node contacts a vertex (node) of the master element (NTV).

These situations are depicted in Figure 8.7. For (NTS) situations we develop the associated contact residuals and tangent stiffness matrices for frictional contact. Matrices for the (NTE) and (NTV) cases are only derived for frictionless contact.

### 8.3.1 Node-to-surface contact element

Many discretizations of solids are based on hexahedral elements. Within such a finite element model the master segment is given by a four-node isoparametric surface element which comes into contact with one slave node  $\mathbf{x}^2$ , see Figure 8.8. Thus, we can also denote this contact discretization as a five-node contact element. The interpolation of the master surface yields

$$\mathbf{x}^1 = \sum_I N_I(\xi^1, \xi^2) \mathbf{x}_I^1, \quad (8.78)$$

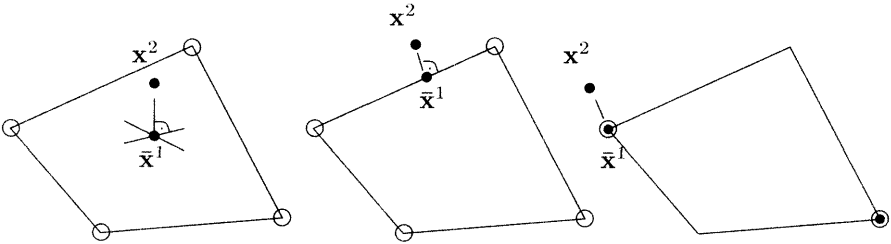
with the standard interpolation functions, see also (6.24),

$$N_I = \frac{1}{4}(1 - \xi_I^1 \xi^1)(1 - \xi_I^2 \xi^2). \quad (8.79)$$

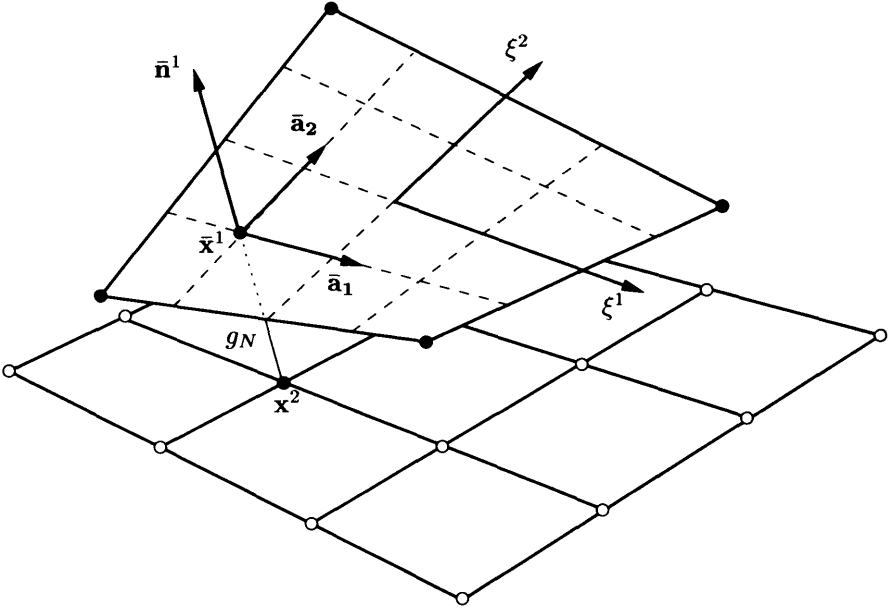
The gap function is given by equation (3.6). In this equation we have to compute the convective coordinates  $\bar{\xi}^1$  and  $\bar{\xi}^2$  to obtain the minimum distance defining the gap  $g_N$

$$g_N = [\mathbf{x}^2 - \mathbf{x}^1(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{n}}^1. \quad (8.80)$$

Since in general the deformed master surface is not flat (it is a hyperbolic surface in the case of the four-node isoparametric interpolation), the coordinates  $\bar{\xi} = (\bar{\xi}^1, \bar{\xi}^2)$  cannot be computed from a closed form expression. To find the minimum distance point related to the slave node, an iterative process has to be applied. To do that, the



**Fig. 8.7** NTS, NTE and NTV contact.



**Fig. 8.8** Geometry of the five-node contact element.

following nonlinear system of equations for  $\xi^1$  and  $\xi^2$  will be solved locally for each slave node  $\mathbf{x}^2$ :

$$[\mathbf{x}^2 - \mathbf{x}^1(\boldsymbol{\xi})] \cdot \mathbf{x}_{,\alpha}^1(\boldsymbol{\xi}) = \left[ \mathbf{x}^2 - \sum_{I=1}^4 N_I(\boldsymbol{\xi}) \mathbf{x}_I^1 \right] \cdot \sum_{K=1}^4 N_{K,\alpha}(\boldsymbol{\xi}) \mathbf{x}_K^1 = 0. \quad (8.81)$$

Using as a starting value the solution for a flat surface, one can apply NEWTON'S method. The associated linearization leads to a linear system of equations for an iteration step  $i$  for  $\Delta \xi_{i+1}^\beta = \xi_{i+1}^\beta - \xi_i^\beta$ :

$$\begin{aligned} \sum_{\beta=1}^2 \left[ \sum_{I=1}^4 \sum_{J=1}^4 N_{I,\alpha} N_{J,\beta} \mathbf{x}_I^1 \cdot \mathbf{x}_J^1 - \sum_{K=1}^4 N_{K,\alpha\beta} \mathbf{x}_K^1 \cdot \left( \mathbf{x}^2 - \sum_{L=1}^4 N_L \mathbf{x}_L^1 \right) \right] \Delta \xi_{i+1}^\beta \\ = \left( \mathbf{x}^2 - \sum_{M=1}^4 N_M \mathbf{x}_M^1 \right) \cdot \sum_{N=1}^4 N_{N,\alpha} \mathbf{x}_N^1, \end{aligned} \quad (8.82)$$

where all shape functions  $N_I$  and their derivatives have to be evaluated at  $\xi_i^\alpha$ , known from the previous step. The solution of (8.81) yields the closest point  $\bar{\xi}$  of the slave node to the master surface. With these coordinates the gap function (8.80) has to be evaluated. In the case of contact ( $g_N \leq 0$ ) the slave node contributes to the weak form.

Again we have to distinguish between frictionless contact and slip and stick in the tangential direction. We start here with the stick part, which is derived according to

Section 8.2, since this leads to the most simple and efficient formulation. Here the total gap function  $\mathbf{g}_s$ ,

$$\mathbf{g}_s = \mathbf{x}^2 - \mathbf{x}^1(\xi_0), \quad (8.83)$$

is used which is evaluated at the fixed stick point  $\xi_0$ . The coordinates  $\xi_0$  are obtained from (8.81) when node  $s$  contacts the master surface for the first time. Otherwise,  $\xi_0$  follows from the previous time step  $\xi_0 = \xi_n$ . The variation of  $\mathbf{g}_s$  is given by

$$\delta \mathbf{g}_s = \boldsymbol{\eta}^2 - \boldsymbol{\eta}^1(\xi_0), \quad (8.84)$$

and we can compute the weak form

$$C_c^{st} = \int_{\Gamma_c} \mathbf{t} \cdot \delta \mathbf{g} dA = \sum_{s=1}^{n_c} \mathbf{t} \cdot \delta \mathbf{g}_s A_s = \sum_{s=1}^{n_c} A_s \boldsymbol{\eta}_s^T \mathbf{G}_s^{st} \quad (8.85)$$

with

$$\boldsymbol{\eta}_s^T = \langle \boldsymbol{\eta}^2, \boldsymbol{\eta}_1^1, \boldsymbol{\eta}_2^1, \boldsymbol{\eta}_3^1, \boldsymbol{\eta}_4^1 \rangle. \quad (8.86)$$

The vector  $\boldsymbol{\eta}_s$  denotes the variations of the displacements related to the contact nodes of the  $s$ th contact element, where  $s$  stands for the slave node  $\mathbf{x}_s$  and  $m$  for the four nodes describing the master element  $\mathbf{x}_1^1, \dots, \mathbf{x}_4^1$ . Furthermore, the vector  $\mathbf{G}_s^{st}$  can be expressed in matrix form as

$$\mathbf{G}_s^{st} = \mathbf{B}_s \mathbf{t} \quad \text{with} \quad \mathbf{B}_s = \begin{Bmatrix} \mathbf{1} \\ -N_1(\xi_0) \mathbf{1} \\ -N_2(\xi_0) \mathbf{1} \\ -N_3(\xi_0) \mathbf{1} \\ -N_4(\xi_0) \mathbf{1} \end{Bmatrix}, \quad (8.87)$$

where  $\mathbf{1}$  is a  $3 \times 3$  unit matrix. Again the linearization is simple, since  $\xi_0$  is fixed and hence does not depend upon the displacement field. Analogous to Section 8.2, the linearization is given for the penalty formulation as

$$\Delta C_c^{st} = \sum_{s=1}^{n_c} A_s \boldsymbol{\eta}_s^T \Delta \mathbf{G}_s^{st} = \sum_{s=1}^{n_c} A_s \boldsymbol{\eta}_s^T \mathbf{K}_s^{st} \Delta \mathbf{u}_s \quad \text{with} \quad \mathbf{K}_s^{st} = \epsilon \mathbf{B}_s \mathbf{B}_s^T. \quad (8.88)$$

In the case of sliding we could either apply the alternative form, developed in Section 8.2, or the classical approach which split the contributions in the contact interface into normal and tangential parts. For more generality concerning the frictional laws, we use the second possibility. With that we start from the weak form

$$C_c = \int_{\Gamma_c} (t_N \delta g_N + \mathbf{t}_T \cdot \delta \mathbf{g}_T^{sl}) dA \approx \sum_{s=1}^{n_c} [t_{Ns}(\bar{\xi}) \delta g_{Ns}(\bar{\xi}) + \mathbf{t}_{Ts} \cdot \delta \mathbf{g}_{Ts}^{sl}(\bar{\xi})] A_s, \quad (8.89)$$

see also (5.28).  $A_s$  is the surface associated with the slave nodes. It is determined by a projection procedure from the current master surface  $\Gamma_c$ . In index notation, for the



tangential part we obtain  $\mathbf{t}_{Ts} \cdot \delta \mathbf{g}_{Ts}^{sl}(\bar{\xi}) = t_{Ts\alpha} \delta \bar{\xi}^\alpha$ . The variations  $\delta g_N(\bar{\xi})$  and  $\delta \bar{\xi}^\alpha$  can be found in (3.29) and (3.30) together with (3.22).

Based on that, we obtain the discretization for the five-node contact element:

$$\mathbf{C}_c^{sl} = \sum_{s=1}^{n_s} A_s \boldsymbol{\eta}_s^T \mathbf{R}_s. \quad (8.90)$$

The vector  $\mathbf{R}_s$  depends only upon the contact contribution. It is given by

$$\mathbf{R}_s = t_{Ns} \mathbf{N}_s - t_{Ts\alpha} \mathbf{D}_s^\alpha, \quad (8.91)$$

with the vectors

$$\mathbf{N}_s = \begin{Bmatrix} \bar{\mathbf{n}}_c^1 \\ -N_1 \bar{\mathbf{n}}^1 \\ -N_2 \bar{\mathbf{n}}^1 \\ -N_3 \bar{\mathbf{n}}^1 \\ -N_4 \bar{\mathbf{n}}^1 \end{Bmatrix}, \quad (8.92)$$

and

$$\mathbf{D}_s^\alpha = H^{\alpha\beta} [\mathbf{T}_\beta - g_{Ns} \mathbf{N}_\beta], \quad (8.93)$$

where

$$\mathbf{T}_\beta = \begin{Bmatrix} \bar{\mathbf{a}}_\alpha \\ -N_1 \bar{\mathbf{a}}_\beta \\ -N_2 \bar{\mathbf{a}}_\beta \\ -N_3 \bar{\mathbf{a}}_\beta \\ -N_4 \bar{\mathbf{a}}_\beta \end{Bmatrix}, \quad \mathbf{N}_\beta = \begin{Bmatrix} \mathbf{0} \\ -N_{1,\beta} \bar{\mathbf{n}}^1 \\ -N_{2,\beta} \bar{\mathbf{n}}^1 \\ -N_{3,\beta} \bar{\mathbf{n}}^1 \\ -N_{4,\beta} \bar{\mathbf{n}}^1 \end{Bmatrix}, \quad (8.94)$$

and  $H^{\alpha\beta} = [H_{\alpha\beta}]^{-1}$  with  $H_{\alpha\beta} = \bar{a}_{\alpha\beta} + g_{Ns} \bar{b}_{\alpha\beta}$ . To derive this result one has to use  $\bar{\xi} = (\bar{\xi}^1, \bar{\xi}^2)$  from (8.82) and  $\delta \bar{\xi}^\alpha$  from (3.19), see also (3.30).

For the contact pressure  $t_{Ns}$  in (8.91) we can either use the LAGRANGE multiplier  $\lambda_N$  or the pressure related to the penalty method,

$$t_{Ns} = \epsilon_N g_{Ns}. \quad (8.95)$$

The frictional stresses  $t_{T\alpha}$  follow from (9.113) and (9.118):

$$t_{T\alpha} = \begin{cases} t_{T\alpha}^{tr} & \text{for } f_s^{tr} \leq 0 \\ \mu |t_N| \frac{t_{T\alpha}^{tr}}{\|t_T^{tr}\|} & \text{for } f_s^{tr} > 0 \end{cases}, \quad (8.96)$$

with the components of the trial tangential stresses  $t_{T\alpha}^{tr}$  within time step  $n+1$

$$t_{T\alpha n+1}^{tr} = \mathbf{t}_{Tn} \cdot \mathbf{a}_{\alpha n+1} + \epsilon_T \bar{a}_{\alpha\beta} \Delta \xi_{n+1}^\beta \quad \text{with} \quad \Delta \xi_{n+1}^\beta = \xi_{n+1}^\beta - \xi_n^\beta. \quad (8.97)$$

With this the residual vector is determined. For a finite element solution where the NEWTON-RAPHSON method is used, we also need the linearizations of the contact contribution. These are derived in Section 3.3, 5.5 for the relative deformations in the contact interface. These results can be applied to linearize the contact stresses, as shown in Sections 8.1 and 8.2.

Based on these results, we can state the discretization of the tangent matrix for the contact contribution in the case of slip as follows:

$$\Delta C_c^{sl}(\mathbf{x}, \boldsymbol{\eta}) \cdot \Delta \mathbf{u} \approx \sum_{s=1}^{n_c} A_s (\boldsymbol{\eta}_s)^T \mathbf{K}_s^{sl} \Delta \mathbf{u}_s. \quad (8.98)$$

The terms  $\Delta \mathbf{u}_s$  denote the increments of the contact node displacements related to slave node  $s$

$$\Delta \mathbf{u}_s^T = \langle \Delta \mathbf{u}^2, \Delta \mathbf{u}_1^1, \Delta \mathbf{u}_2^1, \Delta \mathbf{u}_3^1, \Delta \mathbf{u}_4^1 \rangle. \quad (8.99)$$

Now we have to define some vectors and matrices to state the matrix form of the contact tangent:

$$\mathbf{T}_{\alpha\beta} = \begin{Bmatrix} \mathbf{0} \\ -N_{1,\beta} \bar{\mathbf{a}}_\alpha \\ -N_{2,\beta} \bar{\mathbf{a}}_\alpha \\ -N_{3,\beta} \bar{\mathbf{a}}_\alpha \\ -N_{4,\beta} \bar{\mathbf{a}}_\alpha \end{Bmatrix}, \hat{\mathbf{T}}_{\alpha\beta} = \begin{Bmatrix} \bar{\mathbf{a}}_{\alpha,\beta} \\ -N_1 \bar{\mathbf{a}}_{\alpha,\beta} \\ -N_2 \bar{\mathbf{a}}_{\alpha,\beta} \\ -N_3 \bar{\mathbf{a}}_{\alpha,\beta} \\ -N_4 \bar{\mathbf{a}}_{\alpha,\beta} \end{Bmatrix}, \mathbf{N}_{\alpha\beta} = \begin{Bmatrix} \mathbf{0} \\ -N_{1,\alpha\beta} \bar{\mathbf{n}}^1 \\ -N_{2,\alpha\beta} \bar{\mathbf{n}}^1 \\ -N_{3,\alpha\beta} \bar{\mathbf{n}}^1 \\ -N_{4,\alpha\beta} \bar{\mathbf{n}}^1 \end{Bmatrix}, \quad (8.100)$$

$$\mathbf{E} = \begin{Bmatrix} \mathbf{1} \\ -N_1 \mathbf{1} \\ -N_2 \mathbf{1} \\ -N_3 \mathbf{1} \\ -N_4 \mathbf{1} \end{Bmatrix}, \hat{\mathbf{E}}_\alpha = \begin{Bmatrix} \hat{\mathbf{0}} \\ -N_{1,\alpha} \mathbf{1} \\ -N_{2,\alpha} \mathbf{1} \\ -N_{3,\alpha} \mathbf{1} \\ -N_{4,\alpha} \mathbf{1} \end{Bmatrix}. \quad (8.101)$$

With this the contact tangent matrix  $\mathbf{K}_s^{sl}$  for the three-dimensional node-to-segment (slave to master) element yields, according to (8.98),

$$\begin{aligned} \mathbf{K}_s^{sl} &= \underbrace{\epsilon_N \mathbf{N}_s \mathbf{N}_s^T + \epsilon_N g_{Ns} \Delta(\delta g_{Ns})}_{\mathbf{K}_{Ns}} - \underbrace{\Delta t_{T\alpha} \delta \xi^\alpha}_{\mathbf{K}_{Ts1}^{sl}} - \underbrace{t_{T\alpha} \Delta(\delta \xi^\alpha)}_{\mathbf{K}_{Ts2}^{sl}}. \\ &= \mathbf{K}_{Ns} - \mathbf{K}_{Ts1}^{sl} - \mathbf{K}_{Ts2}^{sl}. \end{aligned} \quad (8.102)$$

The part due to normal contact can be formulated as

$$\mathbf{K}_{Ns} = \epsilon_N \left\{ \mathbf{N}_s \mathbf{N}_s^T + g_N \left[ \mathbf{N}_\alpha \mathbf{D}_s^{\alpha T} + a^{\beta\alpha} \mathbf{T}_\alpha \left( \mathbf{N}_\beta^T - \mathbf{D}_s^{\gamma T} (\bar{\mathbf{n}}^1 \cdot \bar{\mathbf{a}}_{\beta,\gamma}) \right) \right] \right\}. \quad (8.103)$$

The tangent matrix due to frictional contact for slip is described by two parts:  $(\mathbf{K}_{Ts1}^{sl})$  and  $(\mathbf{K}_{Ts2}^{sl})$ . For the first matrix we obtain

$$\begin{aligned} \mathbf{K}_{Ts1}^{sl} &= t_{T\alpha} H^{\alpha\eta} \left[ (\mathbf{T}_{\eta\beta} + \mathbf{T}_{\beta\eta} + \hat{\mathbf{T}}_{\eta\beta}) \mathbf{D}^{\beta T} + \mathbf{D}^\beta (\mathbf{T}_{\eta\beta}^T + \mathbf{T}_{\beta\eta}^T + \hat{\mathbf{T}}_{\eta\beta}^T) \right. \\ &\quad - (\bar{\mathbf{a}}_\eta \cdot \bar{\mathbf{a}}_{\beta,\gamma} + \bar{\mathbf{a}}_\beta \cdot \bar{\mathbf{a}}_{\eta,\gamma} + \bar{\mathbf{a}}_{\eta,\beta} \cdot \bar{\mathbf{a}}_\gamma + g_N \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{a}}_{\eta,\beta\gamma}) \mathbf{D}^\beta \mathbf{D}^{\gamma T} \\ &\quad \left. - \mathbf{E} \mathbf{E}_\eta^T - \mathbf{E}_{,\eta} \mathbf{E}^T + g_N \hat{\mathbf{N}}_{\eta\beta} \mathbf{D}^{\beta T} + g_N \mathbf{D}^\beta \hat{\mathbf{N}}_{\eta\beta}^T \right], \end{aligned} \quad (8.104)$$

and for the second part we have

$$\mathbf{K}_{Ts2}^{sl} = \mu \mathbf{D}^\alpha \left\{ -\epsilon_N \frac{t_{T\alpha}^{tr}}{\|\mathbf{t}_T^{tr}\|} \mathbf{N}^T + \frac{|\mathbf{t}_N|}{\|\mathbf{t}_T^{tr}\|} \left[ -\epsilon_T a_{\beta\alpha} \mathbf{D}^{\beta T} \right. \right.$$

$$\begin{aligned}
& + \epsilon_T \Delta \xi_{n+1}^\beta \left( \mathbf{T}_{\beta\alpha}^T + \mathbf{T}_{\alpha\beta}^T - (\mathbf{a}_{\alpha,\vartheta} \cdot \mathbf{a}_\beta + \mathbf{a}_\alpha \cdot \mathbf{a}_{\beta,\vartheta}) \mathbf{D}^{\vartheta T} \right) \Big] \\
& + \frac{|t_N|}{\|\mathbf{t}_T^{\text{tr}}\|^3} t_{T\alpha}^{\text{tr}} t_T^{\text{tr}\beta} \left[ \mathbf{t}_T^{\text{tr}} \cdot \Delta \mathbf{u}_{,\beta} - \epsilon_T \Delta \xi_{n+1}^\gamma \left( \mathbf{T}_{\beta\gamma}^T + \mathbf{T}_{\gamma\beta}^T \right) \right. \\
& \left. + \mathbf{D}^{\vartheta T} \left( \mathbf{t}_T^{\text{tr}} \cdot \mathbf{a}_{\beta,\vartheta} + \epsilon_T a_{\vartheta\beta} + \epsilon_T \Delta \xi_{n+1}^\gamma (\mathbf{a}_{\gamma,\vartheta} \cdot \mathbf{a}_\beta + \mathbf{a}_\gamma \cdot \mathbf{a}_{\beta,\vartheta}) \right) \right] \Big\}.
\end{aligned} \quad (8.105)$$

With this the matrix form of the five-node contact element is given for frictional contact. The formulation is quite complicated, which stems from the necessity to use convective coordinates for the description of contact.

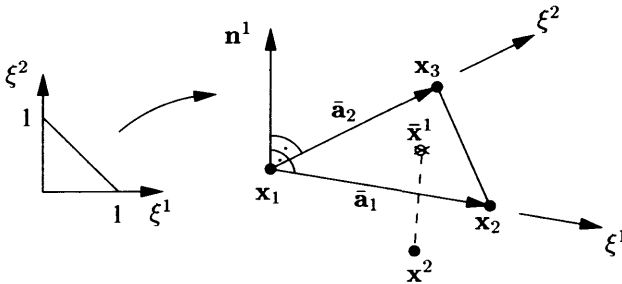
#### 8.4 THREE-NODE MASTER SEGMENT FOR FRICTIONLESS CONTACT

The discretization of the master surface by three-node elements yields simpler equations, since the surface is always flat. The geometric description of such an element is depicted in Figure 8.9. The tangent vectors can be defined as

$$\begin{aligned}
\bar{\mathbf{a}}_1 &= \mathbf{x}_2^1 - \mathbf{x}_1^1 & \text{and} & \quad \mathbf{a}_1 = \frac{\mathbf{x}_2^1 - \mathbf{x}_1^1}{\|\mathbf{x}_2^1 - \mathbf{x}_1^1\|} \\
\bar{\mathbf{a}}_2 &= \mathbf{x}_3^1 - \mathbf{x}_1^1 & \text{and} & \quad \mathbf{a}_2 = \frac{\mathbf{x}_3^1 - \mathbf{x}_1^1}{\|\mathbf{x}_3^1 - \mathbf{x}_1^1\|}.
\end{aligned} \quad (8.106)$$

$\mathbf{a}_\alpha$  are unit vectors. Note that the tangent vectors do not depend upon  $\xi$  when defined in this way. This also means that a closed form expression can be derived for the closest point projection (3.2) when using (8.81):

$$[\mathbf{x}^2 - \mathbf{x}^1(\bar{\xi})] \cdot \bar{\mathbf{a}}_\alpha = \left[ \mathbf{x}^2 - \sum_{I=1}^3 N_I(\bar{\xi}) \mathbf{x}_I^1 \right] \cdot \bar{\mathbf{a}}_\alpha = 0. \quad (8.107)$$



**Fig. 8.9** NTS-contact element, definition of contact normal.

These two equations can be solved directly for  $\bar{\xi}^1$  and  $\bar{\xi}^2$  when the isoparametric shape functions (6.22) are applied for the interpolation. This yields the equation system

$$\begin{bmatrix} \bar{\mathbf{a}}_1 \cdot \bar{\mathbf{a}}_1 & \bar{\mathbf{a}}_1 \cdot \bar{\mathbf{a}}_2 \\ \bar{\mathbf{a}}_2 \cdot \bar{\mathbf{a}}_1 & \bar{\mathbf{a}}_2 \cdot \bar{\mathbf{a}}_2 \end{bmatrix} \begin{Bmatrix} \bar{\xi}^1 \\ \bar{\xi}^2 \end{Bmatrix} = \begin{Bmatrix} (\mathbf{x}^2 - \mathbf{x}_1^1) \cdot \bar{\mathbf{a}}_1 \\ (\mathbf{x}^2 - \mathbf{x}_1^1) \cdot \bar{\mathbf{a}}_2 \end{Bmatrix}. \quad (8.108)$$

We should like to state here the formulation related to the penalty method. A three-node frictionless contact element using the LAGRANGE multiplier method can be developed according to the general formulation in Section 5.3.1 and the discretization techniques in 7.2 and 7.3. Let us recall that for frictionless contact the penalty method is given by (5.33) as

$$C_c = \int_{\Gamma_c} \epsilon_N g_N \delta g_N dA, \quad (8.109)$$

with the local gap function (3.7)  $g_{Ns} = [\mathbf{x}^2 - \bar{\mathbf{x}}^1(\boldsymbol{\xi})] \cdot \bar{\mathbf{n}}^1$  and its variation, see (3.29)

$$\delta g_{Ns} = [\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1(\boldsymbol{\xi})] \cdot \bar{\mathbf{n}}^1. \quad (8.110)$$

The linearization of the contact contribution yields (see Section 5.5)

$$\Delta C_c = \epsilon_N \int_{\Gamma_c} (\Delta g_{Ns} \delta g_{Ns} + g_{Ns} \Delta \delta g_{Ns}) dA \quad (8.111)$$

with

$$\Delta(\delta g_{Ns}) = -\bar{\boldsymbol{\eta}}_{,\alpha}^1 \cdot \Delta \bar{\xi}^\alpha + [\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1(\boldsymbol{\xi})] \cdot \Delta \bar{\mathbf{n}}^1, \quad (8.112)$$

where the differentiation of the variation  $\bar{\boldsymbol{\eta}}^1$  with respect to  $\xi^\alpha$  yields, for the interpolation (6.22),

$$\bar{\boldsymbol{\eta}}_{,1}^1 = \boldsymbol{\eta}_2^1 - \boldsymbol{\eta}_1^1 \quad \text{and} \quad \bar{\boldsymbol{\eta}}_{,2}^1 = \boldsymbol{\eta}_3^1 - \boldsymbol{\eta}_1^1.$$

The normal of the master surface is perpendicular to the tangent vectors, hence we derive

$$\Delta \bar{\mathbf{n}}^1 = \Delta \frac{\bar{\mathbf{a}}_1 \times \bar{\mathbf{a}}_2}{\|\bar{\mathbf{a}}_1 \times \bar{\mathbf{a}}_2\|} = -\bar{\mathbf{a}}_\alpha \bar{a}^{\alpha\beta} (\Delta \bar{\mathbf{a}}_\beta \cdot \bar{\mathbf{n}}^1) = -\bar{a}^{\alpha\beta} [\mathbf{a}_\alpha \otimes \bar{\mathbf{n}}^1] \Delta \bar{\mathbf{a}}_\beta. \quad (8.113)$$

Furthermore, the linearization of the tangent vectors  $\bar{\mathbf{a}}_\alpha$  yields

$$\Delta \bar{\mathbf{a}}_1 = \Delta \mathbf{u}_2 - \Delta \mathbf{u}_1 \quad \text{and} \quad \Delta \bar{\mathbf{a}}_2 = \Delta \mathbf{u}_3 - \Delta \mathbf{u}_1.$$

Note that  $\bar{\mathbf{a}}_{\alpha,\beta} \cdot \bar{\mathbf{n}}_c^1$  which represents the curvature is zero, since the surface of the three-node element is flat. Using these results, the linearization of the gap variation follows:

$$\Delta(\delta g_{Ns}) = -\Delta \bar{\xi}^\alpha \bar{\boldsymbol{\eta}}_{,\alpha}^1 \cdot \bar{\mathbf{n}}^1 - \bar{a}^{\alpha\beta} (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot [\mathbf{a}_\alpha \otimes \bar{\mathbf{n}}^1] \Delta \bar{\mathbf{a}}_\beta \quad (8.114)$$

for the three-node master surface discretization. The linearization of  $\bar{\xi}^\alpha$  can be computed from (3.19). For the flat three-node master surface we have  $\bar{b}_{\alpha\beta} = 0$ , and for the linearization of  $\bar{\xi}^\alpha$  we obtain

$$\Delta \bar{\xi}^\alpha = \bar{a}^{\alpha\beta} [(\Delta \mathbf{u}^2 - \Delta \bar{\mathbf{u}}^1) \cdot \bar{\mathbf{a}}_\beta + g_{Ns} \bar{\mathbf{n}}^1 \cdot \Delta \bar{\mathbf{a}}_\beta]. \quad (8.115)$$

Combining (8.114) and (8.115) finally yields the linearization of the contact contribution:

$$\begin{aligned} \Delta C_c = & \int_{\Gamma_c} \epsilon_N \{ (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot [\bar{\mathbf{n}}^1 \otimes \bar{\mathbf{n}}^1] (\Delta \mathbf{u}^2 - \Delta \bar{\mathbf{u}}^1) \\ & - g_{Ns} [\bar{\boldsymbol{\eta}}_{,\alpha}^1 \cdot (\bar{\mathbf{n}}^1 \otimes \bar{\mathbf{a}}^\alpha) \cdot (\Delta \mathbf{u}^2 - \Delta \bar{\mathbf{u}}^1) + (\boldsymbol{\eta}^2 - \bar{\boldsymbol{\eta}}^1) \cdot (\bar{\mathbf{a}}^\alpha \otimes \bar{\mathbf{n}}^1) \cdot \bar{\mathbf{u}}_{,\alpha}^1] \\ & - g_{Ns}^2 \bar{a}^{\alpha\beta} \bar{\boldsymbol{\eta}}_{,\alpha}^1 \cdot (\bar{\mathbf{n}}^1 \otimes \bar{\mathbf{n}}^1) \cdot \bar{\mathbf{u}}_{,\beta}^1 \} dA. \end{aligned} \quad (8.116)$$

Here the contravariant base vectors  $\bar{\mathbf{a}}_\beta \bar{a}^{\alpha\beta} = \bar{\mathbf{a}}^\alpha$  were introduced to shorten notation. Note that (8.116) is symmetric in  $\Delta \mathbf{u}$  and  $\boldsymbol{\eta}$ .

The residual vector and the tangent stiffness matrix for the three-node frictionless contact element can be stated using the following matrices:

$$\mathbf{N}_s = \begin{Bmatrix} \bar{\mathbf{n}}^1 \\ -N_1 \bar{\mathbf{n}}^1 \\ -N_2 \bar{\mathbf{n}}^1 \\ -N_3 \bar{\mathbf{n}}^1 \end{Bmatrix}, \quad \mathbf{N}_\alpha = \begin{Bmatrix} \mathbf{0} \\ -N_{1,\alpha} \bar{\mathbf{n}}^1 \\ -N_{2,\alpha} \bar{\mathbf{n}}^1 \\ -N_{3,\alpha} \bar{\mathbf{n}}^1 \end{Bmatrix}, \quad \mathbf{T}_\alpha = \begin{Bmatrix} \bar{\mathbf{a}}^\alpha \\ -N_1 \bar{\mathbf{a}}^\alpha \\ -N_2 \bar{\mathbf{a}}^\alpha \\ -N_3 \bar{\mathbf{a}}^\alpha \end{Bmatrix}, \quad (8.117)$$

and

$$\boldsymbol{\eta}_s = \begin{Bmatrix} \boldsymbol{\eta}^2 \\ \boldsymbol{\eta}_1^1 \\ \boldsymbol{\eta}_2^1 \\ \boldsymbol{\eta}_3^1 \end{Bmatrix}, \quad \Delta \mathbf{u}_s = \begin{Bmatrix} \Delta \mathbf{u}^2 \\ \Delta \mathbf{u}_1^1 \\ \Delta \mathbf{u}_2^1 \\ \Delta \mathbf{u}_3^1 \end{Bmatrix}. \quad (8.118)$$

Then the contact in the normal direction contributes to the residual by

$$C_c \approx \sum_{s=1}^{n_c} \boldsymbol{\eta}_s^T \mathbf{R}_s, \quad (8.119)$$

with the contact force

$$\mathbf{R}_s = \varepsilon_{NA} g_{Ns} \mathbf{N}_s. \quad (8.120)$$

It is assumed that the parameter  $\varepsilon_{NA} = \varepsilon_N A_s$  contains the area  $A_s$  surrounding the slave node. Thus  $f_N = t_N A_s = \varepsilon_{NA} g_{Ns}$  represents the reaction force of the slave-node due to penetration.

Finally, the contribution of the contact element to the tangent stiffness matrix is given by

$$\Delta C_c \approx \sum_{s=1}^{n_c} \boldsymbol{\eta}_s^T \mathbf{K}_s \Delta \mathbf{u}_s, \quad (8.121)$$

with the tangent matrix related to one slave-node

$$\begin{aligned} \mathbf{K}_s = & \varepsilon_{NA} \left\{ \mathbf{N}_s \mathbf{N}_s^T - g_{Ns} \left[ \mathbf{N}_1 \mathbf{T}_1^T + \mathbf{N}_2 \mathbf{T}_2^T + \mathbf{T}_1 \mathbf{N}_1^T + \mathbf{T}_2 \mathbf{N}_2^T \right. \right. \\ & \left. \left. + g_{Ns} \left( a^{11} \mathbf{N}_1 \mathbf{N}_1^T + a^{12} (\mathbf{N}_1 \mathbf{N}_2^T + \mathbf{N}_2 \mathbf{N}_1^T) + a^{22} \mathbf{N}_2 \mathbf{N}_2^T \right) \right] \right\}. \end{aligned} \quad (8.122)$$

The tangential stiffness matrix for normal contact  $\mathbf{K}_s$  is symmetric. It has the same structure as the tangent matrix (8.36) in the two-dimensional case for frictionless contact.

### 8.4.1 Matrices for Node-To-Edge (NTE) elements

In certain situations, when the slave node  $\mathbf{x}^s$  slides from one master element to the next during the deformation process, it comes in contact with the edge lying between the master elements. The discretization needed in such a case differs from the previous sections. This is mainly due to the fact that the cross product of the two base vectors cannot be used to define the normal on the edge, since the edge lies between two surfaces which have different tangents. Therefore, we compute the normal from the penetration function, see (3.2),

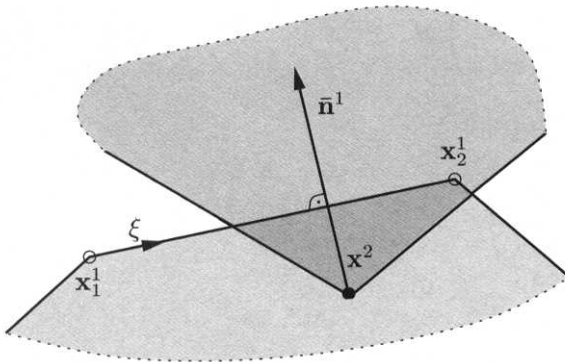
$$\bar{\mathbf{n}}^1 = \frac{\bar{\mathbf{x}}^1 - \mathbf{x}^2}{\|\bar{\mathbf{x}}^1 - \mathbf{x}^2\|} = \frac{1}{g_N^+} (\bar{\mathbf{x}}^1 - \mathbf{x}^2), \quad (8.123)$$

with the gap function  $g_N^+ := \|\bar{\mathbf{x}}^1 - \mathbf{x}^2\|$  and  $\bar{\mathbf{x}}^1 = (1 - \bar{\xi}) \mathbf{x}_1^1 + \bar{\xi} \mathbf{x}_2^1$ . This is shown graphically in Figure 8.10. Based on these definitions, we derive the variation of the gap function  $g_N^+$ ,

$$\delta g_N^+ = \frac{1}{g_N^+} (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot (\bar{\boldsymbol{\eta}}^1 + \bar{\mathbf{x}}_{,\xi}^1 \delta \xi - \boldsymbol{\eta}^2). \quad (8.124)$$

The term containing  $\delta \xi$  disappears in this result, since the first term in the scalar product is the normal vector, see (8.123). This result has to be inserted into the weak form of the contact contribution for  $n_e$  edge elements. It leads in the case of the penalty method to

$$C_c = \int_{\Gamma_c} \epsilon_N g_N^+ \delta g_N^+ dA \approx \sum_{s=1}^{n_e} \epsilon_{NA} (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot (\bar{\boldsymbol{\eta}}^1 - \boldsymbol{\eta}^2), \quad (8.125)$$



**Fig. 8.10** NTE-contact element, definition of contact normal.

which can be stated in matrix form as

$$C_c \approx \sum_{s=1}^{n_e} \eta_s^T \mathbf{R}_s \quad \text{with} \quad \mathbf{R}_s = \epsilon_{NA} \mathbf{X}_s, \quad (8.126)$$

with the vectors

$$\mathbf{X}_s = \begin{Bmatrix} -(\bar{\mathbf{x}}^1 - \mathbf{x}^2) \\ N_1(\bar{\xi}) (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \\ N_2(\bar{\xi}) (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \end{Bmatrix} \quad \text{and} \quad \eta_s^T = \langle \eta^2, \eta_1^1, \eta_2^1 \rangle,$$

where  $\mathbf{1}$  is a  $2 \times 2$  matrix.

The linearization of (8.125) yields

$$\begin{aligned} \Delta C_c = \sum_{s=1}^{n_e} \epsilon_{NA} [ & (\Delta \bar{\mathbf{u}}^1 + \bar{\mathbf{x}}_{,\xi}^1 \Delta \xi - \Delta \mathbf{x}^2) \cdot (\bar{\boldsymbol{\eta}}^1 + \bar{\mathbf{x}}_{,\xi}^1 \delta \xi - \boldsymbol{\eta}^2) \\ & (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot (\bar{\boldsymbol{\eta}}_{,\xi}^1 \Delta \xi + \Delta \bar{\mathbf{u}}_{,\xi}^1 \delta \xi + \bar{\mathbf{x}}_{,\xi\xi}^1 \delta \xi \Delta \xi + \bar{\mathbf{x}}_{,\xi}^1 \Delta \delta \xi) ]. \end{aligned} \quad (8.127)$$

In this equation the term  $\bar{\mathbf{x}}_{,\xi\xi}^1$  is zero for a straight edge; furthermore,  $(\bar{\mathbf{x}}^1 - \mathbf{x}^2)$  shows in the direction of the normal vector hence  $(\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot \bar{\mathbf{x}}_{,\xi}^1 = 0$ . With these results the last two terms in (8.127) disappear. Once  $\delta \xi$  and  $\Delta \xi$  are known, the matrix form of the linearization can be stated. Both result from the condition  $\bar{\mathbf{x}}_{,\xi}^1 \cdot (\bar{\mathbf{x}}^1 - \mathbf{x}^2) = 0$ . The variation or linearization of this expression yields

$$\begin{aligned} \Delta \xi &= -\frac{1}{l^2} [ (\Delta \bar{\mathbf{u}}^1 - \Delta \mathbf{u}^2) \cdot \bar{\mathbf{a}} + (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot \Delta \bar{\mathbf{u}}_{,\xi}^1 ], \\ \delta \xi &= -\frac{1}{l^2} [ (\bar{\boldsymbol{\eta}}^1 - \boldsymbol{\eta}^2) \cdot \bar{\mathbf{a}} + (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot \bar{\boldsymbol{\eta}}_{,\xi}^1 ]. \end{aligned} \quad (8.128)$$

To derive the final form of the linearizations the last two equations have to be combined

$$\begin{aligned} \Delta C_c &= \sum_{s=1}^{n_e} \epsilon_{NA} (\Delta \bar{\mathbf{u}}^1 - \Delta \mathbf{u}^2) \cdot (\bar{\boldsymbol{\eta}}^1 - \boldsymbol{\eta}^2) \\ &\quad - \frac{\epsilon_{NA}}{l^2} \{ [(\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot \bar{\boldsymbol{\eta}}_{,\xi}^1 + (\bar{\boldsymbol{\eta}}^1 - \boldsymbol{\eta}^2) \cdot \bar{\mathbf{a}}] \\ &\quad [(\bar{\mathbf{x}}^1 - \mathbf{x}^2) \cdot \Delta \bar{\mathbf{u}}_{,\xi}^1 + (\Delta \bar{\mathbf{u}}^1 - \Delta \mathbf{u}^2) \cdot \bar{\mathbf{a}}] \}. \end{aligned} \quad (8.129)$$

By introducing the matrices

$$\mathbf{1}_s = \begin{Bmatrix} -\mathbf{1} \\ N_1(\bar{\xi}) \mathbf{1} \\ N_2(\bar{\xi}) \mathbf{1} \end{Bmatrix}, \quad \mathbf{T}_s = \begin{Bmatrix} -\bar{\mathbf{a}} \\ N_1(\bar{\xi}) \bar{\mathbf{a}} \\ N_2(\bar{\xi}) \bar{\mathbf{a}} \end{Bmatrix}, \quad \mathbf{X}_m = \begin{Bmatrix} \mathbf{0} \\ N_1(\bar{\xi}) (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \\ N_2(\bar{\xi}) (\bar{\mathbf{x}}^1 - \mathbf{x}^2) \end{Bmatrix}$$

and

$$\Delta \mathbf{u}_s^T = \langle \Delta \mathbf{u}^2, \Delta \mathbf{u}_1^1, \Delta \mathbf{u}_2^1 \rangle,$$

a compact form of the tangent stiffness for NTE-contact can be stated:

$$\mathbf{K}_s = \epsilon_{NA} \left[ \mathbf{1}_s \mathbf{1}_s^T - \frac{1}{l^2} (\mathbf{T}_s + \mathbf{X}_m) (\mathbf{T}_s + \mathbf{X}_m)^T \right]. \quad (8.130)$$

### 8.4.2 Matrices for Node-To-Node (NTV) elements

If a slave node is close to a vertex node of the master surface then it might be algorithmically useful to formulate the contact as a node-to-node contact. In that case, the contact normal  $\mathbf{n}^1$  is given by (see Figure 8.11)

$$\mathbf{n}^1 = \frac{\mathbf{x}^1 - \mathbf{x}^2}{\|\mathbf{x}^1 - \mathbf{x}^2\|} \quad (8.131)$$

as in the node-to-edge case. Note, however, that one does not have to distinguish between  $\bar{\mathbf{x}}^1$  and  $\mathbf{x}^1$ , since the projection (8.107) returns to the vertex node  $\mathbf{x}^m$ . The gap function and its variation can be stated as for the node-to-edge element:

$$\begin{aligned} g_N^+ &= \|\mathbf{x}^1 - \mathbf{x}^2\|, \\ \delta g_N^+ &= \frac{1}{g_N^+} (\mathbf{x}^1 - \mathbf{x}^2) \cdot (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^2). \end{aligned} \quad (8.132)$$

Thus; the residual vector for  $n_v$  vertices being in contact has the same form as for the NTE element, see (8.125),

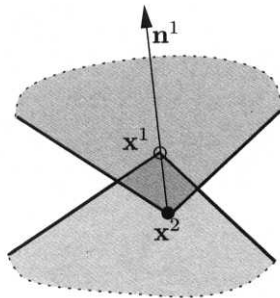
$$C_c \approx \sum_{s=1}^{n_v} \epsilon_{NA} (\mathbf{x}^1 - \mathbf{x}^2) \cdot (\boldsymbol{\eta}^1 - \boldsymbol{\eta}^2). \quad (8.133)$$

The matrix form can thus be computed:

$$C_c \approx \sum_{s=1}^{n_v} \boldsymbol{\eta}_s^T \mathbf{R}_s \quad \text{with} \quad \mathbf{R}_s = \epsilon_{NA} \mathbf{X}_s, \quad (8.134)$$

with

$$\mathbf{X}_s = \begin{Bmatrix} -(\mathbf{x}^1 - \mathbf{x}^2) \\ (\mathbf{x}^1 - \mathbf{x}^2) \end{Bmatrix} \quad \text{and} \quad \boldsymbol{\eta}_s = \begin{Bmatrix} \boldsymbol{\eta}^2 \\ \boldsymbol{\eta}^1 \end{Bmatrix}.$$



**Fig. 8.11** NTV-contact element, definition of contact normal.



The linearization of the residuum follows according to equation (8.127) by setting  $\delta\bar{\xi} = \Delta\bar{\xi} = 0$ :

$$\Delta C_c \approx \sum_{s=1}^{n_v} \epsilon_{NA} (\eta^1 - \eta^2) \cdot (\Delta \mathbf{u}^1 - \Delta \mathbf{u}^2) = \sum_{s=1}^{n_v} \eta_s^T \mathbf{K}_s \mathbf{u}_s, \quad (8.135)$$

with  $\mathbf{u}_s^T = \langle \Delta \mathbf{u}^2, \Delta \mathbf{u}^1 \rangle$ . Thus the tangential stiffness has the simple form

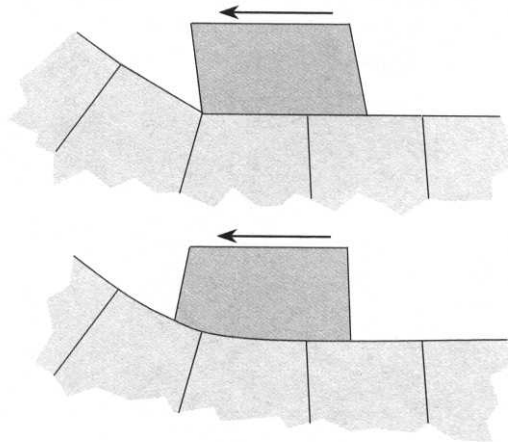
$$\mathbf{K}_s = \epsilon_{NA} \mathbf{1}_s \mathbf{1}_s^T, \quad \text{with} \quad \mathbf{1}_s = \langle -1, 1 \rangle^T, \quad (8.136)$$

where  $\mathbf{1}^T = \langle 1, 1, 1 \rangle$ .

## 8.5 SMOOTH CONTACT DISCRETIZATION

Contact problems undergoing finite deformations are often associated with a large amount of sliding within the contact interface. If the discretizations, discussed in the sections above, are applied, then from one segment to the adjacent one there is no continuity in the normal vector. Thus in the corner between two segments the normal is not uniquely defined, which needs special algorithmic treatment (see also the last two sections). Furthermore, this jump in the normal directions might lead to convergence problems, especially when parts of the master surface change their curvature drastically. These cases often appear in forming tools, for example, see Figure 8.12. In such cases the discretization yields sharp corners. Hence it can be that the movement of the upper body locks in a corner, which then yields incorrect results or a non-convergent solution.

Undefined or not uniquely defined normal vectors can be locally treated as special cases, as in Heegaard and Curnier (1993) and Bittencourt and Creus (1998). Further-



**Fig. 8.12** Non-smooth and smooth surface interpolations.

more, the normal can be averaged as in Papadopoulos and Taylor (1992) and Wang and Nakamachi (1999). Also, as described in Liu et al. (1999), a continuous normal change with the normal vector that is not perpendicular to the contact surface can be introduced.

To overcome the above-mentioned problems, discretizations have recently been derived which allow for a smooth sliding of contacting nodes on the master surface. This leads to discretizations in which higher order interpolations are used to describe the master surface of the contacting bodies. Treatment of 2D- and 3D-smooth contact between deformable bodies can be found for special discretizations in Eterovic and Bathe (1991), Pietrzak (1997), Padmanabhan and Laursen (2001) or Wriggers et al. (2001), for example. Within this approach, a Hermitian, Spline or BÉZIER interpolation is used to discretize the master surface. This leads to a  $C^1$  or even  $C^2$  continuous interpolation of the surface. The latter is not needed for a smooth normal field, but might be important when the contact problem involves dynamics, and thus accelerations are present which can jump at the segment interfaces if the interpolation is not  $C^2$  continuous. Let us note that these interpolations lead in general to a more robust behaviour of the iterative solution algorithms for contact.

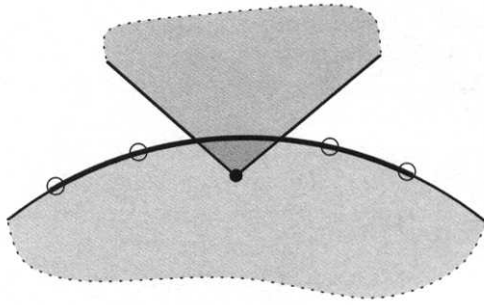
The special case of the contact of a body with a rigid obstacle is treated in Wriggers and Imhof (1993) or Heegaard and Curnier (1993). In such a case, various  $C^1$  surfaces can be defined resulting directly from CAD models; see also Heege and Alart (1996) or Hansson and Klarbring (1990).

Discretizations which lead to smooth node-to-segment contact elements are discussed below. There are different possibilities to derive such contact elements which depend upon the choice of interpolation of the master surface. We concentrate first on two-dimensional interpolations using Hermitian, Spline and BÉZIER polynomials. Also, to make the derivations clearer, only the frictionless case is discussed first for Hermite and BÉZIER interpolations. For these discretizations all matrices are derived explicitly. After that, frictional contact is considered, though the matrices are not stated explicitly. They have been automatically coded by a *Mathematica* package *AceGen* designed by J. Korelc, e.g. see Korelc and Wriggers (1997) or Wriggers et al. (1999).

### 8.5.1 Hermite interpolation for frictionless contact

Again, we consider the contact of one slave node  $\mathbf{x}_s$  with a master segment  $\mathbf{x}^1(\xi)$ . However, this time the master segment is  $C^1$  continuous. In this section a cubic Hermitian interpolation is chosen to interpolate the master surface following Taylor and Wriggers (1999). When applying a cubic Hermitian polynomial, then four nodes are necessary for the interpolation, see Figure 8.13. Using the penalty method, the energy related to a contact constraint, see (5.31), is enforced for each slave node

$$\Pi_\varepsilon \approx \sum_{s=1}^{n_c} \frac{\varepsilon_N}{2} g_{N_s}^2 A_s. \quad (8.137)$$



**Fig. 8.13** Smooth node-to-segment contact element.

Here  $n_c$  is the number of active constraints and  $A_s$  is an area related to the contact node  $s$ . Note that the penalty term can be written as

$$\begin{aligned}\Pi_\varepsilon &\approx \sum_{s=1}^{n_c} \frac{\varepsilon_N}{2} g_{N_s} \bar{\mathbf{n}}^1 \cdot g_{N_s} \bar{\mathbf{n}}^1 A_s \\ &= \sum_{s=1}^{n_c} \frac{\varepsilon_N}{2} (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) \cdot (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) A_s,\end{aligned}\quad (8.138)$$

where now the normal no longer appears. The variation of this expression yields

$$C_c \approx \sum_{s=1}^{n_c} \varepsilon_N (\eta_s^2 - \bar{\eta}^1 - \bar{\mathbf{x}}_{,\xi}^1 \delta \xi) \cdot (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) A_s. \quad (8.139)$$

Again, since  $(\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) = g_{N_s} \bar{\mathbf{n}}^1$ , the third term in the first bracket disappears, and for the variation of  $\Pi_\varepsilon$  we have

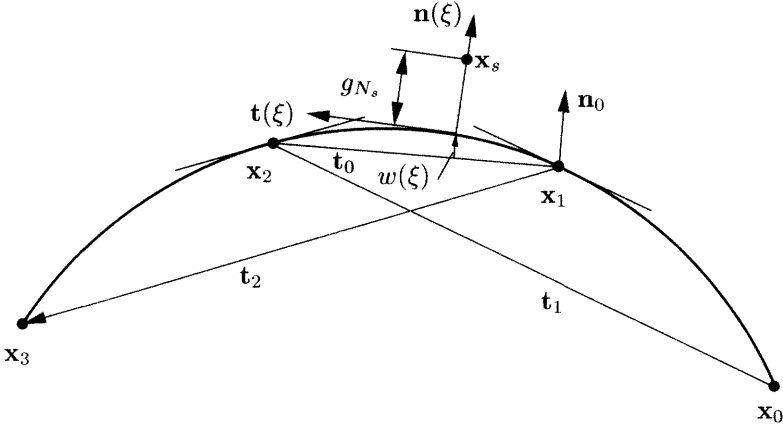
$$C_c \approx \sum_{s=1}^{n_c} \varepsilon_N (\eta_s^2 - \bar{\eta}^1) \cdot (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) A_s. \quad (8.140)$$

The linearization  $\Delta C_c$  starts from (8.139), and leads to the symmetric form

$$\begin{aligned}\Delta C_c &\approx \sum_{s=1}^{n_c} \varepsilon_N A_s [(\eta_s^2 - \bar{\eta}^1 - \bar{\mathbf{x}}_{,\xi}^1 \delta \xi) \cdot (\Delta \mathbf{u}_s^2 - \Delta \bar{\mathbf{u}}^1 - \bar{\mathbf{u}}_{,\xi}^1 \Delta \xi) \\ &\quad - (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) \cdot (\Delta \bar{\eta}^1 + \bar{\eta}_{,\xi}^1 \Delta \xi + \Delta \bar{\mathbf{u}}_{,\xi}^1 \delta \xi + \bar{\mathbf{x}}_{,\xi\xi}^1 \cdot \bar{\mathbf{n}}^1 \delta \xi \Delta \xi)]\end{aligned}\quad (8.141)$$

where the term  $\Delta \bar{\eta}^1$  results from the special Hermitian interpolation, see below. Expression (8.141) can be put into a mixed form by using vector notation. After collecting terms, we arrive at

$$\begin{aligned}\Delta C_c &= \sum_{s=1}^{n_c} \varepsilon_N A_s \{ \langle \eta_s, \bar{\eta}^1, \delta \xi \rangle \begin{bmatrix} \mathbf{I} & -\mathbf{I} & -\bar{\mathbf{x}}_{,\xi}^1 \\ -\mathbf{I} & \mathbf{I} & \bar{\mathbf{z}}^1 \\ -\bar{\mathbf{x}}_{,\xi}^{1T} & \bar{\mathbf{z}}^{1T} & H_{\xi\xi} \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{u}_s^2 \\ \Delta \bar{\mathbf{u}}^1 \\ \Delta \xi \end{Bmatrix} \\ &\quad - (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1)^T \Delta \bar{\eta}^1 \},\end{aligned}\quad (8.142)$$



**Fig. 8.14**  $C^1$  continuous interpolation of contact surface.

where the abbreviations

$$\begin{aligned} H_{\xi\xi} &= (\bar{\mathbf{x}}^1_{,\xi})^T \bar{\mathbf{x}}^1_{,\xi} - g_{N_s} (\bar{\mathbf{n}}^1)^T \bar{\mathbf{x}}^1_{,\xi\xi} \\ \bar{\mathbf{z}}^1 &= \bar{\mathbf{x}}^1_{,\xi} - g_{N_s} \bar{\mathbf{n}}^1 \end{aligned}$$

have been used. Since the variation (8.140) does not depend upon  $\delta\xi$ , we can use a static condensation to eliminate  $\Delta\xi$  in (8.142). This yields

$$\Delta C_c = \sum_{s=1}^{n_c} \varepsilon A_s [(\boldsymbol{\eta}_s, \bar{\boldsymbol{\eta}}^1) \mathbf{A}(\bar{\xi}) \left\{ \begin{array}{c} \Delta \mathbf{u}_s^2 \\ \Delta \bar{\mathbf{u}}^1 \end{array} \right\} - (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1)^T \Delta \bar{\boldsymbol{\eta}}^1], \quad (8.143)$$

where the matrix

$$\mathbf{A}(\bar{\xi}) = \begin{bmatrix} \mathbf{I} - H_{\xi\xi}^{-1} \bar{\mathbf{x}}^1_{,\xi} \bar{\mathbf{x}}^1_{,\xi}^T & -\mathbf{I} + H_{\xi\xi}^{-1} \bar{\mathbf{x}}^1_{,\xi} \bar{\mathbf{z}}^1{}^T \\ -\mathbf{I} + H_{\xi\xi}^{-1} \bar{\mathbf{z}}^1 \bar{\mathbf{x}}^1_{,\xi}^T & \mathbf{I} - H_{\xi\xi}^{-1} \bar{\mathbf{z}}^1 \bar{\mathbf{z}}^1{}^T \end{bmatrix} \quad (8.144)$$

has been introduced.

Now we have to discretize the master surface of body  $\mathcal{B}^1$ , which then yields the matrix form for the smooth contact element. An interpolation which provides a normal field that does not have jumps when going from one segment to the next has to be  $C^1$  continuous. This is provided by the Hermitian interpolation. For the cubic Hermitian polynomials we need four nodal points on the master surface (see Figure 8.14) to define the interpolation. This interpolation will be defined on the master segment between the nodes  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . To obtain continuous tangents from one segment to the next we define the tangent vectors  $\mathbf{t}_1 = \mathbf{x}_2 - \mathbf{x}_0$  and  $\mathbf{t}_2 = \mathbf{x}_3 - \mathbf{x}_1$ .

Let us further introduce a tangent and a normal vector. Both form the local frame, for the segment under consideration:

$$\mathbf{t}_0 = \mathbf{x}_2 - \mathbf{x}_1 \quad \text{and} \quad \mathbf{n}_0 = -\mathbf{e}_3 \times \mathbf{t}_0 = \mathbf{T} \mathbf{t}_0, \quad (8.145)$$

where  $\mathbf{e}_3$  is the unit base vector perpendicular to the plane; thus, the cross product can be expressed by the skew matrix  $\mathbf{T}$

$$\mathbf{T} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (8.146)$$

Based on these definitions, we can introduce the surface interpolation as a linear interpolation between nodes 1 and 2 and a cubic Hermitian interpolation with respect to the local frame

$$\mathbf{x}^1(\xi) = N_1(\xi) \mathbf{x}_1 + N_2(\xi) \mathbf{x}_2 + w(\xi) \mathbf{n}_0, \quad (8.147)$$

where the cubic interpolation is given by  $w(\xi) = H_1(\xi) B_1 + H_2(\xi) B_2$ .  $N_\alpha(\xi)$  being the standard linear shape functions and  $H_\alpha(\xi)$  are the Hermitian polynomials, defined as

$$\begin{aligned} N_1(\xi) &= \frac{1}{2} (1 - \xi), & N_2(\xi) &= \frac{1}{2} (1 + \xi), \\ H_1(\xi) &= \frac{1}{4} (\xi^2 - 1)(\xi - 1), & H_2(\xi) &= \frac{1}{4} (\xi^2 - 1)(\xi + 1). \end{aligned} \quad (8.148)$$

The value  $B_\alpha$  is given in terms of the tangent vectors  $\mathbf{t}_\alpha$  as

$$B_\alpha = \frac{1}{2} \frac{\mathbf{t}_\alpha^T \mathbf{n}_0}{\mathbf{t}_\alpha^T \mathbf{t}_0}, \quad (8.149)$$

and denotes the angle between the tangent  $\mathbf{t}_\alpha$  and the local frame defined by  $(\mathbf{t}_0, \mathbf{n}_0)$ . Now we can summarize the interpolation of the surface within the segment between nodes 1 and 2 as

$$\mathbf{x}^1(\xi) = \sum_{\alpha=1}^2 [N_\alpha(\xi) \mathbf{x}_\alpha + H_\alpha(\xi) B_\alpha \mathbf{n}_0]. \quad (8.150)$$

Based on this interpolation, the closest point projection of the slave node  $\mathbf{x}_s^2$  onto the master surface can be computed, which yields  $\bar{\mathbf{x}}^1 = \mathbf{x}^1(\bar{\xi})$ , and is needed to define the gap function. Contrary to the linear node-to-segment element in Section 8.1, here one has to solve the following nonlinear equation:

$$[\mathbf{x}_s^2 - \mathbf{x}^1(\bar{\xi})] \cdot \mathbf{x}_{,\xi}^1(\bar{\xi}) = 0. \quad (8.151)$$

Equation (8.151) is a polynomial of fifth order in  $\xi$  which is solved using NEWTON'S method, as described for the three-dimensional case in (8.82). For this algorithm, the derivative of  $\mathbf{x}(\xi)$  with respect to the surface coordinate  $\xi$  is needed, and is hence given next:

$$\mathbf{x}^1(\xi)_{,\xi} = \frac{1}{2} (\mathbf{x}_2 - \mathbf{x}_1) + \sum_{\alpha=1}^2 H_{\alpha,\xi}(\xi) B_\alpha \mathbf{n}_0. \quad (8.152)$$

A good starting value for (8.151) is obtained by the projection of the slave node onto  $\mathbf{t}_0$ , which can be computed explicitly:

$$\bar{\xi} = \frac{1}{l^2} [2 \mathbf{x}_s - (\mathbf{x}_2 + \mathbf{x}_1)] \cdot (\mathbf{x}_2 - \mathbf{x}_1) \quad \text{with} \quad l^2 = (\mathbf{x}_2 - \mathbf{x}_1) \cdot (\mathbf{x}_2 - \mathbf{x}_1).$$

Now we have to express the variation and linearization of the gap associated with one slave node  $\mathbf{x}_s$ . The variation of  $\boldsymbol{\eta}^1(\xi)$  yields

$$\boldsymbol{\eta}^1(\xi) = \sum_{\alpha=1}^2 [N_{\alpha}(\xi) \boldsymbol{\eta}_{\alpha} + H_{\alpha}(\xi) (\delta B_{\alpha} \mathbf{n}_0 + B_{\alpha} \delta \mathbf{n}_0)]. \quad (8.153)$$

By exploiting the structure of  $\delta B_{\alpha} = \frac{1}{2} \delta \left( \frac{\mathbf{t}_{\alpha}^T \mathbf{n}_0}{\mathbf{t}_{\alpha}^T \mathbf{t}_0} \right)$  we arrive at the final form:

$$\delta B_{\alpha} = \frac{1}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} \left[ (\mathbf{t}_{\alpha}^T \mathbf{T}^T - B_{\alpha} \mathbf{t}_{\alpha}^T) \delta \mathbf{t}_0 + (\mathbf{n}_0^T - B_{\alpha} \mathbf{t}_0^T) \delta \mathbf{t}_{\alpha} \right]. \quad (8.154)$$

Now we write the variation of  $B_{\alpha}$  in matrix form. For this we define

$$\begin{aligned} \mathbf{p}_1^{(\alpha)} &= \frac{H_{\alpha}}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} [\mathbf{T} \mathbf{t}_{\alpha} - B_{\alpha} \mathbf{t}_{\alpha}], \\ \mathbf{p}_2^{(\alpha)} &= \frac{H_{\alpha}}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} [\mathbf{n}_0 - B_{\alpha} \mathbf{t}_0], \end{aligned} \quad (8.155)$$

and obtain

$$\delta B_{\alpha} = \langle \delta \mathbf{t}_0, \delta \mathbf{t}_{\alpha} \rangle \left\{ \begin{matrix} \mathbf{p}_1^{(\alpha)} \\ \mathbf{p}_2^{(\alpha)} \end{matrix} \right\}. \quad (8.156)$$

Since  $\delta \mathbf{t}_0 = \boldsymbol{\eta}_2 - \boldsymbol{\eta}_1$  and with (8.153), (8.154) and (8.155), the variation  $(\boldsymbol{\eta}_s^2 - \bar{\boldsymbol{\eta}}^1)$  needed in (8.140) can be expressed in matrix form as  $(\boldsymbol{\eta}_s^2 - \bar{\boldsymbol{\eta}}^1) = \boldsymbol{\eta}_s^T \mathbf{B}_s(\bar{\xi})$ , where

$$\boldsymbol{\eta}_s^T \mathbf{B}_s(\bar{\xi}) = \boldsymbol{\eta}_s^T \left\{ \begin{matrix} \mathbf{I} \\ -N_1(\bar{\xi}) \mathbf{I} + \sum_{\alpha} [H_{\alpha}(\bar{\xi}) B_{\alpha} \mathbf{T}^T - \mathbf{p}_1^{(\alpha)} \mathbf{n}_0^T] \\ -N_2(\bar{\xi}) \mathbf{I} - \sum_{\alpha} [H_{\alpha}(\bar{\xi}) B_{\alpha} \mathbf{T}^T - \mathbf{p}_1^{(\alpha)} \mathbf{n}_0^T] \\ \mathbf{p}_2^{(1)} \mathbf{n}_0^T \\ \mathbf{p}_2^{(2)} \mathbf{n}_0^T \end{matrix} \right\}, \quad (8.157)$$

with  $\boldsymbol{\eta}_s^T = \langle \boldsymbol{\eta}_s, \boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \delta \mathbf{t}_1, \delta \mathbf{t}_2 \rangle$ . Thus, the matrix form of the variation of the penalty energy  $\Pi_{\varepsilon}$  in (8.137) of all contact contributions for the active contact constraints in (8.140) is finally given by

$$C_c = \sum_{s=1}^{n_c} \varepsilon A_s \boldsymbol{\eta}_s^T \mathbf{B}_s(\bar{\xi}) (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1). \quad (8.158)$$

The linearization follows with (8.157) from (8.143), as

$$\Delta C_c \sum_{s=1}^{n_c} \varepsilon_N A_s \left[ \boldsymbol{\eta}_s^T \mathbf{B}_s(\bar{\xi}) \mathbf{A}(\bar{\xi}) \mathbf{B}_s^T(\bar{\xi}) \Delta \mathbf{u}_s - (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1)^T \Delta \bar{\boldsymbol{\eta}}^1 \right], \quad (8.159)$$

where  $\Delta \mathbf{u}_s^T = \langle \Delta \mathbf{u}_s, \Delta \mathbf{u}_1, \Delta \mathbf{u}_2, \Delta \mathbf{t}_1, \Delta \mathbf{t}_2 \rangle$ . The last term has to be derived as a function of the unknown variables. Note that in (8.159) the linearization

$$\Delta(\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) = (\Delta \mathbf{u}_s^2 - \Delta \bar{\mathbf{u}}^1) = \mathbf{B}_s^T(\bar{\xi}) \Delta \mathbf{u}_s$$

has been used according to (8.157).

The last term in (8.159) results from the linearization of the variation (8.153) with regard to the variables. This yields

$$(\mathbf{x}_s^2 - \bar{\mathbf{x}}^1)^T \Delta \bar{\boldsymbol{\eta}}^1 = (\mathbf{x}_s^2 - \bar{\mathbf{x}}^1)^T \sum_{\alpha=1}^2 H_{\alpha}(\xi) [\Delta \delta B_{\alpha} \mathbf{n}_0 + \Delta B_{\alpha} \delta \mathbf{n}_0 + \delta B_{\alpha} \Delta \mathbf{n}_0]. \quad (8.160)$$

In this expression, the term  $\delta B_{\alpha}$  is already known and  $\Delta B_{\alpha}$  has the same structure. Thus, the term which has to be investigated in detail is  $\Delta \delta B_{\alpha}$ :

$$\Delta \delta B_{\alpha} = \Delta \left\{ \frac{1}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} [(\mathbf{t}_{\alpha}^T \mathbf{T} - B_{\alpha} \mathbf{t}_{\alpha}^T) \delta \mathbf{t}_0 + (\mathbf{n}_0^T - B_{\alpha} \mathbf{t}_0^T) \delta \mathbf{t}_{\alpha}] \right\}. \quad (8.161)$$

Using the results already obtained in (8.156), and by defining the matrices

$$\begin{aligned} \mathbf{m}_{11}^{(\alpha)} &= -\frac{H_{\alpha}}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} [\mathbf{T} \mathbf{t}_{\alpha} \mathbf{t}_{\alpha}^T + \mathbf{t}_{\alpha} \mathbf{t}_{\alpha}^T \mathbf{T}^T - 2 B_{\alpha} \mathbf{t}_{\alpha} \mathbf{t}_{\alpha}^T], \\ \mathbf{m}_{12}^{(\alpha)} &= \frac{H_{\alpha}}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} [\mathbf{n}_0 \mathbf{t}_{\alpha}^T - \mathbf{t}_0 \mathbf{t}_{\alpha}^T \mathbf{T}^T - 2 B_{\alpha} \mathbf{t}_0 \mathbf{t}_{\alpha}^T], \\ \mathbf{m}_{22}^{(\alpha)} &= \frac{H_{\alpha}}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} [\mathbf{n}_0 \mathbf{t}_0^T + \mathbf{t}_0 \mathbf{n}_0^T - 2 B_{\alpha} \mathbf{t}_0 \mathbf{t}_0^T], \end{aligned} \quad (8.162)$$

we arrive at

$$H_{\alpha} \Delta \delta B_{\alpha} = \langle \delta \mathbf{t}_0, \delta \mathbf{t}_{\alpha} \rangle \mathbf{M}_{\alpha} \begin{Bmatrix} \Delta \mathbf{t}_0 \\ \Delta \mathbf{t}_{\alpha} \end{Bmatrix}. \quad (8.163)$$

In this expression the matrix  $\mathbf{M}_{\alpha}$  has the structure

$$\mathbf{M}^{(\alpha)} = \frac{1}{2 \mathbf{t}_{\alpha}^T \mathbf{t}_0} \begin{bmatrix} \mathbf{m}_{11}^{(\alpha)} & -H_{\alpha} (B_{\alpha} \mathbf{I} - \mathbf{T}) - \mathbf{m}_{12}^{(\alpha)T} \\ -H_{\alpha} (B_{\alpha} \mathbf{I} - \mathbf{T}^T) - \mathbf{m}_{12}^{(\alpha)} & -\mathbf{m}_{22}^{(\alpha)} \end{bmatrix}. \quad (8.164)$$

Furthermore, let us define with (8.155) the  $2 \times 2$  matrices which are needed to describe the last two terms in (8.160):

$$\begin{aligned} \mathbf{P}_1^{(\alpha)} &= \mathbf{T}(\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) \mathbf{p}_1^{(\alpha)T}, \\ \mathbf{P}_2^{(\alpha)} &= \mathbf{T}(\mathbf{x}_s^2 - \bar{\mathbf{x}}^1) \mathbf{p}_2^{(\alpha)T}. \end{aligned} \quad (8.165)$$

Now all terms have been derived, and the final matrix form of the linearization can be stated:

$$(\mathbf{x}_s^2 - \bar{\mathbf{x}}^1)^T \Delta \boldsymbol{\eta} = \langle \delta \mathbf{t}_0, \delta \mathbf{t}_1, \mathbf{t}_2 \rangle \mathbf{M} \begin{Bmatrix} \Delta \mathbf{t}_0 \\ \Delta \mathbf{t}_1 \\ \Delta \mathbf{t}_2 \end{Bmatrix}. \quad (8.166)$$

Here the matrix  $\mathbf{M}$  is given, with (8.164) and (8.165), as

$$\mathbf{M} = \begin{bmatrix} \sum_{\beta=1}^2 (\mathbf{M}_{11}^{(\beta)} + \mathbf{P}_1^{(\beta)} + \mathbf{P}_1^{(\beta)T}) & \mathbf{P}_2^{(1)} + \mathbf{M}_{12}^{(1)} & \mathbf{P}_2^{(2)} + \mathbf{M}_{12}^{(2)} \\ \mathbf{P}_2^{(1)T} + \mathbf{M}_{21}^{(1)} & \mathbf{M}_{22}^{(1)} & \mathbf{0} \\ \mathbf{P}_2^{(2)T} + \mathbf{M}_{21}^{(2)} & \mathbf{0} & \mathbf{M}_{22}^{(2)} \end{bmatrix}. \quad (8.167)$$

Finally, we have to express the vectors  $\mathbf{t}_0$ ,  $\mathbf{t}_1$  and  $\mathbf{t}_2$  in terms of the nodal values  $\mathbf{x}_0$ ,  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  and  $\mathbf{x}_3$ . For this we define the transformation

$$\begin{Bmatrix} \delta \mathbf{t}_0 \\ \delta \mathbf{t}_1 \\ \delta \mathbf{t}_2 \end{Bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\eta}_s \\ \boldsymbol{\eta}_0 \\ \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \end{Bmatrix}, \quad (8.168)$$

where  $\boldsymbol{\eta}_s$  has been added for completeness since it is already used in (8.159). This transformation can now be applied in (8.166), and together with (8.167) yields the final matrix form for the linearization of the normal gap in (8.159).

## 8.5.2 Bézier interpolation for frictionless contact

In the next discretization we use, the contact constraint is enforced for each slave node and the master surface is discretized using BÉZIER polynomials. Note that the derivation of the residual and tangent again follows the derivations performed in the previous chapters. These can be applied within any kind of surface interpolation for the segment. Thus, we can also use such formulations for  $C^1$  and for  $C^2$  continuous interpolations.

The BÉZIER polynomials which are introduced here to obtain a continuous normal field are cubic functions. As the Hermitian functions these are defined by four points on the master surface, though in a different manner, see Figure 8.15. The BÉZIER interpolation for the segment described by nodes 1 and 2 yields

$$\mathbf{x}^1(\xi) = B_1(\xi)\mathbf{x}_1 + B_2(\xi)\mathbf{x}_{1+} + B_3(\xi)\mathbf{x}_{2-} + B_4(\xi)\mathbf{x}_2, \quad (8.169)$$

where the BÉZIER interpolation functions are defined as

$$\begin{aligned} B_1(\xi) &= \frac{1}{8}(1-\xi)^3, & B_2(\xi) &= \frac{3}{8}(1-\xi)^2(1+\xi), \\ B_3(\xi) &= \frac{3}{8}(1-\xi)(1+\xi)^2, & B_4(\xi) &= \frac{1}{8}(1+\xi)^3. \end{aligned} \quad (8.170)$$

Observe that the interpolation lies in the convex hull spanned by the nodes  $\mathbf{x}_1$ ,  $\mathbf{x}_{1+}$ ,  $\mathbf{x}_{2-}$  and  $\mathbf{x}_2$ , see Figure 8.15.

Our main requirement for the interpolation is that the tangent vectors of adjacent segments have to be equal to maintain  $C^1$  continuity over segment boundaries. This condition can be applied to compute the interior points of the segment  $\mathbf{x}_{1+}$ ,  $\mathbf{x}_{2-}$ . By defining the tangent vectors at nodes 1 and 2 as in the previous section, we obtain

$$\mathbf{t}_1 = \frac{\alpha}{2}(\mathbf{x}_2 - \mathbf{x}_0) \quad \text{and} \quad \mathbf{t}_2 = \frac{\alpha}{2}(\mathbf{x}_3 - \mathbf{x}_1). \quad (8.171)$$



Now we take the derivative of (8.169) and evaluate this at the end points  $\xi = -1$  and  $\xi = +1$ . By setting this equal to the tangent vectors  $\mathbf{t}_\alpha$ , we obtain

$$\begin{aligned}\mathbf{x}_{1+} &= \mathbf{x}_1 - \frac{\alpha}{2} (\mathbf{x}_2 - \mathbf{x}_0), \\ \mathbf{x}_{2-} &= \mathbf{x}_2 + \frac{\alpha}{2} (\mathbf{x}_3 - \mathbf{x}_1).\end{aligned}\quad (8.172)$$

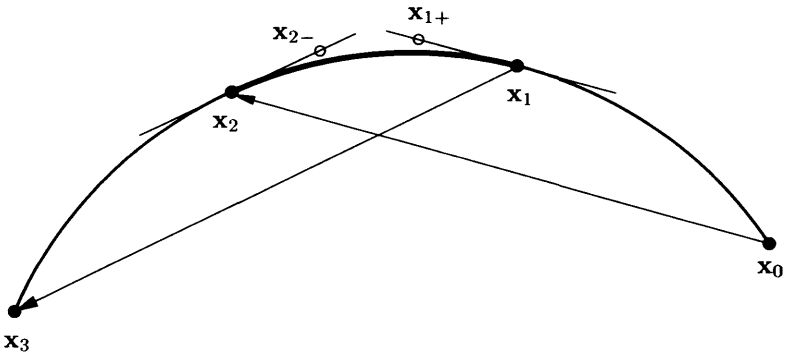
The parameter  $\alpha$  specifies how far nodes  $\mathbf{x}_{1+}$  and  $\mathbf{x}_{2-}$  are away from nodes  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , respectively. For different  $\alpha$  the shape of the surface interpolation changes. In the limit for  $\alpha \rightarrow 0$  we obtain an almost flat segment, however the corner region between adjacent segments is still  $C^1$  continuous. Since the shape of the surface changes during the finite deformation process,  $\alpha$  might be adapted within the calculation. However, a good choice for  $\alpha$  is  $\alpha = \frac{1}{3}$ , see also Pietrzak (1997).

With (8.172) we can rewrite the interpolation (8.169). This leads to

$$\mathbf{x}^1(\xi) = \sum_{i=0}^3 \bar{B}_i(\xi) \mathbf{x}_i, \quad (8.173)$$

with

$$\begin{aligned}\bar{B}_0(\xi) &= \frac{\alpha}{2} B_2(\xi), \\ \bar{B}_1(\xi) &= B_1(\xi) + B_2(\xi) - \frac{\alpha}{2} B_3(\xi), \\ \bar{B}_2(\xi) &= B_3(\xi) + B_4(\xi) - \frac{\alpha}{2} B_2(\xi), \\ \bar{B}_3(\xi) &= \frac{\alpha}{2} B_3(\xi).\end{aligned}\quad (8.174)$$



**Fig. 8.15**  $C^1$  continuous BÉZIER interpolation of contact surface.

Now we compute the first and second derivatives of  $\mathbf{x}^1(\xi)$  with respect to the surface coordinate  $\xi$  for later use:

$$\begin{aligned}\mathbf{x}_{,\xi}^1(\xi) &= \sum_{i=0}^3 \bar{B}_{i,\xi}(\xi) \mathbf{x}_i, \\ \mathbf{x}_{,\xi\xi}^1(\xi) &= \sum_{i=0}^3 \bar{B}_{i,\xi\xi}(\xi) \mathbf{x}_i.\end{aligned}\quad (8.175)$$

The expression for the variation of the gap (3.29) using this interpolation is now

$$\delta g_{N_s} = \left[ \boldsymbol{\eta}_s - \sum_{i=0}^3 \bar{B}_i(\bar{\xi}) \boldsymbol{\eta}_i \right] \cdot \bar{\mathbf{n}}^1, \quad (8.176)$$

which is easily expressed in matrix form as

$$\delta g_{N_s} = \hat{\boldsymbol{\eta}}_s^T \mathbf{B}_n(\bar{\xi}) = \langle \boldsymbol{\eta}_s^T, \boldsymbol{\eta}_0^T, \boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \boldsymbol{\eta}_3^T \rangle \begin{Bmatrix} -\bar{B}_0(\bar{\xi}) \bar{\mathbf{n}}^1 \\ -\bar{B}_1(\bar{\xi}) \bar{\mathbf{n}}^1 \\ -\bar{B}_2(\bar{\xi}) \bar{\mathbf{n}}^1 \\ -\bar{B}_3(\bar{\xi}) \bar{\mathbf{n}}^1 \end{Bmatrix}. \quad (8.177)$$

Thus the residuum connected with the smooth BÉZIER formulation can be stated. For frictionless contact based on (8.12) for the penalty method, it yields

$$C_c \approx \sum_{s=1}^{n_c} \hat{\boldsymbol{\eta}}_s^T [\varepsilon_N A_s g_{N_s} \mathbf{B}_n(\bar{\xi})]. \quad (8.178)$$

The linearization of (8.178) is derived next. Let us recall that it is given by

$$\Delta C_c \approx \sum_{s=1}^{n_c} \varepsilon_N A_s (\delta g_{N_s} \Delta g_{N_s} + g_{N_s} \Delta \delta g_{N_s}). \quad (8.179)$$

In this expression the linearization of the variation of the gap function is needed. It can be derived from (5.105):

$$\begin{aligned}\Delta \delta g_N &= -(\bar{\boldsymbol{\eta}}_{,\xi}^1 \Delta \xi + \Delta \bar{\mathbf{u}}_{,\xi}^1 \delta \xi + \bar{\mathbf{x}}_{,\xi\xi}^1 \Delta \xi \delta \xi) \cdot \bar{\mathbf{n}}^1 \\ &\quad + g_N \bar{\mathbf{n}}^1 \cdot (\bar{\boldsymbol{\eta}}_{,\xi}^1 + \bar{\mathbf{x}}_{,\xi\xi}^1 \delta \xi) \bar{a}^{11} (\Delta \bar{\mathbf{u}}_{,\xi}^1 + \bar{\mathbf{x}}_{,\xi\xi}^1 \delta \xi) \cdot \bar{\mathbf{n}}^1.\end{aligned}\quad (8.180)$$

To specify this for the BÉZIER interpolation, we have to express  $\delta \xi$ , see (3.21), and  $\Delta \xi$ , see (5.98), in matrix form as well as  $\bar{\boldsymbol{\eta}}_{,\xi}^1$  and  $\Delta \bar{\mathbf{u}}_{,\xi}^1$ . All these quantities were derived in Section 5.5.1. Let us first compute the normal component of the variation:

$$\bar{\boldsymbol{\eta}}_{,\xi}^1 \cdot \bar{\mathbf{n}}^1 = \hat{\boldsymbol{\eta}}_s^T \mathbf{B}_{n,\xi}(\bar{\xi}) = \langle \boldsymbol{\eta}_s^T, \boldsymbol{\eta}_0^T, \boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \delta \boldsymbol{\eta}_3^T \rangle \begin{Bmatrix} 0 \\ \bar{B}_{0,\xi}(\bar{\xi}) \bar{\mathbf{n}}^1 \\ \bar{B}_{1,\xi}(\bar{\xi}) \bar{\mathbf{n}}^1 \\ \bar{B}_{2,\xi}(\bar{\xi}) \bar{\mathbf{n}}^1 \\ \bar{B}_{3,\xi}(\bar{\xi}) \bar{\mathbf{n}}^1 \end{Bmatrix}. \quad (8.181)$$

Note that the same structure can be used for computation of the normal component of  $\Delta \bar{\mathbf{u}}_{,\xi}^1$ . Furthermore, we define the matrix form of  $(\boldsymbol{\eta}_s^2 - \bar{\boldsymbol{\eta}}^1) \cdot \bar{\mathbf{x}}_{,\xi}^1$ , which is needed to compute  $\delta \xi$ , see (3.32):

$$(\boldsymbol{\eta}_s^2 - \bar{\boldsymbol{\eta}}^1) \cdot \bar{\mathbf{x}}_{,\xi}^1 = \hat{\boldsymbol{\eta}}_s^T \mathbf{B}_t(\bar{\xi}) = \langle \boldsymbol{\eta}_s^T, \boldsymbol{\eta}_0^T, \boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \boldsymbol{\eta}_3^T \rangle \begin{Bmatrix} \bar{\mathbf{x}}_{,\xi}^1 \\ -\bar{B}_0(\bar{\xi}) \bar{\mathbf{x}}_{,\xi}^1 \\ -\bar{B}_1(\bar{\xi}) \bar{\mathbf{x}}_{,\xi}^1 \\ -\bar{B}_2(\bar{\xi}) \bar{\mathbf{x}}_{,\xi}^1 \\ -\bar{B}_3(\bar{\xi}) \bar{\mathbf{x}}_{,\xi}^1 \end{Bmatrix}. \quad (8.182)$$

The matrix form of the variation of the surface coordinate follows with (8.182) and (8.181) from (3.32):

$$\delta \xi = \hat{\boldsymbol{\eta}}_s^T \left\{ H_{\xi\xi}^{-1} [\mathbf{B}_t(\bar{\xi}) + g_{N_s} \mathbf{B}_{n,\xi}(\bar{\xi})] \right\} = \hat{\boldsymbol{\eta}}_s^T \mathbf{B}_\xi(\bar{\xi}), \quad (8.183)$$

where  $H_{\xi\xi}^{-1} = (\bar{\mathbf{x}}_{,\xi}^1 \cdot \bar{\mathbf{x}}_{,\xi}^1 - g_{N_s} \bar{\mathbf{n}}^1 \cdot \bar{\mathbf{x}}_{,\xi\xi}^1)^{-1}$ . Again, the same structure can be used for the linearization  $\Delta \bar{\xi}$ .

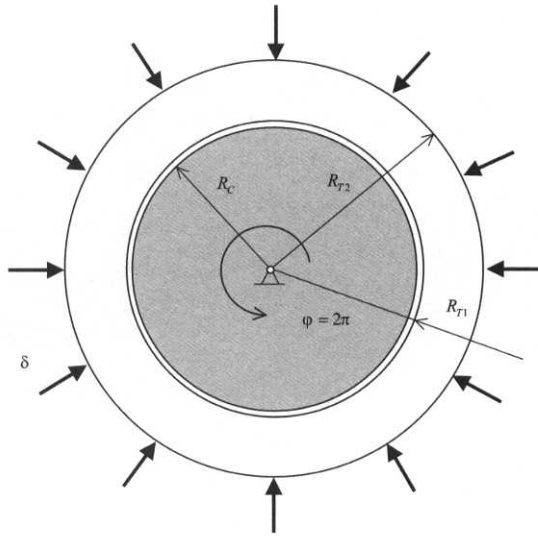
The matrix form of the linearization of the gap function (8.180) can be expressed with (8.181) and (8.183). Thus, for (8.179) with  $\bar{b}_{\xi\xi}^1 = \bar{\mathbf{x}}_{,\xi\xi}^1 \cdot \bar{\mathbf{n}}^1$ , we finally obtain

$$\begin{aligned} \Delta C_c \approx & \sum_{s=1}^{n_c} \varepsilon_N A_s \hat{\boldsymbol{\eta}}_s^T \{ \mathbf{B}_n(\bar{\xi}) \mathbf{B}_n(\bar{\xi})^T \\ & - g_{N_s} [\mathbf{B}_{n,\xi}(\bar{\xi}) \mathbf{B}_\xi(\bar{\xi})^T + \mathbf{B}_\xi(\bar{\xi}) \mathbf{B}_{n,\xi}(\bar{\xi})^T + \bar{b}_{\xi\xi}^1 \mathbf{B}_\xi(\bar{\xi}) \mathbf{B}_\xi(\bar{\xi})^T \\ & - \frac{g_{N_s}}{\|\bar{\mathbf{x}}_{,\xi}^1\|^2} [\mathbf{B}_{n,\xi}(\bar{\xi}) + \bar{b}_{\xi\xi}^1 \mathbf{B}_\xi(\bar{\xi})] [\mathbf{B}_{n,\xi}(\bar{\xi}) + \bar{b}_{\xi\xi}^1 \mathbf{B}_\xi(\bar{\xi})]^T \} \Delta \bar{\mathbf{u}}_s, \end{aligned} \quad (8.184)$$

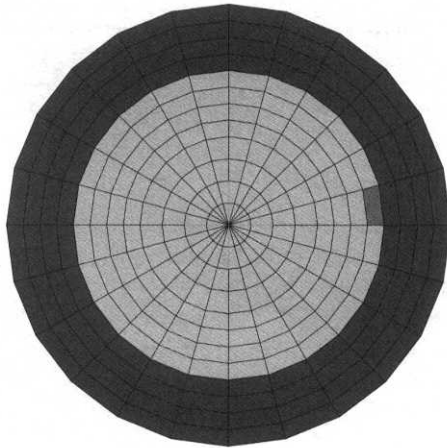
which denotes the tangent matrix of the smooth BÉZIER formulation for frictionless contact. Observe that the formulation for the BÉZIER contact element is much simpler than that for the Hermitian element, which needs special treatment of the rotation  $B_\alpha$  within the linearization.

**Example.** As an example of both formulations, the rotation of a shaft in a cylindrical bearing is considered. The geometry is shown in Figure 8.16. The shaft is modelled by an elastic constitutive equation with modulus of compression  $K = 1.75 \cdot 10^5$  MPa and shear modulus  $G = 8.08 \cdot 10^4$  MPa. The cylindrical tube is characterized by an elastic constitutive equation as well, with the same material data. The radius of the cylindrical shaft is  $R_C = 70$  mm. The tube is defined by the radii  $R_{T1} = 70.1$  mm and  $R_{T2} = 100$  mm. The load is applied in two steps. First a radial displacement  $\delta$  is prescribed at the outer boundary of the tube such that contact between the tube and the shaft occurs. After that, the shaft is rotated for  $\varphi = 360^\circ$ . The finite element mesh is depicted in Figure 8.17.

For this problem a non-smooth interpolation yields unsatisfactory results; since the reaction, here the angular moment is used, oscillates very much, it is not shown here. Even when the Hermitian or BÉZIER interpolations are used, an oscillation

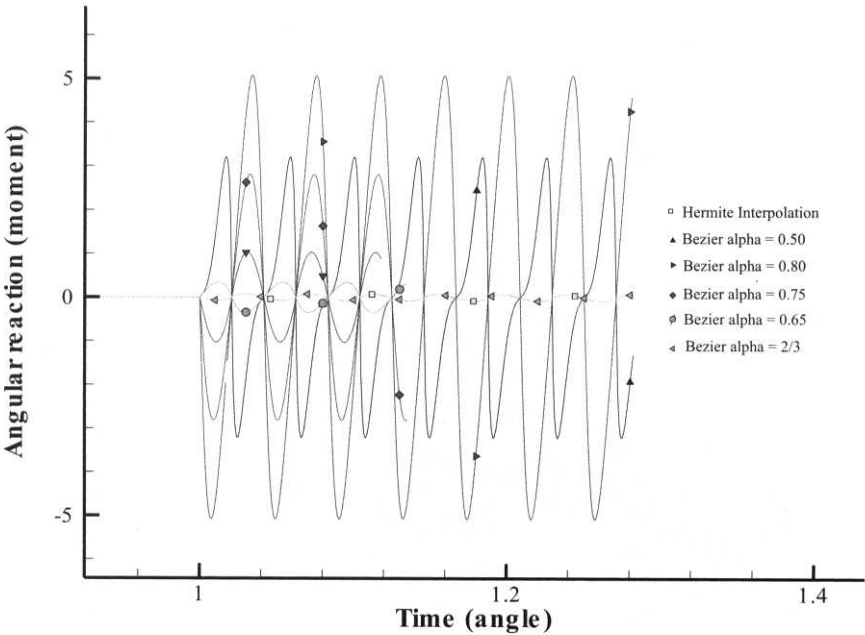


**Fig. 8.16** Cylindrical bearing.



**Fig. 8.17** Finite element mesh of rotating shaft.

occurs which is depicted in Figure 8.18 (here the value  $\bar{\alpha} = 1 - \alpha$  was used to describe the position of the points  $\mathbf{x}_{1+}$  and  $\mathbf{x}_{2-}$ ). This oscillation stems from the choice of parameter  $\alpha$  in (8.172), only for  $\alpha = \frac{1}{3} \equiv \bar{\alpha} = \frac{2}{3}$  the same good solution is obtained as for the Hermitian interpolation. The reason for this is that  $\alpha$  influences the geometrical description of the surface. It is smooth for any value of  $\alpha$ , though the

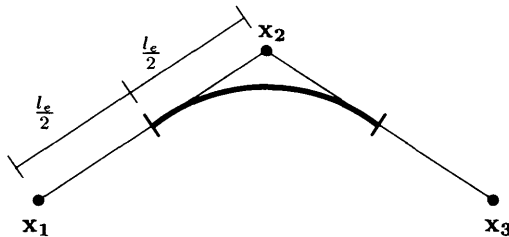


**Fig. 8.18** Angular moment for smooth (SNTS) interpolation with different values of  $\bar{\alpha}$ .

best approximation for a constant curvature which is associated with a circular shaft is only obtained for a special value of  $\alpha$ .

### 8.5.3 Bezier interpolation for frictional contact

To show that there is a certain variety in choosing the boundary discretization for continuous contact formulations, we introduce for the frictional  $C^1$  interpolation another discretization. It is based only on two segments, as described in Wriggers et al. (1999), instead of the three used in the previous section, see Figure 8.19. This

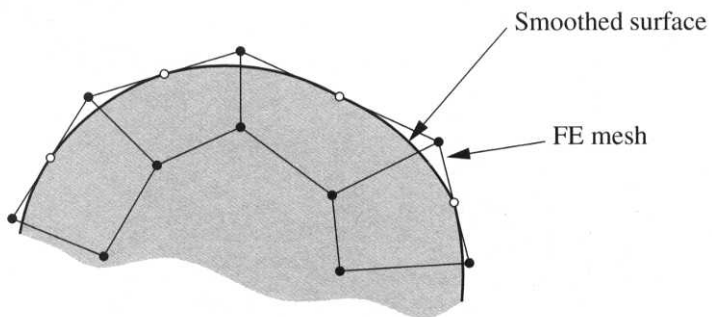


**Fig. 8.19** Contact discretization using a three-node master element.

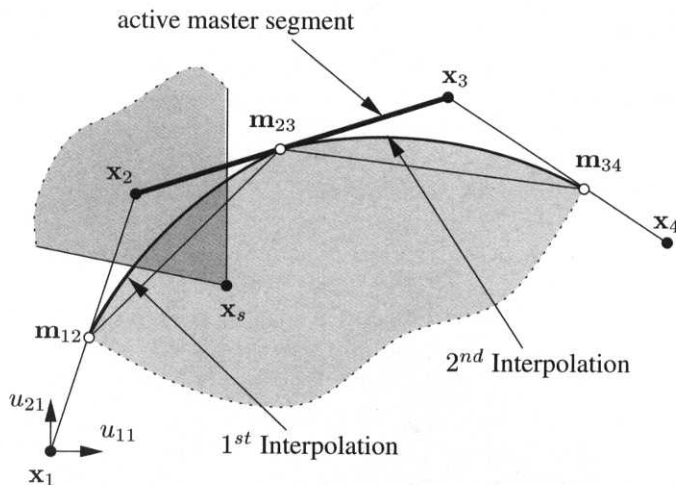
leads to a description of the master surface by only three nodes, which then results in a four-node contact element when the slave node is added. The bandwidth associated with the smooth four-node contact element is smaller than for the five-node contact element in Section 8.5.2, which yields smaller global equation systems for contact problems.

The geometric description of the contact surfaces is different for four and five node contact elements. When the four-node element is constructed following Figure 8.19, then the surface lies inside the polygon spanned by the straight segments. Hence, not all nodes describing the boundary of the finite element mesh are part of the surface used to establish the contact constraints; see node  $\mathbf{x}_2$  in Figures 8.19 and 8.20. In the case of the five-node discretization, all nodes are part of the contact surface, see Figures 8.14 or 8.15. We note that in the limit case, for contact elements with  $l_e \rightarrow 0$ , both discretizations converge to the correct surface.

To develop the matrix formulation for the four node contact element, we define two interpolating polynomials (see Figure 8.21) by two mid-nodes (mid point between two master surface nodes) and two tangent vectors. Mid-nodes  $\mathbf{m}_{12}$  and  $\mathbf{m}_{23}$  represent end-points of the polynomial while the tangent vectors,  $\mathbf{x}_2 - \mathbf{x}_1$  and  $\mathbf{x}_3 - \mathbf{x}_2$  are defined by a line between master surface nodes. The so defined geometry, when also applied in the same way for the neighbouring segments, ensures  $C^1$  continuity between adjacent contact segments, and hence  $C^1$  continuity on the entire master contact surface. For each active contact segment, two interpolations are evaluated (in Figure 8.21 this would be the interpolation defined by  $\mathbf{m}_{12}$  and  $\mathbf{m}_{23}$  and the interpolation defined by  $\mathbf{m}_{23}$  and  $\mathbf{m}_{34}$ ). The polynomial which has the minimum distance to the slave node  $\mathbf{x}_s$  has to be chosen as the active one. It is used for calculation of the contact residual and the associated tangent matrix. For simplicity, we suppose that the first polynomial defined by  $\mathbf{m}_{12}$  and  $\mathbf{m}_{23}$  is closer to the slave node, thus all vectors and matrices are described with respect to this interpolation. Evaluation of all quantities for the second interpolation is similar as for the first one. According to Figure 8.21, active segment nodes are represented by indices 2 and 3, while neighbouring segment nodes are represented by indices 1 and 4. The slave node



**Fig. 8.20** Finite element discretization using the four-node smooth contact element.



**Fig. 8.21** Description of the “active segment”, nodal coordinates and nodal displacements.

$\mathbf{x}_s$  is represented in the following derivations by index 5. With this notation we can define the nodal displacement vector for a single two-dimensional contact element as

$$\mathbf{u} = \{u_{11}, u_{21}, u_{12}, u_{22}, u_{13}, u_{23}, u_{14}, u_{24}, u_{15}, u_{25}\}, \quad (8.185)$$

where the first index describes the direction with respect to a cartesian coordinate system  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ . The second index is the nodal number. With the nodal displacement vector we obtain a relation between the current configuration  $\mathbf{x}$  and the initial configuration  $\mathbf{X}$  of a body, i.e.  $\mathbf{x} = \mathbf{X} + \mathbf{u}$ .

Now two  $C^1$  cubic BÉZIER curves  $\mathbf{x}^1(\xi)$  can be stated in terms of Bernstein polynomials:

$$B_i^m(\xi) = \binom{m}{i} \xi^i (1 - \xi)^{m-i} \quad i = 0, \dots, m, \quad (8.186)$$

where  $m$  is the order of the polynomial. For the case  $m = 3$ , the surface is parameterized by

$$\mathbf{x}^1(\xi) = \mathbf{b}_0 B_0^3(\xi) + \mathbf{b}_1 B_1^3(\xi) + \mathbf{b}_2 B_2^3(\xi) + \mathbf{b}_3 B_3^3(\xi), \quad (8.187)$$

where the explicit form of the polynomials is given by (8.170). Vectors  $\mathbf{b}_i$  are the BÉZIER points that define the polygon. The BÉZIER curve is always lying inside the polygon. This property is called the *convex hull property* of BÉZIER curves, e.g. see Farin (1993), which ensures numerical stability of the interpolation.

When interpolation (8.187) is applied to discretize the contact segments, the end points  $\mathbf{b}_0$  and  $\mathbf{b}_3$  are defined by the mid-nodes (see Figure 8.22). The BÉZIER points  $\mathbf{b}_1$  and  $\mathbf{b}_2$  have to be positioned on the tangents  $\mathbf{x}_2 - \mathbf{x}_1$  and  $\mathbf{x}_3 - \mathbf{x}_2$ , respectively.

However, the distance from the end points  $\mathbf{b}_0$  and  $\mathbf{b}_3$  is still arbitrary and can vary. The choice of this distance plays an important role in the description of rolling contact, and will be discussed in detail at the end of this section; see also the example in the last section.

For simplicity, let points  $\mathbf{b}_1$  and  $\mathbf{b}_2$  be defined by quarter-nodes, i.e. the points between nodes and mid-nodes. This choice ensures  $C^1$ -continuity, see Farin (1993).

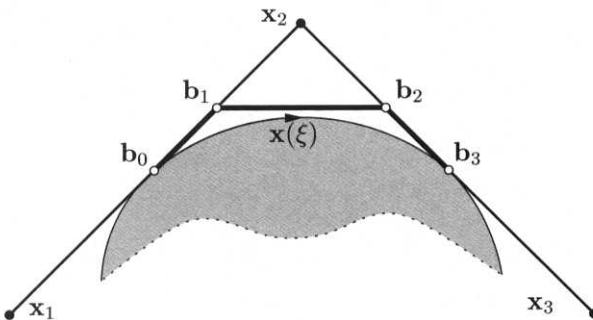
In the case of rolling contact of cylindrical objects we have to consider circular cross-sections which move on a given surface, such as wheels or tyres. Smoothing of the contact discretization is then absolutely necessary, since sliding and rolling of such bodies are significantly influenced by the contact surface geometry. The element based on the cubic BÉZIER interpolation shows a strong sensitivity with respect to the location of the second  $\mathbf{b}_1$  and the third  $\mathbf{b}_2$  BÉZIER point (quarter-nodes; see Figure 8.22). To obtain an optimal surface discretization, for the moment we define the positions of these BÉZIER points as variable (see Figure 8.23) using the parameter  $\alpha \in [0, 1]$ :

$$\mathbf{b}_1 = \mathbf{b}_0 + (\mathbf{x}_2 - \mathbf{b}_0) \alpha, \quad (8.188)$$

$$\mathbf{b}_2 = \mathbf{x}_2 + (\mathbf{b}_3 - \mathbf{x}_2) (1 - \alpha). \quad (8.189)$$

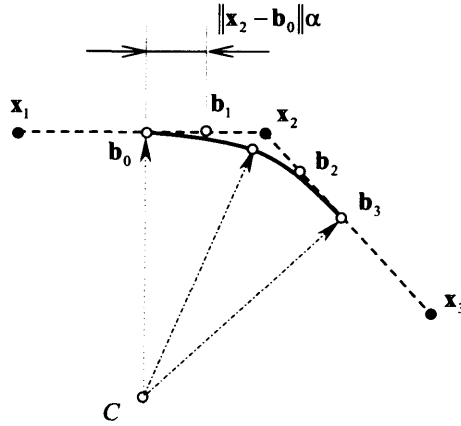
For a circular cross-section, when only a small number of elements is used, special care has to be taken when choosing  $\alpha$ , since otherwise the circular geometry might be very poorly represented, see also Figure 8.18. Different possibilities can be explored to obtain a better approximation of the geometry in the case of the BÉZIER interpolation:

1. When all master segments have the same length, the parameter  $\alpha$  can be computed from the requirement that the point  $\mathbf{x}^1$  ( $\xi = 1/2$ ) for a BÉZIER interpolation should be the same as the point  $\mathbf{x}^1$  ( $\xi = 1/2$ ) for a Hermite interpolation. This results in parameter  $\alpha = 2/3$ , which is completely independent from the problem geometry. It is easy to see that in this case, BÉZIER and Hermitian



**Fig. 8.22** BÉZIER interpolation and points.





**Fig. 8.23** Definition of the parameter  $\alpha$  for variable Bézier points.

interpolation coincide. Note, however, that this is only valid for the case of master segments with equal length.

2. Another way of computing the parameter  $\alpha$  is based on the requirement that the radius, defined as the distance from the circle centre  $C$  to the point on the polynomial (Figure 8.23), has to be the same for points  $\mathbf{x}$  ( $\xi = 0$ ) and  $\mathbf{x}$  ( $\xi = 1/2$ ). This approximates the curvature of the interpolation pointwise. It yields a BÉZIER interpolation which is closer to the ideal arc shape than the interpolation with  $\alpha = 2/3$ . In this formulation, in comparison with the previous one, the parameter  $\alpha$  depends upon the geometry, and has to be recalculated for different surface geometries. This is especially inconvenient, hence in finite deformation problems the geometry description has to be defined based on the current configuration.
3. Finally, one can approximate the curvature of the contact element by an average procedure. This means that the average curvature of the contact segment interpolation has to be equal to the radius of curvature  $1/R$  given by the real geometry. This yields the integral

$$R \approx \frac{1}{S} \int_{\Gamma_e} \|\mathbf{x}_{,\xi\xi}^1\| d\Gamma, \quad (8.190)$$

where  $\mathbf{x}_{,\xi\xi}^1$  is the second derivative of the BÉZIER interpolation with respect to the path parameter  $\xi$ , and  $S$  denotes the length of the segment. The evaluation of this integral using an interpolation with (8.188) and (8.189) yields, for each contact segment, a value for  $\alpha$ . Again,  $\alpha$  depends upon the current geometry.

The significance of smoothing the circular geometry, as well as the correct choice of parameter  $\alpha$ , is demonstrated further by means of the examples below.

Now we derive the residual vector and the tangent matrix associated with the contact element described. The normal gap function is defined in (3.6). For the tangential gap we have to distinguish between the stick and slip phases with the relative tangential displacements given in (3.14) and (3.22). The stick condition was stated in (8.5) for the node-to-segment element with linear segment geometry which cannot be used here in this form. Instead, we have to compute the path on the surface, see Figure 8.24. As the cubic BÉZIER polynomial is defined explicitly, the path length given with (3.26) by

$$s(\bar{\xi}) = \int_0^{\bar{\xi}} \|\mathbf{x}_{,\xi}^1(\xi)\| d\xi = \int_0^{\bar{\xi}} \sqrt{\left(\frac{\partial x_1^1(\xi)}{\partial \xi}\right)^2 + \left(\frac{\partial x_2^1(\xi)}{\partial \xi}\right)^2} d\xi, \quad (8.191)$$

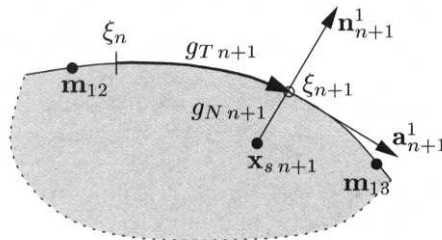
where  $x_i^1$  represents the  $i$ th component of the vector  $\mathbf{x}^1(\xi)$ . There is no explicit analytical solution of the integral (8.191), hence numerical (e.g. GAUSS) integration has to be applied. This leads to the definition of the relative tangential gap function:

$$\Delta \mathbf{g}_{Tn+1}^{st} = [s(\xi_{n+1}) - s(\xi_n)] \mathbf{a}^1(\xi_{n+1}). \quad (8.192)$$

Considering the fact that the tangential stress vector is a function of the displacement  $\mathbf{u}_{n+1}$  and the parameter  $\bar{\xi}_{n+1}$ , one has to use the chain rule to derive the expression for  $i$ th component of the residual vector related to the displacement  $u_{in+1}$ . For stick with a tangential contact stress computed from the elastic part (see (4.21)) this yields

$$\begin{aligned} \{R_{n+1}^{st}\}_i &= \varepsilon_N g_{Nn+1} \left[ \frac{\partial g_{Nn+1}}{\partial u_{in+1}} + \frac{\partial g_{Nn+1}}{\partial \bar{\xi}_{n+1}} \frac{\partial \bar{\xi}_{n+1}}{\partial u_{in+1}} \right] + \\ &+ \mathbf{t}_{Tn+1}^{tr} \cdot \left[ \frac{\partial \Delta \mathbf{g}_{Tn+1}^{st}}{\partial u_{in+1}} + \frac{\partial \Delta \mathbf{g}_{Tn+1}^{st}}{\partial \bar{\xi}_{n+1}} \frac{\partial \bar{\xi}_{n+1}}{\partial u_{in+1}} \right], \quad (8.193) \end{aligned}$$

where index  $i$  has the range  $i = 1, \dots, 10$ . One can show, based on the expression for  $g_N$  in (3.7), that  $\partial g_N / \partial \xi$  is equal to zero, and hence it can be neglected in (8.193).



**Fig. 8.24** Sliding path of the slave node on the master surface.

However from a formal point of view, the latter result is not known *a priori*. Hence it was included when we used the Mathematica package *AceGen* Korelc (1997) for the automatic derivation of matrix formulae. The partial derivative of the path parameter with respect to the displacements is given by

$$\frac{\partial \bar{\xi}_{n+1}}{\partial u_{i n+1}} = - \frac{\partial \|\mathbf{r}_{n+1}(\bar{\xi}_{n+1})\|}{\partial u_{i n+1}} / \frac{\partial \|\mathbf{r}_{n+1}(\bar{\xi}_{n+1})\|}{\partial \bar{\xi}_{n+1}}. \quad (8.194)$$

The tangent matrix is obtained by the linearization of the residual vector

$$\{K_{n+1}^{st}\}_{ij} = \frac{\partial \{R_{n+1}^{st}\}_i}{\partial u_{j n+1}} + \frac{\partial \{R_{n+1}^{st}\}_i}{\partial \bar{\xi}_{n+1}} \frac{\partial \bar{\xi}_{n+1}}{\partial u_{j n+1}}, \quad (8.195)$$

where index  $j = 1, \dots, 10$ . In the finite element code, when calculating the contact contribution of a single slave node, the residual vector is calculated for every iteration within the current time interval.

When sliding occurs, the residual vector is obtained by inserting the tangential stress stemming from the time integration of COULOMB'S law, (9.118), into (5.33):

$$\begin{aligned} \{R_{n+1}^{sl}\}_i &= \varepsilon_N g_{N n+1} \left[ \frac{\partial g_{N n+1}}{\partial u_{i n+1}} + \frac{\partial g_{N n+1}}{\partial \bar{\xi}_{n+1}} \frac{\partial \bar{\xi}_{n+1}}{\partial u_{i n+1}} \right] + \\ &+ \mathbf{t}_{T n+1}^{sl} \cdot \left[ \frac{\partial \mathbf{g}_{T n+1}^{sl}}{\partial u_{i n+1}} + \frac{\partial \mathbf{g}_{T n+1}^{sl}}{\partial \bar{\xi}_{n+1}} \frac{\partial \bar{\xi}_{n+1}}{\partial u_{i n+1}} \right]. \end{aligned} \quad (8.196)$$

The tangent matrix then follows as

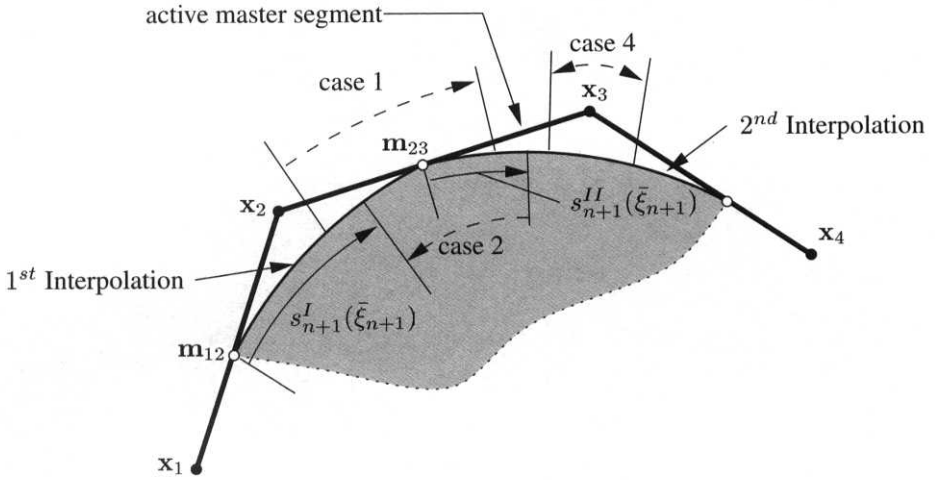
$$\{K_{n+1}^{sl}\}_{ij} = \frac{\partial \{R_{n+1}^{sl}\}_i}{\partial u_{j n+1}} + \frac{\partial \{R_{n+1}^{sl}\}_i}{\partial \bar{\xi}_{n+1}} \frac{\partial \bar{\xi}_{n+1}}{\partial u_{j n+1}}. \quad (8.197)$$

For COULOMB'S frictional law the relative tangential gap follows from (9.119) for the slip case, and is given by

$$\mathbf{g}_{T n+1}^{sl} = \mathbf{g}_{T n}^{sl} + \frac{1}{c_T} (\|\mathbf{t}_{t n+1}^{tr}\| - \mu p_{N n+1}) \mathbf{n}_{T n+1}^{tr},$$

which has to be inserted into (8.196). Equations (8.193), (8.195), (8.196) and (8.197) provide the basis for the automatic code generation using *AceGen*.

**Continuity of history variables between adjacent segments in the slip case.** For both polynomial formulations described in the previous sections,  $C^1$  continuity of adjacent interpolations is ensured. In the frictionless case, where no history variables are needed, there is no influence on the residual vector and the tangent matrix if sliding occurs over several adjacent interpolations and segments. For frictional problems, where path length and traction vector are history variables, extra considerations are needed when sliding of the slave node over adjacent segments and interpolations occurs.



**Fig. 8.25** Description of the path length modification.

According to Figure 8.25, let an interpolation which is defined between nodes  $x_1$  and  $x_3$  be called the *first* interpolation. Furthermore, an interpolation which is defined between  $x_2$  and  $x_4$  is called the *second* interpolation. For both interpolations the path length is calculated via the integral from zero to  $\bar{\xi}_{n+1}$ . Let the path length for the *first* interpolation be denoted by the symbol  $s_{n+1}^I(\bar{\xi}_{n+1})$ , and the path length for the *second* interpolation by the symbol  $s_{n+1}^{II}(\bar{\xi}_{n+1})$ . These values are calculated according to equation (8.191) for each iteration, and saved as a variable which describes the history of the sliding path of the slave node.

The history variable  $s_n(\bar{\xi}_n)$  that represents the path length for the last converged state is constant during the iterations within a time increment. If, within the same active segment, the slave node slides from the *first* interpolation to the *second*, or vice versa, then the path length  $s_{n+1}(\bar{\xi}_{n+1})$  used in equation (8.192) has to be modified. The following possibilities have to be considered when updating the history variable describing the path length:

1. If sliding from the *first* to the *second* interpolation occurs, then  $s_{n+1}(\bar{\xi}_{n+1})$ , used for calculation of  $(s_{n+1}(\bar{\xi}_{n+1}) - s_n(\bar{\xi}_n))$ , is modified according to

$$s_{n+1}(\bar{\xi}_{n+1}) = s_{n+1}^{II}(\bar{\xi}_{n+1}) + s_{n+1}^I(\xi_{n+1} = 1). \quad (8.198)$$

2. If, within the same active segment, the slave node slides from the *second* to the *first* interpolation, then the path length  $s_{n+1}(\bar{\xi}_{n+1})$  has to be modified according to

$$s_{n+1}(\bar{\xi}_{n+1}) = s_{n+1}^I(\bar{\xi}_{n+1}) - s_{n+1}^{II}(\xi_{n+1} = 1). \quad (8.199)$$

3. If there is no sliding between interpolations of the same active segment, then for the *first* interpolation we set

$$s_{n+1}(\bar{\xi}_{n+1}) = s_{n+1}^I(\bar{\xi}_{n+1}), \quad (8.200)$$

or for the *second* interpolation,

$$s_{n+1}(\bar{\xi}_{n+1}) = s_{n+1}^{II}(\bar{\xi}_{n+1}). \quad (8.201)$$

4. If the slave node slides between two adjacent segments, there is no need for a modification of the current path length  $s_{n+1}(\bar{\xi}_{n+1})$ . Hence equations (8.200) and (8.201) are used. For such a case, procedures described by (8.198) and (8.199) could be activated unnecessarily. This can happen, for example, when an interpolation that was the *second* one for the active segment changes to the *first* interpolation, and vice versa. To suppress the use of procedures (8.198) and (8.199) in such a case, another additional history variable (switch) is introduced to monitor whether an active master segment number has changed or not.

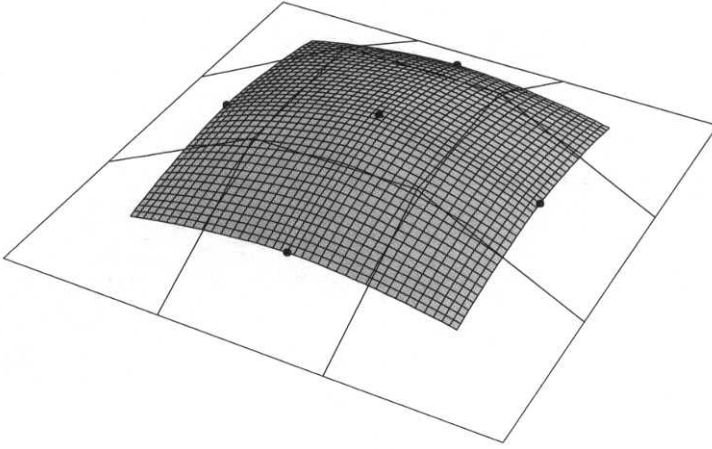
As mentioned before, history variables which need to be used in the next time steps are not influenced by modifications (8.198) and (8.199).

### 8.5.4 Three-dimensional contact discretization

In the last few years, several groups have worked on the design of three-dimensional finite contact elements based on smooth surface interpolations. It is a well known fact from finite element formulations based on KIRCHHOFF plate theory that elements which fulfil  $C^1$  continuity are very hard to design. This is also true for triangular as well as for quadrilateral elements. Hence, the generation of  $C^1$  smooth surfaces is not as straightforward as in two-dimensional applications (see the previous sections). However, there are different possibilities to construct  $C^1$  continuous surface approximations; some will be discussed in the following. Before doing so, we stress that these approximations discretize contact surfaces in the current configuration which are not known *a priori*, since we want to allow for finite deformations and sliding. For rigid surfaces it is much simpler to develop a smooth surface interpolation. Here all CAD tools can be applied immediately; for such interpolations, see Wriggers and Imhof (1993), Heegaard and Curnier (1993), Heege and Alart (1996) or Hansson and Klarbring (1990).

For the  $C^1$  interpolation of the current surface geometry, the approaches were developed for triangular and hexahedral surface meshes.

Interpolations based on hexahedral elements were developed in, for example, Pietrzak (1997), Pietrzak and Curnier (1997) and Dalrymple (1999). Basically, a tensor product representation of one-dimensional BÉZIER or HERMITE polynomials is used to interpolate the master surfaces in the contact interface. In that case, a three-dimensional quadrilateral BÉZIER surface of the order  $m$  is defined in explicit

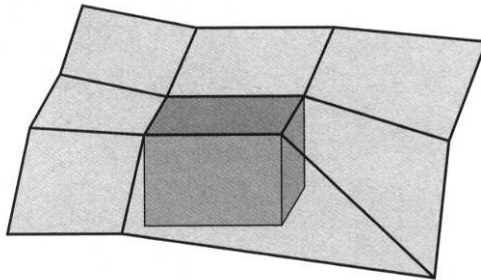


**Fig. 8.26** Eight neighbouring surfaces for brick elements.

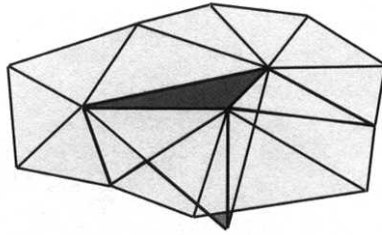
form, see Farin (1993), as

$$\mathbf{x}(\xi^1, \xi^2) = \sum_{i=0}^m \sum_{j=0}^m \mathbf{b}_{i,j} B_i^m(\xi^1) B_j^m(\xi^2). \quad (8.202)$$

$B_i^m(\xi^1)$  and  $B_j^m(\xi^2)$  are the BERNSTEIN polynomials which depend upon the convective surface coordinates  $\xi^1$  and  $\xi^2$ . They have already been defined in equation (8.186) for one-dimensional applications. This definition needs a discretization in which the quadrilateral contact surface has to be surrounded by eight neighbouring quadrilateral contact surfaces, as shown in Figure 8.26. Hence a structured mesh must be provided for the contact analysis. Thus, smoothing of hexahedral



**Fig. 8.27** Interpolation with less than eight neighbouring surfaces.

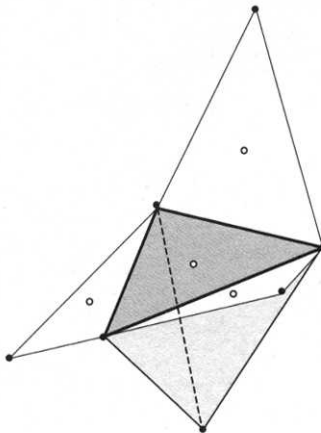


**Fig. 8.28** Triangular contact surfaces for the tetrahedral element.

elements by (8.202) has a limitation since automatic mesh generation for arbitrary three-dimensional bodies does not lead to such a regular form of the mesh.

Hence this approach cannot easily handle a case related to an unstructured mesh, as shown in Figure 8.27. For such cases, a smooth interpolation could be developed, but this would lead in general to a special treatment of different patches with, for example, 5, 6, 7 or 8 adjacent elements, and therefore would lead to very complex coding.

The use of the tetrahedral elements (Figure 8.28) in the finite element analysis of bodies with complex geometry has a significant advantage that more robust automatic mesh generators exist. This is also true for re-meshing when compared with programs for hexahedral elements. The interpolation of the contact surface for triangular elements is based in Krstulovic-Opara et al. (2002) on the three adjacent elements, as shown in Figure 8.29. The  $C^1$  continuous discretization presented is based on a smoothing of the active triangular contact surface (i.e. master surface where contact



**Fig. 8.29** Tetrahedral contact element.

occurs) by using six quartic BÉZIER surfaces. All six quartic BÉZIER surfaces are  $C^1$  continuous with respect to each other, thus there are no jumps in surface normal field between these surfaces. However, directly at the vertices of the element the  $C^1$  continuity is lost pointwise.

The formulation of the sliding path for the frictional problems becomes complex for such surfaces. To overcome this problem, a mapping of the current solution in the last converged configuration is used in Krstulovic-Opara et al. (2002). In the last converged configuration a cubic curve is introduced as an approximation of the path length. Such a definition of sliding, which was first developed in Agelet de Saracibar (1997), enables the description of sliding for the cases of large steps, i.e. sliding of a slave node over several segments during one loading step.

Due to the high order interpolation of the surface by quartic BÉZIER interpolations, the variations and linearizations of geometrical quantities like the normal and relative tangential gap are very complicated. Thus, a general symbolic description of all the relations needed to derive the contact element was applied. It was used in Krstulovic-Opara et al. (2002) to be able to introduce the linearization of the contact residual on a high abstract mathematical level. For this task a computer algebra system for symbolic derivation, automatic differentiation and code generation was used to numerically generate the proposed element routines. The system consists of two major components: (i) the *Mathematica* package *AceGen* that automatically derives formulae and generated general code, see Korelc (1997) and Korelc (2000a)), and (b) the *Mathematica* package *Computational Templates* Korelc (2000b), with prearranged modules for creation of the finite element codes. For a detailed treatment of these subjects the reader should consult the references cited.

In conclusion, research work for a better and simpler way to represent smooth surfaces is still needed. Some ideas, based on patchwise formulations and subdivisions which were proposed in Cirak et al. (2000) for thin shells, could be a step in the right direction for contact analysis with smooth surface interpolations.

## 8.6 NUMERICAL EXAMPLES

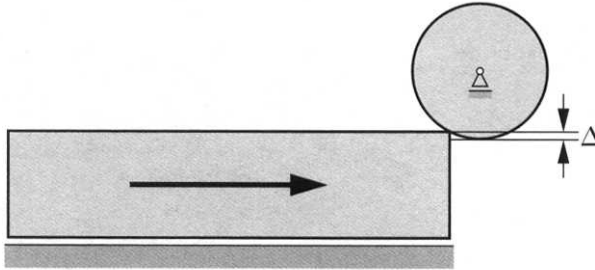
The interpolations derived above have been implemented in the current version of the finite element analysis program FEAP, see Zienkiewicz and Taylor (1989). To show the performance of the two different smooth contact discretizations, we compare the smooth approach with the node-to-segment contact formulation developed in Section 8.1 for straight segments.

### 8.6.1 The sheet/plate rolling simulation

The advantage of the smooth contact elements is evident when this discretization is applied within sheet/plate metal forming simulations using a LAGRANGE formulation. This is particularly true when deformations and stresses of the roller are not of primary interest. In such a situation, smoothing of the cylindrical roller, here defined



as the master surface, enables a reduction in the required number of elements needed for discretization of the roller. In Figure 8.30 the geometry of a sheet/plate forming

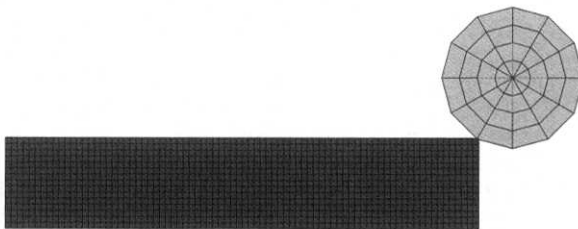


**Fig. 8.30** Problem description.

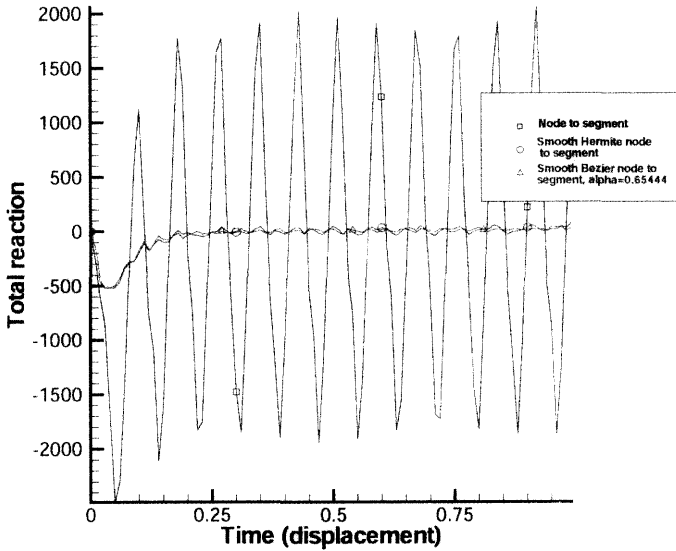
process is shown. A displacement is applied on the horizontal plate. The fixed steel roller of diameter 140 mm rotates around its axis. The plate has a size of  $478 \times 85$  mm. It is supported in the vertical direction on its lower edge. Since we only want to show the usefulness of the smooth contact discretization, both parts are modelled by a hyperelastic constitutive law of the neo-HOOKE type. The material behaviour of the roller is modelled with a bulk modulus of  $K = 1.75 \cdot 10^5$  MPa and a shear modulus of  $G = 8.08 \cdot 10^4$  MPa. The plate is characterized by a bulk modulus of  $K = 55.6$  MPa and a shear modulus of  $G = 3.4$  MPa. The cylinder is discretized by the relatively coarse mesh, as depicted in Figure 8.31. The overlapping between roller and plate is assumed to be approximately  $\Delta \approx 5$  mm.

The contact is modelled using the simple node-to-segment (NTS), the Hermite node-to-segment, and the Bézier node-to-segment contact element, with the parameter  $\alpha = 0.65444$ . This parameter is evaluated from the requirement for the same distance between the circle's centre and the points  ${}^{n+1}\mathbf{x}$  ( ${}^{n+1}\xi = 0$ ) and  ${}^{n+1}\mathbf{x}$  ( ${}^{n+1}\xi = 1/2$ ).

Note that there is a difference in the master surface geometry between the standard and smooth node-to-segment elements due to the different surface approximations (see Figures 8.2 and 8.22). Hence, when the reactions for the central cylinder node are computed, there is a difference in the reaction forces due to the differences in the



**Fig. 8.31** Finite-element-model.



**Fig. 8.32** Horizontal reactions for the central node.

master surface geometry of both approaches. Thus, we only compare the horizontal reaction at the central node of the roller, which has to be zero in the case of rolling.

Figure 8.32 shows the horizontal reaction force. We observe that the solution of using smooth discretization is much better than that obtained with the standard node-to-segment element.

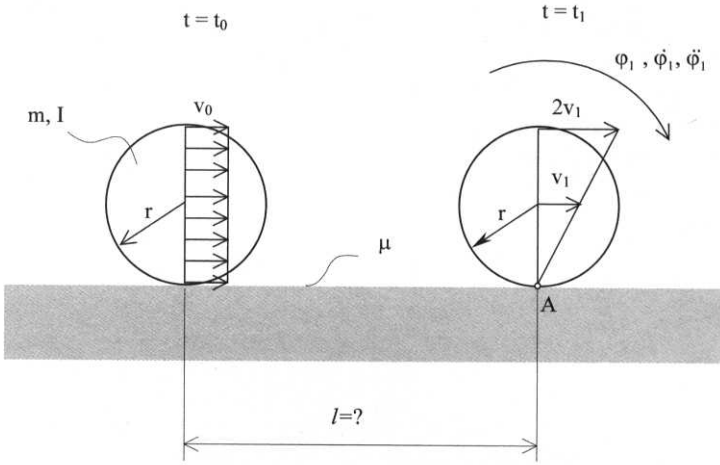
### 8.6.2 Simulation of a sliding and rolling wheel

This example of a disk which is put on a plane with a given contact initial velocity in the horizontal direction is suggested as a benchmark test for two-dimensional contact elements with friction. It has the advantage that the analytical solution is known.

Here a disk with radius  $r$ , density  $\rho$ , COULOMB friction coefficient  $\mu$ , and elasticity parameters for a neo-HOOKE material is put on a flat surface under gravity loading. When the initial velocity  $v_0$  for all points of the disk is applied, the disk starts to slide without rolling. Within the sliding process the disk starts to roll until finally pure rolling occurs, see Figure 8.33.

The analytical solution yields the time  $t_R$  and distance  $l_R$  after pure rolling starts:

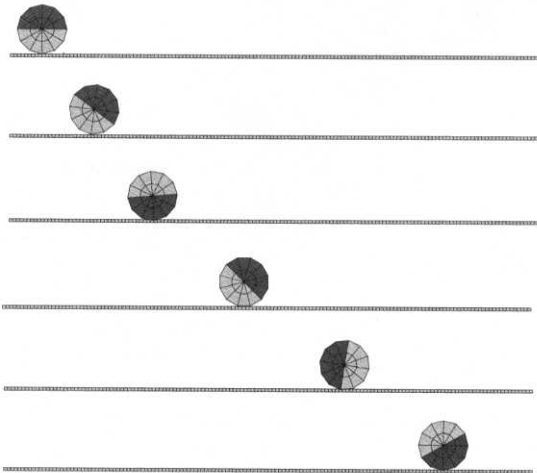
$$t_R = \frac{1}{3} \frac{v_0}{\mu g} \quad l_R = \frac{5}{18} \frac{v_0^2}{\mu g}. \quad (8.203)$$



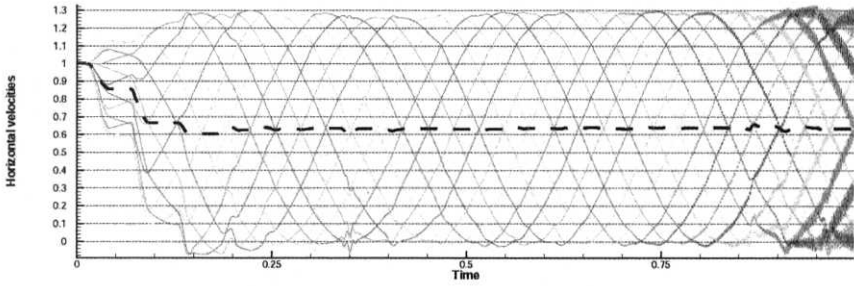
**Fig. 8.33** Description of the rolling wheel problem.

With a frictional coefficient of  $\mu = 0.3$ , the gravitational acceleration  $g = 9.81 \text{ m/s}^2$ , and an initial velocity of  $v_0 = 1 \text{ m/s}$ , it follows from equation (8.203) that the length and time, when pure rolling starts, are  $l_R = 0.094 \text{ m}$  and  $t_R = 0.113 \text{ s}$ .

The FE simulation of the sliding/rolling disk is performed for a steel disk with radius  $r = 0.04 \text{ m}$ , density  $\rho = 7850 \text{ kg/m}^3$ , frictional coefficient  $\mu = 0.3$  and elasticity parameters  $K = 1.75 \cdot 10^5 \text{ MPa}$ ,  $G = 8.08 \cdot 10^4 \text{ MPa}$ . The dynamical



**Fig. 8.34** FE simulation of rolling disk with contact.



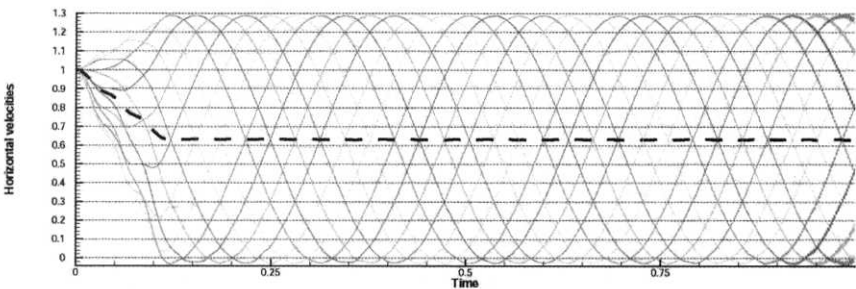
**Fig. 8.35** Straight node-to-segment (NTS) frictional contact element.

problem is solved using the NEWMARK method (see Section 9.2.2) with parameters  $\beta = 0.25$  and  $\gamma = 0.5$ . Figure 8.34 shows the discretization for the disk and the plane surface, and different stages of the sliding/rolling motion of the disk.

Figures 8.35 and 8.36 show the horizontal velocities of the master nodes at the surface and the central nodes of the disk. The time when pure rolling starts can be detected from these diagrams by comparison with the results of the analytical solution. Once the horizontal velocity of a master node is zero, while the velocity of the opposite master node (here plotted with the same grey scale) is as twice as big as the velocity of the central node (dashed line), pure rolling starts. From Figure 8.35 we obtain the result that pure rolling occurs at the distance  $l_R = 0.124$  m for the straight NTS contact. The overall solution behaviour for this discretization is characterized by jumps and non-physical separations from the base surface.

Figure 8.36 depicts the response for smooth discretization using HERMITE and BÉZIER interpolations, the latter with  $\alpha = 2/3$ . In these simulations pure rolling started at  $l_R = 0.099$  m which is close to the analytical solution in (8.203).

As can be observed by comparing Figures 8.35 and 8.36, the straight contact element discretization leads to erroneous results, whereas the smooth interpolation yields better results, close to the analytical solution.



**Fig. 8.36** Smooth HERMITE and BÉZIER ( $\alpha = 2/3$ ) frictional contact elements.

Both examples presented demonstrate the importance of contact surface smoothing in certain applications. Sudden normal changes, that are common for the standard node-to-segment contact formulations, introduce significant errors in the finite element discretization of contact phenomena, especially when rolling is concerned.

In comparison with the HERMITE interpolation in Section 8.5.1, the BÉZIER approach in Section 8.5.2 is simpler. However, one has to bear in mind that the parameter  $\alpha$  has to be chosen within the BÉZIER approach for special cases, as discussed in Section 8.5.3.

# 9

---

## *Solution Algorithms*

In this chapter we consider algorithms which are essential for the treatment of contact problems. These are applied to the discretized problem, which are derived using the formulations in Chapters 5, 6 and 7.

The algorithms for detecting contact are of utmost importance, since in complex engineering applications like a car crash or fragmentation of brittle materials, contact occurs at non-predictable places on the surfaces of the bodies involved. In such situations, contact search has to be performed in every time or load step of the numerical simulation. Additionally, when fractured materials are considered, many new contact surfaces are generated due to the splitting of elements into pieces once fracture occurs. All together, these simulations require complex software which has to be designed to be most efficient. This is necessary since the number of operations needed to check  $O(N)$  surfaces of a discretized domain is  $O(N^2)$ . If such a check has to be carried out at each iteration step, then the search will dominate the overall time for the computation. Hence algorithms have to be developed which reduce the frequency of the search and its complexity.

In general, we have to distinguish between global and local algorithms:

- global algorithms
  1. contact search
  2. solution of the variational inequalities
- local algorithms
  1. contact detection
  2. update of constitutive equations and stresses.

The global algorithms are related to purely geometrical considerations when the contact search has to be done. There are many different approaches for contact detection which will be discussed in section 9.1.

In Sections 9.2 and 9.3 solution methods for non-constraint and constraint problems are discussed. Here, several aspects related to the design of solution methods have to be considered. These are associated with

- the existence of solutions in the domain of interest, and
- the number of solutions in that domain.

The theoretical analysis can be found in the mathematical literature. For general results regarding unconstraint problems, see Vainberg (1964) or Ortega and Rheinboldt (1970). Associated results for contact problems are reported in Hlavacek et al. (1988), Kikuchi and Oden (1988) or Curnier et al. (1992).

The approximation of the solution has to be computed by using adequate algorithms. A direct solution of the equation system resulting from the finite element discretization  $G(\mathbf{u}) = \mathbf{0}$  under the inequality constraints  $C(\mathbf{u}) \geq 0$  due to contact is not possible, due to the nonlinearity of the problem. Hence, iterative algorithms are needed which can be distinguished in the following manner:

- methods based on linearizations,
- minimization methods, or
- reduction methods which yield simpler nonlinear equations.

Before one chooses an algorithm, one has to check the following questions in order to obtain a fast, efficient and robust scheme:

- Does the algorithm converge to the solution?
- How fast is convergence? Does the rate of convergence depend upon the problem size?
- How efficient is the algorithm?
  - number of operations within an iteration,
  - total number of iterations, and
  - memory used.

The first question is related to the global convergence properties of the algorithm. It is essential for the user of such a scheme, since he needs a robust and reliable tool for the solution of his problem. In contact mechanics this is a not completely solved problem, especially when implicit methods are applied to friction problems. However, the other questions raised are essential, in the sense that most of the engineering problems are represented by a three-dimensional discretization which leads to very large number of unknowns. In this respect, the memory used is relevant to keep the process in the CPU, and hence avoid time consuming access to the hard disk. Furthermore, one

should apply solution methods which only need  $O(N)$  operations in order to keep the solution times to a minimum. All these aspects define a vivid research field in which new fast algorithms are constructed based on different mathematical approaches and computer architectures. One cannot, therefore, conclude that an optimal solution procedure exists at the moment for contact problems.

The bandwidth of global algorithms for the solution of variational inequalities is very broad. We should like to mention the simplex method, mathematical programming, active set strategies using LAGRANGE multipliers, sequential quadratic programming, penalty and augmented LAGRANGE techniques, as well as barrier methods. The basis for these different methods with regard to continuum mechanics has been given in Chapter 5. All these techniques have advantages and disadvantages concerning efficiency, accuracy or robustness, and thus have to be applied according to the problem at hand. In Section 9.2 we discuss methods to solve unconstrained problems or problems with equality constraints which are applied within the global algorithms for the solution of variational inequalities. Thereafter, in Section 9.3, we sketch some of the global algorithms which are mainly applied to contact problems.

Local algorithms have to be employed once the global contact search has found possible contact partners. A more accurate local search has to be carried out in order to find the correct faces and elements in contact, and in the case of contact, the correct projection of nodes onto the surface of the contacting partner has to be determined. The latter algorithms have already been described in Chapter 7. The local contact search will be discussed in Section 9.1.2. The update algorithms for the contact stresses, especially the tangential stresses due to friction, have been settled. In this case, the so-called projection methods or return mapping schemes yield the most efficient and robust treatment. Due to the fact that an algorithmic tangent operator can be constructed, this technique can be incorporated into a NEWTON-RAPHSON scheme; see Section 9.5 for further details.

Furthermore, algorithms also have to be devised for coupled problems, which may be necessary in the case of thermo-mechanical coupling or for fluid-structure interaction problems. Algorithms for coupled problems, like staggered schemes, depend upon the type of coupling, and thus have to be designed with special care regarding robustness and efficiency. Algorithms for thermo-mechanical contact problems can be found in Section 10.5.

## 9.1 CONTACT SEARCH

One of the biggest problems in computational contact mechanics is the search for contact between solids. This is especially true when the problem is such that the solids can break, and hence during the solution process several thousands of discrete elements will originate from the initial set up. Examples for these types of dynamic and impact problems can be found in Munjiza and Owen (1993), Munjiza et al. (1995), Camacho and Ortiz (1996) or Kane et al. (1999). In such cases, as also in discrete element or rigid body simulations, most of the computing time has to be devoted to the search, since these computations are usually based on explicit techniques. Thus,

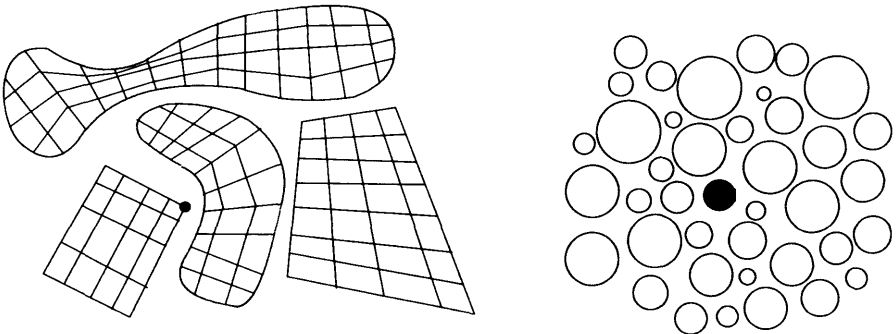


fast algorithms for the detection of contact are needed. The most recent methods are presented in the overview paper by Williams and O'Connor (1999), see also the references therein.

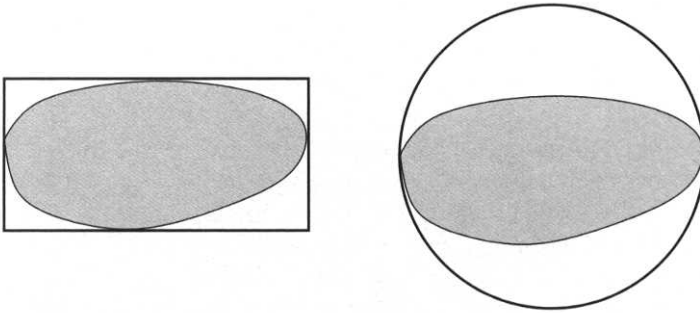
Static contact problems with large deformations also need fast and reliable search algorithms, since contact areas are not known *a priori* and can change within a load step considerably. Furthermore, self-contact has to be included, too. The search for an active set of contact constraints is not trivial in this case, since a surface point of a body may contact any portion of the surface of another body. Such a point can even come into contact with a part of the surface of its own body. Thus the search for the correct contact location eventually needs considerable effort, depending on the problem. An implementation where each node of a surface is checked against each element surface in the mesh is too exhaustive and computationally inefficient, thus refined algorithms have to be constructed. This is especially true when the contact of more than two bodies has to be considered, or when self-contact is possible. A relatively simple situation is depicted in Figure 9.1 (a), where four different finite element discretizations might come into contact. One can immediately observe that it does not make sense to check the contact of one node against each node on the surface of the other discretizations. This effect is even more pronounced in the set of spheres shown in Figure 9.1 (b). Here the possible contact partners of the black sphere have to be found. It is clear that only the spheres in the vicinity of the black sphere should be checked. One thus has to search for the neighbouring objects first, and then perform the local search for real contact.

Thus, in general, contact search has to be split into two phases:

- (I) the spatial search for objects/finite elements which might possibly come into contact; and
- (II) the determination of pairs of objects/finite elements which actually intersect and hence are in contact (contact detection).



**Fig. 9.1** (a) Discrete contact FEM, (b) Discrete contact spheres.

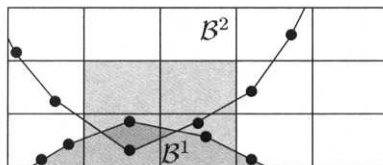


**Fig. 9.2** Bounding box and surrounding sphere of an object.

In phase (I) one orders the finite elements lying on the surface of the solid by a sorting algorithm. As discussed in Williams and O'Connor (1995) it is sufficient in this phase to represent the objects by a coarse description, e.g. bounding boxes or surrounding spheres, see Figure 9.2. The main point when using such simple descriptions is that one can use these to design fast and inexpensive (computationally efficient) sort and search routines. Within the global search a hierarchical structure can be set up to find out which bodies, part of the bodies, surfaces or parts of the surfaces are able to come into contact within a given time step or displacement increment, e.g. see Zhong and Nilsson (1989), Zhong (1993) or Williams and O'Connor (1995).

One of the first implementations, applied within the finite element method to find solutions to large deformation contact problems, see Benson and Hallquist (1990), was called the bucket search. The space is subdivided into cells or buckets, as shown in Figure 9.3. Only the grey bucket is searched, which leads with careful coding to a scheme with  $O(N)$  operations.

Up to now, several different methods have been developed, and applied to determine the possible contact partners. Recently, a considerable impact has come from discrete finite element methods where several thousand particles have to be included in the contact search. Methods like space cell decomposition have been considered by Belytschko and Neal (1991); a combination with the binary tree search can be found in Munjiza et al. (1995); whereas Williams and O'Connor (1995) rely on heapsort algorithms for the global search. More advanced algorithms are the NBS algorithm,



**Fig. 9.3** Bucket search for contact.

see Munjiza and Andrews (1998), the alternating digital tree method from Bonnet and Peraire (1991), a space filling curve technique, see Diekmann et al. (2000), or the Double-Ended Spatial Sorting (DESS) algorithm constructed in Perkins and Williams (1995).

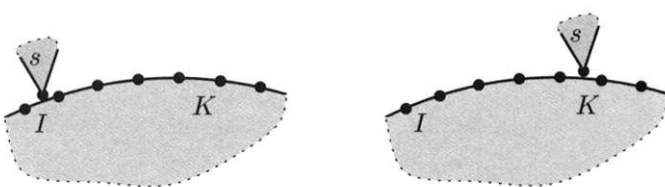
In phase (II) the intersection of the objects is checked, and when intersection occurs the actual contact point, the associated penetration and the local slip are determined using equations (3.3), (3.7) and (3.19). In the special case of contact between a deformable and a rigid body, the rigid body can be described by implicit functions such as superquadrics, see Williams and Pentland (1992). This leads to a simple and efficient contact check for points lying on the surface of the deformable body. For the special case of cylinders or ellipses, see also Hallquist et al. (1992). However, the evaluation of functions describing the superquadrics involves the computation of trigonometric or exponential functions, and is thus expensive. Other representations have therefore been constructed which are based on discrete functions, see Williams and O'Connor (1999). Other methods are the node-to-segment algorithm, see Hallquist (1979), or the pinball technique by Belytschko and Neal (1991).

Now we discuss some techniques which can be applied for the spatial search in phase (I) and the contact detection in phase (II).

### 9.1.1 Spatial search, phase (I)

In the spatial search we have to distinguish between problems in which the evaluation of the deformation is slow or predictable, and cases where arbitrary deformations occur. In the first problem class, we distinguish for finite elements the cases of small deformations and large deformations:

- The problem involves only small deformations. Hence, the position of the solids does not change much, and we can use a linear description for the contact kinematics and assume that the topology of the system is fixed, see Figure 7.1. Therefore, all possible connections and neighbouring pairs are known, and only contact detection is necessary, which itself can be evaluated in an efficient way.
- Large deformation occurs with a smooth movement of a contact point over adjacent elements of the contacting body. In that case (see Figure 9.4) we have to use spatial search to locate the neighbouring segment to a slave node  $s$  which has changed during the iteration from segment  $I$  to  $K$ .

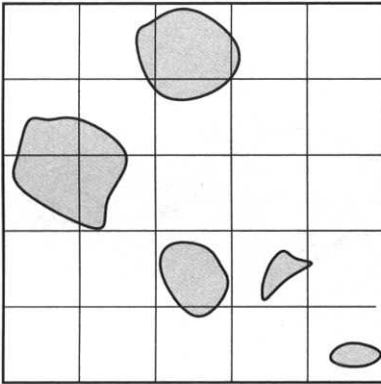


**Fig. 9.4** Large sliding of a node over a surface.

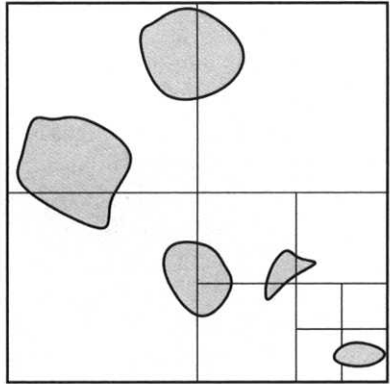
In the second problem class we always have to check possible contact for the present deformation state; see the simple situation depicted in Figure 9.4. This means every object/finite element has to be checked against every other object/finite element. When the contact problem is discretized by  $N$  finite elements, the order of contact checks is  $N^2$ . Associated algorithms result, even for not too large numbers of  $N$ , in very slow algorithms, and thus are called exhaustive algorithms. Hence one has to construct fast algorithms for this task. The following schemes have been developed to date:

- Grid cell algorithms, based on a subdivision of the space which contains the discrete solids (simulation space) into uniform cells, see Figure 9.5(a). This algorithm type works well when the elements are equally distributed within the cells. If clustering of the elements in a few cells occurs, then this algorithm

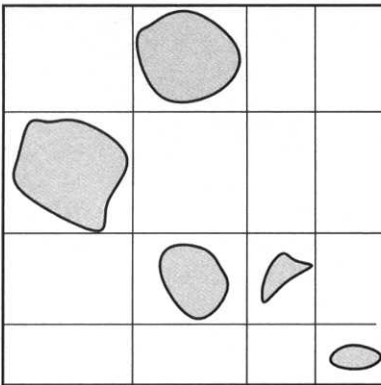
(a) Grid cells



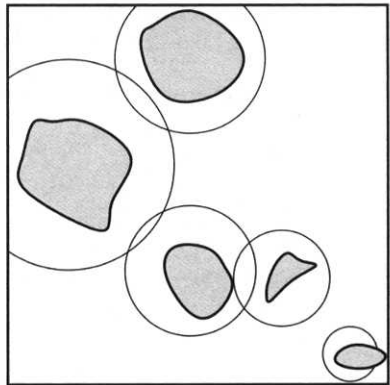
(b) Quadtree, binary tree



(c) Adaptive grid



(d) Body-based cells



**Fig. 9.5** Different sorting strategies.

has no advantages. A variant of this algorithm is to use adaptive grids which accounts for the heterogeneity of the element distribution in space, see Figure 9.5(c). However, in this scheme, the cost of computing the adaptive grid is not negligible.

- A fast method for spatial search is the octree method, which again is based on a grid of rectilinear cells, see Figure 9.5(b), but in this case only cells which contain finite elements are kept in the structure of the tree. As pointed out in Williams and O'Connor (1999), the time for search depends upon the first construction of the octree. With special techniques like balancing the tree branches and minimization of the octree depth, one can reduce the search time, see e.g. Knuth (1973). The time to create such an octree is  $O(N \log N)$  operation, and its evaluation time is of the same order.
- Another method developed by Williams and O'Connor (1995) is the spatial heapsort algorithm, in which a list of the elements is sorted by increasing co-ordinates along the axes of the simulation space. This algorithm is often combined with body-based cells, see Figure 9.5(d). It has a performance in time of the order  $O(N \log N)$  when it is used to sort an unordered set of  $N$  elements. The advantages of this algorithm are that no special data structure is required, it is insensitive to the spatial distribution of the object/elements, and it only needs an array of  $O(N)$  elements to store the necessary data. Thus, it requires about 10 times less storage than the octree method, and less effort is needed to implement it.
- Recently, algorithms have been developed in Perkins and Williams (1995) which are insensitive to the object size. These algorithms are thus very well suited for discrete finite element analysis in which the size of the finite element chunk broken off is not predictable beforehand. The associated algorithm has a performance of  $O(N \log N)$ , where  $N$  defines the number of objects which can come into contact.

### 9.1.2 Contact detection, phase (II)

Several cases have to be distinguished when algorithms for contact detection are constructed. These include contact between rigid bodies, between a rigid body and deformable bodies, and between deformable bodies. Since geometrical contact conditions have to be formulated with respect to the current configuration (see Chapter 3) different algorithms apply to the cases mentioned. Here the case of contact between rigid bodies will not be considered; associated formulations may be found in Pfeiffer and Glocker (1996).

**Contact between rigid and deformable bodies.** In this case, one has to find a mathematical model for the rigid body. This body then also defines the contact normal, hence it will be used as *master* body or reference body. One possibility to

describe rigid objects is given by an implicit function as introduced below:

$$f(X, Y, Z) = \left( \frac{X - X_0}{R_1} \right)^{n_1} + \left( \frac{Y - Y_0}{R_2} \right)^{n_2} + \left( \frac{Z - Z_0}{R_3} \right)^{n_3} - 1 = 0. \quad (9.1)$$

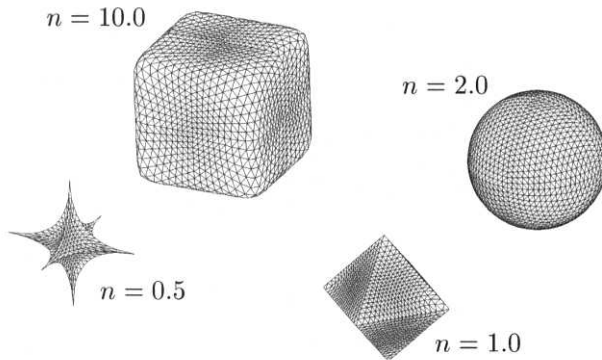
Some cases of bodies resulting from (9.1) are shown in Figure 9.6 for  $n_1 = n_2 = n_3 = n$  and  $R_1 = R_2 = R_3 = 1$ . Observe that quite a variety of different shapes can be created. Due to the construction of the implicit function (9.1), the local contact check as to whether a point (slave node)  $\mathbf{x}_s = \{x_s, y_s, z_s\}^T$  is inside or outside of the rigid body just follows from a pure function evaluation:  $f(x_s, y_s, z_s)$ . Point  $\mathbf{x}_s$  is outside the rigid body for  $f > 0$  and inside for  $f < 0$ . In the case  $f = 0$ , the contacting point lies on the surface of the rigid body. The normal related to the surface of the rigid body is easily obtained by

$$\mathbf{n} = \frac{\text{grad} f}{|\text{grad} f|}. \quad (9.2)$$

It can be used to define the local components of the contact stresses in the contact interface for the point  $\mathbf{x}_s$  by evaluating  $\mathbf{n}$  at  $\mathbf{x}_s$ .

Another implicit superquadric for the description of rigid bodies is presented in Hogue (1998). In that paper a generalization is also provided which allows the representation of arbitrary geometries by a discrete function representation. In the general situation the surface of a body is discretized into a grid of nodes. In between the nodes superquadrics are applied to describe the surface, which leads to a representation that is both discrete and analytical.

**Contact between deformable bodies.** We assume that the global search was successful, and lead to two neighbouring finite element meshes  $\Omega^1$  and  $\Omega^2$ . Now situations, as depicted in Figure 9.7 for the two-dimensional case, have to be investigated in which two bodies can possibly be in contact but it is not clear which slave



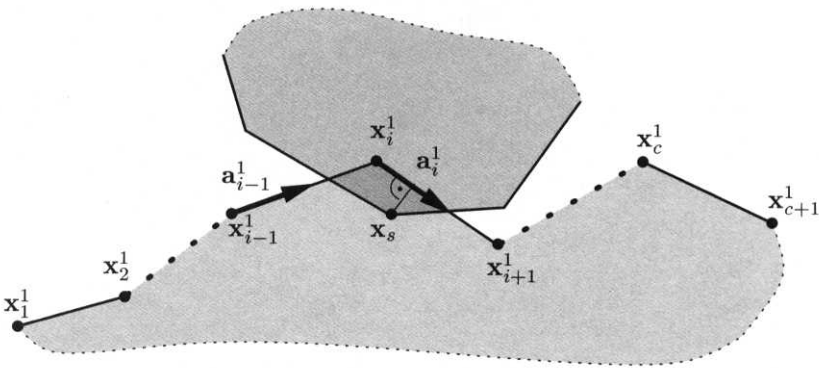
**Fig. 9.6** Representation of bodies by superquadrics for different exponents  $n$ .

node contacts which master segment. There are several possibilities to deal with this problem. One approach which also works in the two- and three-dimensional case is described in Benson and Hallquist (1990). It splits the detection of local contact for a given slave node  $\mathbf{x}_s$  (see Figure 9.7) into three different phases, which consist of

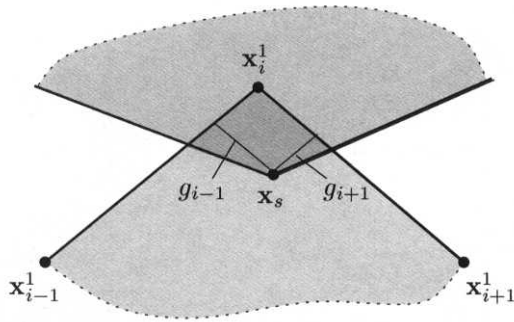
1. Identification of the closest master node  $\mathbf{x}_k^1$  for  $\mathbf{x}_s$ . This can be performed locally by searching for the minimum distance of master nodes  $\mathbf{x}_k^1$  on surface  $\partial\Omega^1$  to the slave node  $\mathbf{x}_s$  as  $\min_{k=1,\dots,c+1} \|\mathbf{x}_s - \mathbf{x}_k^1\|$ . In Figure 9.7 the result of this search for all nodes  $\mathbf{x}_k^1$  of the master surface is node  $\mathbf{x}_i^1$ .
2. Check all finite elements which are adjacent to  $\mathbf{x}_i^1$  and determine the element face which contains the projection  $\bar{\mathbf{x}}^1$ , see (3.3). In the case depicted we obtain the side defined by  $\mathbf{a}_i^1$ .
3. Compute the convective coordinates  $\bar{\xi}$  of the projection using the algorithms stated in Chapter 8 for different discretizations.

Details of actual implementations can be found in Benson and Hallquist (1990). Note that determination of the closest master node and search for the associated elements in phases 1 and 2 is local in nature when appropriately implemented. Hence the algorithms are of the  $O(N)$ -type. The effort which one has to spend in phase 3 depends upon the contact discretization. For the node-to-segment discretization in two dimensions, a closed form solution is obtained, see (8.3). The same also holds for three dimensions when the master surface is interpolated using linear triangles, see (8.107). In all other cases, a NEWTON iteration has to be applied to find  $\bar{\xi}$ , e.g. see (8.82).

Note that in phase 2 of the procedure stated above, we can have situations in which the projection onto the master surface is not uniquely defined, see also Remark 3.1. Such a situation is depicted in Figure 9.8, where a projection of the slave node  $\mathbf{x}_s$  onto two master surfaces,  $\mathbf{x}_i^1 - \mathbf{x}_{i-1}^1$  or  $\mathbf{x}_{i+1}^1 - \mathbf{x}_i^1$ , is possible. This yields two different values of the gap function:  $g_{i-1}$  and  $g_{i+1}$ . In cases like that, one has to decide within



**Fig. 9.7** Contact detection for the two-dimensional case



**Fig. 9.8** Non-unique projection to master surface.

the algorithm which master segment is the correct one. This can be done by trial and error methods, or by smoothing of the local segment geometry by a  $C^1$  function, as described in Section 8.5. Another sound way to tackle this problem is to apply strategies known from optimization theory which have already been successfully used in multi-surface plasticity, e.g. see Laursen and Govindjee (1994).

A different method for the detection of penetration is described in Kane et al. (1999) for finite element discretizations using a triangulation of the bodies. The idea is based on the fact that the interiors of two deformed boundary segments intersect in case of penetration. This observation can be cast for two-dimensional problems into a mathematical formulation by computing the area of the boundary pairs. This area check results in negative areas when penetration occurs. Hence, the contact constraint (3.6) can be put in the form

$$A_s \geq 0. \quad (9.3)$$

In the case of a linear segment, defined in Section 8.1, the area can be computed for a counter-clockwise numbering of surface nodes by

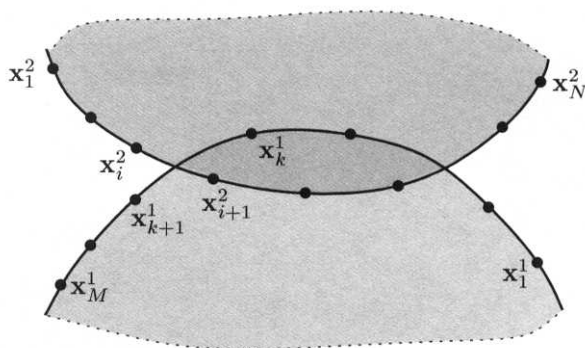
$$A_s = \frac{[(\mathbf{x}_i^2 - \mathbf{x}_k^1) \times (\mathbf{x}_i^2 - \mathbf{x}_{i+1}^2)] \cdot [(\mathbf{x}_{k+1}^1 - \mathbf{x}_k^1) \times (\mathbf{x}_{k+1}^1 - \mathbf{x}_{i+1}^2)]}{[(\mathbf{x}_i^2 - \mathbf{x}_{i+1}^2) \times (\mathbf{x}_{k+1}^1 - \mathbf{x}_k^1)] \cdot \mathbf{e}_3}, \quad (9.4)$$

where  $\mathbf{x}_i^2$ ,  $\mathbf{x}_{i+1}^2$ ,  $\mathbf{x}_k^1$  and  $\mathbf{x}_{k+1}^1$  are the coordinates describing the deformed configuration of the intersection segments, see Figure 9.9. The segments which have to be tested are again found by methods derived for the spatial search, see Section 9.1.1. We note that the area can also be used as a constraint function, as has been proposed in Kane et al. (1999), since  $A_s \geq 0$  excludes possible penetrations. Since the denominator of (9.4) is always positive, the authors noted that the polynomial constraint function

$$g_{As} = [(\mathbf{x}_i^2 - \mathbf{x}_k^1) \times (\mathbf{x}_i^2 - \mathbf{x}_{i+1}^2)] \cdot [(\mathbf{x}_{k+1}^1 - \mathbf{x}_k^1) \times (\mathbf{x}_{k+1}^1 - \mathbf{x}_{i+1}^2)] \quad (9.5)$$

could be employed instead of (9.4), which yields much simpler constraint equations, and hence is algorithmically advantageous.





**Fig. 9.9** Detection of local contact by area check

In the case of three-dimensional applications, the intersection of boundary faces has to be considered, which leads to a volume check. Explicit formulas for this check can be found in Kane et al. (1999) who also use the inequality constraint  $V_s \geq 0$  for three-dimensional contact problems.

**REMARK 9.1:** In some applications it is possible that a body comes into contact with itself. This is depicted in Figure 9.10, where the straight line is bent over, and hence the last node  $x_N^1$  in the deformed configuration comes into contact with the segment associated with node  $x_i^1$ . Thus we have a master node which in the deformed configuration has to be treated like a slave node. Such behaviour, which is called self-contact, often occurs in crash simulations or in large deformation analysis of rubber bearings. In that case, the global search has to be extended such that contact of one surface with itself can be detected. This is more complicated than the standard contact detection between two different bodies. Associated algorithms can be found in Benson and Hallquist (1990).

## 9.2 SOLUTION METHODS FOR UNCONSTRAINED NONLINEAR PROBLEMS

Some of the contact formulations stated in Chapter 5 reduce the variational inequalities to variational equations. This means that within an iterative step of the solution of the variational inequality, one can apply solution methods which were developed for unconstrained problems. These are discussed in this section. Since most applications in engineering also require the consideration of nonlinear effects such as finite deformations or nonlinear constitutive relations, one needs solution methods which can cope with these different phenomena. In case finite element discretizations are used to approximate the physical problem, the discretization process leads to a discrete system of nonlinear equations, see (6.53). In the following, we distinguish between time-independent and time-dependent problems.

In the following we discuss several algorithmic aspects for unconstrained problems, which are often also basis for the solution of constraint problems.

### 9.2.1 Algorithms for time-independent problems

Finite element approximations using the interpolations described in Chapter 6 lead to the above-mentioned system of nonlinear equations. For the construction of nonlinear solution algorithms, we rewrite (6.53):

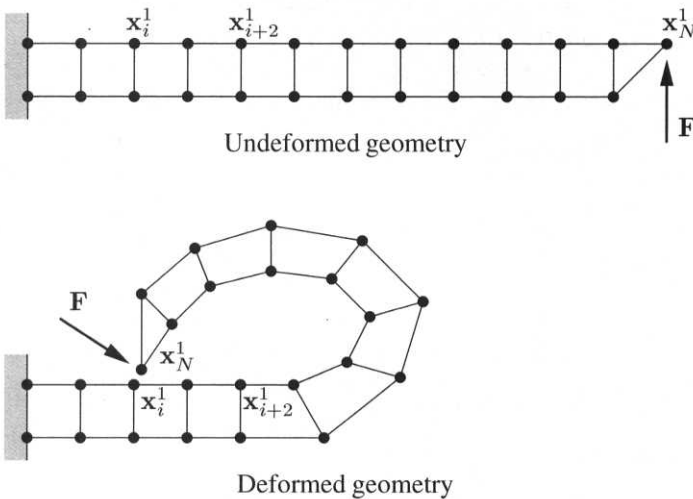
$$\mathbf{G}(\mathbf{u}, \kappa) = \mathbf{R}(\mathbf{u}) - \kappa \mathbf{P} = \mathbf{0}, \quad \mathbf{u} \in \mathbb{R}^N. \quad (9.6)$$

The scaling factor  $\kappa$ , also known as the load parameter, was introduced to be able to apply the load stepwise. Usually, the scaling factor is a problem given quantity, but it can also be useful to view  $\kappa$  as additional variable when special algorithms are used to follow a highly nonlinear load path.

To solve (9.6) by an iterative method, the following standard schemes for unconstrained problems can be applied:

- fix point method,
- NEWTON-RAPHSON method,
- Quasi-NEWTON methods, and
- arc-length methods.

From the list of algorithms, given above, we only state the NEWTON-RAPHSON method. The other methods are discussed in detail in the literature, e.g. see Matthies



**Fig. 9.10** Self-contact of a line.

and Strang (1979), Luenberger (1984) or Bazaraa et al. (1993) for quasi-NEWTON methods, and Riks (1972), Keller (1977), Ramm (1981), Crisfield (1981), Riks (1984), Schweizerhof and Wriggers (1986), Wagner and Wriggers (1988), Wagner (1991) or Crisfield and Shi (1991) for the arc-length methods. Overviews are given in Bathe (1996) or Crisfield (1997).

In all algorithms a sequence of linear equation systems has to be solved within the iterations. Hence equation solvers have a major contribution to the efficiency of the algorithms. While classical direct methods are still sufficient for two-dimensional problems, one has to apply iterative solvers like conjugate gradient schemes, multi-grid techniques or refined sparse-direct solvers for large three-dimensional problems, e.g. see Hackbusch (1991), Schwetlick and Kretschmar (1991), Axelsson (1994), Duff et al. (1989), Kicking (1996), Boersma and Wriggers (1997), Wriggers and Boersma (1998) and Davis and Duff (1999).

The method which is used most frequently for the solution of nonlinear systems of equations is the NEWTON-RAPHSON method. It starts from a TAYLOR series expansion of (9.6) at a known state  $\mathbf{u}_k$ ,

$$\mathbf{G}(\mathbf{u}_k + \Delta \mathbf{u}, \bar{\kappa}) = \mathbf{G}(\mathbf{u}_k, \bar{\kappa}) + D\mathbf{G}(\mathbf{u}_k, \bar{\kappa}) \Delta \mathbf{u} + \mathbf{r}(\mathbf{u}_k, \bar{\kappa}). \quad (9.7)$$

The load parameter  $\bar{\kappa}$  reflects the load level for which the solution has to be computed.  $D\mathbf{G} \cdot \Delta \mathbf{u}$  is the directional derivative which is obtained by linearization of  $\mathbf{G}$  at  $\mathbf{u}_k$ ; for details see Chapters 6 or 8. The linearization of  $\mathbf{G}$  yields a matrix, also known as a HESSE, JACOBI or tangent matrix. It will be abbreviated by  $\mathbf{K}_T$  (see Chapter 6). By neglecting the second order term defined by vector  $\mathbf{r}$  in (9.7), one can derive the iterative algorithms, stated in Box 1, to solve (9.6) for a given load increment at level  $\kappa = \bar{\kappa}$ .

The algorithm is graphically described in Figure 9.11 for the simplified equation  $\hat{R}(u) - \bar{\kappa} = 0$ . The convergence rate of this method is characterized by the inequality  $\|\mathbf{u}_{k+1} - \mathbf{u}\| \leq C \|\mathbf{u}_k - \mathbf{u}\|^2$ , where  $\mathbf{u}$  is solution of (9.6), e.g. see (Isaacson and Keller (1966)). This behaviour manifests the quadratical convergence rate of the NEWTON-RAPHSON method in the vicinity of the solution. Since the quadratic

Initialize algorithm: set  $\mathbf{u}_0 = \mathbf{u}_k$

LOOP over iterations :  $i = 0, \dots$ , convergence

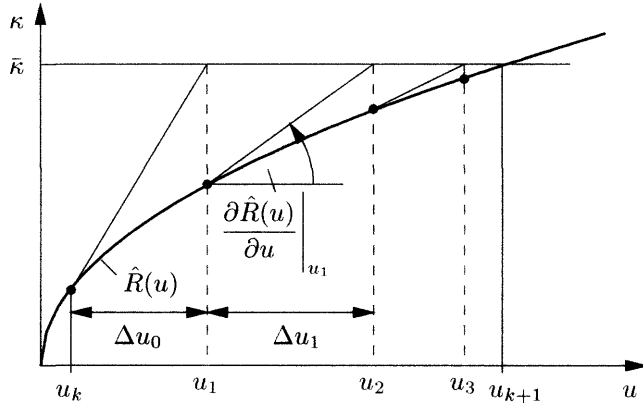
Solve:  $\mathbf{K}_T(\mathbf{u}_i) \Delta \mathbf{u}_{i+1} = -\mathbf{G}(\mathbf{u}_i, \bar{\kappa})$

update displacements:  $\mathbf{u}_{i+1} = \mathbf{u}_i + \Delta \mathbf{u}_{i+1}$

Check for convergence:  $\|\mathbf{G}(\mathbf{u}_{i+1}, \bar{\kappa})\| \leq TOL \Rightarrow \text{STOP}$

END LOOP

Box 1. NEWTON-RAPHSON algorithm.



**Fig. 9.11** NEWTON-RAPHSON method.

rate of convergence is obtained only locally, one has to enhance the algorithms so that they also become globally convergent. However, this cannot be achieved in all cases. Possible methods which yield globally convergent schemes are line search techniques, e.g. see Luenberger (1984), Crisfield (1991) or Crisfield (1997), or the arc-length approach, which was discussed above.

## 9.2.2 Algorithms for time-dependent problems

The discrete form of equations of motion for a solid were derived in Chapter 6, leading to

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{R}[\mathbf{u}(t), t] = \mathbf{P}(t), \quad (9.8)$$

where  $\mathbf{M}$  is the mass matrix,  $\mathbf{R}[\mathbf{u}(t), t]$  represents the stress divergence term and  $\mathbf{P}(t)$  contains the time-dependent applied loads.  $\mathbf{u}(t)$  is the time-dependent solution. To simplify notation, we omit the reference to time dependency in the following equations.

By also including damping (for the background see e.g. Hughes (1987), Bathe (1996) or Zienkiewicz and Taylor (2000a)) the discrete equations of motions are given by

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{R}(\mathbf{u}) = \mathbf{P}. \quad (9.9)$$

This equation can be transformed into a first order algebraic differential equation by introducing as independent variables  $\dot{\mathbf{u}} = \mathbf{v}$  and  $\ddot{\mathbf{u}} = \dot{\mathbf{v}}$ . This leads to

$$\begin{aligned} \dot{\mathbf{u}} &= \mathbf{v}, \\ \dot{\mathbf{v}} &= \mathbf{M}^{-1} [\mathbf{P} - \mathbf{C}\mathbf{v} - \mathbf{R}(\mathbf{u})]. \end{aligned} \quad (9.10)$$

To describe the algorithms we introduce the letter  $\mathbf{a}$  for accelerations  $\ddot{\mathbf{u}}$  and the letter  $\mathbf{v}$  for velocities  $\dot{\mathbf{u}}$ . With this the balance of momentum (9.9) has the discrete form at

time  $t_{n+1}$ :

$$\mathbf{M} \mathbf{a}_{n+1} + \mathbf{C} \mathbf{v}_{n+1} + \mathbf{R}(\mathbf{u}_{n+1}) = \mathbf{P}_{n+1}. \quad (9.11)$$

The subscript  $(\cdot)_{n+1}$  means that the quantity has to be evaluated at time  $t_{n+1}$ .

For the final definition of an initial value problem, we need to introduce the initial conditions for displacements and velocities  $\bar{\mathbf{u}}$  and  $\bar{\mathbf{v}}$ , respectively, at time  $t = t_0$  (often we assume  $t_0 = 0$ ):

$$\begin{aligned} \mathbf{u}_0 &= \bar{\mathbf{u}}, \\ \mathbf{v}_0 &= \bar{\mathbf{v}}. \end{aligned} \quad (9.12)$$

The choice of the time integration procedure to determine the motion of a body  $\mathbf{u}(t)$  depends upon the problem at hand. These are explicit and implicit methods. They have the following general properties:

- **Explicit** time integration methods are easy to implement, since the solution at time  $t_{n+1}$  depends only upon known variables at  $t_n$ . These methods are extremely efficient when the mass matrix in (9.8) is approximated by a lumped mass matrix which is diagonal. Explicit methods are conditionally stable, which means that the time step size is governed by the COURANT criterion.
- **Implicit** time integration schemes approximate time derivatives by quantities which also depend upon the last time step  $t_n$  and upon the still unknown values at time  $t_{n+\alpha}$ . These methods require a solution of a nonlinear equation at each time step. They are much more expensive, since they have to be combined with, for example, the NEWTON-RAPHSON procedures discussed in the last section. However, implicit schemes can be constructed so that they are unconditionally stable, and hence can be applied with a far bigger time step than explicit schemes.

The time step size depends strongly upon the physical process which is simulated. In case shock waves are present (e.g. in car-crash or penetration simulations) then a small time step is necessary. This naturally leads to the use of explicit codes. Large time steps are sufficient when the response of a structure is governed by the low frequency modes, like standard vibration problems.

**Explicit time integration.** In finite element analysis of time-dependent problems, the central difference scheme is mostly applied. In this algorithm the velocities  $\mathbf{v}$  and the accelerations  $\mathbf{a}$  at time  $t_n$  are approximated by

$$\begin{aligned} \mathbf{v}_n &= \frac{\mathbf{u}_{n+1} - \mathbf{u}_{n-1}}{2 \Delta t}, \\ \mathbf{a}_n &= \frac{\mathbf{u}_{n+1} - 2 \mathbf{u}_n + \mathbf{u}_{n-1}}{(\Delta t)^2}. \end{aligned} \quad (9.13)$$

Insertion of this value into (9.9) at time  $t_n$  yields, with

$$\mathbf{M}(\mathbf{u}_{n+1} - 2 \mathbf{u}_n + \mathbf{u}_{n-1}) + \frac{\Delta t}{2} \mathbf{C}(\mathbf{u}_{n+1} - \mathbf{u}_{n-1}) + (\Delta t)^2 \mathbf{R}(\mathbf{u}_n) = (\Delta t)^2 \mathbf{P}_n, \quad (9.14)$$

a system of equations for the unknown displacements  $\mathbf{u}_{n+1}$  at time  $t_{n+1}$ :

$$\left( \mathbf{M} + \frac{\Delta t}{2} \mathbf{C} \right) \mathbf{u}_{n+1} = (\Delta t)^2 [\mathbf{P}_n - \mathbf{R}(\mathbf{u}_n)] + \frac{\Delta t}{2} \mathbf{C} \mathbf{u}_{n-1} + \mathbf{M}(2\mathbf{u}_n - \mathbf{u}_{n-1}). \quad (9.15)$$

In this system  $\mathbf{M}$  and  $\mathbf{C}$  do not change, and hence can be computed once and for all at the beginning of the simulation. Also, the triangularization of the coefficient matrix  $\mathbf{M} + \frac{\Delta t}{2} \mathbf{C}$  has to be calculated only once. All nonlinearities only enter via the vector  $\mathbf{R}(\mathbf{u}_n)$  on the right-hand side. In case  $\mathbf{M}$  and  $\mathbf{C}$  are diagonal matrices, the factorization of  $\mathbf{M} + \frac{\Delta t}{2} \mathbf{C}$  is trivial. Hence, only vector operations are needed to evaluate (9.15).

The definition of the starting values for explicit schemes requires special treatment since at time  $t_0$  values for  $\mathbf{u}_{-1}$  are needed which have to be determined from the initial conditions  $\mathbf{u}_0$  and  $\mathbf{v}_0$ . By using a TAYLOR series expansion at time  $t_{-1}$ , we obtain

$$\mathbf{u}_{-1} = \mathbf{u}_0 - \Delta t \mathbf{v}_0 + \frac{(\Delta t)^2}{2} \mathbf{a}_0. \quad (9.16)$$

Here the accelerations at  $t_0$  follow from (9.11):

$$\mathbf{a}_0 = \mathbf{M}^{-1} [-\mathbf{C} \mathbf{v}_0 - \mathbf{R}(\mathbf{u}_0) + \mathbf{P}_0]. \quad (9.17)$$

A variation of the explicit scheme can be found in Wood (1990). The following approximations for displacements and velocities are used:

$$\begin{aligned} \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta t \mathbf{v}_n + \frac{(\Delta t)^2}{2} \mathbf{a}_n \\ \mathbf{v}_{n+1} &= \mathbf{v}_n + \frac{1}{2} \Delta t (\mathbf{a}_n + \mathbf{a}_{n+1}), \end{aligned} \quad (9.18)$$

together with equation (9.11) to determine the accelerations. This leads to the equation system

$$\left( \mathbf{M} + \frac{\Delta t}{2} \mathbf{C} \right) \mathbf{a}_{n+1} = \mathbf{P}_{n+1} - \mathbf{R} \left( \mathbf{u}_n + \Delta t \mathbf{v}_n + \frac{(\Delta t)^2}{2} \mathbf{a}_n \right) - \frac{\Delta t}{2} \mathbf{C} \mathbf{a}_n. \quad (9.19)$$

Its right-hand side depends, besides the known loading function  $\mathbf{P}$ , only upon vectors which have to be evaluated at time  $t_n$ . Thus the initial conditions (9.12) can be applied directly when starting the algorithm. Displacements and velocities follow after the solution of (9.19) from (9.18). In (9.19) the same coefficient matrix appears as in (9.15).

For linear problems the following estimate is valid for the critical time step size (COURANT criterion):

$$\Delta t \leq \frac{T_N}{\pi}. \quad (9.20)$$

The value  $T_N$  characterizes the smallest period of the finite element discretization which can be estimated on an element basis, e.g. see Bathe (1982). For nonlinear problems an estimation for the critical time step can be found in, for example,

Belytschko et al. (1976):

$$\Delta t \leq \delta \frac{h}{c_L}. \quad (9.21)$$

This estimate depends upon the characteristic length of the smallest element  $h$ , the wave speed of a compression wave in the solid  $c_L$  (in a linear elastic medium we have  $c_L = \frac{3K(1-\nu)}{\rho(1+\nu)}$  with modulus of compression  $K$ , POISSON ratio  $\nu$  and density  $\rho$ ). The constant  $\delta$  ( $0.2 < \delta < 0.9$ ) is a reduction factor which has to be chosen empirically for the problem.

**Implicit time integration.** Among many integration schemes the NEWMARK method is the most well known to solve the equation of motion (9.9), see Newmark (1959). It is based on the following approximations of displacements and velocities at time  $t_{n+1}$ :

$$\begin{aligned} \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta t \mathbf{v}_n + \frac{(\Delta t)^2}{2} [(1 - 2\beta) \mathbf{a}_n + 2\beta \mathbf{a}_{n+1}] \\ \mathbf{v}_{n+1} &= \mathbf{v}_n + \Delta t [(1 - \gamma) \mathbf{a}_n + \gamma \mathbf{a}_{n+1}]. \end{aligned} \quad (9.22)$$

Here the displacements and velocities depend upon values at time  $t_n$ , but also on the accelerations at time  $t_{n+1}$ . The parameters  $\beta$  and  $\gamma$  are constants which could be chosen freely. However, they determine the order and accuracy of the method, and thus have to be chosen with care. The intervals from which one can choose the parameters are  $0 \leq \beta \leq 0.5$  and  $0 \leq \gamma \leq 1$ , e.g. see Hughes (1987). Note that the explicit central difference scheme (9.18) follows for the parameter set  $\gamma = 0.5$  and  $\beta = 0$ .

The accelerations  $\mathbf{a}_{n+1}$  follow from (9.11) when the approximations of displacements  $\mathbf{u}_{n+1}$  and velocities  $\mathbf{v}_{n+1}$  (see (9.22)) are inserted. This leads to a nonlinear algebraic system of equations for  $\mathbf{a}_{n+1}$ :

$$(\mathbf{M} + \gamma \Delta t \mathbf{C}) \mathbf{a}_{n+1} + \mathbf{R}(\mathbf{a}_{n+1}; \mathbf{u}_n; \mathbf{v}_n; \mathbf{a}_n) = \mathbf{P}_{n+1} - \bar{\mathbf{G}}(\mathbf{u}_n; \mathbf{v}_n; \mathbf{a}_n). \quad (9.23)$$

All terms which result from inserting (9.22) into (9.11) are combined in  $\bar{\mathbf{G}}$ . Equation (9.23) can now be solved with, for example, the NEWTON-RAPHSON scheme. Once the accelerations are known, the displacements and velocities follow from (9.22).

Often the NEWMARK method is formulated with displacements as primary variables. Solving (9.22) for  $\mathbf{a}_{n+1}$  and  $\mathbf{v}_{n+1}$  yields

$$\begin{aligned} \mathbf{a}_{n+1} &= \alpha_1 (\mathbf{u}_{n+1} - \mathbf{u}_n) - \alpha_2 \mathbf{v}_n - \alpha_3 \mathbf{a}_n, \\ \mathbf{v}_{n+1} &= \alpha_4 (\mathbf{u}_{n+1} - \mathbf{u}_n) + \alpha_5 \mathbf{v}_n + \alpha_6 \mathbf{a}_n, \end{aligned} \quad (9.24)$$

with the following definition of the constants:

$$\begin{aligned} \alpha_1 &= \frac{1}{\beta (\Delta t)^2}, \quad \alpha_2 = \frac{1}{\beta \Delta t}, \quad \alpha_3 = \frac{1 - 2\beta}{2\beta}, \\ \alpha_4 &= \frac{\gamma}{\beta \Delta t}, \quad \alpha_5 = \left(1 - \frac{\gamma}{\beta}\right), \quad \alpha_6 = \left(1 - \frac{\gamma}{2\beta}\right) \Delta t. \end{aligned} \quad (9.25)$$

Inserting (9.24) into the balance of momentum (9.11) leads to a nonlinear algebraic equation for the unknown displacements  $\mathbf{u}_{n+1}$ :

$$\begin{aligned} \mathbf{G}(\mathbf{u}_{n+1}) = & \mathbf{M} [\alpha_1 (\mathbf{u}_{n+1} - \mathbf{u}_n) - \alpha_2 \mathbf{v}_n - \alpha_3 \mathbf{a}_n] \\ & + \mathbf{C} [\alpha_4 (\mathbf{u}_{n+1} - \mathbf{u}_n) + \alpha_5 \mathbf{v}_n + \alpha_6 \mathbf{a}_n] \\ & + \mathbf{R}(\mathbf{u}_{n+1}) - \mathbf{P}_{n+1} = \mathbf{0}. \end{aligned} \quad (9.26)$$

When using NEWTON'S method to solve this equation for the displacements at time  $t_{n+1}$ , by introducing the tangential stiffness matrix

$$\mathbf{K}_T(\mathbf{u}_{n+1}^i) = \left. \frac{\partial \mathbf{R}}{\partial \mathbf{u}_{n+1}} \right|_{\mathbf{u}_{n+1}^i}, \quad (9.27)$$

one can state the following algorithm (see also Box 1):

$$\begin{aligned} [\alpha_1 \mathbf{M} + \alpha_4 \mathbf{C} + \mathbf{K}_T(\mathbf{u}_{n+1}^i)] \Delta \mathbf{u}_{n+1}^{i+1} &= -\mathbf{G}(\mathbf{u}_{n+1}^i), \\ \mathbf{u}_{n+1}^{i+1} &= \mathbf{u}_{n+1}^i + \Delta \mathbf{u}_{n+1}^{i+1}. \end{aligned} \quad (9.28)$$

This iteration has to be performed with respect to index  $i$  in each time step  $t_{n+1}$ . As the initial value for the next time step the converged displacement from the previous time step  $\mathbf{u}_{n+1}^0 = \mathbf{u}_n$  is used. The iteration is stopped when the criterion stated in Box 1 is fulfilled.

For linear problems one can analyse the NEWMARK method with respect to accuracy and stability; results can be found in Hughes (1987) or Wood (1990). For parameters  $\gamma \leq 0.5$ , the accuracy of NEWMARK'S method is of the order  $O(\Delta t^2)$ . Often one wishes to damp out higher frequency responses. This can be done using parameters  $\gamma > 0.5$ , however this leads to a reduction of the order of accuracy. Hence methods have been developed which retain the order but have the damping properties, see e.g. Hilber et al. (1977) or Wood et al. (1981). The latter method changes the momentum equation (9.11) as follows:

$$(1 - \alpha) \mathbf{M} \mathbf{a}_{n+1} + \alpha \mathbf{M} \mathbf{a}_n + \mathbf{C} \mathbf{v}_{n+1} + \mathbf{R}(\mathbf{u}_{n+1}) - \mathbf{P}_{n+1} = \mathbf{0}, \quad (9.29)$$

while retaining the approximations for displacements and velocities (9.22).

The method in Hilber et al. (1977) was developed for linear elasto-dynamical problems. It introduces a different momentum equation in which the displacements are weighted. Its nonlinear extension yields, instead of (9.29),

$$\mathbf{M} \mathbf{a}_{n+1} + (1 - \alpha) [\mathbf{C} \mathbf{v}_{n+1} - \mathbf{P}_{n+1}] + \alpha [\mathbf{C} \mathbf{v}_n - \mathbf{P}_n] + \mathbf{R}[(1 - \alpha) \mathbf{u}_{n+1} + \alpha \mathbf{u}_n] = \mathbf{0}. \quad (9.30)$$

Again, the displacements and velocities are computed at time  $t_{n+1}$  according to (9.22). However, this method requires an evaluation of the residual vector  $\mathbf{R}$  at the intermediate time  $t_{n+\alpha} = (1 - \alpha) t_{n+1} + \alpha t_n$ . This turns out to be nontrivial when complex nonlinear constitutive equations are considered. The method damps high frequencies for  $0.5 < \alpha < 1$ .



From a continuum mechanics point of view, all integration schemes should be constructed such that they fulfil the basic laws, e.g. momentum, moments of momentum and conservation of energy. However, this is not the case for all schemes. Due to the fact that considerable effort has been put into the development of integration schemes which preserve the above-mentioned conservation laws. Formulations can be found in Simo and Tarnow (1992), Crisfield (1997), Sansour et al. (1997) or Betsch and Steinmann (2000), for instance.

### 9.3 GLOBAL SOLUTION ALGORITHMS FOR CONTACT

The efficient and robust solution of contact problems mainly relies on, besides a good discretization of the contact interface, the algorithmic part. Here a wide diversity of possible methods can be applied which have advantages and disadvantages with respect to different problems.

Algorithms which are applied in many standard finite element programs are related either to the penalty method or to the LAGRANGE multiplier method; see also Sections 5.3.2 and 5.3.1, respectively. Often the first approach is implemented mainly due to its simplicity. When proper estimates for the penalty parameters are known, see e.g. Nour-Omid and Wriggers (1987) or update formulas provided in Bertsekas (1984), then penalty schemes provide robust algorithms for many applications. The penalty method is mostly combined with an active set strategy. The global set of equations is given in (8.20) and (8.36). However, often the LAGRANGE multiplier method provides better and more stable results in confined situations when a process is deformation driven; see also Remark 5.4.

Algorithms for solving variational inequalities are given by mathematical programming, active set strategies or sequential quadratic programming methods, to name only a few. Each of these methods is well known from optimization theory, e.g. see Luenberger (1984), Bertsekas (1984) or Bazaraa et al. (1993). The application of mathematical programming for the solution of contact problems can be found in Conry and Seireg (1971), Panagiotopoulos (1975), Klarbring (1986), Klarbring and Björkman (1988) or Holmberg (1990), for example. The sequential quadratic programming approach has been considered by Barthold and Bischoff (1988), and with application to large strain elasticity by Björkman et al. (1995), recently also within contact-impact analysis, see Raous (1999) or Kane et al. (1999).

New algorithms which are based on iterative methods that can be applied in a parallel computing environment are currently in development. Iterative methods for single processor computers based on multi-grid techniques can be found in Belsky (1993), Hoppe and Kornhuber (1994) or Kornhuber (1997), for instance. Other methods which apply conjugated gradient type iterative solvers are discussed in Heinsteins and Laursen (1999). Several contact algorithms were developed for multi-processor machines using multi-grid techniques, e.g. see Schöberl (1998) or Kornhuber and Krause (2000). Another approach is pursued in Wohlmuth (2000a) and Krause and Wohlmuth (2001), who apply a DIRICHLET-NEUMANN solver for contact problems in combination with mortar techniques.

In the following sections, different global solution algorithms for contact problems will be discussed, starting with the most common ones for frictionless contact.

### 9.3.1 Basic notation

For the discussion of different algorithms the following matrix notation is introduced. As derived before, the weak form of the equilibrium equations for the solid coming into contact is given by

$$\delta \mathbf{u}^T \mathbf{G}(\mathbf{u}) = \delta \mathbf{u}^T [\mathbf{R}(\mathbf{u}) - \mathbf{P}] = 0, \quad (9.31)$$

where  $\mathbf{u}$  is the global displacement vector and  $\delta \mathbf{u}$  denotes a vector containing the variations. Note that we do not distinguish between the two or more different bodies which come into contact to shorten the notation.  $\mathbf{R}(\mathbf{u})$  is the matrix form of the stress divergence term, and vector  $\mathbf{P}$  contains the loading terms, where we have assumed that the loads are conservative and hence do not depend upon the deformation. This scalar equation yields, for arbitrary variations  $\delta \mathbf{u}$ , the vector form

$$\mathbf{G}(\mathbf{u}) = \mathbf{R}(\mathbf{u}) - \mathbf{P} = \mathbf{0}. \quad (9.32)$$

The linearization of (9.31) at a given deformation state  $\bar{\mathbf{u}}$  leads to the expression

$$\delta \mathbf{u}^T \mathbf{G}(\bar{\mathbf{u}}) + \delta \mathbf{u}^T \mathbf{K}_T(\bar{\mathbf{u}}) \Delta \mathbf{u} \quad (9.33)$$

with the tangential stiffness matrix  $\mathbf{K}_T$ ; for details see Chapter 6.

Since the algorithms have to be designed for contact problems, the contact constraints also have to be expressed in a general form depending on the formulation stated in Section 5.3. A detailed description of how contact constraints are incorporated into finite element discretization schemes can be found in Chapters 7 and 8 for small and large deformations, respectively. Here a matrix formulation is stated which will stand for the contact constraints, and which has to be specialized according to the derivations in Chapters 7 and 8 for specific discretizations.  $\mathcal{J}_C$  denotes all possible contact contributions which can be special segments, nodes or elements in the contact interface. For these different cases, a general equation describing the constraint inequality for normal contact will be introduced according to Chapter 8 as follows:

$$\mathbf{G}_s^c(\mathbf{u}) \geq 0, \quad \forall s \in \mathcal{J}_C. \quad (9.34)$$

The single contributions for nodes or segments  $s$  are all combined in the matrix  $\mathbf{G}^c(\mathbf{u})$ . Note that this constraint equation is nonlinear with respect to the deformation. In the case of a linear dependence on the displacement field, we can express (9.34) for normal contact as

$$\mathbf{G}_s^{cL}(\mathbf{u}) = \mathbf{N}_s \mathbf{u} + g_X \geq 0 \quad \forall s \in \mathcal{J}_C, \quad (9.35)$$

where  $\mathbf{N}_s$  stands for the normal vector associated with the node or segment  $s$ .

The contact problem can now be stated as follows. Find the correct deformation  $\mathbf{u}$  and the number of active contact constraints  $s \in \mathcal{J}_A$  such that

$$\boxed{\mathbf{R}(\mathbf{u}) - \mathbf{P} = \mathbf{0} \quad \text{and} \quad \mathbf{G}_s^c(\mathbf{u}) \geq 0} \quad (9.36)$$

is fulfilled.

Any continuum problem with hyperelastic material can be formulated as a minimization problem. This is based on the fact that the total potential energy assumes a minimum at the solution point

$$\Pi(\mathbf{u}) \rightarrow MIN. \quad (9.37)$$

In general, the solution has to be computed for a given set of equality constraints which represent the boundary conditions. In the case of contact the minimization is restricted by the contact constraints (9.34); see also Sections 1.1 or 5.1. Hence, we have the problem

$$\boxed{\begin{array}{ll} \text{Minimize} & \Pi(\mathbf{u}) \\ \text{subject to} & \mathbf{G}^c(\mathbf{u}) \geq \mathbf{0}. \end{array}} \quad (9.38)$$

As already described in previous chapters, one can now incorporate the constraint equations using the LAGRANGE multiplier or penalty methods, see (5.26) and (5.31), respectively. This leads in matrix notation to the formulation of a saddle point problem

$$\Pi^{LM}(\mathbf{u}, \boldsymbol{\Lambda}) = \Pi(\mathbf{u}) + \boldsymbol{\Lambda}^T \mathbf{G}^c(\mathbf{u}) \rightarrow STAT, \quad (9.39)$$

together with the KUHN-TUCKER-KARUSH conditions

$$\mathbf{G}^c(\mathbf{u}) \geq \mathbf{0}, \quad \boldsymbol{\Lambda} \leq \mathbf{0}, \quad \mathbf{G}^c(\mathbf{u}) \boldsymbol{\Lambda} = \mathbf{0} \quad (9.40)$$

in the case of the LAGRANGE multiplier method. For the penalty formulation a minimization problem

$$\Pi^P(\mathbf{u}) = \Pi(\mathbf{u}) + \frac{\epsilon}{2} \mathbf{G}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) \rightarrow MIN \quad (9.41)$$

can be stated. In the latter case it was assumed that a known set of active constraint  $\mathcal{J}_A$  is given which fulfil (9.34). This choice of a set of active constraints is used later in the iterative algorithm to solve (9.38).

The variation of (9.39) yields

$$\delta \Pi^{LM} = \delta \Pi + \delta \boldsymbol{\Lambda}^T \mathbf{G}^c(\mathbf{u}) + \delta \mathbf{u}^T \mathbf{C}^c(\mathbf{u})^T \boldsymbol{\Lambda} = 0, \quad (9.42)$$

where  $\mathbf{C}^c(\mathbf{u}) \delta \mathbf{u}$  is the matrix form of  $\delta \mathbf{G}^c(\mathbf{u})$ , since the variation of a nonlinear constraint is always linear in  $\delta \mathbf{u}$ . For details on how the matrix  $\mathbf{C}^c(\mathbf{u})$  is constructed for different discretizations, see Chapter 8. From (9.42) the nonlinear equation system

$$\begin{aligned} \mathbf{G}(\mathbf{u}) + \mathbf{C}^c(\mathbf{u})^T \boldsymbol{\Lambda} &= \mathbf{0} \\ \mathbf{G}^c(\mathbf{u}) &= \mathbf{0} \end{aligned} \quad (9.43)$$

follows for arbitrary variations. In the same notation, for the penalty formulation (9.41) we obtain

$$\delta \Pi^P = \delta \Pi + \epsilon \delta \mathbf{u}^T \mathbf{C}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) = 0. \quad (9.44)$$

Here the nonlinear equation system

$$\mathbf{G}(\mathbf{u}) + \epsilon \mathbf{C}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) = \mathbf{0} \quad (9.45)$$

results, which can be applied to obtain a solution for a given set of active constraints.

Linearizations of both formulations are needed in the following algorithms. Using the notation of the previous chapters, we can derive the linearization at a fixed state  $(\bar{\mathbf{u}}, \bar{\Lambda})$  for the LAGRANGE multiplier method

$$\begin{aligned} \Delta [\delta \Pi^{LM}(\bar{\mathbf{u}}, \bar{\Lambda})] &= \Delta [\delta \Pi(\bar{\mathbf{u}})] + \delta \Lambda^T \mathbf{C}^c(\bar{\mathbf{u}}) \Delta \mathbf{u} + \delta \mathbf{u}^T \mathbf{C}^c(\bar{\mathbf{u}})^T \Delta \Lambda \\ &\quad + \delta \mathbf{u}^T \mathbf{K}_T^c(\bar{\mathbf{u}}, \bar{\Lambda}) \Delta \mathbf{u}. \end{aligned} \quad (9.46)$$

For arbitrary  $d\mathbf{u}$  and  $\delta\Lambda$ , equation (9.46) yields the linearized system of equations according to (9.33):

$$\begin{bmatrix} \mathbf{K}_T(\bar{\mathbf{u}}) + \mathbf{K}_T^c(\bar{\mathbf{u}}, \bar{\Lambda}) & \mathbf{C}^c(\bar{\mathbf{u}})^T \\ \mathbf{C}^c(\bar{\mathbf{u}}) & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{u} \\ \Delta \Lambda \end{Bmatrix} = - \begin{Bmatrix} \mathbf{G}(\bar{\mathbf{u}}) + \mathbf{C}^c(\bar{\mathbf{u}})^T \bar{\Lambda} \\ \mathbf{G}^c(\bar{\mathbf{u}}) \end{Bmatrix}. \quad (9.47)$$

All matrices have already been defined in this section besides the matrix  $\mathbf{K}_T^c(\bar{\mathbf{u}}, \bar{\Lambda})$ , which stems from the linearization of the last term in (9.42) with respect to the displacement field; for details see Chapter 8. Note that for a linear problem, equation (9.47) reduces to

$$\begin{bmatrix} \mathbf{K}_T & \mathbf{N}^T \\ \mathbf{N} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \Lambda \end{Bmatrix} = - \begin{Bmatrix} -\mathbf{P} \\ \mathbf{G}^{cL}(\bar{\mathbf{u}}) \end{Bmatrix}, \quad (9.48)$$

where the linear constraints are given in (9.35).

By introducing  $\mathbf{w} = \{\mathbf{u}, \Lambda\}^T$ , equation (9.47) can be recast as

$$\mathbf{K}_T^{LM}(\bar{\mathbf{w}}) \Delta \mathbf{w} = -\mathbf{G}^{LM}(\bar{\mathbf{w}}), \quad (9.49)$$

which simplifies the notation.

For the discretized weak form (9.44), stemming from the penalty method, the linearization leads to

$$\Delta [\delta \Pi^P(\bar{\mathbf{u}})] = [\delta \Pi(\bar{\mathbf{u}})] + \epsilon \delta \mathbf{u}^T \left[ \mathbf{C}^c(\bar{\mathbf{u}})^T \mathbf{C}^c(\bar{\mathbf{u}}) + \mathbf{K}_T^P(\bar{\mathbf{u}}) \right] \Delta \mathbf{u}, \quad (9.50)$$

which results in the linearized equation system at the known displacement state  $\bar{\mathbf{u}}$ :

$$\left[ \mathbf{K}_T(\bar{\mathbf{u}}) + \epsilon \left( \mathbf{K}_T^P(\bar{\mathbf{u}}) + \mathbf{C}^c(\bar{\mathbf{u}})^T \mathbf{C}^c(\bar{\mathbf{u}}) \right) \right] \Delta \mathbf{u} = - [\mathbf{G}(\bar{\mathbf{u}}) + \mathbf{C}^c(\bar{\mathbf{u}})^T \mathbf{G}^c(\bar{\mathbf{u}})]. \quad (9.51)$$

Again, the matrix  $\mathbf{K}_T^{cP}(\bar{\mathbf{u}})$ , which stems from the linearization of  $\mathbf{C}(\mathbf{u})$  in (9.44), disappears in linear problems. Furthermore, we can combine all matrices and vectors in (9.51) to simplify notation:

$$\mathbf{K}_T^P(\bar{\mathbf{u}}) \Delta \mathbf{u} = -\mathbf{G}^P(\bar{\mathbf{u}}). \quad (9.52)$$

The structures of the matrices introduced in (9.47) and (9.51) are presented for different discretizations in Chapter 6 for solids and in Chapter 8 for contact contributions.

In many applications the contact region is small compared to the domain of the structure. This fact can be used within contact algorithms to reduce the computational effort. Here we wish to state the reduction for linear analysis. In such cases the equilibrium equation (9.48) can be rearranged as

$$\begin{bmatrix} \mathbf{K}_{II} & \mathbf{K}_{IC} & \mathbf{0} \\ \mathbf{K}_{IC}^T & \mathbf{K}_{CC} & \mathbf{N}^T \\ \mathbf{0} & \mathbf{N} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{u}_I \\ \mathbf{u}_C \\ \Lambda \end{Bmatrix} = \begin{Bmatrix} \mathbf{P}_I \\ \mathbf{P}_C \\ -\mathbf{G}^{cL}(\bar{\mathbf{u}}) \end{Bmatrix}, \quad (9.53)$$

where  $C \in \mathcal{J}_C$  denotes all possible contact nodes and  $I$  are the remaining nodes of the finite element discretization. The displacements  $\mathbf{u}_I$  depend only indirectly upon  $\Lambda$ , hence they can be eliminated from (9.53) by static condensation, which yields a full but smaller system of equations

$$\begin{bmatrix} \hat{\mathbf{K}}_{CC} & \mathbf{N}^T \\ \mathbf{N} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{u}_C \\ \Lambda \end{Bmatrix} = \begin{Bmatrix} \hat{\mathbf{P}}_C \\ -\mathbf{G}^{cL}(\bar{\mathbf{u}}) \end{Bmatrix}, \quad (9.54)$$

with

$$\begin{aligned} \hat{\mathbf{K}}_{CC} &= \mathbf{K}_{CC} - \mathbf{K}_{IC}^T \mathbf{K}_{II}^{-1} \mathbf{K}_{IC}, \\ \hat{\mathbf{P}}_C &= \mathbf{P}_C - \mathbf{K}_{IC}^T \mathbf{K}_{II}^{-1} \mathbf{P}_I. \end{aligned}$$

The SCHUR complement  $\hat{\mathbf{K}}_{CC}$  is obtained at an intermediate step of the GAUSS elimination process when solving (9.53). Therefore its evaluation requires no additional computational effort. The advantage of static condensation is that a smaller system of equations has to be solved during the contact iteration. This static condensation process, here applied for the LAGRANGE multiplier method, can also be used within the penalty approach.

### 9.3.2 Dual formulation

There are many algorithms known in optimization which start from a dual formulation of the minimization problem (9.38). To develop the dual form we start from the LAGRANGE multiplier formulation with the assumption of a linear problem. By assuming that the LAGRANGE multipliers are known, one can solve (9.48), which then reduces to the problem that the gradient of (9.43) has to vanish:

$$\mathbf{K}_T \mathbf{u} + \mathbf{N}^T \Lambda = \mathbf{P}. \quad (9.55)$$

Then a dual problem can now be defined as

$$\begin{aligned} & \text{Maximize} \quad \frac{1}{2} \mathbf{u}^T \mathbf{K}_T \mathbf{u} - \mathbf{P}^T \mathbf{u} + \mathbf{\Lambda}^T (\mathbf{N} \mathbf{u} - \mathbf{G}_X) \\ & \text{subject to} \quad \mathbf{K}_T \mathbf{u} + \mathbf{N}^T \mathbf{\Lambda} = \mathbf{P} \quad \text{and} \quad \mathbf{\Lambda} \leq \mathbf{0}. \end{aligned} \quad (9.56)$$

Now (9.55) can be pre-multiplied by  $\mathbf{u}$  which yields

$$-\mathbf{u}^T \mathbf{K}_T \mathbf{u} = \mathbf{\Lambda}^T \mathbf{N} \mathbf{u} - \mathbf{P}^T \mathbf{u}. \quad (9.57)$$

This result can be inserted into (9.56) to simplify the maximization problem:

$$\begin{aligned} & \text{Maximize} \quad -\frac{1}{2} \mathbf{u}^T \mathbf{K}_T \mathbf{u} - \mathbf{G}_X^T \mathbf{\Lambda} \\ & \text{subject to} \quad \mathbf{K}_T \mathbf{u} + \mathbf{N}^T \mathbf{\Lambda} = \mathbf{P} \quad \text{and} \quad \mathbf{\Lambda} \leq \mathbf{0}. \end{aligned} \quad (9.58)$$

In standard applications the stiffness matrix  $\mathbf{K}_T$  is positive definite. Then we can solve (9.55) for  $\mathbf{u}$ ,

$$\mathbf{u} = -\mathbf{K}_T^{-1} (\mathbf{N}^T \mathbf{\Lambda} - \mathbf{P}), \quad (9.59)$$

and introduce this result into (9.58), leading to the dual problem

$$\begin{aligned} & \text{Maximize} \quad -\frac{1}{2} \mathbf{\Lambda}^T \mathbf{D} \mathbf{\Lambda} + \mathbf{\Lambda}^T \mathbf{d} - \frac{1}{2} \mathbf{P}^T \mathbf{K}_T^{-1} \mathbf{P} \\ & \text{subject to} \quad \mathbf{\Lambda} \leq \mathbf{0}. \end{aligned} \quad (9.60)$$

The definitions of matrix  $\mathbf{D}$  and vector  $\mathbf{d}$  are as follows:

$$\mathbf{D} = \mathbf{N} \mathbf{K}_T^{-1} \mathbf{N}^T, \quad (9.61)$$

$$\mathbf{d} = \mathbf{G}_X - \mathbf{N} \mathbf{K}_T^{-1} \mathbf{P}. \quad (9.62)$$

Note that the last term in (9.60) is constant, and hence can be neglected in many cases.

The same procedure can now be applied to the incremental form (9.47) stemming from the nonlinear problem. By assuming that the LAGRANGE multipliers are known, one can solve (9.47) at a fixed state  $(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$ :

$$[\mathbf{K}_T(\bar{\mathbf{u}}) + \mathbf{K}_T^c(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})] \Delta \mathbf{u} + \mathbf{C}^c(\bar{\mathbf{u}})^T \Delta \mathbf{\Lambda} + [\mathbf{G}(\bar{\mathbf{u}}) + \mathbf{C}^c(\bar{\mathbf{u}})^T \bar{\mathbf{\Lambda}}] = \mathbf{0}, \quad (9.63)$$

and with  $\hat{\mathbf{K}}_T(\bar{\mathbf{u}}) = \mathbf{K}_T(\bar{\mathbf{u}}) + \mathbf{K}_T^c(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$  and  $\hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) = \mathbf{G}(\bar{\mathbf{u}}) + \mathbf{C}^c(\bar{\mathbf{u}})^T \bar{\mathbf{\Lambda}}$  obtain the dual problem

$$\begin{aligned} & \text{Maximize} \quad \frac{1}{2} \Delta \mathbf{u}^T \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \Delta \mathbf{u} + \hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})^T \Delta \mathbf{u} + \Delta \mathbf{\Lambda}^T \mathbf{C}^c(\bar{\mathbf{u}}) \\ & \text{subject to} \quad \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \Delta \mathbf{u} + \mathbf{C}^c(\bar{\mathbf{u}})^T \Delta \mathbf{\Lambda} = -\hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \\ & \text{and} \quad (\bar{\mathbf{\Lambda}} + \Delta \mathbf{\Lambda}) \leq \mathbf{0}, \end{aligned} \quad (9.64)$$

which yields, by using the same manipulation which leads to (9.57), the incremental maximization problem:

$$\begin{aligned} & \text{Maximize} \quad -\frac{1}{2} \Delta \mathbf{u}^T \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \Delta \mathbf{u} \\ & \text{subject to} \quad \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \Delta \mathbf{u} + \mathbf{C}^c(\bar{\mathbf{u}})^T \Delta \mathbf{\Lambda} = -\hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \\ & \text{and} \quad (\bar{\mathbf{\Lambda}} + \Delta \mathbf{\Lambda}) \leq \mathbf{0}. \end{aligned} \quad (9.65)$$

The solution of (9.63) with respect to the incremental displacements yields

$$\Delta \mathbf{u} = -\hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})^{-1} [ \mathbf{C}^c(\bar{\mathbf{u}})^T \Delta \mathbf{\Lambda} + \hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) ]. \quad (9.66)$$

This result can be inserted into (9.65), which then yields the incremental dual problem

$$\begin{aligned} \text{Maximize} \quad & -\frac{1}{2} \Delta \mathbf{\Lambda}^T \mathbf{D}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \Delta \mathbf{\Lambda} + \Delta \mathbf{\Lambda}^T \mathbf{d}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \\ & -\frac{1}{2} \hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})^T \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})^{-1} \hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \\ \text{subject to} \quad & \bar{\mathbf{\Lambda}} + \Delta \mathbf{\Lambda} \leq \mathbf{0}. \end{aligned} \quad (9.67)$$

Here the definition of the matrix  $\mathbf{D}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$  and the vector  $\mathbf{d}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$ , which now depend upon the state  $\bar{\mathbf{u}}$ , have been used:

$$\mathbf{D}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) = \mathbf{C}^c(\bar{\mathbf{u}}) \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})^{-1} \mathbf{C}^c(\bar{\mathbf{u}})^T, \quad (9.68)$$

$$\mathbf{d}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) = -\mathbf{C}^c(\bar{\mathbf{u}}) \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})^{-1} \hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}). \quad (9.69)$$

Hence for a given state  $(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$  we can set up an incremental dual problem which uses the tangent matrix evaluated at  $\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}$ . Note that the constant last term in (9.68) can be omitted in most algorithms. The dual problem yields LAGRANGE multipliers which can then be used to compute the associated incremental displacements from (9.66). This result can be applied within an iterative algorithm to solve the fully nonlinear problem. Algorithms based on the primal and dual formulations are discussed in the following sections.

### 9.3.3 Penalty method

The most widely used method to solve contact problems is the penalty method. This is because only the primal displacement variables enter the formulation; see Section 5.3.2 for the continuum formulation and Chapters 7 and 8 for the associated discretization techniques.

The algorithm for the penalty method can be summarized for frictionless contact in Box 2. Usually, the solution of  $\mathbf{G}_c(\mathbf{u}) = \mathbf{0}$  is performed by a NEWTON-RAPHSON iteration, leading to

$$\begin{aligned} \mathbf{K}_T^P(\mathbf{u}_i^k) \Delta \mathbf{u}_i^{k+1} &= -\mathbf{G}^P(\mathbf{u}_i^k) \\ \mathbf{u}_i^{k+1} &= \mathbf{u}_i^k + \Delta \mathbf{u}_i^{k+1}, \end{aligned} \quad (9.70)$$

where  $\mathbf{K}_T^P(\mathbf{u}_i^k)$  is the tangent matrix resulting from the linearization of  $\mathbf{G}^P(\mathbf{u}_i^k)$ . The iteration index  $k$  is related to the NEWTON loop to solve  $\mathbf{G}^P(\mathbf{u}_i) = \mathbf{0}$  in Box 2. Often the active set strategy, stated in Box 2, is accelerated in such a way that the update of the active set of contact constraints is performed within each step in the NEWTON iteration. In this case the iteration (9.71) yields

$$\begin{aligned} \mathbf{K}_T^P(\mathbf{u}_i) \Delta \mathbf{u}_{i+1} &= -\mathbf{G}^P(\mathbf{u}_i) \\ \mathbf{u}_{i+1} &= \mathbf{u}_i + \Delta \mathbf{u}_{i+1}, \end{aligned} \quad (9.71)$$

which is considerably faster.

**REMARK 9.2**

1. Both procedures might not converge for all cases, and thus have to be applied with care. The problem which occurs is called *jamming* in the optimization literature. Jamming, also known as *zig-zagging*, means that the algorithm jumps in consecutive iterations between two values. One possibility to avoid this is to keep a contact constraint active as long as the new active set is not a subset of the old active set. This will be changed once the new active set is a subset of the old active set, or the increment  $\Delta \mathbf{u}_{i+1}$  is zero for the active set; see also the quadratic programming algorithm in Box 9.
2. Within the algorithm of Box 2 an increase of the penalty parameter is necessary when the final result shows visible penetration, and thus does not fulfil the constraint equation  $g_N = 0$  in a correct way; see the example in Section 5.4, Figure 5.5. On the other hand, a penalty parameter which has been chosen too large can lead to ill-conditioning of the equation system, and thus has to be reduced to avoid this. One possibility for the choice of  $\epsilon_N$  is to relate the penalty parameter to the bulk modulus of the contacting bodies. Ill-conditioning can be so bad that it affects the direct solver, but it is even worse when the solution of the incremental equation system in Box 2 is performed by an iterative solution method in which the condition number of the tangent matrix might change drastically during the solution process. In such a case, special pre-conditioners have to be applied, e.g. see Schöberl (1997b), Klawonn (1998) or Dostal (1999). However, since it is quite hard to estimate the penalty parameter for all cases, it makes sense to apply the augmented LAGRANGE technique.

One estimate for the choice of the penalty parameter was reported in Nour-Omid and Wriggers (1987), which relies on an error analysis taking into account roundoff errors as well as errors due to the penalty approach. It leads to the simple formula for finite element discretizations of continuum problems:

$$\epsilon_N \leq \frac{k_{min}}{\sqrt{N}t}, \quad (9.72)$$

Initialize algorithm: set  $\mathbf{u}_1 = \mathbf{0}$ ,  $\epsilon_N = \epsilon_0$

LOOP over iterations :  $i = 1, \dots$ , convergence

Check for contact:  $g_{Nsi} \leq 0 \rightarrow$  active node, segment or element

Solve:  $\mathbf{G}^P(\mathbf{u}_i) = \mathbf{G}(\mathbf{u}_i) + \epsilon_N \mathbf{C}^c(\mathbf{u}_i)^T \mathbf{G}^c(\mathbf{u}_i) = \mathbf{0}$

Check for convergence:  $\|\mathbf{G}^P(\mathbf{u}_i)\| \leq TOL \Rightarrow$  STOP

END LOOP

Eventually update penalty parameter:  $\epsilon_N$

Box 2. Contact algorithm using the penalty method.



where  $N$  is the number of unknowns of the equation system,  $k_{\min}$  is the smallest stiffness coefficient in the tangent matrix which is modified by  $\epsilon_N$ , and  $t$  represents the roundoff error which is  $\approx 10^{-17}$  for double precision analysis.

3. In the case of frictional contact problems, the algorithm works as described in Box 2. The only difference is that the local algorithm to integrate the constitutive behaviour due to friction is applied within the NEWTON loop to obtain  $\mathbf{G}^P(\mathbf{u}_i) = \mathbf{0}$  for a given set of active constraints.

### 9.3.4 Lagrange multiplier method

Another method which is applied to solve contact problems is the LAGRANGE multiplier method. This method fulfils the contact constraints exactly, but by introducing additional variables, the LAGRANGE multipliers (see Section 5.3.1 for the continuum formulation, and again Chapters 7 and 8 for associated discretization techniques).

One algorithm for the LAGRANGE multiplier method is stated in Box 3 for frictionless contact, where the notation introduced in (9.49) was used. The solution of  $\mathbf{G}^{LM}(\mathbf{w}) = \mathbf{0}$  is obtained by the NEWTON-RAPHSON iteration, leading to

$$\begin{aligned} \mathbf{K}^{LM}(\mathbf{w}_i^k) \Delta \mathbf{w}_i^{k+1} &= -\mathbf{G}^{LM}(\mathbf{w}_i^k) \\ \mathbf{w}_i^{k+1} &= \mathbf{w}_i^k + \Delta \mathbf{w}_i^{k+1}, \end{aligned} \quad (9.73)$$

where the structure of the tangent matrix  $\mathbf{K}_T^{LM}(\mathbf{w}_i^k)$  can be found in (9.47). The iteration index  $k$  is related to the NEWTON loop to solve  $\mathbf{G}^{LM}(\mathbf{w}_i) = \mathbf{0}$  in Box 3. An acceleration of the solution procedure can be obtained by the following iteration in Box 3:

$$\begin{aligned} \mathbf{K}^{LM}(\mathbf{w}_i) \Delta \mathbf{w}_{i+1} &= -\mathbf{G}^{LM}(\mathbf{w}_i) \\ \mathbf{w}_{i+1} &= \mathbf{w}_i + \Delta \mathbf{w}_{i+1}, \end{aligned} \quad (9.74)$$

Initialize algorithm: set  $\mathbf{w}_1 = \mathbf{0}$

LOOP over iterations :  $i = 1, \dots$ , convergence

Check gap:  $g_{Nsi} \leq 0 \rightarrow$  active node, segment or element

Check pressure:  $\lambda_{Nsi} \geq 0 \rightarrow$  node, segment or element not active

Solve:  $\mathbf{G}^{LM}(\mathbf{w}_i) = \mathbf{0}$

Check for convergence:  $\|\mathbf{G}^{LM}(\mathbf{w}_i)\| \leq TOL \Rightarrow \text{STOP}$

END LOOP

Box 3. Contact algorithm for the LAGRANGE multiplier approach.

which at each iteration step includes the change of constraints and hence is faster than the basic algorithm in Box 3. Remarks 9.2.1 and 9.2.3 also apply here. One should note, however, that this procedure does not guarantee convergence of the method. A more reliable but slower algorithm is discussed in Section 9.3.8, see Box 9.

**REMARK 9.3:** When using the LAGRANGE multiplier approach, one has to take special care for rigid body modes which can occur in one or more contacting bodies. In such cases the following procedures can be employed to side-step this difficulty, which leads to singular tangent matrices:

1. Introduction of as many springs in the finite element model as there are rigid body modes. The spring stiffness has to be as soft as possible to avoid an influence of the stiffness on the overall solution.
2. Solve the problem using specified displacements instead of applied forces. This approach only works if the force is applied on the body with rigid body modes.
3. Replace zero diagonal elements in  $\mathbf{D}$  which appear during the factorization of  $\mathbf{K}_T$  in  $\mathbf{L}\mathbf{D}\mathbf{L}^T$  by a small number of the order of required accuracy for the problem. This, in fact, removes the rigid body modes.

### 9.3.5 Augmented Lagrange method, Uzawa algorithm

A combination of the penalty and the LAGRANGE multiplier techniques leads to the so-called augmented LAGRANGE methods, which try to combine the merits of both approaches. A general discussion of these techniques can be found in Glowinski and Le Tallec (1984), and with special attention also to inequality constraints in Bertsekas (1984) and Pietrzak and Curnier (1999). A matrix formulation for frictionless contact follows from the augmented functional (5.58) in Section 5.3.8:

$$\Pi^{AM}(\mathbf{u}) = \Pi(\mathbf{u}) + \mathbf{\Lambda}^T \mathbf{G}^c(\mathbf{u}) + \frac{\epsilon_N}{2} \mathbf{G}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) - \frac{1}{2\epsilon_N} \boldsymbol{\lambda}^T \boldsymbol{\lambda}, \quad (9.75)$$

where  $\mathbf{\Lambda}$  contains all contributions related to contact nodes fulfilling  $\lambda_{s m+1} = \lambda_{s m} + \epsilon_N g_{Ns} \leq 0$ , and  $\boldsymbol{\lambda}$  is related to the nodes with  $\lambda_{s m+1} > 0$ . As discussed in Section 5.3.8, this mixture with the perturbed LAGRANGE method ensures  $C^1$  differentiability of the functional (9.75).

Its variation with respect to displacements and LAGRANGE multipliers yields the nonlinear equation system

$$\begin{aligned} \mathbf{G}(\mathbf{u}) + \mathbf{\Lambda}^T \mathbf{C}^c(\mathbf{u}) + \epsilon_N \mathbf{C}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) &= \mathbf{0} \\ \mathbf{G}^c(\mathbf{u}) &= \mathbf{0} \\ -\frac{1}{\epsilon_N} \boldsymbol{\lambda} &= \mathbf{0}. \end{aligned} \quad (9.76)$$

As in the previous sections, we can employ NEWTON's method to solve (9.76), which leads to the incremental equation system at the state  $(\mathbf{u}_m, \mathbf{\Lambda}_m)$ :

$$\begin{bmatrix} \hat{\mathbf{K}}_T + \epsilon_N \mathbf{C}^{cT} \mathbf{C}^c & \mathbf{C}^{cT} & \mathbf{0} \\ \mathbf{C}^c & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\frac{1}{\epsilon_N} \mathbf{I} \end{bmatrix}_m \begin{Bmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{\Lambda} \\ \Delta \boldsymbol{\lambda} \end{Bmatrix}_{m+1} = - \begin{Bmatrix} \hat{\mathbf{G}} \\ \mathbf{G}^c \\ -\frac{1}{\epsilon_N} \boldsymbol{\lambda} \end{Bmatrix}_m. \quad (9.77)$$

Here all dependencies of the matrix elements with respect to  $\mathbf{u}$  and  $\mathbf{\Lambda}$  have been omitted to shorten notation. This system of equations can, together with (9.76), be used within a NEWTON type algorithm, as discussed in Box 3. It can be seen from (9.77) that the last term in (9.75) affects the LAGRANGE multiplier  $\lambda$  so that it is zero in the next iteration step when  $\lambda_{s\ m+1} > 0$  was detected.

Augmented LAGRANGE techniques are usually applied together with USZAWA type algorithms, see Bertsekas (1984), Glowinski and Le Tallec (1984), which lead to an inner loop for the contact and an outer loop for the update of the LAGRANGE parameters. This iteration procedure increases the total number of iterations, but yields an algorithm which can be implemented easily. For applications of augmented LAGRANGE techniques to contact problems within the finite element method, see Wriggers et al. (1985) and Simo and Laursen (1992), or for a symmetrical treatment of the frictional part Laursen and Simo (1993a) or Zavarise et al. (1995). In the case of "high contact precision", when constitutive interface laws are employed, special augmented LAGRANGE techniques are needed, since often ill-conditioning of the problem may occur, see Wriggers and Zavarise (1993a).

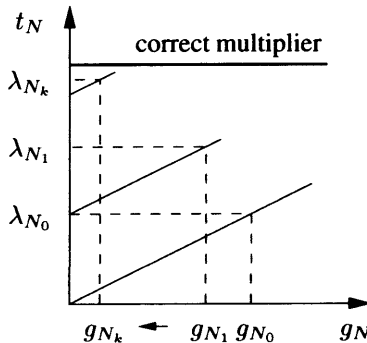
The matrix formulation of the USZAWA algorithm for frictionless contact problems starts from the augmented LAGRANGE functional with constant LAGRANGE multipliers  $\bar{\mathbf{\Lambda}}$

$$\Pi^{UA}(\mathbf{u}) = \Pi(\mathbf{u}) + \bar{\mathbf{\Lambda}}^T \mathbf{G}^c(\mathbf{u}) + \frac{\epsilon_N}{2} \mathbf{G}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) \rightarrow MIN. \quad (9.78)$$

The new LAGRANGE multiplier is computed at the solution point of  $\Pi^{UA}(\mathbf{u}, \bar{\mathbf{\Lambda}}_k)$  within the augmentation loop over the loop index  $k$  (see also Box 4) using the first order update formula

$$\bar{\lambda}_{s\ k+1} = \bar{\lambda}_{s\ k} + \min[\epsilon_N g_{Ns}(\mathbf{u}_{k+1}), \bar{\lambda}_{s\ k}], \quad (9.79)$$

which is written for the LAGRANGE parameter of the contact nodes or contact segments  $s$ . This update is visualized in Figure 9.12. Here the update process for the



**Fig. 9.12** Schematic update process of LAGRANGE multiplier.

LAGRANGE multiplier in the normal direction is considered. One can observe that  $\lambda_{Nk}$  approaches the correct value of the LAGRANGE multiplier for  $g_{Nk} \rightarrow 0$  in the updating procedure. Note that the slope of the linear relation between  $g_N$  and  $\lambda_N$  is  $\epsilon_N$ . From this, it is clear that the convergence behaviour of the outer augmentation loop in Box 4 does strongly depend upon the penalty parameter.

The variation of (9.78) only needs to be performed with respect to displacements

$$\mathbf{G}^{UA}(\mathbf{u}, \bar{\Lambda}) = \mathbf{G}(\mathbf{u}) + \bar{\Lambda}^T \mathbf{C}^c(\mathbf{u}) + \epsilon_N \mathbf{C}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) = 0. \quad (9.80)$$

For a known fixed LAGRANGE multiplier  $\bar{\Lambda}$ , NEWTON'S method is used to solve the nonlinear problem (9.80). This results after linearization in the equation system at the state  $(\mathbf{u}_i, \bar{\Lambda})$

$$\left[ \hat{\mathbf{K}}_T(\mathbf{u}_i, \bar{\Lambda}) + \epsilon_N [\mathbf{C}^c(\mathbf{u}_i)]^T \mathbf{C}^c(\mathbf{u}_i) \right] \Delta \mathbf{u}_{i+1} = -\hat{\mathbf{G}}^{UA}(\mathbf{u}_i, \bar{\Lambda}). \quad (9.81)$$

The global augmented LAGRANGE algorithm is shown in Box 4. Here we again use the discrete formulation (8.20) and (8.36), which has to be adjusted to incorporate the fixed LAGRANGE parameters  $\bar{\Lambda}_{Ns}$ ; see (8.13) for the node-to-segment discretization. Hence, the vector  $\mathbf{G}^{UA}$  contains all contributions of the active contact element  $s$ .

Let us note that it is standard practice in augmented LAGRANGE iterations to update the penalty number  $\epsilon_N$  in order to obtain good convergence, see Bertsekas (1984). This is due to the fact that a small penalty parameter leads to very slow convergence, since the update formula (9.79) is of first order and the contact forces

Initialize algorithm: set  $\mathbf{u}_0 = \mathbf{0}$ ,  $\bar{\Lambda}_0 = 0$ ,  $\epsilon_N = \epsilon_{N0}$

LOOP over augmentations:  $k = 1, \dots$ , convergence

    LOOP over iterations :  $i = 1, \dots$ , convergence

        Solve:  $\mathbf{G}^{UA}(\mathbf{u}_i, \bar{\Lambda}_k) = \mathbf{0}$

        Check for convergence:  $\|\mathbf{G}^{UA}(\mathbf{u}_i)\| \leq TOL \Rightarrow \text{END LOOP}$

    END LOOP

    LOOP over contact nodes :  $s = 1, \dots, n_c$

        Update:  $\bar{\lambda}_{sk+1}$  according to (9.79)

        Update:  $\epsilon_{Nk+1}$  according to (9.82)

        Check for convergence:  $\|g_{Ns}(\mathbf{u}_i)\| \leq TOL \Rightarrow \text{STOP}$

    END LOOP

END LOOP

Box 4. USZAWA algorithm.

due to a small penalty number are small. Thus, it makes sense to increase the penalty parameter within a contact element  $s$  according to an update scheme, see Bertsekas (1984). The update scheme with a starting value for  $\epsilon_{N0}$  being computed from (9.72), for example, yields

$$\epsilon_{Nn+1} = \begin{cases} 10 \cdot \epsilon_{Nn} & \text{for } |g_N(\mathbf{u}_{n+1})| > \frac{1}{4} \cdot |g_N(\mathbf{u}_n)| \\ \epsilon_{Nn} & \text{for } |g_N(\mathbf{u}_{n+1})| \leq \frac{1}{4} \cdot |g_N(\mathbf{u}_n)| \end{cases} \quad (9.82)$$

In relation (9.82) a stopping criterion for the update of the penalty parameter has also been introduced to avoid ill-conditioning of the equation system (9.81).

### 9.3.6 Partitioning method

Since the number of displacement unknowns is usually much higher than the unknowns related to the LAGRANGE multipliers, it is advantageous to reformulate contact problems such that the unknowns are only the LAGRANGE multipliers. This can be achieved easily for geometrically linear elasticity problems by a substructuring method, see Section 9.3.2. However, when a numerical simulation of contact problems involving large strains and inelastic materials has to be considered, only an incremental dual formulation can be established, since the matrix from the SCHUR complement depends upon the deformation.

When the LAGRANGE multiplier method in Section 9.3.4 is used, special care has to be taken during the factorization due to zero elements on the diagonal of the coefficient matrix. Furthermore, since the contact problem is nonlinear, the iteration in Box 3 has to be applied and the factorization has to be carried out in each NEWTON step. Here a method will be stated which takes advantage of the fact that the part of the coefficient matrix associated with the degrees of freedom that do not come into contact remains unchanged during the iteration process. This is, of course, only true for elastic structures with small deformations which will be considered first. However, a nonlinear version of this approach can also be developed.

From the dual formulation in (9.60) a system of equations for the unknown LAGRANGE multipliers can be derived:

$$\mathbf{N} \mathbf{K}_T^{-1} \mathbf{N}^T \boldsymbol{\Lambda} = -\mathbf{G}_X + \mathbf{N} \mathbf{K}_T^{-1} \mathbf{P}. \quad (9.83)$$

Furthermore, from (9.59) we obtain an equation for the unknown displacements  $\mathbf{u}$

$$\mathbf{u} = \mathbf{K}_T^{-1} (\mathbf{P} - \mathbf{N}^T \boldsymbol{\Lambda}), \quad (9.84)$$

which still depends upon the LAGRANGE multipliers.

The matrix  $\mathbf{N} \mathbf{K}_T^{-1} \mathbf{N}^T$  in (9.83) is symmetric, positive definite and full. The size of this matrix depends upon the number of nodes  $s \in \mathcal{J}_A$  that are actually in contact, which is in general much smaller than the total number of degrees of freedom  $N$ . The system of equations (9.83) may be solved using direct procedures. However, this approach requires the evaluation of the coefficient matrix which is not known explicitly. Therefore, an iterative method was advocated in Wriggers and Nour-Omid

Initialize algorithm: use initial approximation for  $\Lambda_0$  and compute

$$\mathbf{p}_0 = \mathbf{r}_0 = \mathbf{N} \mathbf{K}_T^{-1} (\mathbf{P} - \mathbf{N}^T \Lambda) - \mathbf{G}_X$$

$$\rho_0 = \mathbf{r}_0^T \mathbf{r}_0$$

LOOP for  $k = 0, 1, \dots$ , until convergence

$$\mathbf{d}_k = \mathbf{N} \mathbf{K}_T^{-1} \mathbf{N}^T \mathbf{p}_k$$

$$\alpha_k = \rho_k / (\mathbf{d}_k^T \mathbf{p}_k)$$

$$\Lambda_{k+1} = \Lambda_k + \alpha_k \mathbf{p}_k$$

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{d}_k$$

$$\rho_{k+1} = \mathbf{r}_{k+1}^T \mathbf{r}_{k+1}$$

Check for convergence:  $\rho_{k+1} \leq TOL \Rightarrow \text{END LOOP}$

$$\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \rho_{k+1} / \rho_k \mathbf{p}_k$$

END LOOP

Box 5. *cg*-algorithm for evaluating the contact forces.

(1984), which does not need computation of the elements of the coefficient matrix. Since the matrix is positive definite, the conjugate gradient method (*cg*-method) can be employed, which is described in Box 5. The step where  $\mathbf{d}_k$  is computed is the most costly one. Here a triangular factorization  $\mathbf{K}_T = \mathbf{L} \mathbf{D} \mathbf{L}^T$  can be used together with the static condensation procedure in (9.54) to minimize the number of operations.

The contact algorithm is now designed in such a way that an inner and an outer iteration are performed. It is thus possible to solve a general problem with material and geometrical nonlinearities. In the inner iteration the set of active contact constraints and the LAGRANGE multipliers are determined based on the dual formulation. This task is performed for a given state of the outer iteration with displacements  $\bar{\mathbf{u}}$ , LAGRANGE multipliers  $\bar{\Lambda}$ , and a set of contact constraints  $\bar{\mathbf{N}}$ . The inner iteration scheme is described in Box 6, where the solution of the equation system in the inner iteration is obtained with the *cg*-method. This iteration yields, with  $\mathbf{N}_i$ , the contact conditions and the LAGRANGE multipliers  $\bar{\Lambda}$  which have to be enforced in the next NEWTON step in the outer iteration. Note that, in the case of a geometrically nonlinear problem, we have to exchange equation (9.35) by (9.34), and to use the tangent matrix  $\mathbf{K}_T$  evaluated at the current state of the outer iteration.

Vector  $\mathbf{N}_i$  depends upon the normal vectors of the master surfaces at the projection points of the slave nodes. In large deformation problems, the change of these normal vectors has to be considered. However, since the state within a NEWTON step is fixed we can neglect the change of the normals in the inner iteration in Box 6.

Let  $\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}, \bar{\mathbf{N}}$  and set

$$\mathbf{u}_0 = \bar{\mathbf{u}} \text{ and } \mathbf{N}_0 = \bar{\mathbf{N}}$$

LOOP for  $i = 0, 1, \dots$ , until convergence

Determine all active contact constraints from (9.35)

$$\mathbf{p}_i = -\mathbf{G}_X + \mathbf{N}_i \mathbf{u}_i$$

If  $\|\mathbf{p}_i\| \leq TOL \|\mathbf{p}_0\| \Rightarrow \text{END LOOP}$

$$\text{Solve: } \mathbf{N}_i \mathbf{K}_T^{-1} \mathbf{N}_i^T \mathbf{\Lambda}_{i+1} = \mathbf{p}_i$$

If an element of  $\mathbf{\Lambda}_{i+1}$  is positive reset it to zero

$$\Delta \mathbf{u}_{i+1} = \mathbf{K}_T^{-1} \mathbf{N}_i^T \mathbf{\Lambda}_{i+1}$$

$$\mathbf{u}_{i+1} = \mathbf{u}_0 + \Delta \mathbf{u}_{i+1}$$

END LOOP

$$\bar{\mathbf{u}} \leftarrow \mathbf{u}_i, \bar{\mathbf{\Lambda}} \leftarrow \mathbf{\Lambda}_i, \bar{\mathbf{N}} \leftarrow \mathbf{N}_i$$

Box 6. Inner contact iteration.

Start from the initial approximation  $\bar{\mathbf{u}}_0, \bar{\mathbf{\Lambda}}_0$  and  $\bar{\mathbf{N}}_0$

LOOP for  $m = 0, 1, \dots$ , until convergence

Compute residual force  $\mathbf{G}_m = \mathbf{R}(\bar{\mathbf{u}}_m) - \mathbf{P}$  in (9.36)

If  $\|\mathbf{G}_m + \bar{\mathbf{N}}_m \bar{\mathbf{\Lambda}}_m\| \leq TOL \|\mathbf{G}_0\| \Rightarrow \text{END LOOP}$

$$\text{Solve: } \mathbf{K}_T(\bar{\mathbf{u}}_m) \Delta \bar{\mathbf{u}}_{m+1} = -\mathbf{G}_m - \bar{\mathbf{N}}_m \bar{\mathbf{\Lambda}}_m$$

$$\bar{\mathbf{u}}_{m+1} = \bar{\mathbf{u}}_m + \Delta \bar{\mathbf{u}}_{m+1}$$

Perform inner iteration, see Box 6.

END LOOP

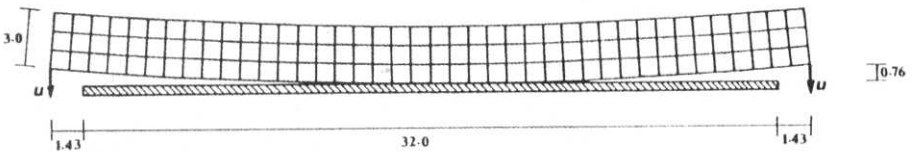
Box 7. Outer contact iteration.

The outer iteration considers all nonlinearities related to material and/or geometry within a modified NEWTON scheme which incorporates the inner iteration. Hence, the outer iteration is not needed for a geometrically linear elastic problem. In that case, the iteration terminates in the first loop for  $m = 0$ . The outer contact iteration is

stated in Box 7. In case rigid body motions of one of the contacting bodies is possible, we refer to the methods discussed in Remark 9.3.

Note that the inner-outer iteration does not require more triangular factorizations than the standard NEWTON method. In many practical cases, the algorithm described in Boxes 6 and 7 will need fewer factorizations of the tangent matrix.

The following example shows the performance of the inner-outer iteration when compared with the standard penalty approach in Section 9.3.3. Figure 9.13 depicts a finite element mesh of a thick circular beam which is discretized by 120 plane stress finite elements. The modulus of elasticity is chosen as  $E = 1000$  and POISSON'S ratio is  $\nu = 0.3$ . In the undeformed configuration the beam is in contact at its centre with the rigid surface. A displacement of  $\bar{u} = 0.8$  is applied at both ends of the model. The contact area therefore changes from the middle of the beam, where we observe uplifting, to its ends. The algorithm has to capture this behaviour. The problem is solved using the partitioning method with one outer and six inner iterations. Only one outer iteration is needed, since the problem is geometrically linear. Also, the penalty method was applied to analyse this problem. Table 9.1 shows the number of iterations and the gaps computed for different penalty parameters. The optimal penalty parameter, obtained from equation (9.72), is  $\epsilon_N = 10^8$ . Here optimality refers to accuracy in fulfillment of the constraint equation; see the last column in Table 9.1. For applications where a lower accuracy may be sufficient, a smaller penalty number can translate into fewer iterations; see the second column in Table 9.1.



**Fig. 9.13** Finite-element-model of contact problem.

**Table 9.1** Influence of penalty parameter on convergence

Penalty Parameter	Iterations	Max. Penetration
$10^1$	4	$5.27 \times 10^{-2}$
$10^2$	5	$2.73 \times 10^{-2}$
$10^3$	6	$7.09 \times 10^{-3}$
$10^4$	8	$9.24 \times 10^{-4}$
$10^6$	8	$9.28 \times 10^{-6}$
$10^8$	8	$9.29 \times 10^{-8}$



### 9.3.7 SQP method

Successive quadratic programming methods employ NEWTON'S method or quasi-NEWTON methods to solve the KUHN-TUCKER conditions of a contact problem. These methods are also known as sequential or recursive quadratic programming (SQP methods).

In the process of the successive quadratic programming, a subproblem is solved which can be stated as the minimization of a quadratic approximation of the nonlinear contact problem

$$\min \Pi(\mathbf{u}) \quad \text{subject to} \quad g_{Ns} \geq 0 \quad \forall s \in \mathcal{J}_C, \quad (9.85)$$

where a linear approximation of the constraints is used. The method can thus be described as a projected LAGRANGE approach. The algorithm is constructed so that a series of quadratic programming subproblems is solved successively. Hence, this algorithm relies on fast and robust methods for generating solutions for the quadratic program, see Goldfarb and Idnani (1983). The sequential quadratic programming approach was derived in Spellucci (1985). Other variants of SQP-algorithms may be found in e.g. Schittkowski (1992).

Based on the derivation in Section 9.3.2, the quadratic approximation of (9.85) yields the quadratic program

$$\begin{aligned} \text{QP}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) : \quad & \text{Minimize} \quad \frac{1}{2} \Delta \mathbf{u}^T \hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}}) \Delta \mathbf{u} + \hat{\mathbf{G}}(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})^T \Delta \mathbf{u} + \Pi(\bar{\mathbf{u}}) \\ & \text{subject to} \quad \mathbf{G}^c(\bar{\mathbf{u}}) + \mathbf{C}^c(\bar{\mathbf{u}}) \Delta \mathbf{u} \geq 0, \end{aligned} \quad (9.86)$$

where all matrices and vectors are defined according to equation (9.64). The values  $(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$  represent a known state within the iterative solution procedure.  $\Delta \mathbf{u}$  is the increment of the displacement. Since this problem represents a second order approximation of the LAGRANGE function  $\Pi(\mathbf{u}) + \mathbf{G}^c(\mathbf{u})^T \mathbf{\Lambda}$ , quadratic convergence can be expected when the iteration starts close to the solution point.

With these preliminary remarks, the SQP algorithm can be derived. It is stated in Box 8. Here the fact is used that if  $\Delta \mathbf{u}_{m+1}$  solves  $\text{QP}(\mathbf{u}_m, \mathbf{\Lambda}_m)$  with the LAGRANGE multipliers  $\mathbf{\Lambda}_{m+1}$  and if  $\Delta \mathbf{u}_{m+1} = \mathbf{0}$ , then the computed displacements and LAGRANGE multipliers fulfil the KUHN-TUCKER conditions of the original problem (9.85).

Several variants of the SQP method are used in applications. The first is to approximate the tangent matrix  $\hat{\mathbf{K}}_T(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$  by a quasi-NEWTON scheme like the BFGS update, e.g. see Matthies and Strang (1979), Luenberger (1984) or Bertsekas (1984). This avoids costly computations of the tangent matrix which have to be performed in the algorithm in Box 8  $m$ -times. However, it is well known that the convergence of the algorithm is only superlinear. Other variants were developed for sparse matrices, see Bartholomew-Biggs and Hernandez (1995) or see the overview paper by Murray (1997). Versions with respect to contact analysis can also be found in Björkman et al. (1995), who employed special diagonalization methods to obtain separable subproblems for the dual problem.

Start from the initial approximation  $\mathbf{u}_0, \mathbf{\Lambda}_0$

LOOP for  $m = 0, 1, \dots$ , until convergence

Solve:  $\text{QP}(\mathbf{u}_m, \mathbf{\Lambda}_m) \rightarrow \Delta \mathbf{u}_{m+1}$  and  $\mathbf{\Lambda}_{m+1}$

If  $\|\Delta \mathbf{u}_{m+1}\| \leq \text{TOL} \Rightarrow \text{END LOOP}$

$\mathbf{u}_{m+1} = \mathbf{u}_m + \Delta \bar{\mathbf{u}}_{m+1}$

END LOOP

### Box 8. Basic SQP-algorithm.

A possible disadvantage of the SQP method in the version stated in Box 8 is the fact that convergence is only guaranteed when the algorithm starts close to the solution point. Often this condition is not matched in practical applications. To obtain a more robust method one can introduce a merit function. One choice of this function could be

$$M(\mathbf{u}) = \min \Pi(\mathbf{u}) + \mu \sum_{s=1}^{n_c} \min [0, g_{Ns}]. \quad (9.87)$$

The merit function is then applied to find the scalar parameter  $\gamma_m \geq 0$  in the update of the displacements  $\mathbf{u}_{m+1} = \mathbf{u}_m + \gamma_m \Delta \mathbf{u}_{m+1}$  in Box 8 by minimization of  $M(\mathbf{u}_m + \gamma_m \Delta \mathbf{u}_{m+1})$ . This procedure is called *line search*. It is also used for nonlinear unconstrained problems like (9.6) to obtain a global convergence of, for example, the NEWTON method. Here the line search has to be performed with respect to a non-differentiable function, hence only methods which do not need derivatives like the Golden Section Method or FIBONACCI search can be employed, e.g. see Luenberger (1984) or Bazaraa et al. (1993).

### 9.3.8 Active set method for quadratic program

The quadratic program in (9.86) with inequality constraints is solved by an active set method. In most applications the tangent matrix  $\hat{\mathbf{K}}_T$  is positive definite, which then leads to a relatively simple solution strategy. This method will be described next. We assume that in iteration  $m$  a deformation state  $\mathbf{u}_m$  is given that satisfies all contact constraints of the active set of contact nodes  $s \in \mathcal{J}_A^m$  (feasible solution). Hence  $\mathbf{G}^c(\mathbf{u}_m)$  is zero in the linearized contact constraint equation in (9.86). The quadratic program is then defined by

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2} \Delta \mathbf{u}_{m+1}^T \hat{\mathbf{K}}_T(\mathbf{u}_m, \mathbf{\Lambda}_m) \Delta \mathbf{u}_{m+1} + \hat{\mathbf{G}}(\mathbf{u}_m, \mathbf{\Lambda}_m)^T \Delta \mathbf{u}_{m+1} \\ &\text{subject to} \quad \mathbf{C}^c(\mathbf{u}_m) \Delta \mathbf{u}_{m+1} = 0, \quad \forall s \in \mathcal{J}_A^m. \end{aligned} \quad (9.88)$$

Start with a feasible solution  $\mathbf{u}_0$  and a working set  $\mathcal{J}_A^0$

LOOP for  $m = 0, 1, \dots$ , until convergence

1. Solve (9.89). If  $\Delta \mathbf{u}_{m+1} = \mathbf{0}$  then go to 5.
2. Compute:  $\gamma_{m+1} = \min_{s \in \mathcal{J}_A^m} \left[ 1, \frac{G_s^c(\mathbf{u}_m) - [\mathbf{C}_s^c(\mathbf{u}_m)]^T \mathbf{u}_m}{[\mathbf{C}_s^c(\mathbf{u}_m)]^T \Delta \mathbf{u}_m} \right]$
3. If  $\gamma_{m+1} < 1$  then  $\mathcal{J}_A^m \rightarrow \mathcal{J}_A^{m+1}$
4. Set  $\mathbf{u}_{m+1} = \mathbf{u}_m + \gamma_{m+1} \Delta \mathbf{u}_{m+1} \rightarrow \text{END LOOP}$
5. Compute:  $\lambda_{max} = \max_{s \in \mathcal{J}_A^m} (\lambda_m s)$ ,

If  $\lambda_{max} \leq 0 \rightarrow \text{STOP}$  else  $\mathcal{J}_A^m \rightarrow \mathcal{J}_A^{m+1}$

END LOOP

#### Box 9. Active set strategy for quadratic programming.

which has only equality constraints. The vector  $\mathbf{C}^c(\mathbf{u}_m)$  contains equality constraints associated with the active contact nodes  $s$ ; see also (9.34). The solution of this problem is obtained from the linear system

$$\begin{bmatrix} \hat{\mathbf{K}}_T(\mathbf{u}_m; \Lambda_m) & \mathbf{C}^c(\mathbf{u}_m)^T \\ \mathbf{C}^c(\mathbf{u}_m) & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{u}_{m+1} \\ \Lambda_{m+1} \end{Bmatrix} = - \begin{Bmatrix} \hat{\mathbf{G}}(\mathbf{u}_m; \Lambda_m) \\ \mathbf{0} \end{Bmatrix}, \quad (9.89)$$

which can be solved using an efficient linear solver for this type of problem.

The active set strategy is described in Box 9. If in step 3 of the algorithm a new inequality constraint is satisfied, then the node associated with this constraint has to be added to the working set of contact constraints. This defines the new working set  $\mathcal{J}_A^{m+1}$ . The same procedure is used in step 5, where the contact node associated with the maximum LAGRANGE multiplier  $\lambda_{max}$  is deleted from the old working set, leading to  $\mathcal{J}_A^{m+1}$ .

It can be shown that the algorithm terminates within a finite number of iteration steps since there is only a finite number of working sets, e.g. the maximum number of contact nodes. This also means that, for three-dimensional problems with a large number of possible contact nodes, many iterations might be needed to solve the problem. Thus special algorithms for solving (9.89) are needed which avoid the computation of the inverse of the system matrix in (9.89), or its factorization. More refined versions for the solution of quadratic programming methods based on the above formulation can be found in Stoer (1971), Gill and Murray (1978), or in the overview by Spellucci (1993). A method which starts from the dual formulation of the quadratic programming problem (for the derivation see Section 9.3.2) is described in Goldfarb and Idnani (1983).

### 9.3.9 Linear complementary problem

Other methods which are well known in optimization theory are based on the formulation of contact with and without friction as a Linear Complementary Problem (LCP), e.g. see Luenberger (1984) or Bazaraa et al. (1993). Applications to contact problems can be found in Björkman (1991) or Björkman et al. (1995). Its theoretical framework for contact mechanics is provided in Klarbring (1999).

The method is based on solving the KARUSH–KUHN–TUCKER optimality conditions of the dual problem. For a linear elasticity problem we depart from (9.60), whereas a general nonlinear problem relies on the dual formulation (9.68). The latter is defined on the tangent space at the known solution point  $(\bar{\mathbf{u}}, \bar{\mathbf{\Lambda}})$ . Hence, for the general nonlinear case we have to use the LCP algorithm within an iterative method, which leads to a similar structure used in the SQP method, where the QP problem is solved successively. Due to this structure, we will describe the LCP method only for the linear case, since the nonlinear case based on (9.68) is straightforward.

Let us recall the dual problem (9.60)

$$\pi(\mathbf{\Lambda}) = -\frac{1}{2} \mathbf{\Lambda}^T \mathbf{D} \mathbf{\Lambda} - \mathbf{\Lambda}^T \mathbf{d}, \quad (9.90)$$

where the sign for the LAGRANGE multiplier has been changed for convenience. This also changes the direction of the inequality:  $\mathbf{\Lambda} \geq \mathbf{0}$ . The optimality conditions for (9.90) can be stated as

$$\begin{aligned} \mathbf{D} \mathbf{\Lambda} + \mathbf{d} &\geq \mathbf{0}, \\ \mathbf{\Lambda} &\geq \mathbf{0}, \\ \mathbf{\Lambda}^T [\mathbf{D} \mathbf{\Lambda} + \mathbf{d}] &= 0. \end{aligned} \quad (9.91)$$

This set of equations is now recast in the form of a linear complementary problem,

$$\mathbf{w} = \mathbf{d} + \mathbf{D} \mathbf{\Lambda} \geq \mathbf{0}, \quad \mathbf{\Lambda} \geq \mathbf{0}, \quad \mathbf{w}^T \mathbf{\Lambda} = 0. \quad (9.92)$$

The new variable  $\mathbf{w}$  which is introduced to convert the inequalities to equalities is called a *slack variable*. Solution of (9.92) can be obtained with standard algorithms. Here we should like to mention the principal pivot algorithm by Cottle and Dantzig (1968) and LEMKE'S method, e.g. see Cottle et al. (1992).

### 9.3.10 Contact algorithm of Dirichlet–Neumann type

Domain decomposition methods are increasingly being applied to the numerical simulation of large engineering problems which do not fit on single processor computers, and hence have to be solved in a parallel computing environment. Parallel computers need data exchange between the processors. Since the contact area changes during an incremental solution procedure, the data structure for the exchange of data also has to be modified. It can thus be advantageous to construct an algorithm which employs a strategy in which the bodies coming into contact are treated separately. Such an

Set initial values:  $\mathbf{v}_0 = \mathbf{0}$ ,  $\mathbf{t}_0 = \mathbf{0}$

Choose damping parameters:  $\omega_D \in ]0, 1]$  and  $\omega_N \in ]0, 1]$

LOOP for  $m = 0, 1, \dots$ , until convergence

Solve NP:  $\mathbf{K}^1 \mathbf{u}_{m+1}^1 = \mathbf{P}^1 - \mathbf{t}_m$

Transfer displacement:  $\mathbf{v}_{m+1} = (1 - \omega_D) \mathbf{v}_m + \omega_D \mathbf{Q} \mathbf{u}_{m+1}^1$

Solve contact problem:  $\frac{1}{2} \mathbf{u}_{m+1}^{2T} \mathbf{K}^2 \mathbf{u}_{m+1}^2 - \mathbf{u}_{m+1}^{2T} \mathbf{P}^2 \rightarrow \text{MIN}$

subject to  $\mathbf{N}^1(\mathbf{v}_{m+1}) \mathbf{u}_{m+1}^2 + \mathbf{G}_X^1(\mathbf{v}_{m+1}) \geq 0$

Compute residual:  $\mathbf{R}_{m+1}^2 = \mathbf{K}^2 \mathbf{u}_{m+1}^2 - \mathbf{P}^2$

Transfer boundary tractions:  $\mathbf{t}_{m+1} = (1 - \omega_N) \mathbf{t}_m + \omega_N \mathbf{Q}^T \mathbf{R}_{m+1}^2$

END LOOP

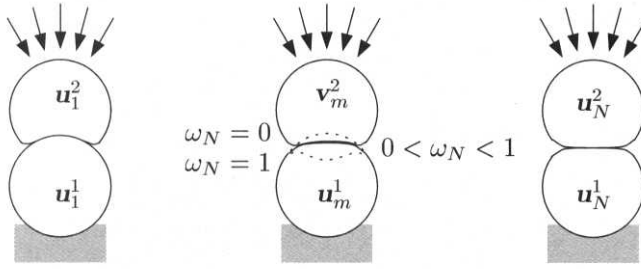
Box 10. DIRICHLET-NEUMANN contact algorithm with different iteration stages.

algorithm was developed in Krause and Wohlmuth (2001). It is based on a nonlinear block GAUSS-SEIDEL method as an iterative solver. From the engineering point of view, it can be interpreted as a DIRICHLET-NEUMANN algorithm for the nonlinear contact problem.

With the notation in Section 9.3.1, we derive the algorithm. The basic idea is that one applies on body  $B^1$  the surface tractions which were computed from a contact problem for body  $B^2$  with a fixed deformed state of body  $B^1$ . This algorithm is described in Box 10 for the case of a linear elasticity problem with constant stiffness matrices  $\mathbf{K}^1$  and  $\mathbf{K}^2$  and given load vectors  $\mathbf{P}^1$  and  $\mathbf{P}^2$  for bodies  $B^1$  and  $B^2$ , respectively. The constraint condition  $\mathbf{N}^1(\mathbf{v}_{m+1}) \mathbf{u}_{m+1}^2 + \mathbf{G}_X^1(\mathbf{v}_{m+1}) \geq 0$  is formulated with respect to the deformed surface of body  $B^1$ . This reflects the dependency of the matrix containing the normal vectors  $\mathbf{N}^1$  and the initial gap vector  $\mathbf{G}_X^1$  on the current displacement  $\mathbf{v}_{m+1}$ .

The matrix  $\mathbf{Q}$  is introduced into the algorithm to transfer the boundary displacements and surface tractions of the finite element mesh of body  $B^1$  to the mesh of body  $B^2$  in case these meshes do not match at the contact interface. The structure of matrix  $\mathbf{Q}$  depends upon the applied discretization scheme. The associated formulations can be found, for example, for the node-to-segment contact in Section 8.1, and for mortar techniques in Section 7.4.2.

The iterative behaviour is depicted in Figure 9.14. Note that for  $\omega_D = 1$ , a choice of  $\omega_N = 1$  does not lead to a convergent scheme, since the solution iterates between the first two solutions. Numerical tests in Krause and Wohlmuth (2001) show that  $\omega_N = \omega_D = 0.7$  yields an algorithm with good convergence behaviour. A more refined algorithm of this type can be found in Wohlmuth and Krause (2002).



**Fig. 9.14** Contact algorithm using a DIRICHLET-NEUMANN technique.

This algorithm can also be applied to frictional contact problems. The only change is that the one-sided contact problem in Box 10 has to be solved for frictional contact. For that task the algorithms described in Section 9.4 can be applied.

### 9.3.11 Algorithm for dynamic contact

Dynamical contact problems are often associated with short process times of only several microseconds, as in the case of care impacts or fragmentation of solids due to blast loading. In such cases, it is advisable to use explicit integration schemes. However, there are also time-dependent engineering problems which require implicit algorithmic schemes, such as earthquake analysis or the vibration of machines. In general, we can state the equations of motions as

$$\mathbf{M}\mathbf{a} + \mathbf{R}(\mathbf{u}) - \mathbf{P} + \mathbf{G}_c(\mathbf{u}) = \mathbf{0}. \quad (9.93)$$

All terms have been defined previously, e.g. see (9.8) and Section 9.3.1.

Algorithms for frictionless and frictional contact-impact have been developed recently in e.g. Laursen and Chawla (1997), Armero and Petocz (1998), Armero and Petocz (1999) and Laursen and Love (2002) using implicit time integration schemes. But explicit methods are also well established for contact-impact analysis, and have a long history, starting with the HEMP-hydrocode. Associated algorithms have been implemented in explicit finite element codes such as DYNA3D or PRONTO3D. For the theoretical background and numerous applications see Hallquist et al. (1985), Benson and Hallquist (1990), Hallquist et al. (1992) or Heinstein et al. (2000).

Here we should like to state in more detail an algorithm which has been developed recently by Kane et al. (1999). It relies on geometric arguments and employs the relation of contact algorithms to the well known radial return algorithms used in elastoplastic finite element analysis. Hence, these authors advocate a scheme which treats the contact force system in an implicit manner while using either implicit or explicit approximations for the terms in the equation of motions which are not affected by contact. Hence the algorithm can be viewed as an explicit/implicit one.

The accelerations are therefore split into two parts, reflecting the contact and internal forces,

$$\mathbf{a} = \mathbf{a}^{int} + \mathbf{a}^{con}. \quad (9.94)$$

Using (9.93), we can define these two accelerations as

$$\begin{aligned} \mathbf{a}^{int} &= \mathbf{M}^{-1} [\mathbf{P} - \mathbf{R}(\mathbf{u})], \\ \mathbf{a}^{con} &= \mathbf{M}^{-1} \mathbf{G}_c(\mathbf{u}). \end{aligned} \quad (9.95)$$

The approximations (9.22) can now be modified slightly by accounting for the split (9.94) as

$$\begin{aligned} \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta t \mathbf{v}_n + (\Delta t)^2 [(1 - 2\beta) \mathbf{a}_n^{int} + \beta \mathbf{a}_{n+1}^{int}] + \frac{(\Delta t)^2}{2} \mathbf{a}_{n+1}^{con} \\ \mathbf{v}_{n+1} &= \mathbf{v}_n + \Delta t [(1 - \gamma) \mathbf{a}_n^{int} + \gamma \mathbf{a}_{n+1}^{int}] + \Delta t \mathbf{a}_{n+1}^{con}. \end{aligned} \quad (9.96)$$

By using the abbreviation

$$\hat{\mathbf{u}}_{n+1} = \mathbf{u}_n + \Delta t \mathbf{v}_n + (\Delta t)^2 (1 - 2\beta) \mathbf{a}_n^{int}, \quad (9.97)$$

evaluated at  $t_n$ , (9.96)<sub>1</sub> can be written as

$$\mathbf{u}_{n+1} = \hat{\mathbf{u}}_{n+1} + (\Delta t)^2 \left[ \beta \mathbf{a}_{n+1}^{int} + \frac{1}{2} \mathbf{a}_{n+1}^{con} \right], \quad (9.98)$$

which then yields, together with (9.93) and (9.95), a nonlinear algebraic system of equations for the unknown displacements  $\mathbf{u}_{n+1}$ :

$$\mathbf{M}(\mathbf{u}_{n+1} - \hat{\mathbf{u}}_{n+1}) + (\Delta t)^2 \beta [\mathbf{R}(\mathbf{u}_{n+1}) - \mathbf{P}] + \frac{(\Delta t)^2}{2} \mathbf{G}_c(\mathbf{u}_{n+1}) = \mathbf{0}. \quad (9.99)$$

The solution of this problem has to be computed by taking into account the contact constraints. Hence any of the methods discussed in the previous sections can be applied. Note that, once  $\mathbf{u}_{n+1}$  is known, we can compute the accelerations related to the internal forces from (9.95) and the acceleration due to contact from (9.98). Finally, the velocities follow from (9.96).

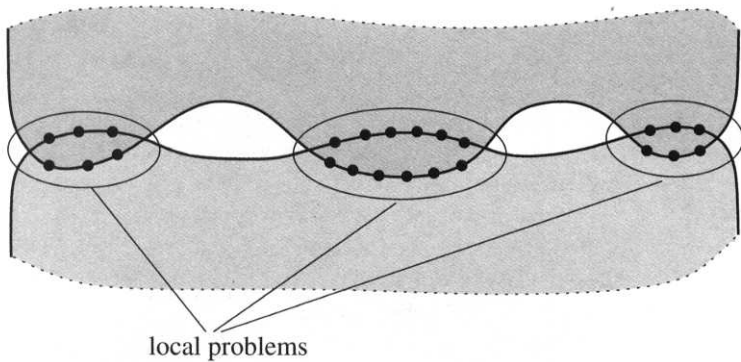
In Kane et al. (1999), the variational structure of equation (9.99) is investigated. By introducing the functional

$$\Pi(\mathbf{u}_{n+1}) = \frac{1}{(\Delta t)^2} (\mathbf{u}_{n+1} - \hat{\mathbf{u}}_{n+1})^T \mathbf{M}(\mathbf{u}_{n+1} - \hat{\mathbf{u}}_{n+1}) + 2\beta [W(\mathbf{u}_{n+1}) - \mathbf{u}_{n+1}^T \mathbf{P}] \quad (9.100)$$

with the strain energy  $W(\mathbf{u}_{n+1})$  (see (2.68)), one can write the nonlinear constraint minimization problem:

$$\min \Pi(\mathbf{u}_{n+1}) \quad \text{subject to} \quad g_{Ns} \geq 0. \quad (9.101)$$

It can be solved using a sequential quadratic programming technique, as described in Section 9.3.7.



**Fig. 9.15** Decoupling of minimization problems for lumped mass matrices

In the case of an explicit solution of the dynamic problem, the functional (9.100) simplifies to

$$\Pi(\mathbf{u}_{n+1}) = \frac{1}{(\Delta t)^2} (\mathbf{u}_{n+1} - \hat{\mathbf{u}}_{n+1})^T \mathbf{M} (\mathbf{u}_{n+1} - \hat{\mathbf{u}}_{n+1}), \quad (9.102)$$

which leads to a reduction of the complexity of the task (9.101) to the solution of a quadratic programming problem.

A considerable simplification of the constraint minimization problems occurs for diagonal mass matrices. In that case, the coupling of all degrees of freedom vanishes, which leads to a number of local optimization problems. These can be set up on the basis of the global/local search algorithms. As can be seen in Figure 9.15, one has to look for disjoint groups of intersecting segments which define the local problem. These problems only consist of a few unknowns, and can be solved separately with great efficiency.

## 9.4 GLOBAL ALGORITHMS FOR FRICTION

Many algorithms have been devised in the last twenty years for the solution of contact problems with friction. This is due to the practical importance of friction when engineering structures have to be designed. Engineering finite element models are becoming increasingly refined, which means that fewer simplifications in the mechanical modelling are made and thus inequality constraints are included. Hence, good algorithms for friction are today even more important.

One of the major difficulties in the numerical simulation of contact problems with friction lies in the fact that the constitutive equations for frictional contact are non-smooth and hence not differentiable at all points. Therefore, non-standard methods have to be applied for contact with friction. Basically, we can distinguish methods which rely on so-called *trial-and-error* methods, methods developed by engineers



which stem from algorithms used in the theory of plasticity and mathematical programming methods from optimization. We note that the last two approaches are based on mathematical principles.

In early treatments, often the *trial-and-error* methods were applied together with small deformation assumptions. Due to the fact that solutions of frictional contact can exhibit features such as non-uniqueness and even non-existence of the solution the *trial-and-error* methods are not reliable for a broad class of frictional problems. However successful applications can be found in the literature, e.g. see Fredriksson (1976) or Valliappan et al. (1984), who used an explicit technique for computation of the frictional forces in finite element applications.

The major break through in terms of convergence behaviour and reliability of the solution algorithms came with the application of the return mapping schemes – known from elasto-plasticity, see Simo and Taylor (1985), – to frictional problems. Its first application can be found in Wriggers (1987) or Giannokopoulos (1989) for geometrically linear problems. This approach provides the possibility to develop algorithmic tangent matrices which are needed to achieve quadratic convergence within NEWTON-type iterative schemes. Due to the non-associativity of the frictional slip these matrices are non-symmetrical. For the case of large deformations, associated formulations have been developed in Ju and Taylor (1988) for a regularized COULOMB friction law, and in Wriggers et al. (1990) for different frictional laws formulated in terms of non-associated plasticity. A three-dimensional formulation can be found in Laursen and Simo (1993b) who also developed an algorithmic symmetrization Laursen and Simo (1993a), see also Zavarise et al. (1995). However, other methods like yield-limited LAGRANGE multiplier formulations were employed also for frictional contact, see Jones and Papadopoulos (2000).

Mathematical programming methods were applied in Alart and Curnier (1991) within the augmented LAGRANGE multiplier technique; see also Section 5.3.8 for the theoretical background. The idea is here to construct a method which is still  $C^1$  differentiable to be able to use NEWTON's method. Related augmented LAGRANGE methods were also developed independently by De Saxe and Feng (1991), Simo and Laursen (1992) and Zavarise et al. (1995). Another closely related approach was introduced by Pang (1990), with the construction of NEWTON methods for so-called B-differentiable equation systems. In Christensen et al. (1998) it is shown that the frictional contact problem is B-differentiable, and hence the system of equations can be solved by an extended NEWTON method. An iterative algorithm based on a GAUSS–SEIDEL scheme can be found in Jourdan et al. (1998). Other possible approaches like interior point methods or barrier methods were investigated in Kloosterman et al. (2001) and Christensen et al. (1998).

Here we describe an algorithm which is implemented in many standard finite element codes for the solution of frictional contact problems. The formulation for frictional contact problems can be found in different sections. Section 3.2 contains the definition of the relative sliding and stick from the geometry point of view. Section 4.2 includes the constitutive equations for friction, and Section 5.3.2 summarizes different weak formulations with regard to solution methods. Discretizations for contact with

frictions are presented in Section 7.2 for the geometrical linear case, and in Sections 8.1 and 8.3.1 for finite deformations.

For a generalization of contact with friction, the following notation is introduced which is based on the matrices introduced in Section 9.3.1. In the case of friction, two different states have to be considered. In the stick case the tangential traction forces follow as reactions from the constraint equations of no tangential slip at the contact interface. Contrary, in slip mode the tangential frictional forces have to be obtained from the constitutive law for frictional slip. Therefore we have to distinguish two different states. Let us introduce two sets, one in which contact nodes  $s$  in sliding conditions are contained,  $\mathcal{F}_{sl}$ , and another one which holds the contact nodes  $s$  being in stick,  $\mathcal{F}_{st}$ . Note that we have the condition  $\mathcal{F}_{sl} \cap \mathcal{F}_{st} = \emptyset$ . Both sets together form the total set of all active nodes being in contact  $\mathcal{F}_{sl} \cup \mathcal{F}_{st} = \mathcal{J}_A$ .

With this notation we can introduce the weak form of equilibrium for all nodes  $s \in \mathcal{J}_A$  which are in frictional contact:

$$\mathbf{G}^f(\mathbf{u}) = \mathbf{G}(\mathbf{u}) + \epsilon \mathbf{C}^c(\mathbf{u})^T \mathbf{G}^c(\mathbf{u}) + \mathbf{t}_T^T(\mathbf{u}) \mathbf{C}_T(\mathbf{u}) = \mathbf{0}, \quad (9.103)$$

where  $\mathbf{t}_T(\mathbf{u})$  is the vector of friction forces which contains  $s \in \mathcal{J}_A$  elements. The  $s$ th component depends upon the stick or slip state at time  $t_{n+1}$  as follows:

$$\mathbf{t}_{Ts} = \begin{cases} \epsilon_T (\mathbf{g}_{Tn+1} - \mathbf{g}_{Tn}^{sl}) & \in \mathcal{F}_{st} \quad \text{for stick,} \\ \mu p_{Nn+1} \mathbf{n}_{Tn+1}^r & \in \mathcal{F}_{sl} \quad \text{for slip;} \end{cases} \quad (9.104)$$

see the algorithm in Box 13 in Section 9.5.2. The linearization of (9.103) at a given state  $\bar{\mathbf{u}}$  yields, according to (9.50), the tangent matrix

$$\mathbf{K}_T^f(\bar{\mathbf{u}}) = \left[ \mathbf{K}_T(\bar{\mathbf{u}}) + \mathbf{K}_T^{cPT}(\bar{\mathbf{u}}) + \epsilon \mathbf{C}^c(\bar{\mathbf{u}})^T \mathbf{C}^c(\bar{\mathbf{u}}) + \mathbf{F}(\bar{\mathbf{u}}) \mathbf{C}_T(\bar{\mathbf{u}}) \right], \quad (9.105)$$

in which  $\mathbf{K}_T^{cPT}$  contains the contributions stemming from the linearization of  $\mathbf{C}^c(\mathbf{u})$  and  $\mathbf{C}_T(\mathbf{u})$  in (9.103); for details of the derivation, see Section 5.5 and Chapter 8.

The vector  $\mathbf{F}(\mathbf{u})$  contains the linearizations of the tangential stresses in (9.104). For the stick case this leads, together with  $\mathbf{C}_T(\mathbf{u})$ , to a symmetric contribution to the tangent stiffness matrix, whereas for the slip state a nonsymmetric matrix is obtained, e.g. see Section 8.1. The latter situation requires a non-symmetric equation solver for frictional contact.

Now we can state an iterative algorithm which is based on the penalty method given in Section 9.3.3; see also Box 2. The algorithm is summarized in Box 11 for a load increment within the global solution. This algorithm can also be devised for the LAGRANGE multiplier method. In that case, the LAGRANGE multiplier method is often only formulated for the normal direction whereas in the tangential direction the treatment of the interface law as introduced in (9.104) is applied.

We note that the algorithm in Box 11 works quite efficiently, as has been shown in Laursen and Simo (1993b) or Agelet de Saracibar (1997). One problem in this approach is that the friction forces in stick depend upon the regularization parameter or penalty value  $\epsilon_T$ , see (9.104). The effect of too small a number for that parameter can be the determination of a stick instead of a sliding state. This behaviour can

Initialize algorithm: set  $\mathbf{u}_1 = \mathbf{0}$

LOOP over iterations :  $i = 1, \dots$ , convergence

Check for contact:  $g_{Nsi} \leq 0 \rightarrow$  active node, segment or element.

IF  $i = 1$ , set all active nodes to state stick  $\mathcal{F}_{st} = \mathcal{J}_A$ .

ELSE update frictional forces, using (9.104) or Box 13.

Solve:  $\mathbf{G}^f(\mathbf{u}_i) = \mathbf{0}$  using NEWTON'S method

Check for convergence:  $\|\mathbf{G}^f(\mathbf{u}_i)\| \leq TOL \Rightarrow$  STOP

END LOOP

Box 11. Frictional contact algorithm using the penalty method.

eventually lead to results which are physically questionable, as in problems where stick-slip effects occur in the contact interface. One way to solve this modelling problem is the use of a large number for the tangential stiffness (penalty parameter  $\epsilon_T$ ). However, often a large number cannot be chosen for  $\epsilon_T$ , since then the problem becomes ill-conditioned or the robustness of the iterative procedure in Box 11 is lost.

To overcome this problem an USZAWA-type algorithm can be applied in which a LAGRANGE multiplier in tangential direction is introduced, e.g. see Alart and Curnier (1991), Laursen and Simo (1993a), Zavarise et al. (1995) or Pietrzak and Curnier (1997). In this formulation the following change of the weak form in (9.103) is made:

$$\mathbf{G}^{fA}(\mathbf{u}) = \mathbf{G}(\mathbf{u}) + [\bar{\Lambda}_N + \epsilon \mathbf{C}^c(\mathbf{u})^T] \mathbf{G}^c(\mathbf{u}) + [\bar{\Lambda}_T + \mathbf{t}_T^T(\mathbf{u})] \mathbf{C}_T(\mathbf{u}) = \mathbf{0}. \quad (9.106)$$

Here the new variables  $\bar{\Lambda}_N$  and  $\bar{\Lambda}_T$  are the fixed LAGRANGE multipliers in the normal and tangential directions, respectively, as introduced in Sections 5.3.8 and 9.3.5. The global iteration procedure with an inner NEWTON and outer augmentation loop is described in Box 12. Due to the fact that in the inner NEWTON loop the update formulas for the frictional forces from Box 13 are used, the tangent matrix is non-symmetric, as in the penalty method given in Box 11.

The update of the normal LAGRANGE multipliers  $\bar{\Lambda}_N$  in Box 12 is performed as described in Section 9.3.5, see (9.79). Special care has to be taken when the tangential LAGRANGE multiplier update is performed, since the update has to fulfil the KARUSH-KUHN-TUCKER conditions (4.28), and with this the slip rule (4.25) in the case of COULOMB friction or (4.30) in more general cases, respectively. Using the update algorithm derived in Section 9.5.2, the LAGRANGE multipliers in the tangential direction are given with (9.113)<sub>1</sub> and (9.115)<sub>1</sub> by

$$\bar{\lambda}_{T_{n+1}}^{k+1} = \bar{\lambda}_{T_{n+1}}^k + \epsilon_T (\Delta \mathbf{g}_{T_{n+1}}^k - \lambda \mathbf{n}_{T_{n+1}}^{k+1}) \quad (9.107)$$

```

Initialize algorithm: set  $\mathbf{u}_1 = \mathbf{0}$ ,  $\bar{\mathbf{\Lambda}}_N^1 = \bar{\mathbf{\Lambda}}_T^1 = \mathbf{0}$ 

LOOP over augmentations :  $k = 1, \dots$ , convergence

    LOOP over iterations :  $i = 1, \dots$ , convergence

        Check for contact:  $g_{Nsi} \leq 0 \rightarrow$  active node, segment.
        IF  $i = 1$ , set all active nodes to state stick  $\mathcal{F}_{st} = \mathcal{J}_A$ .
        ELSE update frictional forces, using (9.104) or Box 13.
        Solve:  $\mathbf{G}^{fA}(\mathbf{u}_i, \bar{\mathbf{\Lambda}}_N^k, \bar{\mathbf{\Lambda}}_T^k) = \mathbf{0}$  using NEWTON'S method
        Convergence:  $\|\mathbf{G}^{fA}(\mathbf{u}_i, \bar{\mathbf{\Lambda}}_N^k, \bar{\mathbf{\Lambda}}_T^k)\| \leq TOL \Rightarrow$  STOP
    END LOOP

    Update LAGRANGE multipliers
     $\bar{\mathbf{\Lambda}}_N^{k+1} = \langle \bar{\mathbf{\Lambda}}_N^k + \epsilon_N \mathbf{G}^c(\mathbf{u}_i) \rangle$ 
     $\bar{\mathbf{\Lambda}}_T^{k+1} \Leftarrow \bar{\mathbf{\Lambda}}_T^k$ , see (9.107)

END LOOP

```

Box 12. Frictional contact algorithm using the USZAWA algorithm.

for the time step  $t_{n+1}$ . The parameter  $\lambda$  describes the slip increment and  $\mathbf{n}_{T_{n+1}}^{k+1}$  the slip direction.

A variant of this algorithm which was developed in Laursen and Simo (1993a) yields a symmetrical tangent matrix. It can be derived by using the LAGRANGE multipliers  $\bar{\lambda}_N^k$  in the update of the local friction forces in Box 13, instead of the actual normal force  $t_N^k = \langle \bar{\lambda}_N^k + \epsilon_N g_N^k \rangle$ . Since  $\lambda_N^k$  is constant during the iteration in the inner NEWTON loop, it does not contribute to the linearization. But it is this term, as shown for example in Section 8.1 (see (8.47)) which is responsible for the non-symmetric tangent. Hence the tangent matrix becomes symmetrical. This approach is numerically more efficient, though in some cases it is not as robust as the scheme described in Box 12.

## 9.5 LOCAL INTEGRATION OF CONSTITUTIVE EQUATIONS IN THE CONTACT AREA

In general, we have to distinguish three cases of constitutive equations in the contact interface. These are related to the normal and the tangential behaviour. For the normal contact a mere function evaluation – as for finite elasticity – can be used to obtain the contact pressure for a given approach. Even if the micromechanical derivation of

the contact compliance involves plastic deformations, the constitutive equations for the normal pressure require only a pure function evaluation in terms of the approach; see (4.10) or (4.12). This is theoretically not satisfactory, but up to now – due to the extremely complex behaviour in the contact interface – it is the only reasonable method for the macroscopic description of normal contact compliance.

In the following sections we discuss the integration of the local constitutive equations at the contact interface for adhesion and friction.

### 9.5.1 Evolution of adhesion

A different situation occurs when adhesion is present in the interface. In such cases, the constitutive behaviour is given by, for example, (4.63) and (4.66); see Section 4.4. Hence the evolution equation

$$\dot{\beta} = -\frac{1}{\eta} (w - C_N g_N^2 \beta) \quad (9.108)$$

has to be integrated in the case of  $w - C_N g_N^2 \beta < 0$ . A EULER backward algorithm leads, within the time step  $\Delta t_{n+1} = t_{n+1} - t_n$ , to

$$\beta_{n+1} = \beta_n - \frac{\Delta t_{n+1}}{\eta} (w - C_N g_{N\,n+1}^2 \beta_{n+1}). \quad (9.109)$$

This equation can be solved for the new adhesion intensity  $\beta_{n+1}$ ,

$$\beta_{n+1} = \left( 1 - \frac{\Delta t_{n+1}}{\eta} C_N g_{N\,n+1}^2 \right)^{-1} \left( \beta_n - \frac{\Delta t_{n+1}}{\eta} w \right). \quad (9.110)$$

To compute the normal contact stress  $t_N$  in the adhesive zone, the current value of the adhesion intensity has to be used in (4.63), which yields

$$p_N = C_N g_{N\,n+1} \left( 1 - \frac{\Delta t_{n+1}}{\eta} C_N g_{N\,n+1}^2 \right)^{-2} \left( \beta_n - \frac{\Delta t_{n+1}}{\eta} w \right)^2. \quad (9.111)$$

This represents a nonlinear function of the normal gap, and hence needs to be linearized.

### 9.5.2 Friction laws

The situation is again different for friction. Here one has to solve an evolution equation for the frictional slip (see Section 4.2.4), which is coupled with the slip condition  $f_s$ , and hence coupled with an inequality. This constitutive formulation needs special algorithms. In early finite element applications, explicit schemes were often applied which might not converge in some cases. Another method which is now standard for numerical simulations involving friction is related to the possibility to recast the frictional interface laws in terms of non-associated plasticity.

The algorithmic update of the tangential stress  $\mathbf{t}_{Tn+1}$  and dissipation  $\mathcal{D}_{n+1}^s$  is performed by the return algorithm based on an objective (backward EULER) integration of the evolution equation for the plastic slip, e.g. see Wriggers (1987), Ju and Taylor (1988), Giannakopoulos (1989), Wriggers et al. (1990). The results can be summarized as follows: integration of (3.22) gives the increment of the total slip within the time step  $\Delta t_{n+1}$ :

$$\Delta \mathbf{g}_{Tn+1} = (\bar{\xi}_{n+1}^\alpha - \bar{\xi}_n^\alpha) \bar{\mathbf{a}}_{\alpha n+1}. \quad (9.112)$$

The total slip  $\mathbf{g}_{Tn+1} = \mathbf{g}_{Tn} + \Delta \mathbf{g}_{Tn+1}$  has to be decomposed into an elastic and a plastic part, see (5.64). From this we can compute the elastic trial state from (4.22) and evaluate the slip criterion (4.25) at time  $t_{n+1}$ :

$$\begin{aligned} \mathbf{t}_{Tn+1}^{tr} &:= c_T (\mathbf{g}_{Tn+1} - \mathbf{g}_{Tn}^s) = \mathbf{t}_{Tn} + c_T \Delta \mathbf{g}_{Tn+1}, \\ f_{sn+1}^{tr} &:= \|\mathbf{t}_{Tn+1}^{tr}\| - \mu p_{Nn+1}. \end{aligned} \quad (9.113)$$

Here the vector  $\mathbf{t}_{Tn} = c_T (\mathbf{g}_{Tn} - \mathbf{g}_{Tn}^s)$  is the tangent traction vector of the last increment at  $t_n$ . A value of the slip criterion which fulfils ( $f_{sn+1}^{tr} \leq 0$ ) indicates stick. Hence no friction takes place, and we have to use the elastic relation (4.21). For  $f_{sn+1}^{tr} > 0$  sliding occurs in the tangential direction, and a return mapping of the trial tractions to the slip surface has to be performed. This return mapping procedure is derived from the time integration algorithm. In case the implicit EULER scheme is applied to approximate (4.27), we obtain

$$\begin{aligned} \mathbf{g}_{Tn+1}^s &= \mathbf{g}_{Tn}^s + \lambda \mathbf{n}_{Tn+1} \quad \text{with} \quad \mathbf{n}_{Tn+1} = \frac{\mathbf{t}_{Tn+1}^{tr}}{\|\mathbf{t}_{Tn+1}^{tr}\|}, \\ g_{vn+1} &= g_{vn} + \lambda. \end{aligned} \quad (9.114)$$

With the standard arguments regarding the projection schemes, e.g. see Simo and Taylor (1985) or Wriggers (1987), we obtain

$$\begin{aligned} \mathbf{t}_{Tn+1} &= \mathbf{t}_{Tn+1}^{tr} - \lambda c_T \mathbf{n}_{Tn+1}, \\ \mathbf{n}_{Tn+1} &= \mathbf{n}_{Tn+1}^{tr} \quad \text{with} \quad \mathbf{n}_{Tn+1}^{tr} = \frac{\mathbf{t}_{Tn+1}^{tr}}{\|\mathbf{t}_{Tn+1}^{tr}\|}. \end{aligned} \quad (9.115)$$

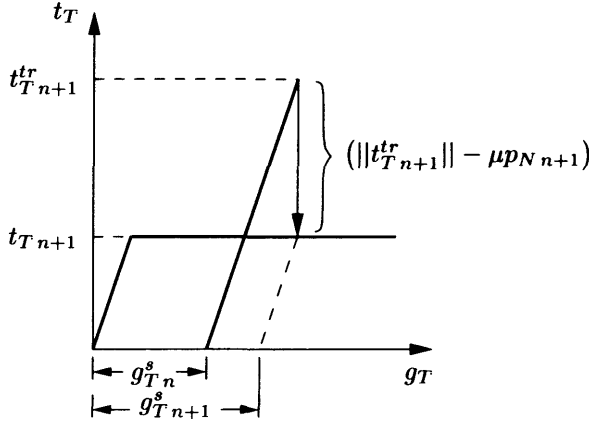
The multiplication of (9.115) by  $\mathbf{n}_{Tn+1}$  yields the condition from which  $\lambda$  can be computed:

$$\kappa(\lambda) = \|\mathbf{t}_{Tn+1}^{tr}\| - \hat{g}_s(p_{Nn+1}, \theta, g_{vn+1}) - c_T \lambda = 0, \quad (9.116)$$

where  $\hat{g}_s$  is in general a nonlinear function of  $\lambda$ . Thus we need an iterative scheme to solve  $\kappa(\lambda)$ . In such circumstances, NEWTON'S method is usually applied. Knowing  $\lambda$ , the stress update follows from (9.115) and the frictional slip from (9.114).

In the case of COULOMB'S model,  $\hat{g}_s = \mu p_{Nn+1}$  does not depend upon  $\lambda$ , and one can solve (9.116) directly for  $\lambda$ ,

$$\lambda = \frac{1}{c_T} (\|\mathbf{t}_{Tn+1}^{tr}\| - \mu p_{Nn+1}). \quad (9.117)$$



**Fig. 9.16** Projection of trial stress onto the slip surface.

The projection of the trial stresses  $\mathbf{t}_{T n+1}^{tr}$  which do not fulfil the slip condition (9.116) is graphically depicted in Figure 9.16 for the two-dimensional case. As can be seen, the quantity of  $\mathbf{t}_{T n+1}^{tr}$  which overshoots the tangent stress that is allowed according to COULOMB'S law,  $\mu p_{N n+1}$ , is used to correct the stresses leading to  $\mathbf{t}_{T n+1}$ . Furthermore, the increase of the slip (inelastic) part of the relative tangential motion  $\mathbf{g}_{T n+1}^s$  is shown graphically. The role of the stick (elastic) part is obvious from Figure 9.16.

Once  $\lambda$  is known, the frictional stresses follow from (9.115) and the total frictional slip from (9.114), which yields the explicit results for COULOMB'S model:

$$\mathbf{t}_{T n+1} = \mu p_{N n+1} \mathbf{n}_{T n+1}^{tr}, \quad (9.118)$$

$$\mathbf{g}_{T n+1}^s = \mathbf{g}_{T n}^s + \frac{1}{c_T} (\|\mathbf{t}_{T n+1}^{tr}\| - \mu p_{N n+1}) \mathbf{n}_{T n+1}^{tr}. \quad (9.119)$$

This update completes the local integration algorithm for the frictional interface law.

For COULOMB'S law an alternative update procedure which is based on geometric arguments was proposed in Section 8.2 for large sliding within large deformation processes. This so-called moving cone algorithm, which directly computes the total force at the contact interface, can be efficiently applied instead of the algorithm stated in Box 13. However, at present it is only advantageous to use the moving cone method for the classical law of COULOMB.

The dissipation due to the plastic slip follows from (4.23) as

$$\mathcal{D}_{n+1}^s = \mathbf{t}_{T n+1} \cdot \dot{\mathbf{g}}_{T n+1}^s,$$

where the relative tangential slip has to be approximated by the implicit EULER scheme

$$\dot{\mathbf{g}}_{T n+1}^s = \frac{1}{\Delta t_{n+1}} (\mathbf{g}_{T n+1}^s - \mathbf{g}_{T n}^s). \quad (9.120)$$

This approximation for the relative tangential slip velocity yields, with (9.114)<sub>1</sub>, in general

$$\mathcal{D}_{n+1}^s = \begin{cases} 0 & \text{for } f_{s n+1}^{trial} \leq 0, \\ \mathbf{t}_{T n+1} \cdot (\lambda / \Delta t_{n+1}) \mathbf{n}_{T n+1}^{trial} & \text{otherwise,} \end{cases} \quad (9.121)$$

and for COULOMB'S law with (9.119)<sub>2</sub>

$$\mathcal{D}_{n+1}^s = \begin{cases} 0 & \text{for } f_{s n+1}^{trial} \leq 0, \\ \mathbf{t}_{T n+1} \cdot (\mathbf{t}_{T n+1}^{trial} - \mathbf{t}_{T n+1}) / c_T \Delta t_{n+1} & \text{otherwise.} \end{cases} \quad (9.122)$$

The general algorithm which was derived above can be summarized as in Box 13. This algorithm has to be applied within the general solution algorithm in Box 11 to obtain the friction forces for the new time increment at time  $t_{n+1}$ . Furthermore, the tangent matrices associated with the implicit integration scheme have to be derived. For the node-to-surface contact elements, the associated tangents can be found in Sections 7.2, 8.1 and 8.3.1.

As an explicit example of how to treat the more general form (9.116), the constitutive equation for friction between soil and concrete is considered. The friction law is defined in (4.54) and (4.55). Again the return mapping algorithm is used to fulfil the inequality restriction for the slip surface  $f_s \leq 0$ , and an implicit Eulerian integration

Starting values:  $\mathbf{g}_{T n+1}$ ,  $\mathbf{g}_{T n}^s$ ,  $p_{N n+1}$

Compute trial state:  $\mathbf{t}_{T n+1}^{tr} = c_T (\mathbf{g}_{T n+1} - \mathbf{g}_{T n}^s)$

Check:  $f_{s n+1}^{tr} = \|\mathbf{t}_{T n+1}^{tr}\| - \mu p_{N n+1}$

If  $f_{s n+1}^{tr} \leq 0$  STOP

Else if  $f_{s n+1}^{tr} > 0$  perform return mapping

Solve for  $\lambda$ :  $\|\mathbf{t}_{T n+1}^{tr}\| - \hat{g}_s(p_{N n+1}, \theta, g_{v n+1}) - c_T \lambda = 0$

Update slip state:  $\mathbf{g}_{T n+1}^s = \mathbf{g}_{T n}^s + \lambda \mathbf{n}_{T n+1}$

$g_{v n+1} = g_{v n} + \lambda$

Compute new friction force:  $\mathbf{t}_{T n+1} = c_T (\mathbf{g}_{T n+1} - \mathbf{g}_{T n+1}^s)$

Box 13. Local integration of friction law.



for the time discretization. Time discretization of (4.54) and (4.55) yields

$$\begin{aligned}
 \mathbf{t}_{Tn+1} &= \epsilon_T p_{Nn+1}^{a_7} (\mathbf{g}_{Tn+1} - \mathbf{g}_{Tn+1}^s) \\
 \mathbf{g}_{Tn+1}^s &= \mathbf{g}_{Tn}^s + \gamma_{n+1} \frac{\mathbf{t}_{Tn+1}}{\|\mathbf{t}_{Tn+1}\|} \\
 w_{n+1} &= w_n + \gamma_{n+1} \\
 f_{sn+1} &= \|\mathbf{t}_{Tn+1}\| - p_{Nn+1}^{a_6} \mu(w_{n+1}),
 \end{aligned} \tag{9.123}$$

with  $\gamma_{n+1} = \lambda_{n+1} \Delta t$ .

For calculation of the tangential stresses  $\mathbf{t}_{Tn+1}$ , first the elastic *trial*-stresses are calculated, assuming stick,

$$\mathbf{t}_{Tn+1}^{tr} = \epsilon_T p_{Nn+1}^{a_7} (\mathbf{g}_{Tn+1} - \mathbf{g}_{Tn}^p) = \mathbf{t}_{Tn} + \epsilon_T p_{Nn+1}^{a_7} (\mathbf{g}_{Tn+1} - \mathbf{g}_{Tn}). \tag{9.124}$$

Inserting this result into the slip function (4.55) yields

$$f_s^{tr} = \|\mathbf{t}_{Tn+1}^{tr}\| - p_{Nn+1}^{a_6} \mu(w_n). \tag{9.125}$$

For  $f_s^{tr} < 0$  no sliding occurs, hence the assumed stick condition is valid and no further computations are needed. For sliding ( $f_s^{tr} \geq 0$ ) a projection on the slip surface  $f_s = 0$  has to be performed. Evaluation of (9.123) and (9.124) using  $\frac{\mathbf{t}_T}{\|\mathbf{t}_T\|} = \frac{\mathbf{t}_T^{tr}}{\|\mathbf{t}_T^{tr}\|}$  leads to

$$\begin{aligned}
 \mathbf{t}_{Tn+1} &= \mathbf{t}_{Tn+1}^{tr} \left( 1 - \epsilon_T \frac{p_{Nn+1}^{a_7}}{\|\mathbf{t}_{Tn+1}^{tr}\|} (w_{n+1} - w_n) \right) \\
 f_{sn+1} &= \|\mathbf{t}_{Tn+1}^{tr}\| - \epsilon_T p_{Nn+1}^{a_7} (w_{n+1} - w_n) - p_{Nn+1}^{a_6} \mu(w_{n+1}) = 0,
 \end{aligned} \tag{9.126}$$

with

$$\begin{aligned}
 \mu(w) &= \frac{a_1 w}{1 + a_2 w + a_3 w^2} + a_4 \arctan(a_5 w), \\
 \mathbf{t}_{Tn+1}^{tr} &= \mathbf{t}_{Tn} + \epsilon_T p_{Nn+1}^{a_7} (\mathbf{g}_{Tn+1} - \mathbf{g}_{Tn}).
 \end{aligned} \tag{9.127}$$

$w_{n+1}$  can now be calculated iteratively using (9.126<sub>2</sub>). Inserting  $w_{n+1}$  into (9.126<sub>1</sub>) leads to the tangential stresses  $\mathbf{t}_{Tn+1}$  at the current time  $t_{n+1}$ .

# 10

---

## *Thermo-mechanical Contact*

In the case of thermo-mechanical contact problems two fields – deformation and temperature – interact, and thus have to be considered within the formulation. In the general setting these fields are coupled, since the constitutive parameters depend upon the temperature. Furthermore, the evolution of the thermal field is related to the deformation, and heat can be generated by dissipative mechanisms like plastic deformations or frictional forces. Applications of thermo-mechanical coupled processes include the cooling of electronic devices, shrink fitting problems, screw connections under temperature loading, and frictional heating in rubber or hot metal forming problems. The technical importance of these models has recently raised interest in these phenomena, thus many contributions can be found in the literature.

Let us note that the movement of skis or skates on snow or ice is also related to thermo-mechanical contact. There a thin water film occurs during motion, and as a lubricant reduces the frictional coefficient considerably between skates and ice. This effect stems from frictional heating due to dissipation. Since this mechanism no longer works at a temperature range below minus 35° Celsius, one then observes a high coefficient of friction, like sand, and skis no longer glide with the usual ease.

In this chapter only research which is directly related to the numerical treatment of contact problems within the finite element method is discussed. In cases where thermo-mechanical contact has to be considered, a “high contact precision” formulation must be applied to account correctly for the pressure dependency of the heat conduction in the contact area. This is due to the fact that the heat conduction depends upon the approach of the two rough surfaces being in contact, see Section 3.3. In this context, models have been discussed for the constitutive behaviour in the normal direction on the basis of statistical methods, e.g. see Cooper et al. (1969) or Song

and Yovanovich (1987). A finite element treatment for thermo-mechanical contact problems can be found in Zavarise (1991), Zavarise et al. (1992b) or Wriggers and Zavarise (1993b), and in combination with frictional heating in Wriggers and Miehe (1992) or , Zavarise et al. (1995). Also, other contact phenomena like wear, e.g. see Johansson and Klarbring (1993), need special constitutive laws which have to be developed in the interface.

From the algorithmic side, a global iterative procedure is used for a stationary thermo-mechanical process. These so-called monolithic schemes can also be applied in transient processes. Here, however, staggered schemes, which treat the deformation and temperature fields separately, can also be computationally more advantageous, see Wriggers and Miehe (1992) or Agelet de Saracibar (1998) for thermo-mechanical contact, and Simo and Miehe (1992) for thermo-mechanical problems without contact.

## 10.1 EQUATIONS FOR THE CONTINUUM

For the description of the continuum we need kinematical relations, equilibrium and constitutive equations. All these are well known, hence we only state the equations which are needed to analyse bodies which are subjected to thermo-elastic deformations.

### 10.1.1 Kinematical relations, multiplicative split

Let  $\mathcal{B} \subset \mathbb{R}^3$  be the reference configuration of the body of interest. Denote by  $\varphi$  the deformation map, and let  $\mathbf{F} = \text{Grad } \varphi$  be the deformation gradient. For the thermo-mechanical problem a multiplicative split of the deformation gradient (2.6) is made,

$$\mathbf{F} = \mathbf{F}_e \mathbf{F}_\theta, \quad (10.1)$$

where the indices  $e$  and  $\theta$  indicate the elastic and the purely thermal part of the local deformation, respectively. This general split can be specified more by the introduction of a split of the deformation gradient  $\mathbf{F}$  in a volume preserving  $\widehat{\mathbf{F}}$  and a volumetric part  $J^{\frac{1}{3}} \mathbf{1}$ , which yields

$$\mathbf{F} = J^{\frac{1}{3}} \widehat{\mathbf{F}}; \quad \text{with} \quad \det \widehat{\mathbf{F}} = 1. \quad (10.2)$$

With the classical assumptions of pure volumetric thermal deformations,  $\widehat{\mathbf{F}}_\theta = \mathbf{1}$ , the following multiplicative split is derived, which is basis for the subsequent derivations,

$$J = J_e J_\theta, \quad \widehat{\mathbf{F}} = \widehat{\mathbf{F}}_e, \quad (10.3)$$

where

$$\mathbf{F}_e = J_e^{\frac{1}{3}} \widehat{\mathbf{F}}_e, \quad \mathbf{F}_\theta = J_\theta^{\frac{1}{3}} \mathbf{1}. \quad (10.4)$$

The right CAUCHY–GREEN tensor associated with thermoelastic deformations is then defined by

$$\mathbf{C} = \mathbf{F}^T \mathbf{F} = (J_e J_\theta)^{\frac{2}{3}} \widehat{\mathbf{F}}^T \widehat{\mathbf{F}}. \quad (10.5)$$

Alternatively, the left CAUCHY–GREEN tensor in the current configuration is given by

$$\mathbf{b} = \mathbf{F} \mathbf{F}^T = (J_e J_\theta)^{\frac{2}{3}} \hat{\mathbf{F}} \hat{\mathbf{F}}^T. \quad (10.6)$$

### 10.1.2 Thermoelastic constitutive law

Constitutive equations for both bodies which come into contact have to be defined. These describe the material behaviour for the mechanical response and the thermal conductivity. The constitutive equations can be different for each of the contacting bodies. In this section only hyperelastic response functions are considered. Since they are assumed to have the same structure for each of the materials, one can drop the superscript  $\alpha$  which is a pointer to body  $\mathcal{B}^\alpha$  to simplify notation. More advanced material descriptions which include finite strain plasticity or visco-elasticity can be found in Simo and Miehe (1992) or Reese (2001).

As a model equation for nonlinear constitutive behaviour isotropic finite elasticity is used. It leads to a nonlinear response function for the stresses, and can be derived from a strain energy function  $W(\mathbf{b})$ , see Section 2.4.1. The simplest example for hyperelasticity is the compressible Neo-Hookian model which can be applied for rubber materials undergoing moderately large strains, for instance. A possible extension of the associated free HELMHOLTZ energy for the thermo-elastic case yields (see Miehe (1988) or Wriggers et al. (1992))

$$W(J_e, \hat{\mathbf{b}}_e, \theta) = \frac{1}{2} \mu (I_{\hat{\mathbf{b}}_e} - 3) + \frac{1}{2} K (\ln J_e)^2 + T(\theta). \quad (10.7)$$

Based on equation (2.75) from (10.7) one obtains a constitutive equation for the KIRCHHOFF stresses  $\boldsymbol{\tau}$  and the thermo-elastic entropy  $\eta_e$ ,

$$\boldsymbol{\tau} = 2\mathbf{b}_e \frac{\partial W}{\partial \mathbf{b}_e} = K \ln J_e \mathbf{1} + \mu \operatorname{dev} \hat{\mathbf{b}}_e \quad \text{and} \quad \eta_e = -\frac{\partial W}{\partial \theta}. \quad (10.8)$$

$K$  is the bulk modulus and  $\mu$  the shear modulus.

Following Lu and Pister (1975) and Miehe (1988), the volumetric thermal contribution to the total deformation can be written as

$$J_\theta = e^{3\alpha_t(\theta - \theta_0)}. \quad (10.9)$$

Here  $\alpha_t$  stands for the linear thermal expansion coefficient, and  $\theta - \theta_0$  is the increase of the absolute temperature from a given reference temperature  $\theta_0$ . Using this constitutive equation, the elastic volume change can be computed via equation (10.3)<sub>1</sub>:

$$J_e = J e^{-3\alpha_t(\theta - \theta_0)}. \quad (10.10)$$

KIRCHHOFF stresses in (10.8) can be related to second PIOLA–KIRCHHOFF stresses via  $\mathbf{S} = \mathbf{F}^{-1} \boldsymbol{\tau} \mathbf{F}^{-T}$ . Furthermore, the connection to the first PIOLA–KIRCHHOFF stress which enters the local equilibrium equation, when formulated with respect to the reference configuration, can be made via  $\mathbf{P} = \mathbf{F} \mathbf{S}$ .

To describe the heat conduction within the solid body, the classical law of FOURIER is assumed for isotropy. It relates the KIRCHHOFF heat flux vector  $\mathbf{q}$ , defined in the current configuration, to the temperature gradient as follows:

$$\mathbf{q} = -k \text{grad } \theta, \quad (10.11)$$

where  $k > 0$  is the heat conduction coefficient. With respect to the undeformed configuration, we obtain  $\mathbf{Q} = -k \text{Grad } \theta$ .

## 10.2 CONSTITUTIVE EQUATIONS FOR THERMO-MECHANICAL CONTACT

For mechanical contact all relations which were developed in Chapter 3 for the contact geometry can be used as well as the contact weak formulation and discretization in Sections 5.3 and 7.

The temperatures of the two surfaces  $\Gamma^\alpha$  are assumed to be given as functions of placement and time in the spatial configuration:  $\theta^\alpha = \theta^\alpha(\mathbf{x}^\alpha, t)$ . In the contact interface one can then use the geometrical relation (3.6) which defines the projection of a point on surface  $\Gamma^2$  onto  $\Gamma^1$  to find the location  $\bar{\mathbf{x}}^1$ , at which  $\bar{\theta}^1$  is given.

The heat flux in the contact zone also needs a constitutive equation for its determination. In cases where thermo-mechanical contact has to be considered, a "high contact precision" formulation must be applied to correctly account for the pressure dependency of the heat conduction in the contact area. We assume the following structure for the constitutive equation for the heat flux, see Figure 10.1:

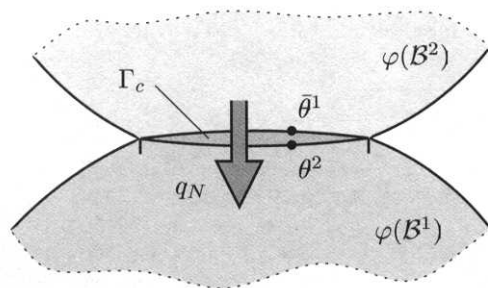
$$q_N = \hat{q}_N(\theta^1, \theta^2, p_N) = \hat{h}(\theta^1, \theta^2, p_N)(\theta^2 - \bar{\theta}^1), \quad (10.12)$$

where  $\theta^\alpha$  are the temperatures of both contact surfaces.  $\bar{\theta}^1$  is defined by the closest point projection used for the static normal contact; see above and equation (3.2).

The heat transfer coefficient  $\hat{h}(\theta^1, \theta^2, p_N)$  depends upon the surface temperatures and the contact pressure. The latter is needed for stability reasons, e.g. see Section 1.3. Thus we have to find a pressure-dependent model which describes the heat flux in the contact interface. From a micromechanical viewpoint, the heat conduction depends upon the approach of the two rough surfaces in contact. In this context, models have been discussed for the constitutive behaviour in the normal direction on the basis of statistical methods, e.g. see Cooper et al. (1969) or Song and Yovanovich (1987).

Due to the technical impossibility of obtaining perfectly plane surfaces, the real contact area is always limited and corresponds to a series of spots (see Figure 10.3). Determination of the true contact area is fundamental for the modelling of mechanical and thermal phenomena. In detail, it is necessary to determine the number of spots in contact, their distribution and their medium size. All these parameters depend upon the applied apparent contact pressure.

The problem can be represented using correlation formulae generated by the fitting of experimental data, or using a theoretical approach. Here a microscopic plastic



**Fig. 10.1** Heat flux in the contact interface.

model by Song and Yovanovich (1987) has been adapted for the development of the macroscopic thermal contact law. Different techniques for computation of the thermal contact resistance, taking into account the dependence on various parameters have been proposed. The resistance is mainly due to the low percentage of surface area which is really in contact. The presence of a reduced set of spots surrounded by micro-cavities characterizes the contact zone, hence heat exchange is possible by heat conduction through the spots, heat conduction through the gas contained in the cavities and radiation between micro-cavity surfaces.

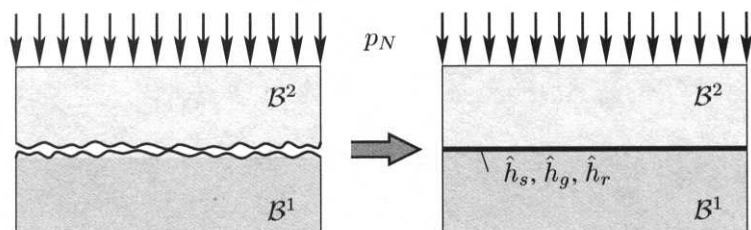
All these results yield a homogenized constitutive relation for the thermo-mechanical behaviour in the contact interface, see Figure 10.2. The assumption that the mentioned mechanisms act in parallel is well accepted, which leads to the following relationship for the thermal contact resistance  $R$  of a representative area element  $A_a$ :

$$\frac{1}{R} = \frac{1}{R_{spots}} + \frac{1}{R_{gas}} + \frac{1}{R_{radiation}} \quad (10.13)$$

or in terms of the conductivity  $\hat{h}$ ,

$$\hat{h} = \hat{h}_s + \hat{h}_g + \hat{h}_r. \quad (10.14)$$

The constitutive relations for heat conductance all depend upon variables which change during the analysis. A nonlinear dependence with regard to the pressure



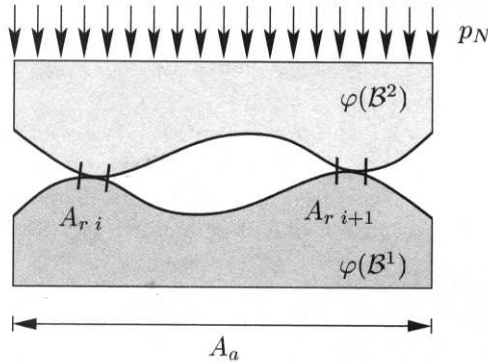
**Fig. 10.2** Homogenization of the constitutive data.

$p_N$  and the surface temperatures  $\theta^\alpha$  has to be considered. Thus, in general, we can write for the contact conductivity

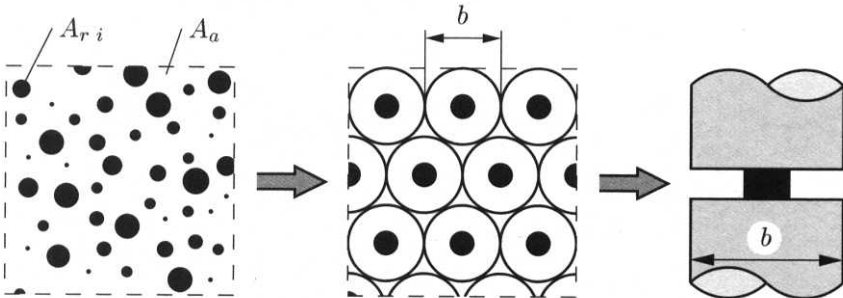
$$\hat{h} = h_c(p_N, \theta^1, \theta^2). \quad (10.15)$$

### 10.2.1 Heat conductance through spots

Considering the contribution of  $i$  spots with area  $A_{si}$  (see Figure 10.4 (a)), the thermal resistance of a representative area  $A_a$  is first analysed (see Figure 10.4 (b)); the knowledge of the number of spots then permits us to compose these resistances in parallel. Around each spot a heat flux tube having a narrowing in the contact zone is considered (see Fig 10.4 (b)). The solution of the heat conduction equation involves some difficulties, thus auxiliary hypotheses are used, and the solution is determined using series expansion. Various shapes of narrowing have been studied and comparisons with experimental tests Shai and Santo (1982) show the best correspondence for the flat disk narrowing model, see Figure 10.4 (b). This model is now well



**Fig. 10.3** Conductance through spots in the contact interface.



**Fig. 10.4** (a) Statistical model, (b) Spot resistance model.

accepted. By combining a mechanical contact model, like (4.10), which gives the representative inner and outer radii of the spots, with the thermal resistance model, a complicated relationship has been obtained in Cooper et al. (1969) by averaging over the area  $A_a$ . Since the relationship involves integrals of exponential functions, no closed form solution is available. This leads to the development of a simpler model, see Yovanovich (1981).

The additional effect of taking into account the hardness variation with the mean planes approach has been suggested in Song and Yovanovich (1987), which yields

$$\hat{h}_s = \frac{1.25k^*\bar{m}}{\sigma} \left[ \frac{p_N}{c_1} \left( 1.6177 \frac{10^6 \sigma}{\bar{m}} \right)^{-c_2} \right]^{\frac{0.95}{1+0.0711c_2}}, \quad (10.16)$$

where  $c_1$  and  $c_2$  are experimental constants governing the hardness variation, and  $k^* = 2k_1k_2/(k_1 + k_2)$  is the harmonic mean thermal conductivity, which depends upon the conductivities  $k_1$  and  $k_2$  of the materials in contact. Two geometric constants are involved;  $\bar{m}$ , the mean absolute asperity slope, and  $\sigma$  the *rms* surface roughness. These constants have to be extrapolated from experimental data. Finally,  $p_N$  is the apparent mechanical contact pressure. We note that the interface laws thus derived are obtained from a curve fitting of a sophisticated theoretical model. Hence, it is not just a simple curve fitting of some particular experimental results.

A simplified model for a pressure-dependent heat conduction through the spots in the contact interface can be given in terms of the VICKERS hardness  $H_v$ , the thermal resistance coefficient  $h_{s0}$  and an exponent  $\omega$ :

$$\hat{h}_s(p_N) = h_{s0} \left( \frac{p_N}{H_v} \right)^\omega. \quad (10.17)$$

This relation only depends upon three parameters which follow from experiments.

### 10.2.2 Heat conductance through gas

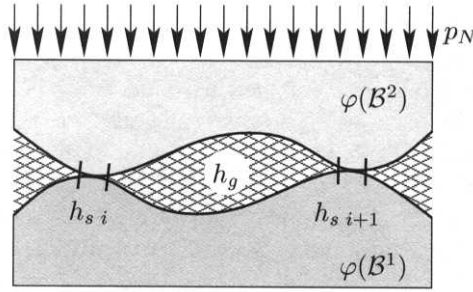
As far as the contribution of the gas, contained in the micro-cavities, to the heat conductance is concerned, interesting reviews of the proposed models are available in Madhusudana and Fletcher (1981). Due to the reduced height of the cavities, convective movements are strongly limited, and thermal exchange takes place mainly by conduction, like in solid materials. Hence thermal conductance can be calculated as

$$\hat{h}_g = \frac{k_g}{d + g_1 + g_2}, \quad (10.18)$$

where  $k_g$  is the gas conductivity, and  $d$  is the effective height of the cavity which corresponds to the mean plane distance. Terms  $g_1$  and  $g_2$  are introduced to reproduce the nonlinear surface effect and the temperature surface discontinuity. These additional distances can be obtained by employing the kinetic theory of gases, e.g. see Kennard (1938). We obtain

$$g_1 + g_2 = G_{pc} \frac{T_g}{p_g}. \quad (10.19)$$





**Fig. 10.5** Conductance through gas in the contact interface.

Here  $G_{pc}$  is a constant collecting various fixed parameters describing the physical properties of the gas,  $T_g$  is the current gas temperature and  $p_g$  denotes the current gas pressure. The explicit expression for the constant  $G_{pc}$  is given below:

$$g_1 + g_2 = \left( \frac{2 - \alpha_1}{\alpha_1} + \frac{2 - \alpha_2}{\alpha_2} \right) \frac{2\gamma}{\gamma + 1} \frac{1}{Pr} \Lambda_0 \frac{p_{g0}}{T_0}, \quad (10.20)$$

where  $\alpha_1$  and  $\alpha_2$  are experimental coefficients related to the surfaces  $\Gamma^1$  and  $\Gamma^2$ , respectively.  $\gamma = \frac{c_p}{c_v}$  indicates the specific heat ratio,  $Pr$  is the PRANDTL number, and  $\Lambda_0$  is the mean free path calculated at reference temperature  $T_0$  and reference gas pressure  $p_{g0}$ .

The basic relation (10.19) is adopted by a great number of authors, who propose different techniques to calculate the effective height  $d$  of the cavity. However, only a few models account for the reduction of cavity height under increasing contact pressure which is included in (10.18). According to the mechanical model previously considered, see Cooper et al. (1969), the dependence of the effective height from the mechanical pressure is given in Yovanovich (1981), and yields

$$d = 1.363\sigma \left[ -\ln \left( 5.589 \frac{p_N}{H_e} \right) \right]^{0.5}, \quad (10.21)$$

where  $H_e$  is the micro hardness distribution. Again, the relationship (10.21) is a high precision curve fitting of the original theoretical equation. Thus (10.19) and (10.21) complete the derivation of a thermal conductance for the gas, and can be inserted into (10.18).

Within this approach the mechanical stiffness and thermal contact conductivities have been calculated based on the assumption that the contacting asperities undergo plastic deformations. Relations that consider pure elastic deformations of the contacting asperities have also been formulated, e.g. see Greenwood and Williamson (1966), and a wide class of problems characterized by smooth surfaces can be studied within that hypothesis.

### 10.2.3 Heat conductance by radiation

Another way to transmit heat from one surface to the other is by electromagnetic radiation. This works in a vacuum as well as in a medium. However, when this medium is a liquid the amount of heat transferred by radiation is often negligible. When the medium is a gas then radiation can be of importance. Hence this effect also has to be considered within the contact interface when a general relation for heat conduction should be established. As shown in Figure 10.6, both surfaces generate radiation waves. The basic relation for the heat flux due to radiation is given in terms of the STEFAN-BOLTZMANN law

$$q_r = F_{12} \sigma_n [(\theta^2)^4 - (\theta^1)^4], \quad (10.22)$$

where  $F_{12}$  is the mutual radiation factor of the surfaces  $\Gamma^1$  and  $\Gamma^2$  on both sides of the contact interface,  $\sigma_n$  is the STEFAN-BOLTZMANN constant and  $\theta^\alpha$  represent the temperatures of the two surfaces, e.g. see Boley and Weiner (1997). This leads, with the form of the heat conductance equation for radiation (10.12)

$$q_r = h_r(\theta^1, \theta^2)(\theta^2 - \bar{\theta}^1),$$

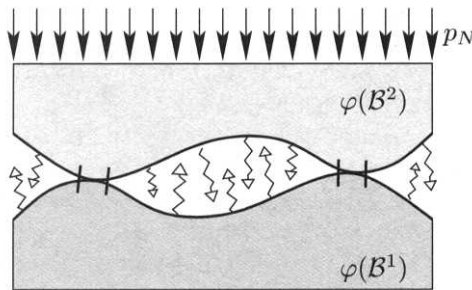
to the heat conductance coefficient for radiation

$$h_r = F_{12} \sigma_n [(\theta^2)^2 + (\bar{\theta}^1)^2](\theta^2 + \bar{\theta}^1), \quad (10.23)$$

which depends in a nonlinear way upon the surface temperatures.

Usually, one has to compute view factors according to the geometry of the surface. Here the assumption has been made that the surfaces are flat and parallel, which is sufficient when the aspect ratio of the asperities is not too high, or when this relation is used as a regularization of the thermal contact conductivity.

Often the radiation effects between the surfaces of the micro-cavities can be neglected, because the small difference of temperature greatly reduces this effect. However the radiation effect might be important to regularize the jump in the thermal resistance between status gap open and status gap closed, e.g. see Wriggers and Zavarise (1993b). Disregarding the radiation effect can lead to an unstable algorithm when the contact pressure is very small.



**Fig. 10.6** Conductance through gas in the contact interface.

### 10.3 INITIAL VALUE PROBLEM FOR THERMO-MECHANICAL CONTACT

In the previous sections we have discussed contact geometry and constitutive equations associated with the contact interface. Let us now formulate the initial-boundary-value problem for nonlinear thermoelasticity combined with thermo-mechanical frictional contact for two bodies. For a more simple presentation, all relations are formulated with respect to the initial configuration of the contacting bodies. Furthermore, to obtain a compact structure we introduce locally at the material point  $\mathbf{X}^\gamma$  of body  $B^\gamma$  a vector of primary variables at time  $t$ ,

$$\mathcal{Z}^\gamma(\mathbf{X}^\gamma, t) = \left\{ \begin{array}{l} \varphi^\gamma(\mathbf{X}^\gamma, t) \\ \mathbf{V}^\gamma(\mathbf{X}^\gamma, t) \\ \theta^\gamma(\mathbf{X}^\gamma, t) \end{array} \right\}, \quad (10.24)$$

which contains the configuration  $\varphi^\gamma$ , the material velocity  $\mathbf{V}^\gamma$  and the material temperature  $\theta^\gamma$ . The superscript  $\gamma = 1, 2$  characterizes the body  $B^1$  or  $B^2$ .

Locally at points  $\mathbf{X}^\gamma \in \Gamma_c^\gamma$  of the contact surfaces a vector of history variables

$$\mathcal{H}(\mathbf{X}^\gamma, t) = \{ \mathbf{g}_T^s(\mathbf{X}^\gamma, t) \} \quad (10.25)$$

is defined, which contains the plastic tangential slip  $\mathbf{g}_T^s$ . Other internal variables in the domains  $B^\gamma$  do not appear, since only thermoelastic constitutive response is considered.

The thermo-mechanical initial-boundary-value problem is governed by the local field equations in the domains  $B^\gamma$ :

$$\begin{aligned} \frac{\partial}{\partial t} \varphi^\gamma &= \mathbf{v}^\gamma + \mathbf{0} \\ \frac{\partial}{\partial t} \mathbf{V}^\gamma &= \frac{1}{\rho_0^\gamma} \text{DIV } \mathbf{P}^\gamma + \frac{1}{\rho_0^\gamma} \bar{\mathbf{b}} + \mathbf{0} \\ \frac{\partial}{\partial t} \theta^\gamma &= 0 - \frac{1}{c^\gamma} \text{DIV } \mathbf{Q}^\gamma + \frac{1}{c^\gamma} S^\gamma, \end{aligned} \quad (10.26)$$

which represent the definition of the material velocity, see (2.28), the balance of linear momentum, see (2.49), and the balance of internal energy in form of the temperature evolution equation following from (2.53).  $\mathbf{P}^\gamma$  is the first PIOLA-KIRCHHOFF stress tensor,  $\mathbf{Q}^\gamma$  denotes the material heat flux vector and  $S^\gamma$  is a heat source which describes in the framework of thermoelasticity the GOUGH-JOULE coupling effect. The field equations (10.26) form a coupled first order evolution system for the primary variable vector  $\mathcal{Z}^\gamma$  introduced above. Hence one can write in short notation

$$\frac{\partial}{\partial t} \mathcal{Z}^\gamma = \hat{\mathbf{A}}^\gamma(\mathcal{Z}^\gamma), \quad (10.27)$$

where the nonlinear evolution operator  $\hat{\mathbf{A}}$  represents the right-hand side of (10.26). In this equation one has to introduce the thermoelastic constitutive equations in the

domains  $\mathcal{B}^\gamma$ , which describe the dependency of stress, heat flux and the heat source on  $\mathcal{Z}^\gamma$ :

$$\begin{aligned}\mathbf{P}^\gamma &= \hat{\mathbf{P}}^\gamma(\mathcal{Z}^\gamma), \\ \mathbf{Q}^\gamma &= \hat{\mathbf{Q}}^\gamma(\mathcal{Z}^\gamma), \\ \mathcal{S}^\gamma &= \hat{\mathcal{S}}^\gamma(\mathcal{Z}^\gamma).\end{aligned}\quad (10.28)$$

Note that usually, instead of stating a constitutive equation for the first PIOLA-KIRCHHOFF stress  $\mathbf{P}^\gamma$ , one introduces a constitutive equation for the second PIOLA-KIRCHHOFF stress  $\mathbf{S}^\gamma$ , which is related to  $\mathbf{P}^\gamma$  via  $\mathbf{P}^\gamma = \mathbf{F}^\gamma \mathbf{S}^\gamma$ .

Next initial conditions for the primary and history variables are given for a time interval  $[t_n, t]$ :

$$\mathcal{Z}^\gamma(\mathbf{X}^\gamma, t_n) = \mathcal{Z}_n^\gamma(\mathbf{X}^\gamma) \quad \text{and} \quad \mathcal{H}^\gamma(\mathbf{X}^\gamma, t_n) = \mathcal{H}_n^\gamma(\mathbf{X}^\gamma). \quad (10.29)$$

Boundary conditions for the deformation are given on  $\Gamma_\varphi^\gamma \subset \partial\mathcal{B}^\gamma$  and for the temperature on  $\Gamma_\Theta^\gamma \subset \partial\mathcal{B}^\gamma$ :

$$\begin{aligned}\varphi^\gamma(\mathbf{X}^\gamma, t) &= \bar{\varphi}^\gamma(\mathbf{X}^\gamma, t) \quad \text{on } \Gamma_\varphi^\gamma, \\ \theta^\gamma(\mathbf{X}^\gamma, t) &= \bar{\theta}^\gamma(\mathbf{X}^\gamma, t) \quad \text{on } \Gamma_\Theta^\gamma.\end{aligned}\quad (10.30)$$

For the traction vector, boundary conditions are prescribed on  $\Gamma_t^\gamma \subset \partial\mathcal{B}^\gamma$ , and for the heat flux vector on  $\Gamma_q^\gamma \subset \partial\mathcal{B}^\gamma$ :

$$\begin{aligned}\mathbf{t}^\gamma(\mathbf{X}^\gamma, t) &= \bar{\mathbf{t}}^\gamma(\mathbf{X}^\gamma, t) \quad \text{on } \Gamma_t^\gamma, \\ Q^\gamma(\mathbf{X}^\gamma, \theta_t^\gamma, t) &= \bar{Q}^\gamma(\mathbf{X}^\gamma, \theta_t^\gamma, t) \quad \text{on } \Gamma_q^\gamma.\end{aligned}\quad (10.31)$$

The initial-boundary-value problem is completed by the thermo-mechanical constitutive equations on the current slave contact surface  $\varphi^1(\Gamma_c^1)$ :

$$\begin{aligned}p_N &= \hat{p}_N(\mathcal{Z}^{1,2}, \mathcal{H},) \\ \mathbf{t}_T &= \hat{\mathbf{t}}_T(\mathcal{Z}^{1,2}, \mathcal{H},) \\ q_N &= \hat{q}_N(\mathcal{Z}^{1,2}, \mathcal{H},) \\ \mathcal{D}^s &= \hat{\mathcal{D}}^s(\mathcal{Z}^{1,2}, \mathcal{H},),\end{aligned}\quad (10.32)$$

for the contact pressure  $p_N$ , the contact tangential stress  $\mathbf{t}_T$ , the contact heat flux  $q_N$  and the frictional dissipation  $\mathcal{D}^s$ .  $\mathcal{Z}^{1,2}$  is introduced here to point out that at the contact surface, fields of both bodies interact, and thus have an input to the associated constitutive relation. Furthermore, we have the evolution equation for the plastic slip, see (4.27), with the structure

$$\frac{\partial}{\partial t} \mathcal{H} = \lambda \hat{\mathcal{E}}(\mathcal{Z}^{1,2}, \mathcal{H}) \quad \text{with } \hat{\mathcal{E}} = \frac{\partial \hat{f}_s}{\partial \mathbf{t}_T}, \quad (10.33)$$

which is constrained by the KUHN-TUCKER loading-unloading conditions

$$\lambda \geq 0; \quad \hat{f}_s(\mathcal{Z}^{1,2}, \mathcal{H}) \leq 0; \quad \lambda \hat{f}_s(\mathcal{Z}^{1,2}, \mathcal{H}) = 0. \quad (10.34)$$

Finally, the contact conditions for the normal contact, see (4.2),

$$g_N \geq 0; \quad \hat{p}_N (\mathcal{Z}^{1,2}, \mathcal{H}) \leq 0; \quad g_N p_N (\mathcal{Z}^{1,2}, \mathcal{H}) = 0, \quad (10.35)$$

completes the formulation of a coupled thermo-mechanical initial-boundary-value problem.

For formulation of the weak form of a pure mechanical and thermo-mechanical contact problem, we have to discuss only the additional terms due to contact in detail. The equations describing the behaviour of the bodies coming into contact do not change. For completeness, the weak forms for the bodies coming into contact are stated next.

#### 10.4 WEAK FORMS IN THERMO-MECHANICAL ANALYSIS

In the case of thermo-mechanical contact problems, two fields – deformation and temperature – interact, and thus have to be considered within the formulation. In the general setting these fields are coupled because the constitutive parameters depend upon the temperature; the evolution of the thermal field is related to the deformation and heat can be generated by dissipative mechanisms like plastic deformations or frictional forces.

In the coupled thermo-mechanical analysis one has to extend the weak form (5.20) for mechanical contact to the coupled case. The weak forms are stated here for a known active set, so that instead of an inequality we can write an equality for the weak form. This yields, in analogy to (5.24), two weak forms in which mechanical and thermal variables are coupled:

$$\begin{aligned} G_M(\varphi^\gamma, \eta^\gamma, \theta^\gamma) &= \sum_{\gamma=1}^2 \left[ \int_{B^\gamma} \boldsymbol{\tau}^\gamma \cdot \nabla^S \boldsymbol{\eta}^\gamma dV - \int_{\Gamma_c^\gamma} \boldsymbol{\eta}^\gamma \cdot \hat{\mathbf{t}}^\gamma da \right] \\ &+ \int_{\varphi^1(\Gamma_c^1)} [\delta g_N p_N + \delta \mathbf{g}_T \cdot \mathbf{t}_T] da = 0. \quad (10.36) \end{aligned}$$

$\eta^\gamma$  denotes the mechanical test function defined on the current configuration, also known as the virtual displacement. The second term in (10.36) describes the contributions due to contact with the variations of the normal gap, see (3.29), and the variations of the relative tangential displacement in the contact interface, see (3.30). The first term in (10.36) contains the standard domain contributions with the KIRCHHOFF stress tensor  $\boldsymbol{\tau}^\gamma$ . It is integrated with respect to the initial configuration, whereas the integration of the contact terms has to be performed with respect to the current configuration.

The weak form of the thermal contribution takes the form

$$G_T(\theta^\gamma, \vartheta^\gamma, \varphi^\gamma) = - \int_{\varphi^1(\Gamma_c^1)} [(\vartheta^2 - \vartheta^1) q_N + \frac{1}{2} (\vartheta^1 + \vartheta^2) \mathcal{D}^s] da +$$

$$\sum_{\gamma=1}^2 \left[ \int_{\mathcal{B}^\gamma} \{ \nabla \vartheta^\gamma \cdot \mathbf{q}^\gamma - \vartheta^\gamma (S^\gamma - c^\gamma \dot{\theta}^\gamma) \} dV - \int_{\varphi^\gamma(\Gamma_q^\gamma)} \vartheta^\gamma \bar{q}^\gamma da \right] = 0. \quad (10.37)$$

Here  $\vartheta^\gamma$  denotes the thermal test function defined on the current configuration, also called the virtual temperature. In the integral concerned with the contact interface, one can see two contributions. The first is due to the heat conduction in the contact area. The second term stems from the frictional heating which is generated by the frictional dissipation, e.g. see (4.23). In a finite element computation, the dissipation can be computed together with the stress update of the tangential stresses, see Section 9.5.

## 10.5 ALGORITHMIC TREATMENT

The main solution strategies for coupled problems are monolithic schemes where equations (10.36) and (10.37) are solved together as a coupled system for the different variables (here deformation and temperature). In large three-dimensional problems this is related to the large number of variables, which is not very efficient. Thus, other strategies known as staggered schemes were developed in which the different variables are computed separately within a given time step (see Wood (1990) for an overview).

Global solution algorithms for coupled thermo-mechanical analysis have been formulated in Argyris and Doltsinis (1981), Miehe (1988), Doltsinis (1990), Simo and Miehe (1992), Agelet de Saracibar (1998) or Lewis and Schrefler (2000), among others. In what follows, the staggered scheme proposed in Simo and Miehe (1992) and for contact in Wriggers and Miehe (1994) is adopted. This scheme is based on an operator split of the global thermo-mechanical evolution operator discussed in Section 10.3. Such a strategy yields an algorithmic decoupling of the thermo-mechanical equations within a time step on the basis of subproblems which are in the frictionless case both symmetric.

The central idea is a split of the evolution operator  $\hat{\mathbf{A}}^\gamma$  into its natural mechanical and thermal parts, as indicated in equation (10.26):

$$\hat{\mathbf{A}}^\gamma(\mathcal{Z}^\gamma) = \hat{\mathbf{A}}_M^\gamma(\mathcal{Z}^\gamma) + \hat{\mathbf{A}}_T^\gamma(\mathcal{Z}^\gamma). \quad (10.38)$$

Within the time step  $\Delta t_{n+1} := t_{n+1} - t_n$ , this split defines two subproblems:

$$(M): \quad \frac{\partial}{\partial t} \bar{\mathcal{Z}}^\gamma = \hat{\mathbf{A}}_M^\gamma(\bar{\mathcal{Z}}^\gamma) \quad \text{and} \quad (T): \quad \frac{\partial}{\partial t} \mathcal{Z}^\gamma = \hat{\mathbf{A}}_T^\gamma(\mathcal{Z}^\gamma); \quad (10.39)$$

a purely mechanical subproblem (M) at frozen temperature along with the mechanical initial and boundary conditions in (10.30)–(10.31) followed by a purely thermal subproblem (T) at frozen configuration along with the thermal initial and boundary conditions in (10.30)–(10.31). Both subproblems are constrained by the evolution of the plastic variables (10.32). We consider the quasistatic problem and integrate both

phases (10.39) with the backward Euler algorithm. Hence, the algorithmic counterpart  $ALGO_M$  for the mechanical subproblem (M) takes the form

$$\text{DIV } \hat{\mathbf{P}}^\gamma(\varphi_{n+1}^\gamma; \theta_n^\gamma) = \mathbf{0} \quad \text{and} \quad \mathbf{v}_{n+1}^\gamma = \frac{1}{\Delta t_{n+1}}(\varphi_{n+1}^\gamma - \varphi_n^\gamma) \quad (10.40)$$

in  $B^\gamma$  at frozen thermal primary variable  $\theta_n^\gamma$ .

The thermal subproblem (T) obtains the algorithmic form  $ALGO_T$

$$\frac{c^\gamma}{\Delta t_{n+1}}(\theta_{n+1}^\gamma - \theta_n^\gamma) = -\text{DIV } \hat{\mathbf{Q}}^\gamma(\theta_{n+1}^\gamma; \varphi_{n+1}^\gamma) + \hat{S}^\gamma(\theta_{n+1}^\gamma; \varphi_{n+1}^\gamma, \mathbf{v}_{n+1}^\gamma) \quad (10.41)$$

in  $B^\gamma$  at frozen mechanical primary variables  $\varphi_{n+1}^\gamma, \mathbf{v}_{n+1}^\gamma$ .

Thus, we first solve within a time step  $\Delta t_{n+1}$  the mechanical problem (10.40) for the actual configuration field  $\varphi_{n+1}^\gamma$ . Next we compute the actual temperature field  $\theta_{n+1}^\gamma$  by solving the thermal problem (10.41). The overall global thermo-mechanical solution algorithm within a typical time step can be regarded as a composition  $ALGO_{TM} = ALGO_T \circ ALGO_M$  of the two subalgorithms  $ALGO_T$  and  $ALGO_M$ . Thus, the operator split algorithm results in an algorithmic decoupling of the coupled thermo-mechanical equations within the time interval. The split (10.39) was proposed as a coupling algorithm by Argyris and Doltsinis (1981), and applied for thermo-mechanical contact by Wriggers and Miehe (1994). It is characterized by an isothermal deformation predictor followed by a heat conduction corrector. Within the work of Argyris and Doltsinis (1981), this algorithm has been interpreted as a one-pass GAUSS-SEIDEL scheme in terms of the variables  $\varphi$  and  $\theta$ .

This algorithm now has to be incorporated in the weak forms (10.36) and (10.37), which are the basis for the finite element method. Since this is again a standard procedure, we do not derive this here explicitly.

## 10.6 DISCRETIZATION TECHNIQUES

Here we discuss two discretizations for thermo-mechanical contact. The first formulation is for frictionless problems with small deformations, hence the linear theory is valid and a node-to-node discretization, as already discussed in Section 7.2 can be applied. The associated finite element formulation can be found in Wriggers and Zavarise (1993b). For the case of large slip in the contact interface, as can be observed in tyre contact or in metal forming processes, one has to use a discretization for large slip as presented in Section 8.1. Furthermore, friction has to be included in the analysis.

### 10.6.1 Node-to-node contact element

In the case of a node-to-node contact discretization, the constraint equation for contact is formulated for each nodal pair, see Figure 7.4.

The geometrical contact constraint condition for the normal contact (3.1) can be formulated for one finite element nodal pair  $I$  as

$$g_{NI} = x_{Ik}^2 - x_{Ik}^1 = (X_{Ik}^2 + u_{Ik}^2) - (X_{Ik}^1 + u_{Ik}^1) \geq 0, \quad (10.42)$$

where  $k$  is the local direction of the normal in the contact interface ( $\mathbf{x}_I^2 \cdot \mathbf{n}_I^1 = x_{Ik}$ ). This has been introduced to simplify notation.  $x^\alpha$  refers to the current and  $X^\alpha$  to the initial configuration.

The temperature jump in the contact interface at node  $I$  is given by

$$g_{\theta I} = \theta_I^2 - \theta_I^1, \quad (10.43)$$

with the current temperature  $\theta^\alpha$  at the two bodies  $B^\alpha$ . The temperature of the gas inside the micro-cavities is computed as the mean temperature of both surfaces,

$$T_{gI} = \frac{1}{2} (\theta_I^2 + \theta_I^1). \quad (10.44)$$

The variation of (10.42) is simply

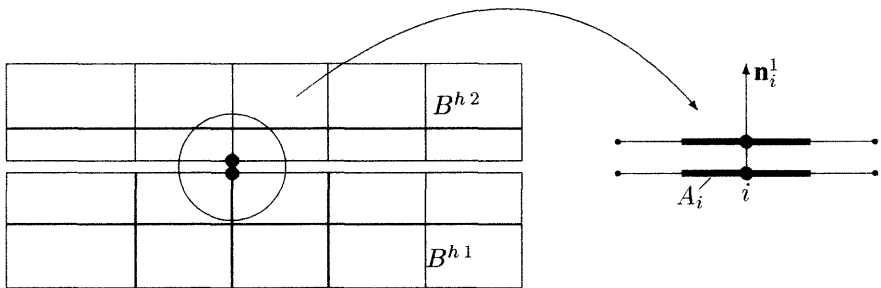
$$\delta g_{NI} = \eta_{Ik}^2 - \eta_{Ik}^1. \quad (10.45)$$

In the same way one derives the variations of  $g_{\theta I}$  and  $T_{gI}$

$$\begin{aligned} \delta g_{\theta I} &= \vartheta_I^2 - \vartheta_I^1, \\ \delta T_{gI} &= \frac{1}{2} (\vartheta_I^2 + \vartheta_I^1). \end{aligned} \quad (10.46)$$

Due to its simplicity, and assuming that the contact constraint is active for  $n_c$  nodes, we can express the integrals (10.36) and (10.37) for the contact contributions in the weak form by a sum over all active contact nodes.

Since the normal contact has to be expressed in the case of thermo-mechanical contact by a constitutive equation, one has to use the formulation of Section 5.3.4. With this one approximates equations (10.36) and (10.37), leading for the contact



**Fig. 10.7** Node-to-node thermal contact element.



part to

$$\hat{G}_M^c = \int_{\Gamma_c} \delta g_N p_N da \rightarrow \sum_{I=1}^{n_c} p_{N I} \delta g_{N I} A_I, \quad (10.47)$$

$$\hat{G}_T^c = \int_{\Gamma_c} (\vartheta^2 - \vartheta^1) q_N da \rightarrow \sum_{I=1}^{n_c} q_{N I} \delta g_{\theta I} A_I. \quad (10.48)$$

The area  $A_i$  is associated with the contact point  $i$ , see Figure 7.4. For computation of the approach due to a mechanical applied pressure in the contact interface, equation (4.12) will be used. This yields a nonlinear relation between the contact pressure and geometrical approach defined in (4.9):

$$p_N = c_N (g_N). \quad (10.49)$$

Note that the constitutive equation for the contact pressure does not depend upon the temperature in the case of frictionless contact.

The heat flux  $q_N$  in the contact interface is assumed to be governed by equations (10.16) for the spot conductance and (10.18) for the gas conductance, leading to the general form of (10.15) for the heat conductance in the contact interface:  $h_c(\theta^1, \theta^2, p_N)$ . Hence, with (10.12) one obtains

$$q_N = h_c(\theta^1, \theta^2, p_N) (\theta^2 - \theta^1) = h_c(T_g, p_N) (\theta^2 - \theta^1). \quad (10.50)$$

Since the heat conduction in the contact interface depends upon the contact pressure, coupling of thermal and mechanical variables is present.

Now, to complete the formulation for the thermo-mechanical contact element, the matrices

$$\boldsymbol{\vartheta}_I = \begin{Bmatrix} \vartheta_I^2 \\ \vartheta_I^1 \end{Bmatrix}, \quad \boldsymbol{\eta}_I = \begin{Bmatrix} \eta_{Ik}^2 \\ \eta_{Ik}^1 \end{Bmatrix}, \quad \mathbf{C}_I = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}, \quad \mathbf{E} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \quad (10.51)$$

are introduced (see also Section 7.2 for the purely mechanical case). With this the weak form of the contact contributions follows from (10.48), together with (10.49) and (10.50), as

$$\hat{G}_M^c = \sum_{I=1}^{n_c} \boldsymbol{\eta}_I^T c_N(g_{N I}) \mathbf{C}_I A_I, \quad (10.52)$$

$$\hat{G}_T^c = \sum_{I=1}^{n_c} \boldsymbol{\vartheta}_I^T h_c(\theta^1, \theta^2, p_N) (\theta^2 - \theta^1) \mathbf{C}_I A_I. \quad (10.53)$$

These terms are nonlinear in the geometrical approach  $g_N$  and the temperatures  $\theta^\alpha$ . The linearization of (10.53) yields the tangent operator which has to be used within a NEWTON type solution procedure, see Section 9.3. For the linearization around a

state  $\bar{g}_N$  and  $\bar{\theta}^\alpha$ , one derives

$$\begin{aligned}
 D G_M \cdot \Delta g_N &= \sum_{I=1}^{n_c} \boldsymbol{\eta}_I^T \left. \frac{\partial c_N}{\partial g_N} \right|_{\bar{g}_N} A_I \mathbf{C}_I \mathbf{C}_I^T \Delta \mathbf{u}_I \\
 D G_T \cdot \Delta g_N &= \sum_{I=1}^{n_c} \boldsymbol{\vartheta}_I^T \left. \frac{\partial h_c}{\partial g_N} \right|_{\bar{g}_N} (\bar{\theta}^2 - \bar{\theta}^1) A_I \mathbf{C}_I \mathbf{C}_I^T \Delta \mathbf{u}_I \\
 D G_T \cdot \Delta \theta &= \sum_{I=1}^{n_c} \boldsymbol{\vartheta}_I^T \left[ h_c \mathbf{C}_I \mathbf{C}_I^T + \frac{1}{2} \left. \frac{\partial h_c}{\partial T_g} \right|_{\bar{T}_g} (\bar{\theta}^2 - \bar{\theta}^1) \mathbf{C}_I \mathbf{E}^T \right] A_I \Delta \boldsymbol{\theta}_I,
 \end{aligned} \quad (10.54)$$

with the increments of the displacements and temperatures

$$\Delta \mathbf{u}_I = \left\{ \begin{array}{c} \Delta u_{Ik}^2 \\ \Delta u_{Ik}^1 \end{array} \right\}, \quad \Delta \boldsymbol{\theta}_I = \left\{ \begin{array}{c} \Delta \theta_I^2 \\ \Delta \theta_I^1 \end{array} \right\}. \quad (10.55)$$

The combination of all contributions from (10.55) then results in the tangent matrix for contact within the thermo-mechanical analysis using node-to-node elements

$$\sum_{I=1}^{n_c} \langle \boldsymbol{\eta}_I^T, \boldsymbol{\vartheta}_I \rangle \mathbf{K}_I^c \left\{ \begin{array}{c} \Delta \mathbf{u}_I \\ \Delta \boldsymbol{\theta}_I \end{array} \right\}, \quad (10.56)$$

with the explicit form of the tangent matrix

$$\mathbf{K}_I^c = \left[ \begin{array}{cc} \frac{\partial c_N}{\partial g_N} \mathbf{C}_I \mathbf{C}_I^T & \mathbf{0} \\ \frac{\partial h_c}{\partial g_N} (\bar{\theta}^2 - \bar{\theta}^1) \mathbf{C}_I \mathbf{C}_I^T & h_c \mathbf{C}_I \mathbf{C}_I^T + \frac{1}{2} \frac{\partial h_c}{\partial T_g} (\bar{\theta}^2 - \bar{\theta}^1) \mathbf{C}_I \mathbf{E}^T \end{array} \right] A_I. \quad (10.57)$$

Note that this matrix is not symmetric. Hence a solver for a non-symmetric matrix system has to be applied, even in the case of frictionless heat transfer.

#### REMARKS 10.1:

1. A linear constitutive relation for the contact pressure as used in the standard penalty formulation, see Section 5.3.2, the term  $\partial c_N / \partial g_n$  yields the penalty parameter and one recovers the penalty formulation, see also (7.27).
2. If no gas is contained in the micro cavities, or if the dependence of the conductivity through the gas is negligible, then the heat conductance in the interface only depends upon the contact pressure  $h_c(\theta^\alpha, p_N) \rightarrow h_s(g_N)$ . For this case, the last submatrix in (10.57) becomes symmetric and takes the form  $h_c \mathbf{C}_I \mathbf{C}_I^T$ .
3. When a staggered scheme is applied in which one first solves the mechanical and then the thermal part, as discussed in Section 10.5, then the coupling term in (10.57) vanishes, since the pressure is fixed when solving the thermal part of the coupled problem.

### 10.6.2 Node-to-segment contact element

Node-to-segment contact elements are used for large deformation contact. For thermo-mechanical contact such elements have been developed by Zavarise (1991), Zavarise

et al. (1992b) or Wriggers and Miehe (1994) for different constitutive equations in the contact interface. In this section only the contributions due to the thermal and coupled thermo-mechanical parts of the contact will be considered, since the purely mechanical NTS-element has been stated in Section 8.1.

Again, the projection of the slave node  $\mathbf{x}_s^2$  onto the master segment (1)–(2) is needed, see Figure 10.8. The coordinates denoting the surface of the master segment are given by a linear interpolation  $\hat{\mathbf{x}}_t^1(\xi) = \mathbf{x}_1^1 + \xi(\mathbf{x}_2^1 - \mathbf{x}_1^1)$ . The evaluation of (8.3) yields the surface coordinate  $\bar{\xi}_{n+1}$ , which denotes the minimum distance point on the master segment.

In the same way, the interpolation for the temperature on the master element surface yields

$$\hat{\theta}_t^1(\xi) = (1 - \xi)\theta_1^1 - \xi\theta_2^1, \quad (10.58)$$

with the nodal values  $\theta_1^1$  and  $\theta_2^1$ .

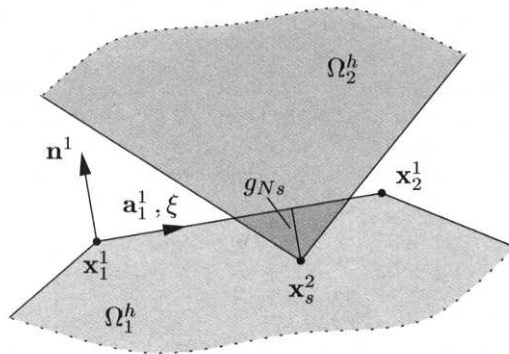
In what follows we formulate the constitutive relationships for the normal force  $P_N$ , tangential force  $T_T$ , the discrete heat flux  $Q_N$  and the discrete dissipation  $\mathcal{D}^p$  at the discrete slave node of the contact element under consideration. Assume that  $\bar{\xi}_{n+1}$  and  $g_{N\,n+1}$  are known from the exploitation of (8.3).

With this one can compute analogous to (4.12), the normal force at the slave node

$$P_{Ns\,n+1} = C_N (g_{N\,n+1})^m. \quad (10.59)$$

The algorithmic update of the tangential force is performed by the return algorithm as described in Section 9.5. Note that only one component appears in the two-dimensional case. The return algorithm then leads for COULOMB'S friction law, with (9.118), to

$$T_{T\,n+1} = \begin{cases} T_{T\,n+1}^{trial} & \text{for } f_{s\,n+1}^{trial} \leq 0, \\ T_{T\,n+1}^{trial} (\mu P_{N\,n+1} / |T_{T\,n+1}^{trial}|) & \text{otherwise.} \end{cases} \quad (10.60)$$



**Fig. 10.8** Node-to-segment thermal contact element.

Using (9.121), the frictional dissipation is obtained by

$$\mathcal{D}_{n+1}^p = \begin{cases} 0 & \text{for } f_{s\,n+1}^{trial} \leq 0, \\ T_{T\,n+1} (T_{T\,n+1}^{trial} - T_{T\,n+1}) / C_T \Delta t_{n+1} & \text{otherwise,} \end{cases} \quad (10.61)$$

where  $C_T$  denotes the tangential stiffness in the contact interface.

Finally, the discrete heat flux takes the form analogous to (10.17)

$$Q_{N\,n+1} = H_{So} \left[ \frac{P_{N\,n+1}/l_{n+1}}{H} \right]^\epsilon [\theta_{s\,n+1}^2 - (1 - \bar{\xi}_{n+1}) \theta_{1\,n+1}^1 - \bar{\xi}_{n+1} \theta_{2\,n+1}^1]. \quad (10.62)$$

Here  $l_{n+1}$  is the current length of the master segment ( $l_{n+1} = \|\mathbf{x}_{2\,n+1}^1 - \mathbf{x}_{1\,n+1}^1\|$ ).

The contributions of the thermo-mechanical contact in the mechanical and thermal weak forms (10.36) and (10.37) take the form

$$G_M^c = \delta g_{N\,n+1} P_{N\,n+1} - \delta g_{T\,n+1} T_{T\,n+1}, \quad (10.63)$$

$$G_T^c = \delta g_{Q\,n+1} Q_{N\,n+1} - \delta g_{\mathcal{D}\,n+1} \mathcal{D}_{n+1}^p, \quad (10.64)$$

for the discrete slave node with the mechanical variations analogous to (8.4) and (8.9) and thermal variations (see interpolation (10.58))

$$\delta g_{Q\,n+1} = [\vartheta_s^2 - (1 - \bar{\xi}_{n+1}) \vartheta_1^1 - \bar{\xi}_{n+1} \vartheta_2^1], \quad (10.65)$$

$$\delta g_{\mathcal{D}\,n+1} = [\vartheta_s^2 + (1 - \bar{\xi}_{n+1}) \vartheta_1^1 + \bar{\xi}_{n+1} \vartheta_2^1] / 2. \quad (10.66)$$

These equations can now be cast into a matrix formulation. For the mechanical part, this is given in (8.14) and (8.15). Thus, the virtual mechanical work of the contact element (10.63) can be written in the matrix formulation  $G_M^c = \mathbf{u}^T \cdot \mathbf{R}_{M\,n+1}^c$  with the mechanical contact element residual

$$\boxed{\mathbf{R}_{M\,n+1}^c := P_{N\,n+1} \mathbf{N}_{n+1} - T_{T\,n+1} \hat{\mathbf{T}}_{n+1}}. \quad (10.67)$$

In this equation the contact normal force  $P_{N\,n+1}$  follows from (10.59), whereas the tangential force  $T_{T\,n+1}$  is given by the return algorithm (10.60). Due to this approach, a pure displacement formulation of the contact problem is possible, which is in contrast to the LAGRANGIAN multiplier technique which is often used to enforce the non-penetration condition. For a global algorithmic treatment using NEWTON's method, we have to linearize equation (10.67). The associated formulation for this discretization can be found in Wriggers et al. (1990).

The matrix formulation of the thermal part (10.64) is similar to the mechanical part. As a consequence of the global operator split algorithm discussed in Section 10.5, and the assumed simplified constitutive equation, the thermal part is linear.

A formulation completely analogous to that of the mechanical part yields the matrix representation  $G_T^c = \boldsymbol{\vartheta} \cdot \mathbf{R}_{T\,n+1}^c$  for the virtual thermal work (10.64), with the thermal contact element residual

$$\boxed{\mathbf{R}_{T\,n+1}^c := Q_{N\,n+1} \mathbf{H}_{n+1} - \mathcal{D}_{n+1}^p \mathbf{H}_{0\,n+1}}. \quad (10.68)$$

Here we have introduced the matrices

$$\boldsymbol{\vartheta} := (\vartheta_s^2 \quad \vartheta_1^1 \quad \vartheta_2^1)^T,$$

and

$$\mathbf{H}_{n+1} = \left\{ \begin{array}{c} 1 \\ -(1 - \bar{\xi}) \\ -\bar{\xi} \end{array} \right\}_{n+1}, \quad \mathbf{H}_{0\,n+1} = \frac{1}{2} \left\{ \begin{array}{c} 1 \\ (1 - \bar{\xi}) \\ \bar{\xi} \end{array} \right\}_{n+1}, \quad (10.69)$$

for the matrix formulation of the thermal variations (10.65) and (10.66). In (10.68) the discrete contact heat flux  $Q_{N\,n+1}$  follows from (10.62), whereas the frictional dissipation  $\mathcal{D}^p_{n+1}$  is given by the return algorithm (10.61). The nodal contact contribution (10.67) and (10.68) have to be added to the global system of equations.

Which terms of the contact contributions have to be linearized now depends upon the global algorithm used to solve the thermo-mechanical contact problem.

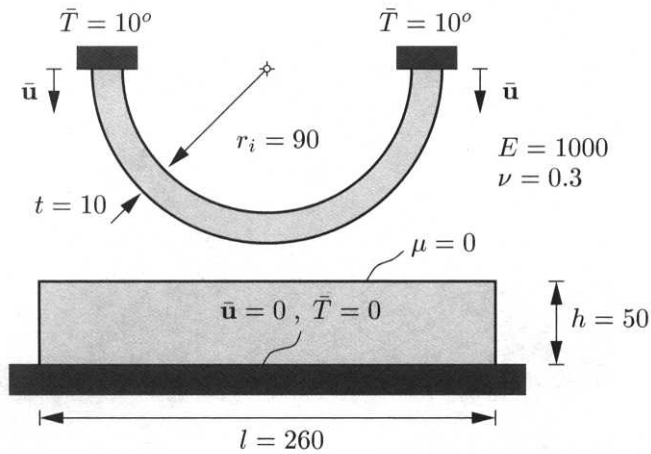
## 10.7 EXAMPLES

Two examples are discussed here to show the performance of the node-to-segment contact formulation for thermo-mechanical problems. Both problems exhibit finite deformations which are modelled by assuming the thermo-elastic constitutive equation based on the HELMHOLTZ energy (10.7) and FOURIER'S law (10.11). For the analysis two different strategies could be followed: the fully coupled or the staggered schemes. The latter are based on the operator split technique, described in detail in Section 10.5. In the fully coupled analysis, all coupling terms are taken into account implicitly, whereas the staggered scheme takes care of the coupling terms explicitly. Here fully coupled treatment is used for both examples. The time integration which is necessary to solve the non-stationary second example is performed by an implicit EULER scheme.

All examples are solved using the finite element software FEAP, see Zienkiewicz and Taylor (2000a). In this code all extensions related to thermo-mechanical contact analysis, which were discussed in the previous chapters, are implemented. The meshes are generated by the software tool DOMESH developed by Rank et al. (1993).

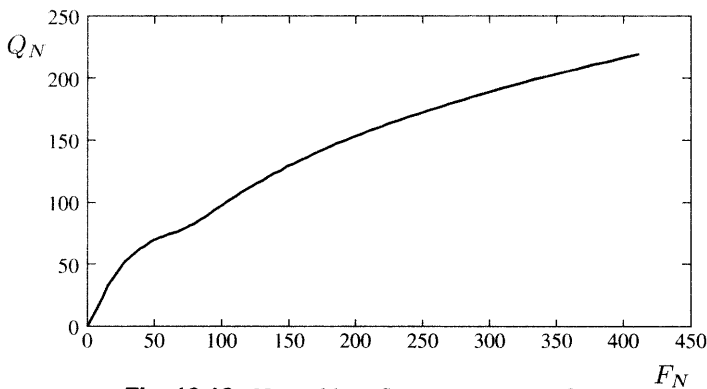
### 10.7.1 Heat transfer at finite deformations

The first example discusses heat transfer through a changing contact surface when finite deformations are present. The problem is depicted in Figure 10.9, which consists of a ring and a foundation. The contact is assumed to be frictionless. During the analysis steady state conditions are assumed for the different loading steps. The material parameters for the mechanical part are  $E = 1000$ ,  $\nu = 0.3$ . For the thermal part the heat conductivity coefficient is given with  $k = 100$ , and the heat capacity is  $c = 1$ . The thermal expansion coefficient has the value  $\alpha = 0.0001$ . The reference temperature is zero.

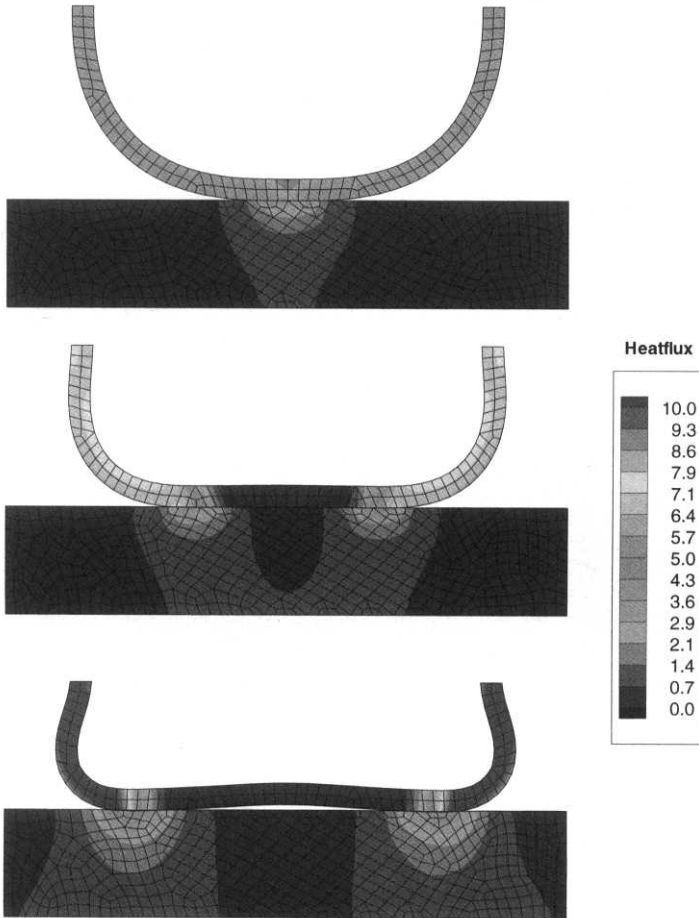


**Fig. 10.9** Elastic ring in contact with a foundation.

In the contact interface we assume frictionless contact. The parameters in the pressure-dependent interface law for the heat conductance (10.17) are given as thermal resistance coefficient  $h_{c0} = 1.0$ , VICKERS hardness  $H_v = 3.0$ , and  $\omega = 1.5$ . A temperature  $\bar{T}$  and a displacement  $\bar{u}$  are prescribed at the top of the ring. Eighty load steps were performed with a prescribed displacement of  $\bar{u} = 0.5$ , see Figure 10.9. In Figure 10.10 the resulting heat flux in the contact interface is plotted versus the total contact force. The heat flux increases in a nonlinear way with respect to the normal contact force. This is related to the change in the contact surface from one contact area in the middle of the ring to two different contact areas at its sides. Furthermore, the contact pressure increases due to the loading process, which increases the prescribed displacement. As a result, a temperature distribution occurs



**Fig. 10.10** Normal heat flux versus contact force.



**Fig. 10.11** Elastic ring in contact with a foundation.

in ring and foundation which depends heavily on the contact pressure at the contact interface and the deformed configuration of the ring, see Figure 10.11.

### 10.7.2 Frictional heating at finite deformations

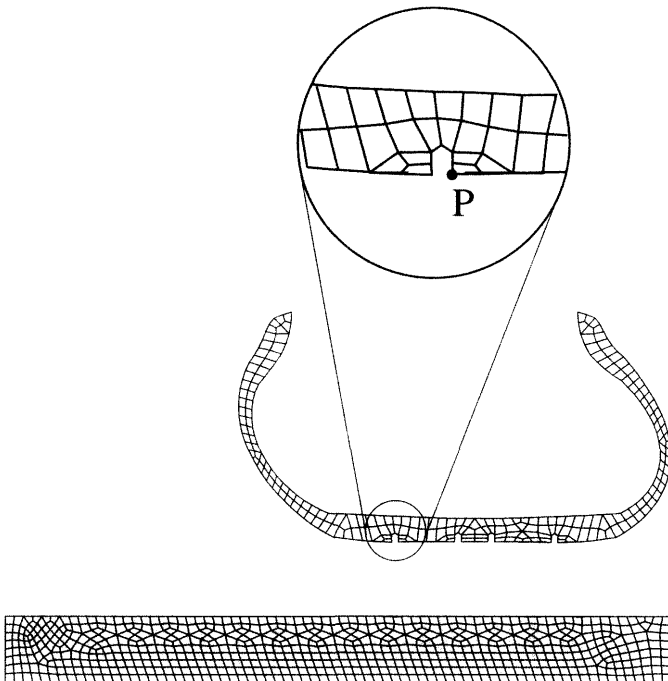
In the next example the frictional contact of a tyre section with a road is investigated. This is a problem in which the temperature field, generated by friction, depends upon the time, leading to non-stationary heat conduction problem. However, we assume that the process is slow enough that no inertia forces have to be included in the

mechanical part of the weak forms. Again, finite deformations are assumed. The analysis is performed in different steps. First the gas pressure in the tyre is increased up to a certain value (inflation of the tyre). Then the tyre is pressed against the road surface, and finally, the tyre is moved with a prescribed velocity at its top horizontally across the road.

The mesh is shown in Figure 10.12 in the undeformed configuration. The mechanical material properties for the tyre are  $E = 1000$  and  $\nu = 0.3$ . The heat conductivity coefficient is  $k = 5$  and the heat capacity is chosen as  $c = 1$ . The thermal coefficient is  $\alpha = 0.0001$  and the reference temperature is zero. The road material has only a different modulus of elasticity  $E = 10000$ .

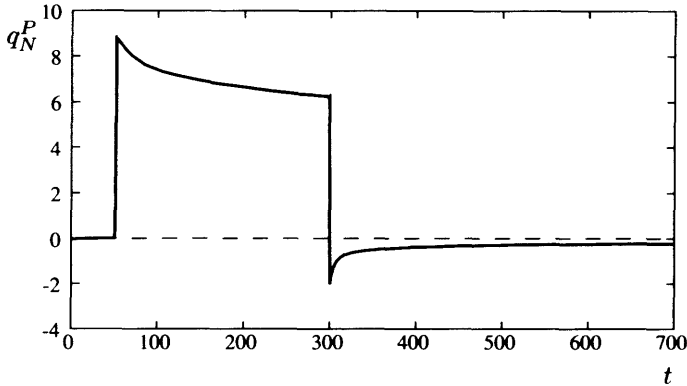
In the contact interface we assume frictional contact with COULOMB'S law and a friction coefficient of  $\mu = 0.5$ . The parameters in the pressure-dependent interface law for the heat conductance (10.17) are given as thermal resistance coefficient  $h_{c0} = 1.0$ , VICKERS hardness  $H_v = 3.0$ , and  $\omega = 1.5$ . The internal pressure in the tyre is  $P = 0.1$ . The prescribed horizontal velocity at the top has a value of 7.

During the deformation process, which was computed using 1500 time steps, frictional sliding occurs in the contact interface, leading to frictional heating. Thus a heat flux is generated in the contact interface.



**Fig. 10.12** Elastic tyre in contact with a road.





**Fig. 10.13** Heat flux at point  $P$  versus time.

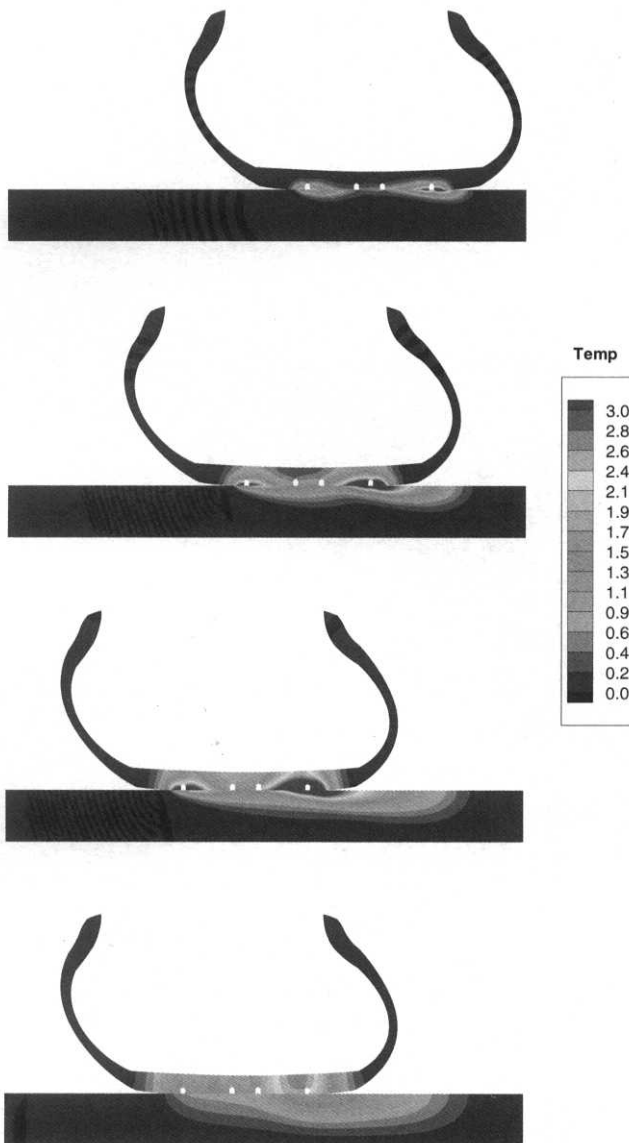
To show the general behaviour of the heat transfer mechanism in this problem, the local heat flux  $q_N^P$  is measured in point  $P$  denoted in Figure 10.12. This quantity is plotted versus the process time in Figure 10.13.

It can be seen that no heat flux occurs in the initial phase ( $0 \leq t < 50$ ) when the tyre is pressed onto the road surface.

During the tangential movement, a positive heat flux with decreasing tendency is observed within  $50 \leq t < 300$ , which means the heat flows into the road surface. This can also be observed in the first two Figures 10.14, which depict the deformed configurations of the tyre and the temperature distributions at time  $t = 150$  and  $t = 250$ , respectively.

At time  $t = 300$  the horizontal movement is stopped, which leads to a jump in the heat flux. Now heat flows back from the road into the tyre (see the last two temperature states at  $t = 350$  and  $t = 400$  in Figure 10.14).

At  $t = 700$  steady state conditions are reached, and hence the heat flux at point  $P$  tends to zero.



**Fig. 10.14** Temperature distribution in road and tyre due to frictional heating.

**This page intentionally left blank**

# 11

---

## *Beam Contact*

Many engineering problems involve contact of beams undergoing large displacement. These include robot parts, woven fabrics or racquet strings. In such cases the beams are either already in contact in the initial configuration, like woven fabrics, or they come into contact during the motion which can happen when parts of robots are moving. The contact formulations discussed in the previous chapters are usually developed for two- or three-dimensional solids. However, the contact between beams has some specialities which are related to the description of the beam as a one-dimensional curve in space. Thus, the formulation of the kinematical contact conditions is different, and hence will be derived in this chapter in detail.

Since the underlying theories for beams in three-dimensional space are very complex we will not describe these here. The interested reader should consult Simo (1985), Simo and Vu-Quoc (1986), Crisfield (1990) or Gruttmann et al. (2000) for finite element formulations of beams undergoing large deflections and rotations.

### **11.1 KINEMATICS**

In this section we describe the kinematical relations which have to be used when contact between two beams occurs. The beams are mathematically described by curves in the three-dimensional space. Therefore, we have to consider basically the contact between two curves. This leads locally to one contact point as long as the beams are not parallel and have a smooth convex cross-section. We now derive the kinematical relations for the contact of two curves in space. Most of the formulation is based on Wriggers and Zavarise (1997), Zavarise and Wriggers (2000) and Litewka

and Wriggers (2002a), which consider the contact of beams in three-dimensional space for the frictionless and frictional cases.

### 11.1.1 Normal contact

Figure 11.1 depicts the geometrical situation of two crossing beams in space. We parameterize the reference curve (beam axis) of beams  $\mathcal{B}_\xi$  and  $\mathcal{B}_\zeta$  by the convective coordinates  $\xi$  and  $\zeta$ . For simplicity, we assume that the beams have a circular cross-section with radii  $r_\xi$  and  $r_\zeta$ . A more advanced treatment in which rectangular cross-sections are considered can be found in Litewka and Wriggers (2002a), however the general formulation does not change because of that. Thus we proceed here with the more simple case of circular cross-sections. The geometrical setup for rectangular beams is discussed in Remark 11.1. With the notation introduced, the position of a point of the beam axis is given in the deformed configuration by

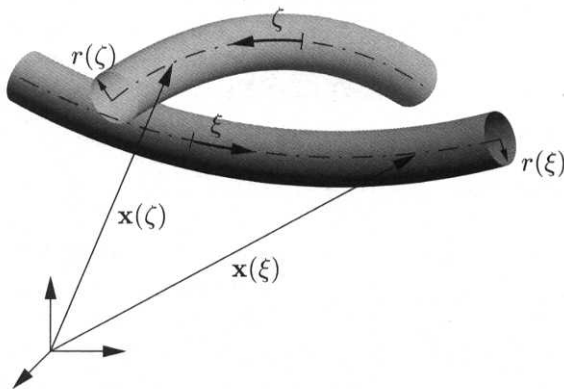
$$\varphi(\mathcal{B})_\xi : \mathbf{x}^\xi(\xi) \quad \text{and} \quad \varphi(\mathcal{B})_\zeta : \mathbf{x}^\zeta(\zeta). \quad (11.1)$$

The contact condition for normal contact can be computed similar to the minimum distance problem (3.2). However, now we search for the minimum distance between the two beams. Thus, we do not fix a point at one beam as in Section 3.1, but search the minimum distance regarding both parameters  $\xi$  and  $\zeta$  as variables. This leads to

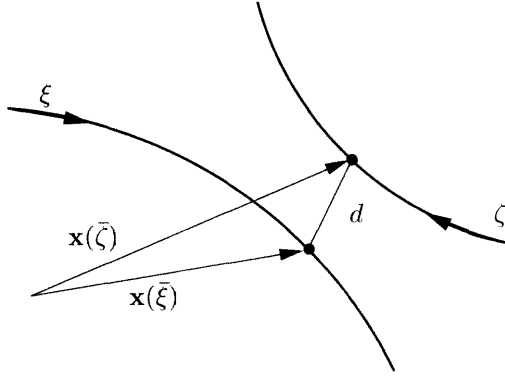
$$d(\xi, \zeta) = \min_{\xi, \zeta} \|\mathbf{x}^\xi(\xi) - \mathbf{x}^\zeta(\zeta)\| = \|\mathbf{x}^\xi(\bar{\xi}) - \mathbf{x}^\zeta(\bar{\zeta})\| = \|\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta\|, \quad (11.2)$$

where  $\bar{\xi}$  and  $\bar{\zeta}$  denote the position of points on the two beam axes which have a minimum distance, see Figure 11.2. Since the contact occurs on the beam surface, we can introduce the gap function

$$g_N = d - (r_\xi + r_\zeta) \quad (11.3)$$



**Fig. 11.1** Geometry of contacting beams.



**Fig. 11.2** Minimum distance of contacting beams.

to measure the minimum distance where the radii of the cross-sections have to be subtracted from the distance function  $d$ . Penetration takes place for negative values of  $g_N$ . Hence, the contact condition for the beams can be written analogous to (3.6) as

$$g_N \geq 0. \quad (11.4)$$

The minimum problem (11.2) is in general for arbitrary curves a nonlinear problem. The minimum distance is computed from the two conditions

$$\begin{aligned} \frac{d}{d\xi} \hat{d}(\xi, \zeta) &= \frac{\mathbf{x}^\xi(\xi) - \mathbf{x}^\zeta(\zeta)}{\|\mathbf{x}^\xi(\xi) - \mathbf{x}^\zeta(\zeta)\|} \cdot \mathbf{x}_{,\xi}^\xi(\xi) = 0, \\ \frac{d}{d\zeta} \hat{d}(\xi, \zeta) &= -\frac{\mathbf{x}^\xi(\xi) - \mathbf{x}^\zeta(\zeta)}{\|\mathbf{x}^\xi(\xi) - \mathbf{x}^\zeta(\zeta)\|} \cdot \mathbf{x}_{,\zeta}^\zeta(\zeta) = 0, \end{aligned} \quad (11.5)$$

which represent a nonlinear system of equations for the two unknown positions  $\bar{\xi}$  and  $\bar{\zeta}$ , see Figure 11.2. The solution of this system of equations is provided by NEWTON'S method leading to the incremental system of equations

$$\mathbf{A}(\xi, \zeta) \begin{Bmatrix} \Delta\xi \\ \Delta\zeta \end{Bmatrix} = - \begin{Bmatrix} [\mathbf{x}^\xi(\xi) - \mathbf{x}^\zeta(\zeta)] \cdot \mathbf{x}_{,\xi}^\xi(\xi) \\ [\mathbf{x}^\zeta(\zeta) - \mathbf{x}^\xi(\xi)] \cdot \mathbf{x}_{,\zeta}^\zeta(\zeta) \end{Bmatrix}, \quad (11.6)$$

with

$$\mathbf{A}(\xi, \zeta) = \begin{bmatrix} \mathbf{x}_{,\xi}^\xi \cdot \mathbf{x}_{,\xi}^\xi + [\mathbf{x}^\xi - \mathbf{x}^\zeta] \cdot \mathbf{x}_{,\xi\xi}^\xi & -\mathbf{x}_{,\xi}^\xi \cdot \mathbf{x}_{,\zeta}^\zeta \\ -\mathbf{x}_{,\xi}^\xi \cdot \mathbf{x}_{,\zeta}^\zeta & \mathbf{x}_{,\zeta}^\zeta \cdot \mathbf{x}_{,\zeta}^\zeta + [\mathbf{x}^\zeta - \mathbf{x}^\xi] \cdot \mathbf{x}_{,\zeta\zeta}^\zeta \end{bmatrix}. \quad (11.7)$$

The nonlinear system (11.5) might have several or no solutions. Here we assume that a unique solution of (11.5) exists. Hence the multiple solution cases are not considered here, although one has to take these into account in the algorithmic treatment within the finite element program.

For straight beams the curvature terms  $\mathbf{x}_{,\xi\xi}^\xi$  and  $\mathbf{x}_{,\zeta\zeta}^\zeta$  disappear, and the tangent vectors  $\mathbf{x}_{,\xi}^\xi$  and  $\mathbf{x}_{,\zeta}^\zeta$  are constant. Then (11.5) reduces to two linear equations for  $\bar{\xi}$  and  $\bar{\zeta}$ , which can be solved directly. This solution can also be used as a starting point when NEWTON'S method (see (11.6)) has to be applied for curved beams.

Note that (11.6) can also be applied to compute the variations and linearizations of  $\xi$  and  $\zeta$  which will be needed in later sections for derivation of the weak form and its linearization.

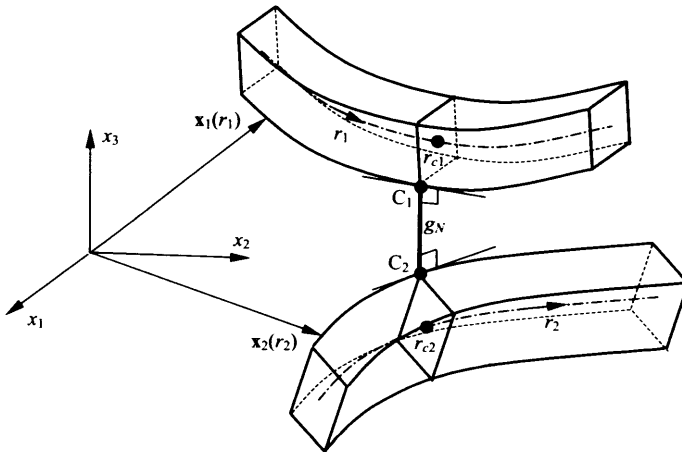
Once the solution of (11.5) is known, we can define the normal vector by

$$\bar{\mathbf{n}} = \mathbf{n}(\bar{\xi}, \bar{\zeta}) = \frac{\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta}{\|\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta\|}, \quad (11.8)$$

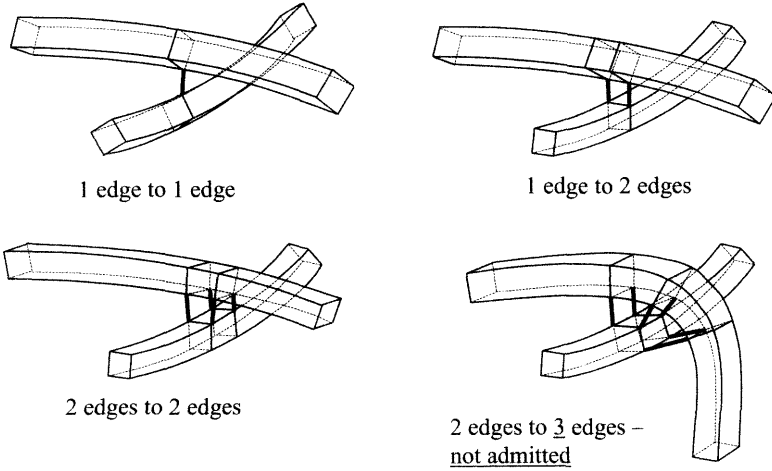
which is normal to both beams at the minimum distance point  $(\bar{\xi}, \bar{\zeta})$ . In the later sections concerning beam contact, we assume that the beams are straight to avoid too complex derivations. This assumption also goes well with the finite element method, since there beams for large deformation problems are in most cases based on straight elements. Thus, locally the assumption is then valid.

**REMARK 11.1:** In the case of rectangular beams the definition of the gap function is more elaborate, since contact can occur at different edges of the beam cross-section. This is depicted in Figure 11.3 for the case of edge to edge contact. Contrary to the contact of beams with a circular cross-section, the normal contact force is eccentric with respect to the beam axes  $r_1$  and  $r_2$ , and hence a torsional moment is generated by the normal contact force with respect to  $r_{c1}$  and  $r_{c2}$ . This is also the case for a force due to friction, which leads to additional bending moments in the beam around  $r_{c1}$  and  $r_{c2}$ .

However, as shown in Figure 11.4, the situation can become even more complex in the case of the contact of beams with rectangular cross-sections, since besides the pointwise contact,



**Fig. 11.3** Gap definition for beams with rectangular cross-sections.



**Fig. 11.4** Gap definition for beams with rectangular cross-sections.

described in Figure 11.4.1, other configurations can occur. These are a line contact of one edge to a surface on the beam, see Figure 11.4.2, and an area contact of the surface of one beam to the surface of the other beam, see Figure 11.4.3. Within the contact analysis we assume, according to classical beam theory, that cross-sections of contacting beams do not undergo any deformation, i.e. they remain plane rectangles, although not necessarily perpendicular to the respective beam axes as the shear deformation is allowed within a beam element. We can exclude the fourth configuration in Figure 11.4.4, since this would involve large finite strains in the beam model. Such a model will not be pursued here, based on the fact that for such cases a fully three-dimensional discretization is more adequate.

### 11.1.2 Tangential contact

In the case of the tangential contact, we have to consider the fact that the contact point also slides on beam  $\mathcal{B}_\xi$  as well as on beam  $\mathcal{B}_\zeta$ , in general. Again, we can compute the sliding distance from the integral (3.16); this leads to two integrals which describe the total sliding distance for beam  $\mathcal{B}_\xi$  and for  $\mathcal{B}_\zeta$  in the current configuration:

$$g_T^\xi = \int_{t_0}^t \|\dot{\xi} \bar{\mathbf{x}}_\xi^\xi\| dt \quad \text{and} \quad g_T^\zeta = \int_{t_0}^t \|\dot{\zeta} \bar{\mathbf{x}}_\zeta^\zeta\| dt. \quad (11.9)$$

Here we have to integrate with respect to the time, see also Section 3.2. Again, for straight beams the tangent vectors are constant, and the integrals in (11.9) can be



evaluated exactly. We obtain in this case

$$\boxed{g_T^\xi = [\xi(t) - \xi(t_0)] l^\xi \quad \text{and} \quad g_T^\zeta = [\zeta(t) - \zeta(t_0)] l^\zeta}, \quad (11.10)$$

where  $l^\xi$  and  $l^\zeta$  denote the beam length of  $\varphi(B)^\xi$  and  $\varphi(B)^\zeta$ , respectively.

The relative velocity of the contact point is now given with respect to both beams as  $\dot{\xi} \bar{\mathbf{x}}_{,\xi}^\xi$  and  $\dot{\zeta} \bar{\mathbf{x}}_{,\zeta}^\zeta$ . This means that both beams slide relative to each other.

## 11.2 VARIATION OF THE GAP IN NORMAL AND TANGENTIAL DIRECTIONS

The normal gap for two beams in contact was given with (11.2) and (11.3) by

$$g_N = \bar{g}_N - (r_\xi + r_\zeta) = \|\mathbf{x}^\xi(\bar{\xi}) - \mathbf{x}^\zeta(\bar{\zeta})\| - (r_\xi + r_\zeta). \quad (11.11)$$

The variation of the normal gap follows with this definition as

$$\delta g_N = \delta \bar{g}_N = \frac{1}{\|\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta\|} (\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta) \cdot (\bar{\boldsymbol{\eta}}^\xi - \bar{\boldsymbol{\eta}}^\zeta).$$

Using the definition of the normal, see (11.8), we obtain a similar result as for the continuum case, see (3.29):

$$\boxed{\delta g_N = (\bar{\boldsymbol{\eta}}^\xi - \bar{\boldsymbol{\eta}}^\zeta) \cdot \bar{\mathbf{n}}}. \quad (11.12)$$

For a straight beam the tangential gap was formulated in (11.10), leading to

$$g_T^\xi = [\xi(t) - \xi(t_0)] l^\xi \quad \text{and} \quad g_T^\zeta = [\zeta(t) - \zeta(t_0)] l^\zeta. \quad (11.13)$$

The variation of these two equations yields

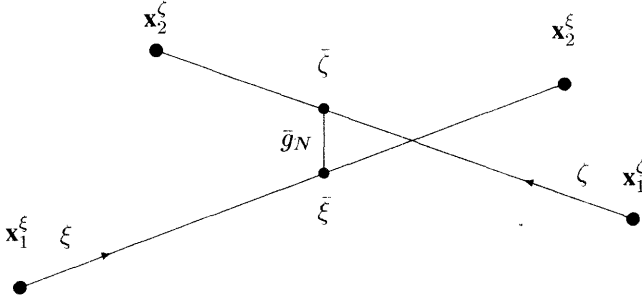
$$\begin{aligned} \delta g_T^\xi &= l^\xi \delta \xi + [\xi(t) - \xi(t_0)] \delta l^\xi, \\ \delta g_T^\zeta &= l^\zeta \delta \zeta + [\zeta(t) - \zeta(t_0)] \delta l^\zeta. \end{aligned} \quad (11.14)$$

In these equations the variations of the convective coordinates  $\xi$  and  $\zeta$  are missing. Also, we have to compute the variation of lengths  $l^\xi$  and  $l^\zeta$ . The latter task can be accomplished by assuming that both lengths  $l$  are given for a straight beam by

$$l^\xi = \|\mathbf{x}_2^\xi - \mathbf{x}_1^\xi\| \quad \text{and} \quad l^\zeta = \|\mathbf{x}_2^\zeta - \mathbf{x}_1^\zeta\|, \quad (11.15)$$

where  $\mathbf{x}_i^\xi$  and  $\mathbf{x}_i^\zeta$  are the end points of both beams (in a finite element discretization these will, for example, be the element nodes), see Figure 11.5. With these definitions we compute

$$\begin{aligned} \delta l^\xi &= \mathbf{t}^\xi \cdot (\boldsymbol{\eta}_2^\xi - \boldsymbol{\eta}_1^\xi) \quad \text{with} \quad \mathbf{t}^\xi = \frac{\mathbf{x}_2^\xi - \mathbf{x}_1^\xi}{\|\mathbf{x}_2^\xi - \mathbf{x}_1^\xi\|}, \\ \delta l^\zeta &= \mathbf{t}^\zeta \cdot (\bar{\boldsymbol{\eta}}_2^\zeta - \bar{\boldsymbol{\eta}}_1^\zeta) \quad \text{with} \quad \mathbf{t}^\zeta = \frac{\mathbf{x}_2^\zeta - \mathbf{x}_1^\zeta}{\|\mathbf{x}_2^\zeta - \mathbf{x}_1^\zeta\|}. \end{aligned} \quad (11.16)$$



**Fig. 11.5** Two straight beams with initial gap.

The derivation of  $\delta\xi$  and  $\delta\zeta$  is more complex. For this we start from the condition that the gap vector is orthogonal to the beam tangents at the closest point, see (11.5):

$$\begin{aligned} (\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta) \cdot \bar{\mathbf{x}}_{,\xi}^\xi &= 0, \\ (\bar{\mathbf{x}}^\zeta - \bar{\mathbf{x}}^\xi) \cdot \bar{\mathbf{x}}_{,\zeta}^\zeta &= 0. \end{aligned}$$

The variation of the first equation yields

$$(\bar{\eta}^\xi + \bar{\mathbf{x}}_{,\xi}^\xi \delta\xi - \bar{\eta}^\zeta - \bar{\mathbf{x}}_{,\zeta}^\zeta \delta\zeta) \cdot \bar{\mathbf{x}}_{,\xi}^\xi + (\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta) \cdot (\bar{\eta}_{,\xi}^\xi + \bar{\mathbf{x}}_{,\xi\xi}^\xi \delta\xi) = 0.$$

The same relation can be written down for the second equation above. This then leads to a system of equations for both variations,

$$\begin{bmatrix} \bar{a}_{\xi\xi} + \bar{g}_N \bar{b}_{\xi\xi} & \bar{a}_{\xi\zeta} \\ \bar{a}_{\xi\zeta} & \bar{a}_{\zeta\zeta} - \bar{g}_N \bar{b}_{\zeta\zeta} \end{bmatrix} \begin{Bmatrix} \delta\xi \\ \delta\zeta \end{Bmatrix} = -\bar{\mathbf{b}}, \quad (11.17)$$

with

$$\begin{aligned} \bar{a}_{\xi\xi} &= \bar{\mathbf{x}}_{,\xi}^\xi \cdot \bar{\mathbf{x}}_{,\xi}^\xi, & \bar{b}_{\xi\xi} &= \bar{\mathbf{n}} \cdot \bar{\mathbf{x}}_{,\xi\xi}^\xi, \\ \bar{a}_{\xi\zeta} &= \bar{\mathbf{x}}_{,\xi}^\zeta \cdot \bar{\mathbf{x}}_{,\zeta}^\zeta, & \bar{b}_{\zeta\zeta} &= \bar{\mathbf{n}} \cdot \bar{\mathbf{x}}_{,\zeta\zeta}^\zeta. \end{aligned} \quad (11.18)$$

Here definitions (11.11) and (11.8) have been used. The right-hand side in (11.17) is given by

$$\bar{\mathbf{b}} = \begin{Bmatrix} (\bar{\eta}^\xi - \bar{\eta}^\zeta) \cdot \bar{\mathbf{x}}_{,\xi}^\xi + (\bar{\mathbf{x}}^\xi - \bar{\mathbf{x}}^\zeta) \cdot \bar{\eta}_{,\xi}^\xi \\ (\bar{\eta}^\zeta - \bar{\eta}^\xi) \cdot \bar{\mathbf{x}}_{,\zeta}^\zeta + (\bar{\mathbf{x}}^\zeta - \bar{\mathbf{x}}^\xi) \cdot \bar{\eta}_{,\zeta}^\zeta \end{Bmatrix}. \quad (11.19)$$

For the contact of straight beams, as shown in Figure 11.5, several terms in (11.18) and (11.19) disappear. These are the second derivatives of the position vectors with respect to the parameters  $\xi$  and  $\zeta$ . Furthermore, the first derivative yields a constant vector, leading to

$$\bar{\eta}_{,\xi}^\xi = \bar{\eta}_2^\xi - \bar{\eta}_1^\xi \quad \text{and} \quad \bar{\eta}_{,\zeta}^\zeta = \bar{\eta}_2^\zeta - \bar{\eta}_1^\zeta. \quad (11.20)$$

Using this, for the coefficients of  $\bar{\mathbf{A}}$  we obtain

$$\bar{a}_{\xi\xi} = (l^\xi)^2, \quad \bar{a}_{\xi\zeta} = l^\xi l^\zeta t_{\xi\zeta} \quad \text{and} \quad \bar{a}_{\zeta\zeta} = (l^\zeta)^2, \quad (11.21)$$

where the definition  $t_{\xi\zeta} = \mathbf{t}^\xi \cdot \mathbf{t}^\zeta$  with the unit tangent vectors from (11.16) was used. Furthermore, for the right hand-side of (11.17) with definitions (11.11) and (11.8), we obtain

$$\bar{\mathbf{b}} = \begin{bmatrix} l^\xi & 0 \\ 0 & l^\zeta \end{bmatrix} \left\{ \begin{array}{l} (\bar{\eta}^\xi - \bar{\eta}^\zeta) \cdot \mathbf{t}^\xi + \frac{\bar{g}_N}{l^\xi} (\bar{\eta}_2^\xi - \bar{\eta}_1^\xi) \cdot \bar{\mathbf{n}} \\ (\bar{\eta}^\zeta - \bar{\eta}^\xi) \cdot \mathbf{t}^\zeta - \frac{\bar{g}_N}{l^\zeta} (\bar{\eta}_2^\zeta - \bar{\eta}_1^\zeta) \cdot \bar{\mathbf{n}} \end{array} \right\} = \mathbf{L} \hat{\mathbf{b}}. \quad (11.22)$$

With these relations for straight beams, equation (11.17) can now be solved for  $\delta\xi$  and  $\delta\zeta$ :

$$\left\{ \begin{array}{c} \delta\xi \\ \delta\zeta \end{array} \right\} = \frac{1}{1 - t_{\xi\zeta}^2} \left[ \begin{array}{cc} \frac{1}{l^\xi} & \frac{1}{l^\xi} t_{\xi\zeta} \\ \frac{1}{l^\zeta} t_{\xi\zeta} & \frac{1}{l^\zeta} \end{array} \right] \hat{\mathbf{b}}. \quad (11.23)$$

We observe that this system of equations does not have a solution for  $\mathbf{t}^\xi \cdot \mathbf{t}^\zeta = 1$ , which means that both beams are parallel. This case has to be excluded here, and needs special treatment. On the other hand, the equation system (11.23) decouples for  $t_{\xi\zeta} = 0$ , which is equivalent to a perpendicular position of the beams.

Finally, we can insert this result together with (11.16) into (11.14), and obtain for the variations of the tangential gap

$$\left\{ \begin{array}{c} \delta g_T^\xi \\ \delta g_T^\zeta \end{array} \right\} = \frac{1}{1 - t_{\xi\zeta}^2} \left[ \begin{array}{cc} 1 & t_{\xi\zeta} \\ t_{\xi\zeta} & 1 \end{array} \right] \hat{\mathbf{b}} + \left\{ \begin{array}{c} [\xi(t) - \xi(t_0)] \mathbf{t}^\xi \cdot (\bar{\eta}_2^\xi - \bar{\eta}_1^\xi) \\ [\zeta(t) - \zeta(t_0)] \mathbf{t}^\zeta \cdot (\bar{\eta}_2^\zeta - \bar{\eta}_1^\zeta) \end{array} \right\}. \quad (11.24)$$

Note that relation (11.23) can also be used to compute the time derivatives  $\dot{\xi}$  and  $\dot{\zeta}$ . One only has to exchange the variations  $\bar{\eta}^\xi$  and  $\bar{\eta}^\zeta$  by the velocities  $\mathbf{v}^\xi$  and  $\mathbf{v}^\zeta$ , respectively.

### 11.3 CONTACT CONTRIBUTION TO WEAK FORM

Within the finite element analysis the weak form expression for the beam contact is needed. Since contact takes place at only one point, as described in the previous section, the contribution of contact to the general weak forms of a beam is given by

$$G_c^B = F_N \delta g_N + \mathbf{F}_T \cdot \delta \mathbf{g}_T + \sum_{j=1}^2 G_j^B, \quad (11.25)$$

where  $F_N$  is the normal contact force and  $\mathbf{F}_T$  is the force in tangential direction of the contacting beam axes. The weak contribution of the beams,  $G_j^B$ , is not considered here in detail; associated formulations can be found in Bathe and Bolourchi (1979),

Simo (1985), Simo and Vu-Quoc (1986), Crisfield (1990) or Gruttmann et al. (2000) for large deflections and rotations. The normal force  $F_N$  is either computed using the LAGRANGE multiplier method or approximated by the penalty approach. The variation of the gap  $g_N$  follows from (11.12). The force vector in the tangential direction of the beams has the two components,  $\mathbf{F}_T = F_T^\xi \mathbf{t}^\xi + F_T^\zeta \mathbf{t}^\zeta$ , with the unit base vectors defined in (11.16). Since the tangent force acts in reality on the surface of the beam, with this definition we have neglected the bending moments due to the eccentricity of the tangent force with respect to the beam axis. Such terms are considered in Litewka and Wriggers (2002b). The variation of the tangential gap  $\delta \mathbf{g}_T$  is given by (11.24), such that the scalar product in (11.25) can be written as

$$\mathbf{F}_T \cdot \delta \mathbf{g}_T = F_T^\xi \delta g_T^\xi + F_T^\zeta \delta g_T^\zeta. \quad (11.26)$$

The tangential force  $\mathbf{F}_T$  follows in the case of friction from a constitutive relation in the contact interface; see Section 4.2 for different possibilities. Here for simplicity we use COULOMB'S law. However, other models for friction can also be applied in the same way.

Let us assume that the normal force  $F_N$  is given by either a LAGRANGE multiplier method  $F_N = \lambda_N$  or by the penalty method  $F_N = \epsilon_N g_N$ . Using the relations derived in Section 9.5.2 for COULOMB friction, the following expressions are obtained for the frictional forces  $F_T^\xi$  and  $F_T^\zeta$ :

$$\text{stick: } F_T^{st\alpha} = \epsilon_T g_T^{\alpha\zeta} \quad \text{slip: } F_T^{sl\alpha} = \mu F_N \text{sign}(F_T^{\alpha tr}). \quad (11.27)$$

The superscript  $\alpha$  stands here either for  $\xi$  or for  $\zeta$ . The trial values  $F_T^{\alpha tr}$  for the tangential forces are computed from

$$F_T^{\alpha tr} = \epsilon_T g_T^{st\alpha} = \epsilon_T (g_T^{\alpha\zeta} - g_T^{sl\alpha}), \quad (11.28)$$

with the total relative tangential deformation  $g_T^{\alpha\zeta}$  at time  $t_{n+1}$  and the relative tangential slip  $g_T^{sl\alpha}$  at time  $t_n$ .

For the application of NEWTON'S method, the linearization of the weak form (11.25) is necessary. This leads to

$$\Delta G_c^B = \Delta F_N \delta g_N + F_N \Delta \delta g_N + \Delta \mathbf{F}_T \cdot \delta \mathbf{g}_T + \mathbf{F}_T \cdot \Delta \delta \mathbf{g}_T + \sum_{j=1}^2 \Delta G_j^B. \quad (11.29)$$

In this equation one has to linearize the force components and also the variation of normal and tangential gaps. The linearization of the normal force yields

$$\Delta F_N = \epsilon_N \Delta g_N \quad \text{with } \Delta g_N = (\Delta \mathbf{u}^\xi - \Delta \mathbf{u}^\zeta) \cdot \bar{\mathbf{n}}, \quad (11.30)$$

where the linearization of the normal gap is computed in the same way as the variation, see (11.12). For the tangential force one has to distinguish between the stick and slip phases. In the stick phase the linearization leads to

$$\Delta F_T^{st\alpha} = \epsilon_T \Delta g_T^{st\alpha}. \quad (11.31)$$

Again, the linearization of the tangential gap has to be computed in the same way as for its variation, see (11.24). In the slip phase, (11.27) has to be linearized

$$\begin{aligned}\Delta F_T^{sl\alpha} &= \mu \Delta F_N \text{sign}(F_T^{\alpha tr}) + \mu F_N \Delta \text{sign}(F_T^{\alpha tr}) \\ &= \mu \epsilon_N \Delta g_N \text{sign}(F_T^{\alpha tr}).\end{aligned}\quad (11.32)$$

Note that the linearization of the second term is zero.

The linearization of the variation of normal,  $\Delta \delta g_N$ , and tangential gap,  $\Delta \delta g_T$ , is straightforward, though quite involved, and hence leads to complex expressions which will not be reproduced here. The associated formulas can be found for the normal gap in Wriggers and Zavarise (1997), and for the tangential gap in Zavarise and Wriggers (2000).

## 11.4 FINITE ELEMENT FORMULATION

The weak form now has to be derived in matrix form for discretization of the contact contribution. Assuming straight beam elements which are formulated using the TIMOSHENKO beam theory, one can simply use linear shape functions for all variables. The simplest approximation, adopted here, is the assumption of a linear function of the current coordinates in the deformed beam element. This leads to

$$\mathbf{x}^\alpha = \mathbf{X}^\alpha + \mathbf{u}^\alpha = \sum_{I=1}^2 N_I(\alpha) \mathbf{x}^\alpha \quad \text{with } N_I(\alpha) = \frac{1}{2} (1 + \alpha_I \alpha), \quad (11.33)$$

where again,  $\alpha$  stands for  $\xi$  or  $\zeta$ . The same ansatz is made for the variations  $\eta^\xi$  and  $\eta^\zeta$ .

Now the matrix form of the variation of the normal and tangential gaps has to be derived to discretize (11.25). From (11.12), with  $\eta^\xi = \frac{1}{2} (1 - \xi) \eta_1^\xi + \frac{1}{2} (1 + \xi) \eta_2^\xi$  and  $\eta^\zeta = \frac{1}{2} (1 - \zeta) \eta_1^\zeta + \frac{1}{2} (1 + \zeta) \eta_2^\zeta$ , we obtain the contribution of the normal force  $F_N$  in  $n_c$  contact points to the weak form:

$$\boxed{G_N = \sum_{c=1}^{n_c} \eta_s^T \mathbf{G}_{Nc} F_{Nc}}, \quad (11.34)$$

with

$$\eta_c^T = \langle \eta_1^\xi, \eta_2^\xi, \eta_1^\zeta, \eta_2^\zeta \rangle \quad \text{and} \quad \mathbf{G}_{Nc} = \left\{ \begin{array}{l} \frac{1}{2} (1 - \bar{\xi}) \bar{\mathbf{n}} \\ \frac{1}{2} (1 + \bar{\xi}) \bar{\mathbf{n}} \\ \frac{1}{2} (1 - \bar{\zeta}) \bar{\mathbf{n}} \\ \frac{1}{2} (1 + \bar{\zeta}) \bar{\mathbf{n}} \end{array} \right\}, \quad (11.35)$$

where  $\bar{\xi}$  and  $\bar{\zeta}$  are the convected coordinates which define the contact point  $c$  in the beams.

For the contribution of the tangential forces to the weak form, the matrices follow from (11.24) and (11.26). By introducing

$$\mathbf{T}_c = \frac{1}{1 - t_{\xi\zeta}^2} \begin{bmatrix} 1 & t_{\xi\zeta} \\ t_{\xi\zeta} & 1 \end{bmatrix}, \quad (11.36)$$

$$\mathbf{A}_c = \begin{bmatrix} \frac{1}{2}(1 - \bar{\xi}) \mathbf{t}^\xi - \frac{g_N}{l^\xi} \mathbf{n} & -\frac{1}{2}(1 - \bar{\xi}) \mathbf{t}^\zeta \\ \frac{1}{2}(1 + \bar{\xi}) \mathbf{t}^\xi + \frac{g_N}{l^\xi} \mathbf{n} & -\frac{1}{2}(1 + \bar{\xi}) \mathbf{t}^\zeta \\ -\frac{1}{2}(1 - \bar{\zeta}) \mathbf{t}^\xi & \frac{1}{2}(1 - \bar{\zeta}) \mathbf{t}^\zeta + \frac{g_N}{l^\zeta} \mathbf{n} \\ -\frac{1}{2}(1 + \bar{\zeta}) \mathbf{t}^\xi & \frac{1}{2}(1 + \bar{\zeta}) \mathbf{t}^\zeta - \frac{g_N}{l^\zeta} \mathbf{n} \end{bmatrix} \quad (11.37)$$

and

$$\mathbf{B}_c = \begin{bmatrix} -(\bar{\xi} - \xi_0) \mathbf{t}^\xi & 0 \\ (\bar{\xi} - \xi_0) \mathbf{t}^\xi & 0 \\ 0 & -(\bar{\zeta} - \zeta_0) \mathbf{t}^\zeta \\ 0 & (\bar{\zeta} - \zeta_0) \mathbf{t}^\zeta \end{bmatrix}, \quad (11.38)$$

for the pair of contact points,  $c$ , the weak form for the tangential contact force can be stated as

$$G_T = \sum_{c=1}^{n_c} \boldsymbol{\eta}_c^T \mathbf{G}_{Tc} \mathbf{F}_{Tc} \quad \text{with} \quad \mathbf{G}_{Tc} = \mathbf{A}_c \mathbf{T}_c + \mathbf{B}_c. \quad (11.39)$$

The tangential contact force is here defined as the vector  $\mathbf{F}_T^T = \langle F_T^\xi, F_T^\zeta \rangle$ , and  $\boldsymbol{\eta}_c$  is the variation of the nodal point values of both beams, as defined in (11.35).

The matrix form of the linearization of the fully nonlinear case will not be reported here. It can be found for the frictionless case in Wriggers and Zavarise (1997), and for the case with friction in Zavarise and Wriggers (2000), for a beam with circular cross-sections. The complete matrix formulation for rectangular cross-sections is provided in Litewka and Wriggers (2002a) and Litewka and Wriggers (2002b) for frictionless and frictional contact, respectively.

In the case of a small deformation response, one can disregard the contribution of the linearization of the gap function  $\Delta\delta g_N$  and  $\Delta\delta \mathbf{g}_T$  and the dependency of  $\mathbf{A}_s$  on the normal gap  $g_N$ . In this case, the tangent matrix for normal contact can be stated for the penalty method with (11.30) as follows:

$$\mathbf{K}_{Tc}^N = \epsilon_N \mathbf{G}_{Nc} \mathbf{G}_{Nc}^T. \quad (11.40)$$

For the tangential contributions one has to distinguish the stick phase, which yields a symmetric matrix, and the slip phase, resulting in a non-symmetric matrix. The tangent matrix for stick follows from (11.31):

$$\mathbf{K}_{Tc}^{stick} = \epsilon_T \mathbf{G}_{Tc} \mathbf{G}_{Tc}^T \quad (11.41)$$

where in  $\mathbf{G}_{Tc}$  the gap distance  $g_N$  has to be set to zero. In the slip case one has to start from (11.32), which is stated for only one component of the friction force. Furthermore, we have to distinguish between sliding in the  $\xi$  direction and stick in  $\zeta$ , sliding in the  $\zeta$  direction and stick in  $\xi$ , or sliding in both directions. The latter case leads to the non-symmetric tangent matrix

$$\mathbf{K}_{Tc}^{slip} = \mu \epsilon_N \mathbf{G}_{Tc} \mathbf{S}_c \mathbf{G}_{Nc}^T, \quad (11.42)$$

with the diagonal matrix

$$\mathbf{S}_c = \begin{bmatrix} \text{sign}(F_{T_{n+1}}^{\xi tr}) & 0 \\ 0 & \text{sign}(F_{T_{n+1}}^{\zeta tr}) \end{bmatrix}. \quad (11.43)$$

## 11.5 CONTACT SEARCH FOR BEAMS

The search for contact in the beam is different from the contact search routines used for two- and three-dimensional solid problems. This stems from the fact that the typical distinction in slave and master elements is not needed. The search is based on the following assumptions:

- contact between beams is pointwise (this excludes beams in parallel contact),
- cross-sections of the contacting beams do not deform, and
- contact occurs along the edges of beams; for special cases see Remark 11.1.

The first step of contact search is to find the pair of finite beam elements which is closest to each other. This step is carried out by considering the mid points  $\mathbf{x}_1^{IC}$  and  $\mathbf{x}_2^{JC}$  of the beams (see Figure 11.6) in the current configuration. The mid points are defined by

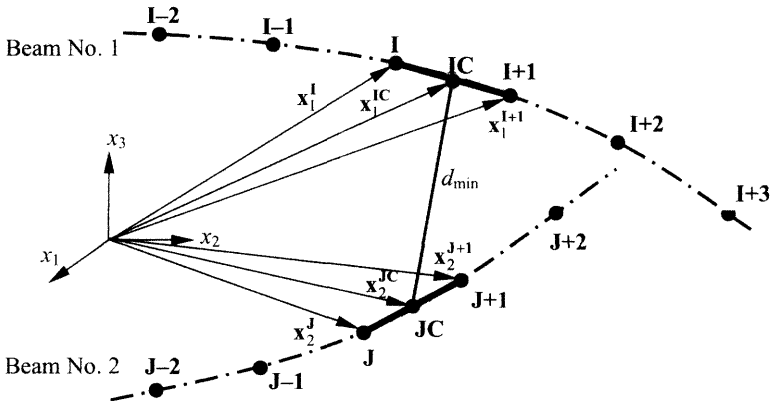
$$\mathbf{x}_1^{IC} = \frac{1}{2} (\mathbf{x}_1^I + \mathbf{x}_1^{I+1}), \quad \mathbf{x}_2^{JC} = \frac{1}{2} (\mathbf{x}_2^J + \mathbf{x}_2^{J+1}). \quad (11.44)$$

Now the pair of beams which fulfils  $d_{min} = \min_{I,J} \|\mathbf{x}_1^{IC} - \mathbf{x}_2^{JC}\|$  is selected as possible contact pair. This preliminary search yields a rough location of the contact point. However, it does not necessarily mean that the actual contact point will be associated with the determined elements.

A complete search which takes into account the local conditions (11.2) is only performed if

$$d_{min} \leq \frac{1}{2} (L^\zeta + L^\xi), \quad (11.45)$$

where  $L^\zeta = \min(l^\zeta, d^\zeta)$  and  $L^\xi = \min(l^\xi, d^\xi)$  are the minimum of beam lengths  $IC$  and  $JC$ , see definition (11.15), and the maximal diagonal of the beam cross-sections:  $d^\zeta$  and  $d^\xi$ . This condition means that an imaginary sphere has been drawn around the mid point of the elements, and only if the two spheres of the beams intersect then a



**Fig. 11.6** First stage of contact search – a pair of closest elements.

local search makes sense. Such a search strategy is similar to the pin-ball algorithm for solids, discussed in Belytschko and Neal (1991).

If (11.45) is fulfilled, then one has to compute the minimum distance in the next step using the local conditions (11.2). In case (11.2) yields a pair of coordinates  $(\bar{\zeta}, \bar{\xi})$  with  $-1 \leq \bar{\zeta} \leq 1$  and  $-1 \leq \bar{\xi} \leq 1$ , see definition (11.33), then the correct element pair in which contact takes place has been found. Otherwise, the local condition (11.2) has to be investigated for the neighbouring elements.

**REMARK 11.2:** For large incremental steps, one has to take care that the beams do not cross each other without contact having been detected. To prevent such cases, the following procedure can be used in which two distance vectors are introduced:

$$\mathbf{d}_n = \mathbf{x}_{1n}^{IC} - \mathbf{x}_{2n}^{JC} \quad \text{and} \quad \mathbf{d}_{n+1} = \mathbf{x}_{1n+1}^{IC} - \mathbf{x}_{2n+1}^{JC}, \quad (11.46)$$

where  $n+1$  denotes the current and  $n$  the previous time step in the incremental loading process. One can now distinguish the following cases:

$$\begin{aligned} \mathbf{d}_n \cdot \mathbf{d}_{n+1} &> 0 \implies \text{axes did not cross} \\ \mathbf{d}_n \cdot \mathbf{d}_{n+1} &< 0 \implies \text{axes have crossed} \\ \mathbf{d}_n \cdot \mathbf{d}_{n+1} &= 0 \implies 90^\circ \text{ rotation.} \end{aligned}$$

This procedure works for incremental steps in which the beams do not rotate more than 90 degrees. However, the latter case is quite unusual, and if it occurs a reduction of the incremental step has to be used.

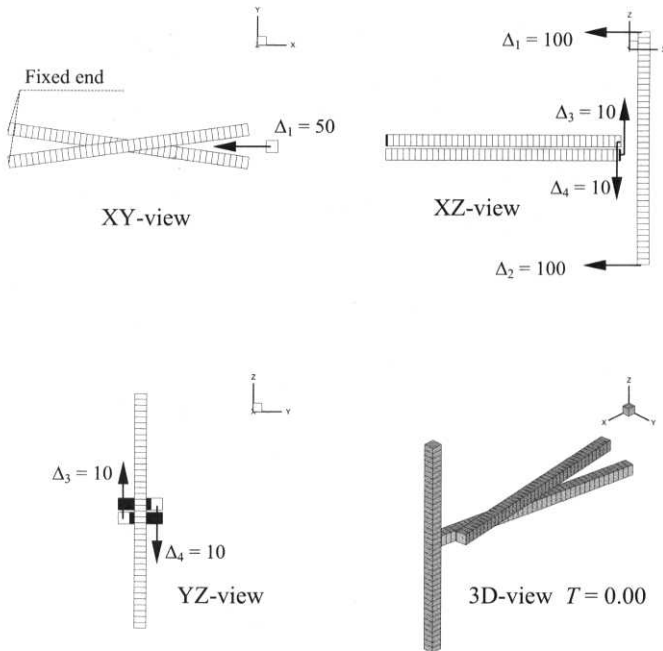


## 11.6 EXAMPLES

Two examples are discussed which also include a comparison with three-dimensional discretizations of beam contact. In the first example, frictionless contact of three beams is considered, whereas the second example shows the behaviour of beams in frictional contact. These examples are also discussed in Litewka and Wriggers (2002a) and Litewka and Wriggers (2002b), and involve the contact of beams with rectangular cross-sections.

### 11.6.1 Three beams in frictionless contact

We consider three beams which come into contact during the deformation process. The beams have an initial configuration as shown in Figure 11.7. One beam is fixed at its upper and lower end, and is moved by a prescribed displacement in a negative  $X$ -direction between the other two crossed beams (crossing angle  $11.5^\circ$ ) which are clamped at their ends. Furthermore, the tips of the crossed beams undergo prescribed displacements in the  $Z$ -direction, as shown in Figure 11.7. The beam elements are described in Litewka and Wriggers (2002a), and the references therein.

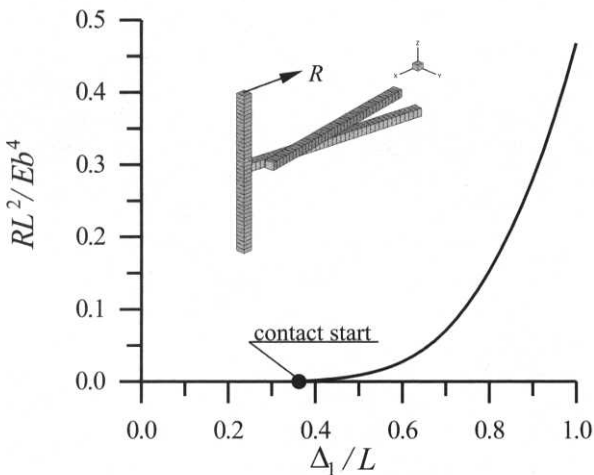


**Fig. 11.7** Initial configuration of the contacting beams.

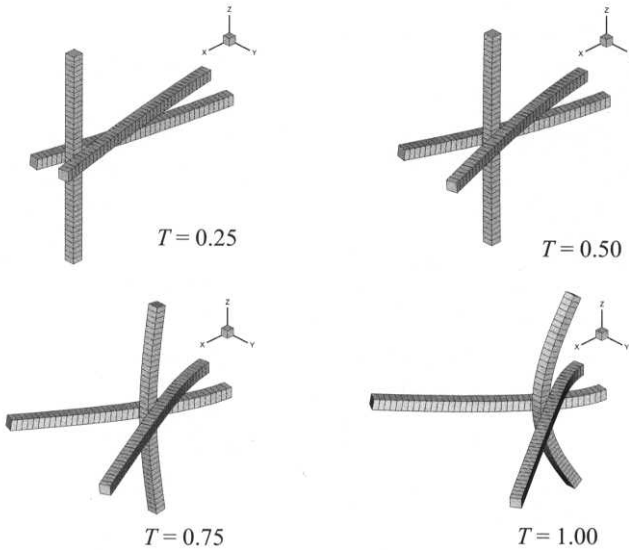
The beams are discretized using 10 beam elements each. All beams have the same data concerning the constitutive behaviour which are stated in the following. Modulus of elasticity:  $E = 2 \times 10^4$ , POISSON ratio:  $\nu = 0.3$ , beam length:  $L = 100$ , and width of the quadratic cross section  $b = 5$ . The initial gap between the crossed beams is  $g_{N0} = 1$ . The penalty method is applied with a penalty parameter of  $\varepsilon = 2.4 \cdot 10^3$  between the crossed beams and  $\varepsilon = 1.25 \cdot 10^4$  between the “free” beam and the other ones. The choice of the penalty parameters was dictated by the requirement that the penetration due to the penalty method should be kept under 1% of the size of the beam cross-section. The prescribed displacement of  $\Delta_1 = \Delta_2 = 100$  for both ends of the “free” beam and  $\Delta_3 = \Delta_4 = 10$  for the tip deflection of the crossed beams is applied within 80 loading steps of  $\Delta T = 0.0125$ . Contact between “free” beam and crossed ones starts at  $T = 0.3625$ , which is marked by the dot in the load-deflection curve shown in Figure 11.8. An increase of the reaction force associated with the upper end of the “free” beam occurs at the end of the process. This is reflected in the deformation of the beams; see Figure 11.9, which depicts the deformation process by showing four stages of the deformation process. One can observe that the beams undergo large displacements and rotations. Within all increments, quadratic convergence is achieved leading to an average of three iterations per step.

### 11.6.2 Two beams in contact with friction

Two beams as shown in Figure 11.10 are considered to compare the frictional response computed with beams to the same analysis based on a three-dimensional contact



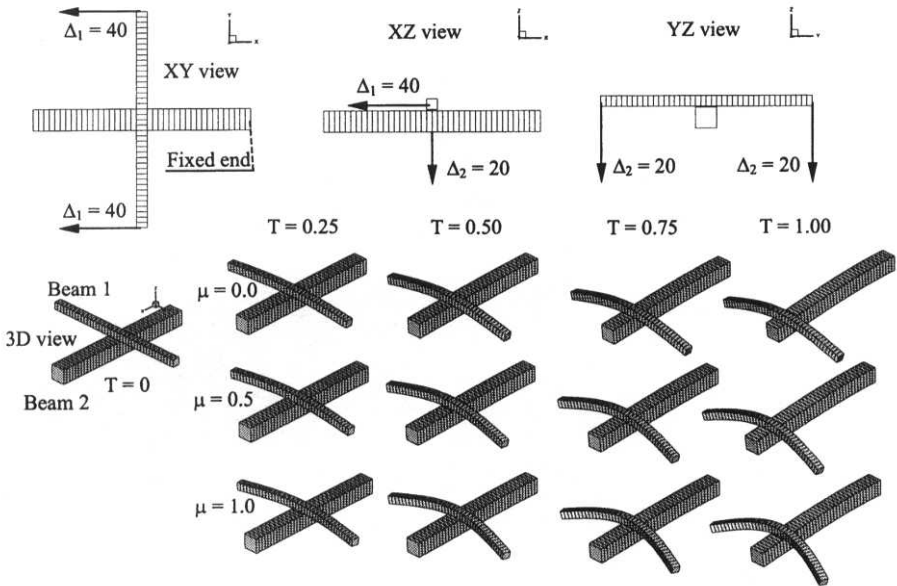
**Fig. 11.8** Load deflection curve for frictionless beam contact.



**Fig. 11.9** Deformed configuration at four stages of the deformation process.

analysis using solid elements. Beam 2 is a cantilever beam which is clamped at the right end. Beam 1 has constrained rotations about the  $Y$ -axis at its ends, as well as the constraints resulting from the symmetry with respect to plane  $X - Z$ . The following data are used: dimensions of Beam 1:  $(5 \times 5 \times 100)$ , dimensions of Beam 2:  $(10 \times 10 \times 100)$ , YOUNG'S moduli:  $E_1 = E_2 = 3 \cdot 10^4$ , POISSON'S ratios:  $\nu_1 = \nu_2 = 0.17$ , initial gap:  $g_{N0} = 0.5$ , penalty parameters:  $\varepsilon_N = 2.5 \cdot 10^4$  and  $\varepsilon_T = 2.5 \cdot 10^2$ . Both beams are discretized using 10 elements. The vertical and horizontal displacements of both ends of Beam 1 are applied in 50 increments using  $\Delta t = 0.02$ . The analysis is performed for the frictionless case and two frictional cases with a COULOMB friction coefficient of  $\mu = 0.5$  and  $\mu = 1.0$ .

Table 11.1 presents the convergence rates and friction status for four selected stages of the process. These stages are also depicted in Figure 11.10. The solution shows in a clear way the influence of friction, which is reflected in the deformation of Beam 1. This beam sticks at first and then starts to slide along Beam 2. All cases are characterized by very good convergence rate with 3 – 5 iterations per step, which is an outcome of the application of NEWTON'S method with a consistent linearization of the residual term in Section 11.4. The results were compared to a full three-dimensional contact analysis using solid elements. To this end, the program ABAQUS was used. The 8-node brick elements C3D8 were chosen to discretize beam 2 with a  $2 \times 2 \times 20$  element mesh. For Beam 1 the number of elements was doubled in the  $X$  direction, leading to a discretization with  $2 \times 2 \times 40$  brick elements.

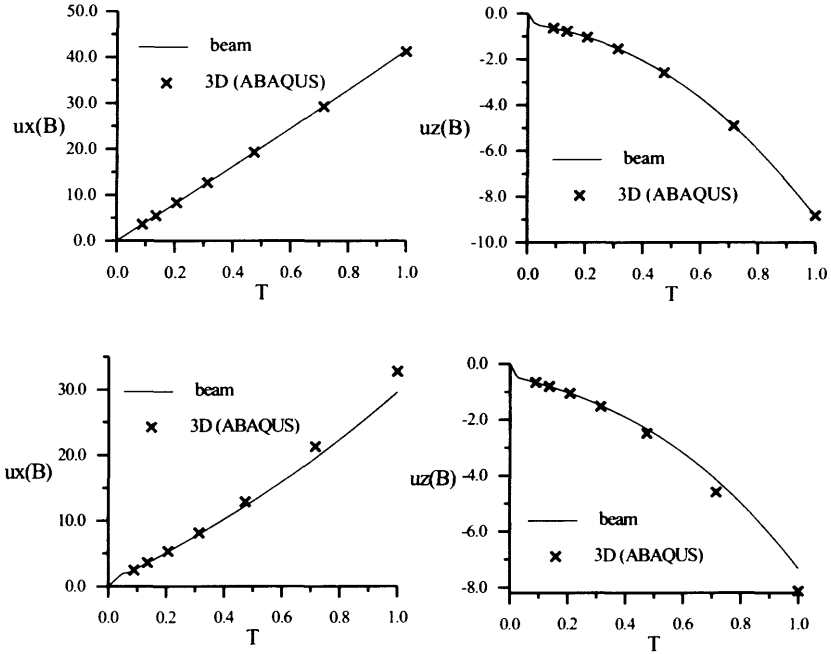


**Fig. 11.10** Initial and deformed configuration at four stages of the deformation process for different values of the friction coefficient.

**Table 11.1** Convergence study and friction status

$T$		0.25	0.5	0.75	1.00
$\mu = 0.0$	Status of beam 2	slip	slip	slip	slip
	Status of beam 1	slip	slip	slip	slip
	Number of iterations	3	3	3	3
$\mu = 0.5$	Status of beam 2	slip	slip	slip	slip
	Status of beam 1	stick	stick	stick	stick
	Number of iterations	3	3	3	3
$\mu = 1.0$	Status of beam 2	stick	slip	slip	slip
	Status of beam 1	stick	stick	stick	stick
	Number of iterations	5	5	3	3

The lower surface of Beam 1, which has the finer mesh, was taken as the slave surface. Comparison of the displacement components  $u_x$  and  $u_z$  at the centre point  $B$  of Beam 1 is depicted in Figure 11.11 for the frictionless case and for the case with frictional coefficient  $\mu = 1.0$ . The comparison shows excellent results for the frictionless case where the difference between both models does not exceed 0.6% for



**Fig. 11.11** Comparison of beam to three-dimensional analysis for  $\mu = 0$  and  $\mu = 1.0$ .

the calculated displacements. The results for friction depict discrepancies. However, in our opinion a maximum of 10% difference in the case of large friction ( $\mu = 1.0$ ) can be considered as evidence of very good performance of the beam model when compared with full three-dimensional analysis. Note the change in the character of the behaviour of displacement component  $u_x$ . For the frictionless case it is practically linear, while for the increasing friction its nonlinearity becomes more pronounced as the sliding is increasingly limited by increasing frictional force.

# 12

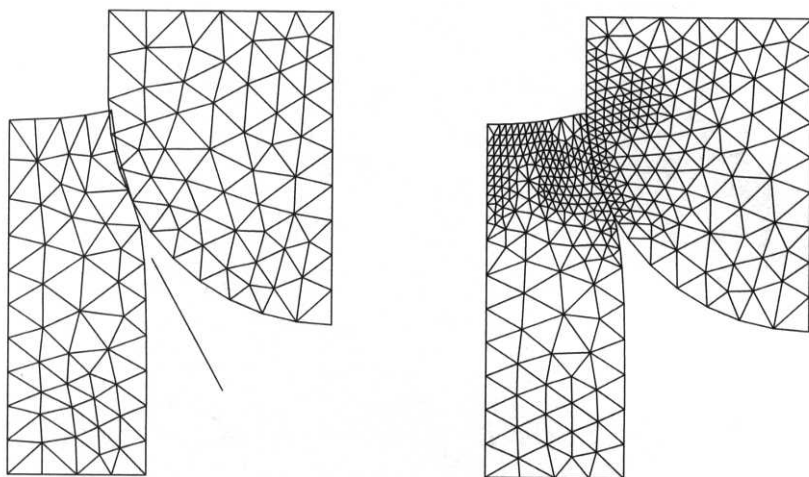
---

## *Adaptive Finite Element Methods for Contact Problems*

In this chapter we focus on a relatively new method in the area of finite element techniques which ensures a successive improvement of the numerical solution via an adaptive mesh refinement. The main idea is depicted in Figure 12.1. It shows the basic ingredients of an adaptive method for contact problems. Part 1 is the determination of the contact, part 2 a solution of the problem with a given mesh, and finally, the adaptive process in part 3 leads to an automatic mesh refinement of the discretization. Thus the objective of adaptive techniques is to obtain a mesh which is optimal, in the sense that the computational costs involved are minimal under the constraint that the error in the finite element solution is beyond a certain limit. Since the computational effort can somehow be linked to the number of unknowns of the finite element mesh, the task is to find a mesh with a minimum number of unknowns or nodes for a given error tolerance.

The adaptive finite element method will be formulated here first for frictionless contact problems in linear elasticity. For this class of problem it is still possible to develop mathematically sound error estimators. Since standard contact applications often include friction, we will also derive an error indicator for frictional contact, and discuss its extension in the case of large deformations. Error measures for thermo-mechanical contact problems have not been established so far, and hence will not be discussed here.

Basically, one has two different possibilities to derive error estimators which can be applied within adaptive methods to refine the finite element mesh. These are the residual-based error estimators and projection or defect correction methods, which rely on super convergence properties. Both techniques will be discussed first for geometrically linear problems, and then in the presence of large deformations.



**Fig. 12.1** Adaptive process for contact problems.

Let  $\mathbf{u}$  denote the exact solution and let  $\mathbf{u}_h$  be the discrete finite element solution. Now we can define the error in the displacement field by

$$\mathbf{e}_u = \mathbf{u} - \mathbf{u}_h. \quad (12.1)$$

In the same way, the error in the stress field can be defined as

$$\mathbf{e}_\sigma = \boldsymbol{\sigma} - \boldsymbol{\sigma}_h. \quad (12.2)$$

During the last ten years, research activities have been focused on adaptive techniques that provide a discretization which is accurate and reliable. Adaptive techniques rely on indicators and/or estimators which are able to predict the error given in (12.1) or (12.2). These quantities display the error distribution of the finite element solution (e.g. see Johnson (1987) and references therein). Based on the error distribution, a new refined mesh can then be constructed which yields a better approximate solution.

The methods rely on error estimators which have been developed so far in different versions. Estimators which are most frequently used for elastic problems in solid mechanics are residual-based error estimators, e.g. see Babuska and Rheinboldt (1978) or Johnson and Hansbo (1992), or error estimators which use superconvergence properties, e.g. see Zienkiewicz and Zhu (1987).

For frictionless contact problems, *a priori* error estimators have been derived for linear elastic bodies, e.g. see Kikuchi and Oden (1988) or Hlavacek et al. (1988). An adaptive method for problems with unilateral constraints has been developed by Lee et al. (1991) who treated as an example a free surface flow problem. In Wriggers et al. (1994), a residual-based error estimator has been developed following an approach

pursued by Johnson and Hansbo (1992) for unilateral membrane problems. However, as shown in Wriggers and Scherf (1998), the  $Z^2$  error indicators, due to Zienkiewicz and Zhu (1991), can also be applied to contact problems. The advantages and disadvantages of residual-based error estimators and  $Z^2$  indicators have been discussed in Wriggers and Scherf (1998) for frictionless contact. Due to this discussion, we give here an overview including residual-based error estimators, projection based error indicators and error indicators based on dual methods. The residual-based error measures will be also employed as indicators for contact with friction.

## 12.1 CONTACT PROBLEM AND DISCRETIZATION

Assume that two bodies come into contact. In that case, the non-penetration condition is given by (3.6) as  $g_N \geq 0$ . Let us recall the variational inequality (5.10)

$$a(\mathbf{u}, \mathbf{v} - \mathbf{u}) \geq f(\mathbf{v} - \mathbf{u}),$$

which describes the frictionless contact problem in the case of small deformations. The problem  $(\mathbf{P})$  is now to find  $\mathbf{u} \in \mathbf{K}$  such that (5.10) is fulfilled for all  $\mathbf{v} \in \mathbf{K}$  with

$$\mathbf{K} = \{\mathbf{v} \in \mathbf{V} \mid (\mathbf{v}^2 - \bar{\mathbf{v}}^1) \cdot \bar{\mathbf{n}}^1 + G_N \geq 0 \text{ on } \Gamma_c\}$$

where  $\mathbf{V}$  is the space of variations or test functions. Due to the inequality constraint on the displacement field, this problem is nonlinear.

Different solution techniques can be applied to solve this problem; we just mention here the LAGRANGE multiplier approach, the penalty method or augmented LAGRANGE techniques. For the rest of the section we focus on the penalty method for the solution of  $(\mathbf{P})$ ; its mathematical background is described in Luenberger (1984). This technique replaces  $(\mathbf{P})$  by an unconstrained problem  $(\mathbf{P}_\varepsilon)$  with regard to the contact constraint (3.1) as follows: Find  $\mathbf{u}_\varepsilon \in \mathbf{V}$  such that

$$a(\mathbf{u}_\varepsilon, \mathbf{v}) + c^-(\mathbf{u}_\varepsilon, \mathbf{v}) = f(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{V}, \quad (12.3)$$

where  $\mathbf{V}$ ,  $a(\mathbf{u}_\varepsilon, \mathbf{v})$  and  $f(\mathbf{v})$  are defined as above, and

$$c^-(\mathbf{u}_\varepsilon, \mathbf{v}) = \int_{\Gamma_c} \varepsilon u_{\varepsilon N}^- v_N d\Gamma. \quad (12.4)$$

$u_{\varepsilon N}^-$  has already been defined in (3.12) and  $v_N = \mathbf{v} \cdot \bar{\mathbf{n}}^1$ . The penalty parameter  $\varepsilon$  is a positive constant. It can be shown, see Kikuchi and Oden (1988), that the solution of  $(\mathbf{P}_\varepsilon)$  will converge to the solution of  $(\mathbf{P})$  as  $\varepsilon$  tends to infinity.

To discretize  $(\mathbf{P}_\varepsilon)$ , we divide  $\Omega$  into non-overlapping finite elements  $T$  of diameter  $h_T$ , and introduce a standard finite element space

$$\mathbf{V}_h = \{\mathbf{v} \in \mathbf{V} \mid \mathbf{v} \in C(\Omega), \mathbf{v}|_T \in [P(T)]^2, \forall T\}, \quad (12.5)$$

where  $P(T)$  is a space of polynomials of degree  $p_T$  on  $T$ , and  $p_T$  is a positive integer. The discrete finite element problem  $(\mathbf{P}_h)$  for  $(\mathbf{P}_\varepsilon)$  is now: Find  $\mathbf{u}_h \in \mathbf{V}_h$  such that

$$a(\mathbf{u}_h, \mathbf{v}) + c^-(\mathbf{u}_h, \mathbf{v}) = f(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{V}_h. \quad (12.6)$$



## 12.2 RESIDUAL BASED ERROR ESTIMATOR FOR FRICTIONLESS CONTACT

Error estimators have been derived for frictionless contact problems of linear elastic bodies, e.g. see Kikuchi and Oden (1988) or Hlavacek et al. (1988). In this section, we follow the approach which was developed in Johnson (1991) for the unilateral membrane problem, and extend it to the case of elastic bodies, see Carstensen et al. (1999). For this purpose a penalty regularization is used to approximate the constraint problem by an unconstrained one. Thereafter, an error estimator is formulated which is an extension of the error estimator in Babuska and Miller (1987).

Let  $\mathbf{u}_\varepsilon$  and  $\mathbf{u}_h$  denote the exact penalty solution of  $(\mathbf{P}_\varepsilon)$  and the discrete finite element solution of  $(\mathbf{P}_h)$ , respectively. With

$$\mathbf{e} = \mathbf{u}_\varepsilon - \mathbf{u}_h \quad (12.7)$$

we define an error measure as follows:

$$r = \alpha \|\nabla \mathbf{e}\|_{L^2(\Omega)}^2 + c^-(\mathbf{e}, \mathbf{e}), \quad \alpha > 0. \quad (12.8)$$

Due to the ellipticity of  $a(\mathbf{e}, \mathbf{e})$ , e.g. see Johnson and Hansbo (1992), we can estimate

$$r \leq a(\mathbf{e}, \mathbf{e}) + c^-(\mathbf{e}, \mathbf{e}). \quad (12.9)$$

By subtracting (12.6) from (12.3), we derive the condition

$$a(\mathbf{e}, \mathbf{v}) + c^-(\mathbf{e}, \mathbf{v}) = 0 \quad \forall \mathbf{v} \in \mathbf{V}_h. \quad (12.10)$$

Adding (12.10) with  $\mathbf{v} = \mathbf{u}_h - \mathbf{w}$  to the right side of (12.9), we obtain

$$r \leq a(\mathbf{e}, \mathbf{u}_\varepsilon - \mathbf{w}) + c^-(\mathbf{e}, \mathbf{u}_\varepsilon - \mathbf{w}). \quad (12.11)$$

Inserting (12.7) and rearranging terms provides

$$r \leq a(\mathbf{u}_\varepsilon, \mathbf{u}_\varepsilon - \mathbf{w}) + c^-(\mathbf{u}_\varepsilon, \mathbf{u}_\varepsilon - \mathbf{w}) - a(\mathbf{u}_h, \mathbf{u}_\varepsilon - \mathbf{w}) - c^-(\mathbf{u}_h, \mathbf{u}_\varepsilon - \mathbf{w}). \quad (12.12)$$

The first two terms on the right side of inequality (12.12) are, according to (12.3), equal to  $f(\mathbf{u}_\varepsilon - \mathbf{w})$ .

Next we choose  $\mathbf{w} = I_h \mathbf{u}_\varepsilon$ , where  $I_h$  is a projection onto  $\mathbf{V}_h$  (the nodal interpolation if  $\mathbf{u}_\varepsilon$  is smooth), and find that  $\mathbf{u}_\varepsilon - \mathbf{w} = \mathbf{e} - I_h \mathbf{e}$ . Elementwise integration by parts of the third term under consideration of the constitutive equation then leads to

$$\begin{aligned} r \leq & \sum_T \int_T \hat{\mathbf{b}} \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Omega + \sum_{\partial T \subseteq \Gamma_\sigma} \int_{\partial T} \hat{\mathbf{t}} \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma \\ & + \sum_T \int_T \operatorname{div} \boldsymbol{\sigma}_h \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Omega + \sum_{\partial T} \int_{\partial T} (-\boldsymbol{\sigma}_h \mathbf{n}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma \\ & - \sum_{\partial T \subseteq \Gamma_\varepsilon} \int_{\partial T} \varepsilon \mathbf{u}_{N_h}^-(\mathbf{e} - I_h \mathbf{e})_N \, d\Gamma. \end{aligned} \quad (12.13)$$

The last part denotes the contribution due to the penalization of the contact constraint equation. Here we have defined the normal projection  $(\mathbf{e} - I_h \mathbf{e})_N = (\mathbf{e} - I_h \mathbf{e}) \cdot \bar{\mathbf{n}}^1$ .

We now introduce  $\Gamma_i$  as the set of faces of elements which are not contained in  $\Gamma$ . Thus we can write

$$\begin{aligned} r \leq & \sum_T \int_T (\operatorname{div} \boldsymbol{\sigma}_h + \hat{\mathbf{b}}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Omega \\ & + \sum_{\partial T \subseteq \Gamma_i} \int_{\partial T} \frac{1}{2} [-\boldsymbol{\sigma}_h \mathbf{n}] \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma + \sum_{\partial T \subseteq \Gamma_\sigma} \int_{\partial T} (\hat{\mathbf{t}} - \boldsymbol{\sigma}_h \mathbf{n}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma \\ & + \sum_{\partial T \subseteq \Gamma_c} \int_{\partial T} (-\boldsymbol{\sigma}_h \mathbf{n}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma - \sum_{\partial T \subseteq \Gamma_c} \int_{\partial T} \varepsilon u_{hn}^- (\mathbf{e} - I_h \mathbf{e})_N \, d\Gamma. \quad (12.14) \end{aligned}$$

The square brackets denote the jump of a function over the element interfaces. Using the symmetry of the inner product  $(\mathbf{e} - I_h \mathbf{e})_N$ , (12.14) can be rewritten as

$$\begin{aligned} r \leq & \sum_T \int_T (\operatorname{div} \boldsymbol{\sigma}_h + \hat{\mathbf{b}}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Omega \\ & + \sum_{\partial T \subseteq \Gamma_i} \int_{\partial T} \frac{1}{2} [-\boldsymbol{\sigma}_h \mathbf{n}] \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma + \sum_{\partial T \subseteq \Gamma_\sigma} \int_{\partial T} (\hat{\mathbf{t}} - \boldsymbol{\sigma}_h \mathbf{n}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma \\ & + \sum_{\partial T \subseteq \Gamma_c} \int_{\partial T} (-\varepsilon u_{hn}^- \bar{\mathbf{n}}^1 - \boldsymbol{\sigma}_h \mathbf{n}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma. \quad (12.15) \end{aligned}$$

The right side of (12.15) can be estimated with the absolute value, and therefore

$$\begin{aligned} r \leq & \sum_T \left| \int_T (\operatorname{div} \boldsymbol{\sigma}_h + \hat{\mathbf{b}}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Omega \right| \\ & + \sum_{\partial T \subseteq \Gamma_i} \left| \int_{\partial T} \frac{1}{2} [\boldsymbol{\sigma}_h \mathbf{n}] \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma \right| + \sum_{\partial T \subseteq \Gamma_\sigma} \left| \int_{\partial T} (\hat{\mathbf{t}} - \boldsymbol{\sigma}_h \mathbf{n}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma \right| \\ & + \sum_{\partial T \subseteq \Gamma_c} \left| \int_{\partial T} (\varepsilon u_{hn}^- \bar{\mathbf{n}}^1 + \boldsymbol{\sigma}_h \mathbf{n}) \cdot (\mathbf{e} - I_h \mathbf{e}) \, d\Gamma \right|. \quad (12.16) \end{aligned}$$

Using the CAUCHY-SCHWARZ inequality and an estimate for the interpolation error  $(\mathbf{e} - I_h \mathbf{e})$ , we obtain

$$\begin{aligned} r \leq & C \left\{ \sum_T h_T^2 \|\operatorname{div} \boldsymbol{\sigma}_h + \hat{\mathbf{b}}\|_{L^2(T)}^2 \right. \\ & + \sum_{\partial T \subseteq \Gamma_i} h_T \|\frac{1}{2} [\boldsymbol{\sigma}_h \mathbf{n}]\|_{L^2(\partial T)}^2 + \sum_{\partial T \subseteq \Gamma_\sigma} h_T \|\hat{\mathbf{t}} - \boldsymbol{\sigma}_h \mathbf{n}\|_{L^2(\partial T)}^2 \\ & \left. + \sum_{\partial T \subseteq \Gamma_c} h_T \|\varepsilon u_{hn}^- \bar{\mathbf{n}}^1 + \boldsymbol{\sigma}_h \mathbf{n}\|_{L^2(\partial T)}^2 \right\} \end{aligned}$$

$$+ \sum_{\partial T \subseteq \Gamma_c} h_T \|\varepsilon u_{hn}^- \bar{\mathbf{n}}^1 + \boldsymbol{\sigma}_h \mathbf{n}\|_{L_2(\partial T)}^2 \Big\}^{\frac{1}{2}} \|\nabla \mathbf{e}\|_{L_2(\Omega)}, \quad (12.17)$$

where  $C$  is a positive constant. Division of (12.17) by  $\|\nabla \mathbf{e}\|_{L_2(\Omega)}$  provides the error estimator

$$\begin{aligned} \alpha \|\nabla \mathbf{e}\|_{L_2(\Omega)} + \frac{c^-(\mathbf{e}, \mathbf{e})}{\|\nabla \mathbf{e}\|_{L_2(\Omega)}} &\leq C \left\{ \sum_T h_T^2 \|\operatorname{div} \boldsymbol{\sigma}_h + \hat{\mathbf{b}}\|_{L_2(T)}^2 \right. \\ &+ \sum_{\partial T \subseteq \Gamma_i} h_T \|\frac{1}{2} [\boldsymbol{\sigma}_h \mathbf{n}]\|_{L_2(\partial T)}^2 + \sum_{\partial T \subseteq \Gamma_\sigma} h_T \|\hat{\mathbf{t}} - \boldsymbol{\sigma}_h \mathbf{n}\|_{L_2(\partial T)}^2 \\ &\left. + \sum_{\partial T \subseteq \Gamma_c} h_T \|\varepsilon u_{hn}^- \bar{\mathbf{n}}^1 + \boldsymbol{\sigma}_h \mathbf{n}\|_{L_2(\partial T)}^2 \right\}^{\frac{1}{2}}. \end{aligned} \quad (12.18)$$

Inequality (12.18) yields an upper bound for the error, which consists of two parts: an error in the strain components; and an error resulting from the not exactly fulfilled contact constraint (3.6).

The error is bounded by the deviation of the discrete solution from equilibrium and the element size. The first and the third terms of the right-hand side contribute to the error bound if local equilibrium is violated. Further, the term  $-\varepsilon u_{hN}^- \bar{\mathbf{n}}^1$  can be interpreted as the contact pressure on  $\Gamma_c$ . Therefore, the second and fourth terms also correspond to a deviation from the equilibrium condition.

*REMARK 12.1:* There are different constants which are related to each of the contributions in (12.17). Thus, using only one constant is not optimal for error control. However, a good estimation of these different constants needs additional effort. Two different techniques have to be mentioned in this context. The first is related to the computation of eigenvalue problems to estimate  $C$ , e.g. see Johnson and Hansbo (1992). The second approach uses the sequence of meshes which is generated within the adaptive refinement process. In this case, the last generated meshes can be used to estimate the constant  $C$ .

So far we have estimated the error in the displacement field. However, it is also interesting to find an error measure for the stresses. According to Johnson and Hansbo (1992) the following residual-based error estimator for linear elastic problems can be found in for the stresses:

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{E^{-1}}^2 \leq \|h C_1 R_1(\boldsymbol{\sigma}_h)\|_{L_2(\Omega)}^2 + \|h C_2 R_2(\boldsymbol{\sigma}_h)\|_{L_2(\Omega)}^2, \quad (12.19)$$

where the quantities are defined on the finite element as follows:

$$R_1(\boldsymbol{\sigma}_h) = |R_1(\boldsymbol{\sigma}_h)| = |\operatorname{div} \boldsymbol{\sigma}_h + \hat{\mathbf{b}}| \quad \text{on } T \quad (12.20)$$

$$R_2(\boldsymbol{\sigma}_h) = \max_{S \in \partial T} \sup_S \frac{1}{2 h_T} |[\boldsymbol{\sigma}_h \mathbf{n}_S]| \quad \text{on } \partial T \quad (12.21)$$

$$\text{or } R_2(\boldsymbol{\sigma}_h) = \frac{1}{h_T} (\hat{\mathbf{t}} - \boldsymbol{\sigma}_h \mathbf{n}) \quad \text{on } \partial T \cap \Gamma_\sigma. \quad (12.22)$$

Here  $\Omega$  denotes the discretized region,  $h_T$  is a characteristic length of an element,  $T$  is the area of a finite element and  $\partial T$  its surface. The norm  $\|\cdot\|_{E^{-1}}$  in (12.19) is the complementary energy norm (written in stress space)

$$\|\sigma - \sigma_h\|_{E^{-1}}^2 = \int_{\Omega} (\sigma - \sigma_h) \cdot \mathbb{C}_0^{-1} [\sigma - \sigma_h] d\Omega. \quad (12.23)$$

In Wriggers et al. (1994) an additional term for the error associated with contact has been introduced

$$R_3(\sigma_h, \mathbf{u}) = |\epsilon_N g_{N-} \bar{\mathbf{n}}^1 - \mathbf{t}_h| \quad \text{on } \partial T \cap \Gamma_c, \quad (12.24)$$

where the term on the right side corresponds to the local equilibrium in the contact interface. The term  $\epsilon_N g_{N-} \bar{\mathbf{n}}^q$  can be interpreted as the contact pressure on  $\Gamma_c$ .

Adding (12.24) to equation (12.19) leads for the linear elastic contact problem to the following *a posteriori* error estimate:

$$\|\sigma - \sigma_h\|_{E^{-1}}^2 \leq \sum_{k=1}^3 \|h C_k R_k(\sigma_h)\|_{L_2(\Omega)}^2. \quad (12.25)$$

A thorough mathematical derivation of the *a posteriori* error estimator can be found in Carstensen et al. (1999). Within the finite element discretization, equation (12.25) has to be evaluated on the element domain, which yields

$$\|\sigma - \sigma_h\|_{E^{-1}}^2 \leq C \sum_T [E_T(h_T, \mathbf{u}_h, \hat{\mathbf{b}}_T)]^2. \quad (12.26)$$

$E_T$  can be computed for each element in the finite element mesh as follows:

$$E_T^2 = h_T^2 \int_T |\operatorname{div} \sigma_h + \hat{\mathbf{b}}|^2 d\Omega + h_T \int_{\partial T \cap \Omega} \frac{1}{2} |[\mathbf{t}_h]|^2 d\Gamma \quad (12.27)$$

$$+ h_T \int_{\partial T \cap \Gamma_\sigma} |\hat{\mathbf{t}} - \mathbf{t}_h|^2 d\Gamma + h_T \int_{\partial T \cap \Gamma_c} |\epsilon_N g_{N-} \bar{\mathbf{n}}^1 - \mathbf{t}_h|^2 d\Gamma. \quad (12.28)$$

Inequality (12.26) yields an upper bound for the error which is bounded by the deviation of the discrete solution from equilibrium and the element size. The first and third terms on the right-hand side contribute to the error bound if the local equilibrium and the traction boundary conditions, respectively, are violated. In (12.28) we have introduced the stress vector  $\mathbf{t}_h = \sigma_h \mathbf{n}$ . Local equilibrium requires that  $[\mathbf{t}_h] = \mathbf{0}$ , which is associated with the second term where  $[\mathbf{t}_h]$  describes the jumps of the tractions over the interface. The fourth term has already been discussed above.

The error estimator described above yields a measure between the exact penalty solution of (12.3) and its finite element approximation (12.6). What is really needed is the error between the exact solution of (5.10) and the approximate finite element solution (12.6). So far there are no computable error bounds for contact problems in

elasticity. But we can make use of a result derived by Kikuchi and Oden (1988) to change the penalty parameter in such a way that an optimal convergence rate of the method is achieved. To this purpose, we state the result of Kikuchi and Oden (1988) which was derived for a perturbed LAGRANGE formulation of the contact problem

$$\| \mathbf{u} - \mathbf{u}_{\varepsilon h} \|_1 + |p_N - p_{N\varepsilon h}|^* \leq C_3 h + C_4 \varepsilon_N^{-1} h^{-1/2}. \quad (12.29)$$

From this equation it is clear that an optimal convergence rate can be obtained for  $\varepsilon_N \approx h^{-3/2}$ ,  $h$  being the characteristic length of an element. According to this relation, we develop now the following update at iteration  $k + 1$  for the penalty parameter in the contact interface:

$$\varepsilon_{Nk+1} = \varepsilon_{N0} \left( \frac{h_{k+1}}{h_0} \right)^{-\frac{3}{2}},$$

where  $\varepsilon_{N0}$  and  $h_0$  are the starting values at the beginning of the adaptive iteration.

### 12.3 ERROR INDICATOR FOR CONTACT BASED ON PROJECTION

Another possibility to derive an error estimator for elastic contact problems starts directly from the complementary elastic energy norm (12.23). A simple, but in many cases efficient, error estimator is now provided by the superconvergent-stress-recovery technique which is due to Zienkiewicz and Zhu (1987). The equivalence of such error measures with the residual-based error estimators of the last section has been shown in Verfürth (1996). The idea to derive these error estimators is based on the fact that many finite element meshes have superconvergence properties, which means that there are points in which the stresses are approximated with higher accuracy. By using a projection procedure, the stresses  $\sigma^*$  can be computed from the superconvergent points. It should be noted in passing that the stress-recovery error estimators also work well if the sampling points are not superconvergence points, see Babuska et al. (1994). An especially efficient projection technique is provided by the lumped  $L^2$  projection, which is described in detail in Zienkiewicz and Taylor (1989). In general, the projection procedures assume that the projected stresses do not have jumps, which needs some special considerations in the contact interface (see below). Denoting by  $\mathbb{P}$  a projection operator, from  $\sigma^*$  we obtain

$$\int_{\Omega} \mathbb{P} [\sigma^* - \sigma_h] d\Omega = \mathbf{0}, \quad (12.30)$$

and can then compute an approximation of the error using (12.23):

$$\| \sigma - \sigma_h \|_{E^{-1}}^2 \leq \int_{\Omega} (\sigma^* - \sigma_h) \cdot \mathbf{C}_0^{-1} [\sigma^* - \sigma_h] d\Omega. \quad (12.31)$$

This error estimator can be evaluated in an efficient way, and has been shown to be robust, see Babuska et al. (1994). Equation (12.31) does not include special

boundary terms for the contact contributions. The estimate for the contact area is included in this case implicitly, since the evaluation of the integral in (12.31) has to be done with respect to  $\Omega = \Omega^1 \cup \Omega^2$ , and thus also includes the contact interface. However, a special projection has to be performed, since it makes sense to treat normal and tangential stress components in the contact interface differently. Hence in the frictionless contact the normal component, given by  $p_N = \bar{\mathbf{n}}^1 \cdot \boldsymbol{\sigma} \bar{\mathbf{n}}^1$ , has to be projected using all elements connected to a point in the contact interface, e.g. from both bodies. Since we deal here with frictionless contact, we can neglect the tangential stresses for the moment. Thus, this special projection scheme first has to be used and then the normal components of the stresses are transformed back to  $\boldsymbol{\sigma}^*$ .

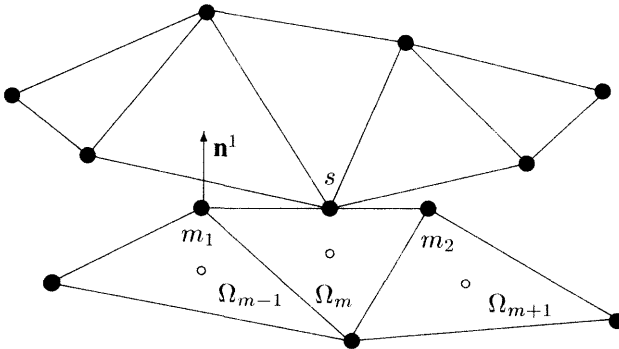
Since the normal stress is continuous over  $\Gamma_c$ , in the frictionless case it is given by  $p_N = \bar{\mathbf{n}}^1 \cdot \boldsymbol{\sigma} \bar{\mathbf{n}}^1$ . We can refer to  $p_N$  as the principal stress on the interface, since all tangential components of the stress tensor are zero in the frictionless case. A simple procedure which yields a projected and a normal field is as follows. We define a patch  $\omega_{m\alpha}$  associated with node  $m_\alpha$  ( $\alpha = 1, 2$ ) on the master surface (see Figure 12.2) by all elements belonging to the node (for  $m_1$  these are  $\Omega_{m-1}$  and  $\Omega_m$  here). Now a projection procedure like  $Z^2$  or another, which is more accurate for boundary nodes can be applied, e.g. see Zienkiewicz and Taylor (1991). It yields the stress tensor at the node  $m_\alpha$

$$\mathbb{P}(\boldsymbol{\sigma}_{\omega_{m\alpha}}) \longrightarrow \boldsymbol{\sigma}_{m\alpha}^*. \quad (12.32)$$

The normal and tangential components of the stress vector can be computed from the stresses  $\boldsymbol{\sigma}_{m\alpha}^*$  via CAUCHY'S theorem by using the normal and tangential vectors defined following (8.2) for the two-dimensional case (see Figure 12.2):

$$p_{N\alpha}^* = (\boldsymbol{\sigma}_{m\alpha}^* \mathbf{n}^1) \cdot \mathbf{n}^1, \quad (12.33)$$

$$t_{T\alpha}^* = (\boldsymbol{\sigma}_{m\alpha}^* \mathbf{n}^1) \cdot \mathbf{a}_1^1. \quad (12.34)$$



**Fig. 12.2** Element patches for stress projection on  $\Gamma_c$ .

In the case of frictionless contact, we are only interested in the normal stress, and can now define with  $p_{N\alpha}^*$  a linear interpolation of the normal contact stress:

$$p_N^*(\xi) = (1 - \xi) p_{N1}^* + \xi p_{N2}^*. \quad (12.35)$$

As in (12.31), this result can now be applied to compute the error associated with the contact stresses. This leads to the additional error in  $\partial\Omega^\gamma \cap \Gamma_c = \sum_T \partial T \cap \Gamma_c$ :

$$\|e\|^{C\gamma_N} = \sum_T \|e\|_{\partial T}^{C\gamma_N}, \quad (12.36)$$

with the error related to the segment  $m_1-m_2$ :

$$\|e\|_{\partial T}^{C\gamma_N} = \left\{ \int_0^1 \frac{1}{\epsilon_N} [p_N^*(\xi) - p_{Nh}]^2 l d\xi \right\}^{\frac{1}{2}}, \quad (12.37)$$

where  $p_{Nh}$  is obtained according to (8.13) for the slave  $s$  node which is in contact with the master surface  $m_1-m_2$ , see Figure 12.2.

The error within the whole domain is then computed by the sum over all elements  $T$ , with  $\Omega$  being the union of all elements. Thus we have

$$\|\sigma - \sigma_h\|_{E^{-1}} = \sum_T \|e\|_T, \quad (12.38)$$

with

$$\|e\|_T = \int_T (\sigma^* - \sigma_h) : \mathbf{C}_0^{-1} : (\sigma^* - \sigma_h) dT. \quad (12.39)$$

Equation (12.37) provides the additional term due to contact which has to be included in case the element includes a contact interface:

$$\|e\|_T^C = \|e\|_T + \|e\|_{\partial T}^{C\gamma_N}. \quad (12.40)$$

**REMARK 12.2:**

1. In the stick case a penalty stiffness is used to prevent relative tangential movements in the contact area. Thus we can derive an error measure for stick using interpolation (12.35), together with (12.34)  $t_T^*(\xi) = (1 - \xi) t_{T1}^* + \xi t_{T2}^*$ , which yields

$$\|e\|_{\partial T}^{C\gamma_T} = \left\{ \int_0^1 \frac{1}{c_T} [t_T^*(\xi) - t_{Th}]^2 l d\xi \right\}^{\frac{1}{2}}, \quad (12.41)$$

which has to be added to (12.40);  $t_{Th}$  follows from (8.13).

2. If the penalty parameter  $\epsilon_N$  has a very large value compared to the elasticity constants of the bodies in contact, then terms (12.38) and (12.41) are small compared to (12.31). In this case it makes sense to scale the penalty parameter such that it has the same magnitude as the elasticity constants in (12.31) of the bodies in contact.

3. *So far the error estimators have been developed for small elastic deformations. In the case of large elastic strains, there is no mathematically sound basis. However, it should be noted that existence results exist for polyconvex materials, e.g. see the overview in Ciarlet (1988). These results have been extended to contact problems in Ciarlet (1988), Kikuchi and Oden (1988) and Curnier et al. (1992). Thus there is at least an existence result available for contact problems. The question of uniqueness, also needed for the derivation of error estimators, can of course not be solved, since problems undergoing large elastic deformations may also exhibit material as geometrical instabilities (e.g. limit points or bifurcations).*

*If we now formulate all equations associated with the variational inequality in the tangent space of a given deformation map, and exclude within this configuration instabilities, then the information from these incremental equations can be used for an error estimate within the incremental step. This means that we can exchange the stresses in (12.31) by the appropriate nonlinear stress measures when using the residual-based error estimator. If we employ the error estimator based on the superconvergent recovery technique, then additionally the incremental constitutive tensors have to be used in (12.31) and (12.37) to compute the error (12.40).*

## 12.4 ERROR ESTIMATORS BASED ON DUAL PRINCIPLES

The adaptive mesh refinement strategies described in the previous sections are based on *a posteriori* error estimates in the global energy or  $L^2$  norm involving local residuals or postprocessed stresses of the computed solution. However, for practical purposes, the energy or  $L^2$  norm may be of minor importance. In general, the variables of interest are displacements or stresses at some particular points. Recently procedures for the estimation of various kinds of local error functionals have been introduced in a general framework by Becker and Rannacher (1996) or Rannacher and Suttmeier (1997). This approach is developed here for contact problems involving large elastic strains. The error estimator is evaluated by solving the set of equations for an additional right-hand side, and applying the classical error estimators (based on residuals or superconvergent patch recovery). The combination of the discretization errors for the initial and the dual problems gives the desired error quantity. The formal approach is given in the following.

### 12.4.1 Displacement error control

As for residual norm error estimators, the starting point is again the differential equation for the discretization error  $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$ :

$$\mathbf{L}_{\hat{x}}(\mathbf{u} - \mathbf{u}_h) = \mathbf{L}_{\hat{x}}\mathbf{e} = \mathbf{b} - \mathbf{L}_{\hat{x}}\mathbf{u}_h := \mathbf{R}_1, \quad (12.42)$$

with the formal differential operator  $\mathbf{L}$  and the element internal residuals  $\mathbf{R}_1$ .



With  $\eta$  denoting the test functions and  $a(\cdot, \cdot)$  as defined above, we obtain the weak form

$$a(\mathbf{e}, \eta) = \sum_T \left\{ \int_{\Omega_T} (\operatorname{div} \hat{\tau}_h + \bar{\mathbf{f}}) \cdot \eta \, d\Omega + \int_{\partial T \cap \Gamma_\sigma} (\bar{\mathbf{t}} - \hat{\tau}_h \mathbf{n}) \cdot \eta \, d\Gamma + \int_{\partial T \cap \Omega} \frac{1}{2} [\mathbf{t}_h] \cdot \eta \, d\Gamma \right\}. \quad (12.43)$$

The first integral represents the virtual work done by element internal residuals whereas the second term represents the virtual work done by jumps in the tractions at the force boundaries.

For edges apart from the structure boundaries, the jumps in the tractions of both neighbouring elements are the edge residuals represented in the third term. The jumps are split by a factor of 1/2 into the two elements sharing the edge considered.

For simplicity, we use  $\mathbf{R}_1$  and  $\mathbf{R}_2$  containing internal and jump residuals, respectively:

$$a(\mathbf{e}, \eta) = \sum_T \{ (\mathbf{R}_1, \eta)_{\Omega_T} + (\mathbf{R}_2, \eta)_{\Gamma_T} \}, \quad (12.44)$$

To estimate the error of a specific displacement in the component  $i$  at point  $\mathbf{x} = \bar{\mathbf{x}}$ , we additionally consider the following dual problem:

$$\operatorname{div} \tau(\mathbf{G}) + \delta_i(\bar{\mathbf{x}}) = \mathbf{0}, \quad (12.45)$$

or in weak form

$$a(\mathbf{G}, \eta) = (\delta_i, \eta), \quad (12.46)$$

where  $\delta_i$  is the DIRAC delta (unit point load vector) in the direction  $i$ , and  $\mathbf{G}$  denotes the Green's function.

Applying the principle of BETTI-MAXWELL to the error problem (12.44) and the dual problem (12.46) yields the following relation:

$$(\mathbf{e}, \delta_i) = \sum_T \{ (\mathbf{R}_1, \mathbf{G})_{\Omega_T} + (\mathbf{R}_2, \mathbf{G})_{\Gamma_T} \}. \quad (12.47)$$

The term on the left-hand side is the work of a unit load with the error function  $\mathbf{e}$  and is equal to the error  $e_i(\bar{\mathbf{x}})$  of the  $i$ th component of the displacement at point  $\bar{\mathbf{x}}$ .

Now inserting  $\mathbf{G}$  in equation (12.44) instead of the test function  $\eta$ , the local error can be expressed by the bilinear form

$$e_i(\bar{\mathbf{x}}) = a(\mathbf{e}, \mathbf{G}). \quad (12.48)$$

Of course, the solution of the dual problem is not known, but it can also be computed numerically based on the same discretization. It is simply another load case.

Using the GALERKIN orthogonality with  $\mathbf{G}_h$  as the finite element approximation of  $\mathbf{G}$

$$e_i(\bar{\mathbf{x}}) = a(\mathbf{e}, \mathbf{G} - \mathbf{G}_h) \quad (12.49)$$

and applying the CAUCHY–SCHWARZ inequality, the local error is estimated by the error energy  $a(\mathbf{e}, \mathbf{e})$  of the primal problem (12.44) weighted by the error energy  $a(\mathbf{G} - \mathbf{G}_h, \mathbf{G} - \mathbf{G}_h)$  of the dual problem (12.46):

$$e_i^2(\bar{\mathbf{x}}) \leq a(\mathbf{e}, \mathbf{e}) a(\mathbf{G} - \mathbf{G}_h, \mathbf{G} - \mathbf{G}_h). \quad (12.50)$$

The second term serves as the weighting function and filters out the influence of the overall residuals over the displacement error of interest. Inequality (12.50) can be computed element wise, applying the estimators (residual-based or post-processed) described in the previous sections on the primal and dual problems, respectively:

$$e_i^2(\bar{\mathbf{x}}) \leq \sum_T a(\mathbf{e}, \mathbf{e})_T a(\mathbf{G} - \mathbf{G}_h, \mathbf{G} - \mathbf{G}_h)_T. \quad (12.51)$$

For instance,  $a(\mathbf{e}, \mathbf{e})_T$  and  $a(\mathbf{G} - \mathbf{G}_h, \mathbf{G} - \mathbf{G}_h)_T$  are computed by applying the stress-recovery error estimator:

$$a(\mathbf{e}, \mathbf{e})_T = \int_{\Omega_T} (\boldsymbol{\tau}^*(\mathbf{u}_h) - \boldsymbol{\tau}(\mathbf{u}_h)) \cdot \mathbf{C}^{-1} [\boldsymbol{\tau}^*(\mathbf{u}_h) - \boldsymbol{\tau}(\mathbf{u}_h)] d\Omega \quad (12.52)$$

and

$$a(\mathbf{G} - \mathbf{G}_h, \mathbf{G} - \mathbf{G}_h)_T = \int_{\Omega_T} (\boldsymbol{\tau}^*(\mathbf{G}_h) - \boldsymbol{\tau}(\mathbf{G}_h)) \cdot \mathbf{C}^{-1} [\boldsymbol{\tau}^*(\mathbf{G}_h) - \boldsymbol{\tau}(\mathbf{G}_h)] d\Omega. \quad (12.53)$$

With the same arguments as in the previous section, the stresses required to evaluate (12.52), (12.53) are obtained from the solution of the corresponding linear problems.

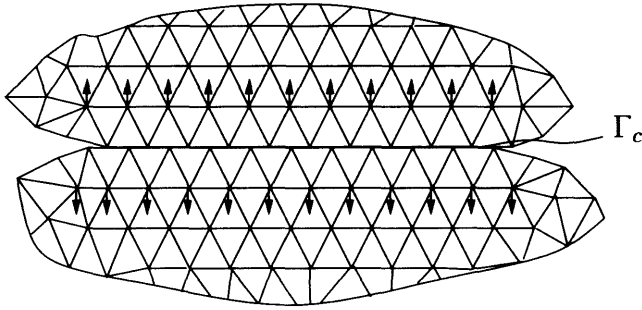
Therefore, we compute  $\bar{\boldsymbol{\tau}}(\boldsymbol{\varphi}_h)$  and  $\bar{\boldsymbol{\tau}}^*(\boldsymbol{\varphi}_h)$  from the increase of displacements of a perturbed loading state  $\hat{\lambda} + \epsilon \lambda$  in the vicinity of the current equilibrium point  $(\hat{\boldsymbol{\varphi}}_h, \hat{\lambda})$ .

In the same way,  $\boldsymbol{\tau}(\mathbf{G}_h)$  and  $\boldsymbol{\tau}^*(\mathbf{G}_h)$  are obtained from the increase of displacements by applying  $\boldsymbol{\delta}_i$  as an additional small load on the current loading state. Since, for a small load, we will not observe any change in the size of the contact area, this procedure is also consistent with the assumption that the contact area  $\Gamma_c^h$  does not change when  $\boldsymbol{\delta}_i$  is applied.

*REMARK 12.3: Although the internal energy of an elastic structure loaded by point loads is infinite, the error for the displacement of a specific finite element node can be computed simply by applying a nodal force in the dual problem. The smearing effect of elements with finite length leads to the simplest way of regularization bounding the energy.*

## 12.4.2 Stress error control

Errors for stress variables at a specific point  $\bar{\mathbf{x}}$  can be estimated in a straightforward manner using the same concept as in the previous section. In contrast to the local



**Fig. 12.3** Dual problem for contact pressure.

displacement error control, a discontinuity on the related displacement variable must be applied in the dual problem:

$$\operatorname{div} \tau(\mathbf{z}) + \frac{\partial}{\partial x_j} \delta_i(\bar{\mathbf{x}}) = \mathbf{0}. \quad (12.54)$$

Applying again the reciprocal theorem, and subsequently the GALERKIN orthogonality, the error in the displacement gradient or respective stress value can be computed by

$$\left( \mathbf{e}, \frac{\partial}{\partial x_j} \delta_i \right) = \frac{\partial e_i(\bar{\mathbf{x}})}{\partial x_j} = \sum_T \{ (\mathbf{R}_1, \mathbf{z})_{\Omega_T} + (\mathbf{R}_2, \mathbf{z})_{\Gamma_T} \} = a(\mathbf{e}, \mathbf{z} - \mathbf{z}_h). \quad (12.55)$$

To apply a discontinuity to an individual displacement in two or three dimensions, some sort of regularization must be applied. The simplest way in the finite element method is to substitute the displacement jump by equilibrated forces acting at two neighbouring nodes to point  $\bar{\mathbf{x}}$ .

The above described concept for local quantities can easily be extended to arbitrary integral variables like reaction forces or pressure quantities on a contact interface. In order to estimate the error in boundary reactions, the displacement discontinuity can be imitated by a line load in the vicinity of the boundary (see Figure 12.3).

## 12.5 ADAPTIVE MESH REFINEMENT STRATEGY

Based on knowledge of the error distribution, a refined mesh can be constructed which yields a better approximate solution. In general, there are several possibilities to obtain refined meshes. One approach is called the *hierarchical method*, where elements of an existing mesh are subdivided at parts where the error measure indicates it. Another method is to define an error density function and to use this for the generation of a

completely new refined mesh. To obtain an optimal mesh, which means we want the error to be of the same order in all mesh parts, it is desirable to design the mesh such that the error contributions of the elements are equi-distributed over the mesh.

The object of an adaptive algorithm is usually stated as a nonlinear optimization problem: construct a mesh such that the associated FEM-solution satisfies

$$\|\sigma - \sigma_h\|_{E^{-1}} \leq C \sum_T [E_T(h_T, \mathbf{u}_h, \hat{\mathbf{b}}_T)]^2 \leq TOL, \quad (12.56)$$

with  $TOL$  being a given tolerance. Furthermore, the expense to compute  $\mathbf{u}_h$  or  $\sigma_h$  should be nearly minimal. The measure  $E_T$  can be either

$$E_{T1}^2 = E_T^2 \text{ from equation (12.28),} \quad (12.57)$$

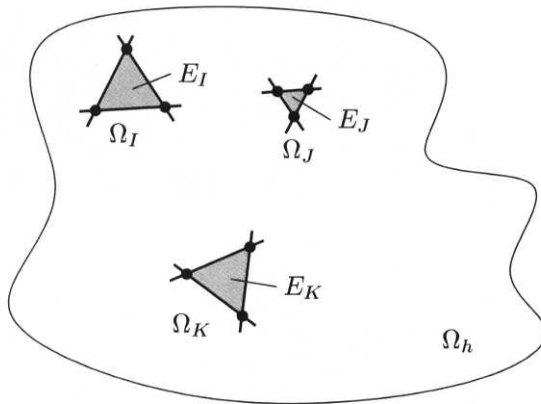
$$E_{T2}^2 = \|\mathbf{e}\|_T^C \text{ from equation (12.40), or} \quad (12.58)$$

$$E_{T3}^2 = e_i^2(\bar{\mathbf{x}}) \text{ from equation (12.51).} \quad (12.59)$$

As a measure of computational work, the total number of degrees of freedom is chosen. Since the exact solution  $\mathbf{u}$  is not known, we demand that the error contributions of all elements yields

$$\sum_T E_{T_k}^2 \leq TOL, \quad (12.60)$$

which guarantees that (12.56) is fulfilled,  $k$  being 1, 2 or 3, depending on the choice of error estimator. Here the constant  $C$  appearing in (12.26) has been included in  $TOL$  for convenience when the error measure  $E_{T1}$  is used. Equation (12.60) serves as a stopping criterion in the adaptive process. To minimize the number of degrees of freedom during refinement, we require that the mesh is an optimal mesh, i.e. that the error  $E_T^2$  is equally distributed between elements (see Figure 12.4). Ideally, we



**Fig. 12.4** Equi-distributed error in finite element mesh.

then have for the elements  $I, J, K$

$$E_I = E_J = E_K. \quad (12.61)$$

With this requirement, for the error contribution of all elements we can write

$$\sum_T E_{T_k}^2 = N E_I^2. \quad (12.62)$$

$N$  denotes the number of elements in the mesh, and  $E_I$  the error which is now associated with all elements according to (12.61). Finally, (12.60) together with (12.56) yields the refinement criterion,

$$E_{T_k}^2 \leq \frac{TOL}{N}. \quad (12.63)$$

Now we state the overall algorithm of the h-adaptive method for contact problems. The algorithm includes the following steps:

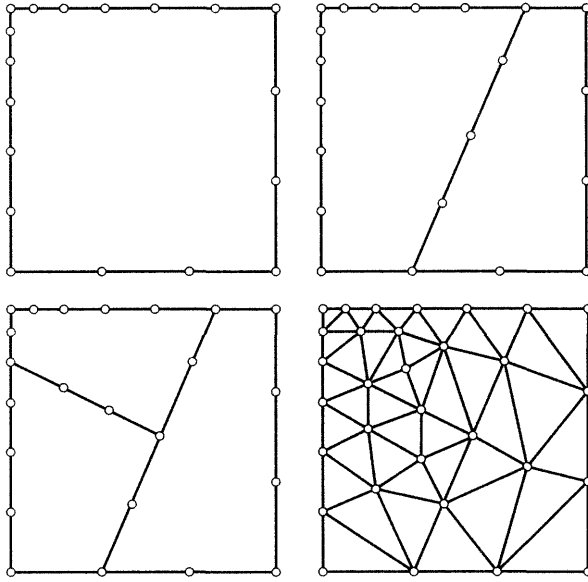
1. Set initial values:  $l = 0, \lambda_0 = 0, \Delta\lambda, i = 0$
2. Generation of start mesh:  $\mathcal{M}_i$
3. Loop over load increments:  $\lambda_{l+1} = \lambda_l + \Delta\lambda$ 
  - (a) IF  $\lambda_{l+1} > \lambda_{max} \Rightarrow$  STOP
  - (b) Iteration loop to solve contact problem
  - (c) Mesh optimization
    - Compute  $E_{T_k}^2$
    - IF  $\sum E_{T_k}^2 < TOL \Rightarrow$  GOTO 3.
    - IF  $E_{T_k}^2 > TOL / N \Rightarrow$  refine element  $T$
    - Set  $i = i + 1$
    - Generate new mesh  $\mathcal{M}_i$ 
      - Delaunay triangularization
      - Smoothing, if necessary
    - Interpolate displacement and history variables of the new mesh
    - GOTO 3 (b)

The new mesh is assumed to be generated by a DELAUNAY triangularization, but different generation techniques like the advancing front method or others can also be applied (see Sloan (1993) or Rank et al. (1993) for two-dimensional mesh generation or Schöberl (1997a) for a three-dimensional meshing tool).

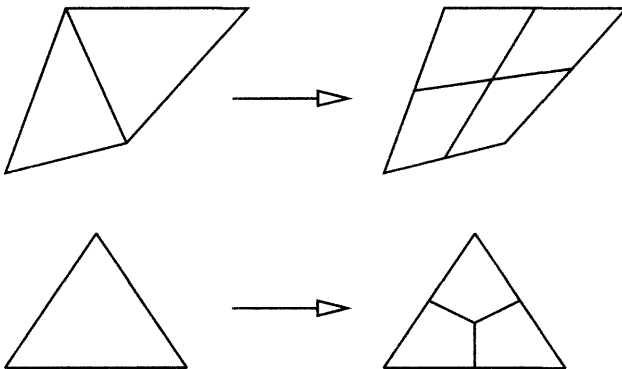
A meshing procedure which combines the DELAUNAY triangularization and a subdivision techniques for the generation of meshes consisting of quadrilateral elements is described next. This approach was also used to generate the meshes of the examples in Sections 10.7.

In the proposed procedure the *recursive region splitting* is applied for mesh generation (Figure 12.5). The region which has to be discretized is subdivided recursively. Within the procedure the nodes are positioned based on a density function. This defines the local element size on the basis of the error indication. The element generation then follows by a DELAUNAY-triangularization, e.g. see Sloan (1993).

The triangular mesh can now be transformed into a mesh consisting completely of quadrilaterals. To achieve this goal, two neighbouring triangles are combined and



**Fig. 12.5** Recursive region splitting.



**Fig. 12.6** Conversion: triangular to a quadrilateral mesh.

subdivided into four quadrilaterals. The left over single triangles are then subdivided into three quadrilaterals, as depicted in Figure 12.6.

It should be noted that when a background density function is applied to define the local element size, distorted element geometries cannot be avoided. To avoid this, a relaxation of the nodes can be applied to improve the shape of the elements; see Rank et al. (1993) for details.

## 12.6 NUMERICAL EXAMPLES

In this section we discuss two examples in which different error indicators and estimators are compared. Furthermore, an example which exhibits large strains is considered. All computations were performed by using the finite element program FEAP, which is described in Zienkiewicz and Taylor (1989). The adaptive meshes for quadrilaterals were constructed with mesh generators developed in Bank (1990) and Rank et al. (1993). Mesh refinements for triangular elements are based on the algorithm described in Sloan (1993). The three-dimensional meshes were created using the mesh generator described in Schöberl (1997a).

### 12.6.1 Hertzian contact problem

In this example we apply all error estimation procedures, described within the previous section, to solve the well known Hertzian problem of an elastic cylinder (Young's modulus  $E = 7000$ , Poisson's ratio  $\nu = 0.3$ ) contacting a planar rigid surface. The cylinder has a radius of  $r = 1$ , and is loaded by an overall load of  $F = 100$ .

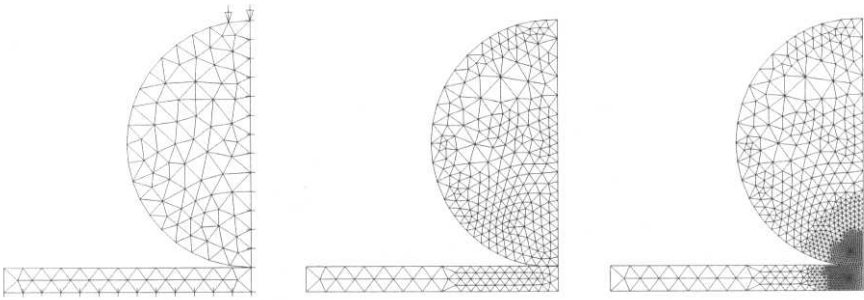
For this problem an exact solution for the contact pressure can be obtained analytically:

$$p_{max} = \sqrt{\frac{F}{\pi r} \frac{E}{(1 + \nu)(1 - \nu)}}.$$

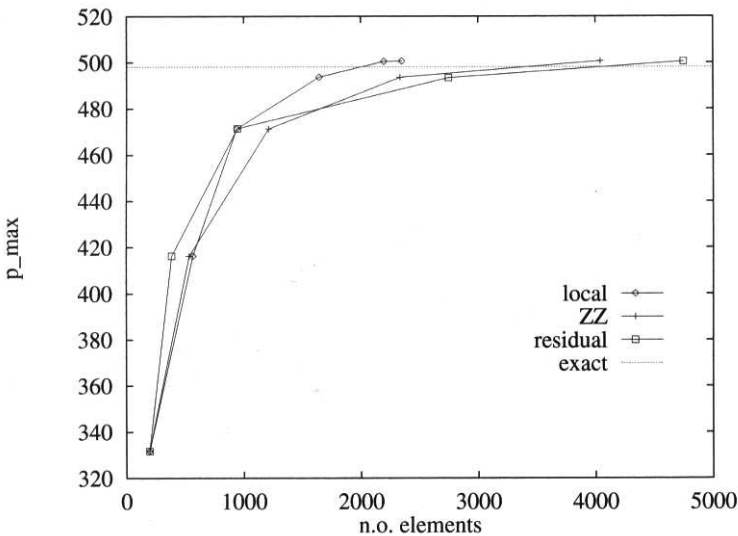
Hence we compare the results computed by the adaptive methods directly to the exact solution. To omit problems related to a point load in elasticity, the load is distributed over a small surface on top of the cylinder. In order to simulate the rigid surface we set Young's modulus to  $E = 100000$  and Poisson's ratio to  $\nu = 0.45$  in the finite element model. The initial mesh is depicted in Figure 12.7a.

As the locally controlled error quantity we use the maximum contact pressure within the method based on duality. Hence, equilibrated vertical forces are applied centrally on the contact interface.

Three different error measures were applied to perform the adaptive computation. These are the residual-based error estimator due to Johnson and Hansbo (1992) (see Section 12.2), the  $Z^2$  indicator of Zienkiewicz and Zhu (1987) (see Section 12.3), and a local error estimator developed in Rannacher and Suttmeier (1997) (see Section 12.4). All error measures have been enhanced to include contact, see Wriggers et al. (2000). For the local error estimator, the error measure chosen was the maximal



**Fig. 12.7** (a) Starting mesh: 258 elements, (b) and mesh refinement: local error estimator.



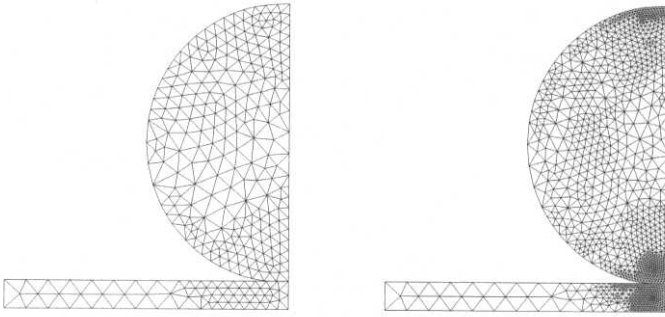
**Fig. 12.8** Convergence behaviour.

contact pressure. Due to (12.55), we then have to apply to equilibrated force systems in the contact interface.

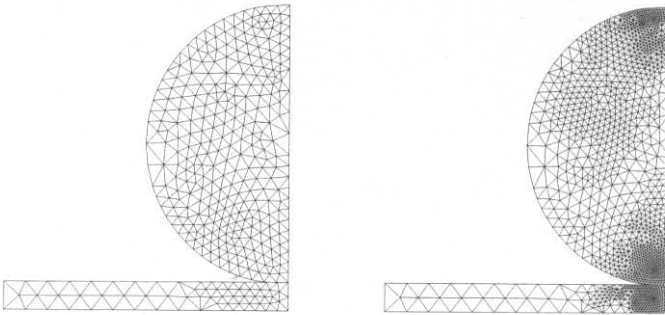
Pictures 12.7(b) and 12.9 to 12.10 show meshes which were obtained using the different error measures. The final mesh refinement belongs to the converged solution.

The maximum contact pressure is depicted in Figure 12.8 for the different adaptive strategies. It is compared with the analytical solution of  $p_{max} = 494, 83$ . It is clearly visible that the local error estimator leads a solution which converges very fast, since it only needs half the number of elements than the other methods. One can also see that this error estimator, by design, yields only a refinement where needed, whereas the other two error measures take the complete solution into account and also refine





**Fig. 12.9** Mesh refinement:  $Z^2$  error indicator.

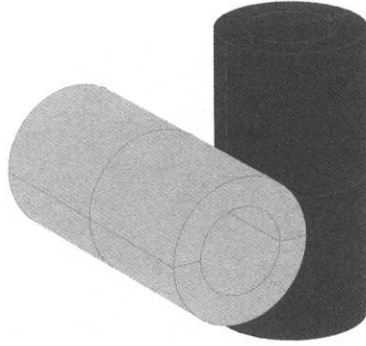


**Fig. 12.10** Mesh refinement: residual-based error estimator.

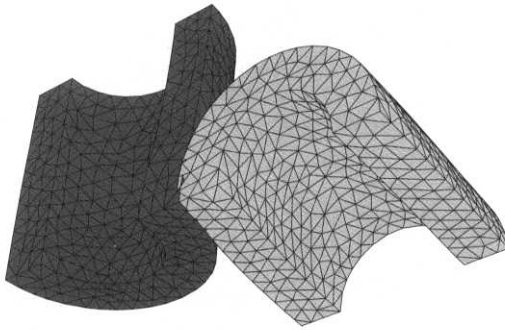
close to the concentrated load at the top of the disk. Thus if only the pressure is of interest, the local error estimator provides the most efficient method. However, if one does not know where maximal values will appear in the structure then one has to look for a method which takes all errors in the domain into account. Hence the residual-based or the  $Z^2$ -indicator are preferable.

### 12.6.2 Crossing tubes

In this example we consider two crossing rubber tubes (see Figure 12.11) with the material data  $E = 3000$  and  $\nu = 0.3$ . Each tube has a length of  $L = 100$  mm, an outer radius of  $r_o = 50$  mm and an inner radius of  $r_i = 30$  mm. Due to symmetry,



**Fig. 12.11** Initial geometry.

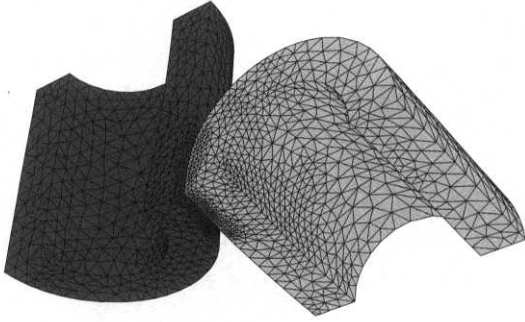


**Fig. 12.12** One fourth of the model, initial mesh.

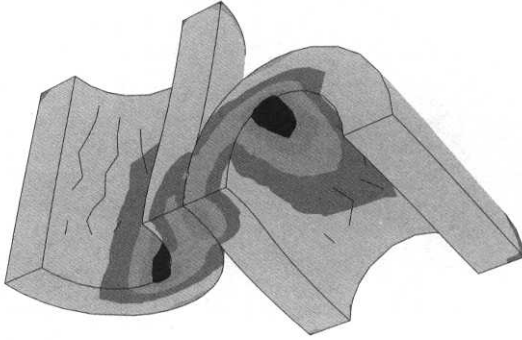
a one fourth model is used. The tube is fixed at one end and loaded so that both tubes come into contact. The undeformed state is depicted in Figure 12.11. The deformed state can be found in Figure 12.12. The mesh adaption is performed with respect to the global error energy norm and with a prescribed relative error tolerance of  $\eta_{err} = 15\%$ . It leads at the given deformation state to the adapted mesh shown in Figure 12.13. We see a very fine mesh developed in the contact zone. That this refinement is necessary can be observed from the stress field. The principal stresses are reported for the initial mesh and the adapted mesh in Figures 12.14 and 12.15.

### 12.6.3 Fractal interface

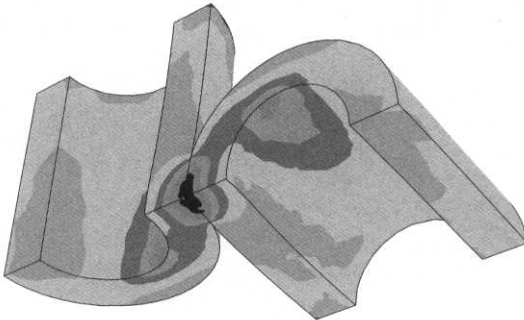
The approximations of the fractal interfaces are combined with a penalty regularization based on the minimization of the potential energy after some appropriate transformations are performed. For this type of contact problem, there are often asperities with corners on the interfaces which lead to possible stress concentrations. Further-



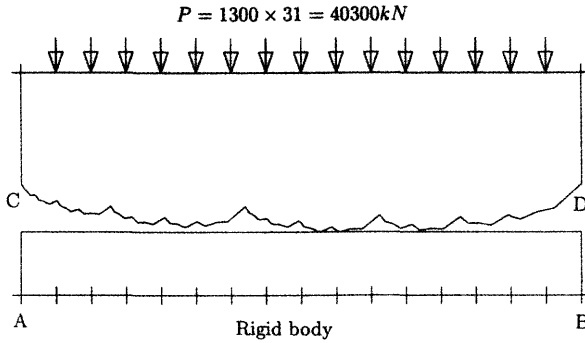
**Fig. 12.13** Adapted mesh.



**Fig. 12.14**  $\sigma_I$ , initial mesh.



**Fig. 12.15**  $\sigma_I$ , adapted mesh.



**Fig. 12.16** Finite elasticity with a fractal interface.

more, the convergence of a finite element solution under a sufficient discretization cannot be determined from the outset.

An adaptive finite element strategy appears to be suitable for such a kind of contact problem, in that it possesses the properties of automatically adjusting the mesh sizes both in the interior of bodies and at the contact zone. Here we use the residual-based error estimator for adaptive analysis. An in-depth treatment of the underlying mathematical structure can be found in Hu et al. (2000).

Let us consider the structure of Figure 12.16 with a fractal boundary. This example was first proposed in Panagiotopoulos et al. (1992), although the solution did not include adaptive mesh refinement. The system is subjected to a gravity load in its plane. Large traction loads on the upper boundary enforce the fractal part to be in contact with the rigid support  $AB$ . The body has been discretized by triangular elements. Finite elasticity, using a Neo-Hookean constitutive equation (see Section 2) and geometrical nonlinearity are assumed. The modulus of elasticity is  $E = 2.1 \times 10^9$  N/m<sup>2</sup> and the Poisson's ratio  $\nu = 0.33$ . The thickness of the plate is taken to be 0.05 m. We study the planification of the fractal surface and the corresponding variations of stress and displacement fields.

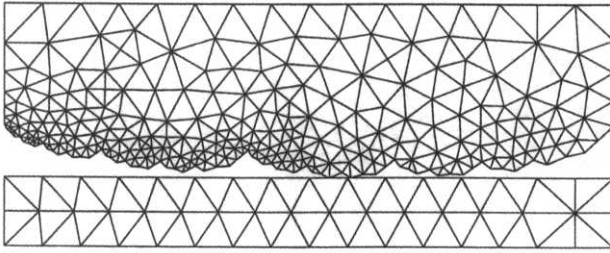
The fractal boundary is defined to be the attractor of the Iterated Function System IFS (see also Section 4.7)  $\{\mathbb{R}^2; w_1, w_2\}$ , where

$$w_1 \begin{Bmatrix} x \\ y \end{Bmatrix} = \begin{bmatrix} 0.40 & 0.0 \\ -0.04 & 0.60 \end{bmatrix} \begin{Bmatrix} x \\ y \end{Bmatrix} + \begin{Bmatrix} 0.0 \\ 0.0 \end{Bmatrix}$$

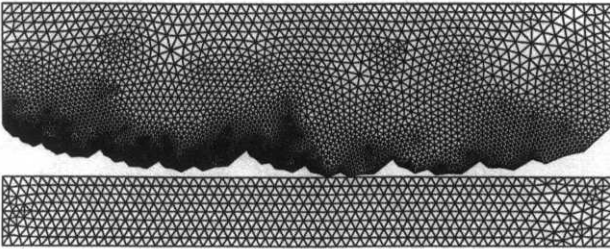
and

$$w_2 \begin{Bmatrix} x \\ y \end{Bmatrix} = \begin{bmatrix} 0.60 & 0.0 \\ 0.04 & 0.80 \end{bmatrix} \begin{Bmatrix} x \\ y \end{Bmatrix} + \begin{Bmatrix} 1.0 \\ -0.1 \end{Bmatrix}.$$

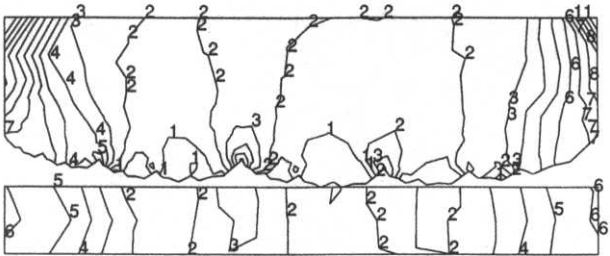
We must note here that the above relations describe the IFS on the assumption that the coordinates of the point  $C$  are  $(x, y) = (0.0, 0.0)$ . With this IFS we calculate the approximations of the fractal interfaces. Note that the stress and displacements fields of the fractal interface become insignificant after the fifth iteration. This is because



**Fig. 12.17** Initial mesh.



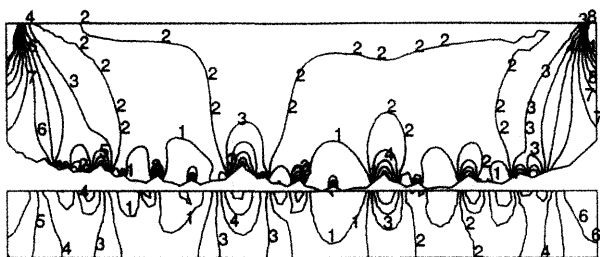
**Fig. 12.18** Second mesh refinement.



**Fig. 12.19** Stress  $\sigma_{33}$  distribution for the initial mesh.

the HAUSDORFF distance between the approximation  $f_5$  and the attractor  $f$  is very small, i.e. the approximation  $f_5$  sufficiently approximates the attractor  $f$ .

The load is applied in one increment on the initial mesh shown in Figure 12.17. After that, the load is kept constant during the refinement stages. Figure 12.18 depicts the mesh of the second refinement, which shows a concentration of the discretization along the fractal contact surface. The vertical stresses are reported in Figures 12.19 and 12.20 for the initial and refined mesh, respectively.



**Fig. 12.20** Stress  $\sigma_{33}$  distribution for the second mesh refinement.

The whole algorithm proved to be stable, although the complicated interfaces introduced by the approximation of the fractal interface do not offer an ideal framework for a unilateral contact problem. The complexity of the higher order approximations of the fractal interface caused an increase in the computing time. It is also important to note that the approximation of the fractal contact boundary  $\Gamma_c$  does not considerably affect the stress and displacement fields inside the body. Of course, this is compatible with ST. VENANT'S principle of classical elastostatics, which actually holds for classical boundaries and for bilateral boundary conditions. The numerical results are quite reliable concerning the consideration of possible stress concentration, which can be seen by comparing Figures 12.19 and 12.20. Indeed, the action of external and re-entrant corners in each approximation has been appropriately taken into account by increasing the number of elements around the singular point.

## 12.7 ERROR INDICATOR FOR FRICTIONAL PROBLEMS

In the previous section, additional terms for the error associated with normal contact have been introduced for frictionless contact. Now we want to establish the same sort of error indicators for contact with friction. Here the mathematical theory is not developed in so much detail, e.g. see Hlavacek et al. (1988), although several attempts to estimate the errors *a posteriori*, e.g. see Lee and Oden (1994). However, for general applications, involving finite deformations, error estimators are still missing. Thus, in the case of frictional contact we use error measures which indicate the error in the contact interface. Again, we can base the error computations on either residual error measures or error indicators based on superconvergence properties. Both will be developed in the following.

Let us recall the residual error estimator for frictionless contact for geometrical linear problems (12.24):

$$R_3(\sigma_h, \mathbf{u}_h) = |\epsilon_N \mathbf{g}_{N_h}^- \bar{\mathbf{n}}^1 - \mathbf{t}_h| \quad \text{on } \partial T \cap \Gamma_c, \quad (12.64)$$

which corresponds to the local equilibrium in the contact interface. The term  $\epsilon_N \mathbf{g}_{N_h}^- \bar{\mathbf{n}}^1$  can be interpreted as the contact pressure on  $\Gamma_c$ . The stress vector is given by  $\mathbf{t}_h = \boldsymbol{\sigma}_h \bar{\mathbf{n}}^1$ .

According to (12.24), we can now introduce an error measure for the frictional part of the contact in the same way:

$$R_4(\boldsymbol{\sigma}_h; \mathbf{u}_h) = |\mathbf{t}_{T_{n+1}h} - \mathbf{t}_h| \quad \text{on } \partial T \cap \Gamma_c. \quad (12.65)$$

In this relation, the frictional stress  $\mathbf{t}_{T_{n+1}h}$  is computed via the update formula (9.113) or (9.114).

Adding (12.24) and (12.65) to equation (12.25) yields the following error indicator for the contact problem:

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{E^{-1}}^2 \leq \|h C_1 R_1(\boldsymbol{\sigma}_h)\|_{L_2(\Omega)}^2 + \sum_{k=2}^4 \|h C_k R_k(\boldsymbol{\sigma}_h)\|_{L_2(\partial\Omega)}^2. \quad (12.66)$$

Within the finite element discretization, equation (12.66) has to be evaluated on the element domain, which yields

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{E^{-1}}^2 \leq C \sum_T [E_T(h_T, \mathbf{u}_h, \hat{\mathbf{b}}_T)]^2. \quad (12.67)$$

$E_T$  can be computed for each element in the finite element mesh as follows (see (12.28) in Section 12.2 and Wriggers et al. (1994)):

$$\begin{aligned} E_T^2 = & h_T^2 \int_T |\operatorname{div} \boldsymbol{\tau}_h + \hat{\mathbf{b}}|^2 d\Omega + h_T \int_{\partial T \cap \Omega} \frac{1}{2} |[\mathbf{t}_h]|^2 d\Gamma \\ & + h_T \int_{\partial T \cap \Gamma_\sigma} |\hat{\mathbf{t}} - \mathbf{t}_h|^2 d\Gamma + h_T \int_{\partial T \cap \Gamma_c} |\epsilon_N \mathbf{g}_{N_h}^- \bar{\mathbf{n}}^1 + \mathbf{t}_{T_{n+1}h} - \mathbf{t}_h|^2 d\Gamma. \end{aligned} \quad (12.68)$$

Since frictional processes are path-, and with this, time-dependent, an estimation for the magnitude of the time step used to integrate the frictional constitutive equations (9.112) to (9.114) is still missing. In this study we neglect this fact, and prescribe the time steps from outside the adaptive algorithm.

Now we also state the error indicator which can be derived using the superconvergence properties of the projected stress values at the boundary. As before, we can define the projected stresses by any procedure at the nodes, see Figure 12.2. With this we obtain the stress tensor at the node  $m_\alpha$ , see (12.32):

$$\mathbb{P}(\boldsymbol{\sigma}_{\omega_{m_\alpha}}) \longrightarrow \boldsymbol{\sigma}_{m_\alpha}^*.$$

The normal and tangential components of the stress vector then follow from (12.34). As in the case of frictionless contact, we can now define with  $p_{N_\alpha}^*$  and  $t_T^*$  a linear interpolation of the contact stresses:

$$\begin{aligned} p_N^*(\xi) &= (1 - \xi) p_{N_1}^* + \xi p_{N_2}^*, \\ t_T^*(\xi) &= (1 - \xi) t_{T_1}^* + \xi t_{T_2}^*. \end{aligned} \quad (12.69)$$

As in (12.38), this result can now be applied to compute the error associated with the contact stresses, and we arrive at an expression for the additional contact error in  $\partial\Omega^\gamma \cap \Gamma_c = \sum_T \partial T \cap \Gamma_c$ :

$$\|\mathbf{e}\|^{C\gamma} = \sum_T \|\mathbf{e}\|_{\partial T}^{C\gamma_N} + \sum_T \|\mathbf{e}\|_{\partial T}^{C\gamma_T}, \quad (12.70)$$

with the error related to the segment  $m_1-m_2$ :

$$\begin{aligned} \|\mathbf{e}\|_{\partial T}^{C\gamma_N} &= \left\{ \int_0^1 \frac{1}{\epsilon_N} [p_N^*(\xi) - p_{N_h}]^2 l d\xi \right\}^{\frac{1}{2}}, \\ \|\mathbf{e}\|_{\partial T}^{C\gamma_T} &= \left\{ \int_0^1 \frac{1}{c_T} [t_T^*(\xi) - t_{T_{n+1}h}]^2 l d\xi \right\}^{\frac{1}{2}}, \end{aligned} \quad (12.71)$$

where  $p_{N_h}$  is obtained according to (8.13) for the slave  $s$  node which is in contact with the master surface  $m_1-m_2$ , see Figure 12.2. The tangential stress  $t_{T_{n+1}h}$  follows from the algorithmic update formula (9.118).

The error within the whole domain including contact is computed from the sum over all elements  $T$ , with  $\Omega$  being the union of all elements. Thus with (12.39) we have

$$\|\sigma - \sigma_h\|_{E^{-1}} = \sum_T \|\mathbf{e}\|_T. \quad (12.72)$$

Equation (12.70) provides the additional term due to normal and tangential contact stress which has to be included (12.72) in case the element includes a contact interface.

$$\|\mathbf{e}\|_T^C = \|\mathbf{e}\|_T + \|\mathbf{e}\|_{\partial T}^{C\gamma}. \quad (12.73)$$

### 12.7.1 Adaptive strategies

As discussed before, we state the adaptive method as a nonlinear optimization problem: construct a mesh such that the associated FEM solution satisfies

$$\|\mathbf{e}_\tau\| = \|\tau - \tau_h\|_{E^{-1}} = \|\mathbf{u} - \mathbf{u}_h\|_E \leq C \sum_T E_T^2 \leq TOL, \quad (12.74)$$

with  $TOL$  being a given tolerance. Furthermore, the expense of computing  $\mathbf{u}_h$  or  $\tau_h$  should be minimal. The measure  $E_T$  in (12.74) can be defined according to (12.68) or (12.73).

Since the optimal mesh changes with the deformation, the data of the current state of the structure has to be transferred from one mesh to the other in order to achieve an optimal usage of the computational resources.

In the last section, mesh refinement techniques were applied which are based on evolving meshes during deformation. As a criterion for re-meshing, the relative error

$$\eta = \left( \frac{\|\mathbf{e}_\tau\|^2}{\|\mathbf{e}_\tau\|^2 + \|\tau\|^2} \right)^{\frac{1}{2}} \quad (12.75)$$



can be introduced. The error control now works so that this relative error has to be limited to a prescribed tolerance  $\eta_{tol}$ . To obtain a mesh fulfilling the required tolerance  $\eta \leq \eta_{tol}$ , the element size is locally predicted, leading to an equally distributed error in the finite element mesh. The element sizes  $h_{new}$  are then computed from

$$h_{new}^e = h_{old}^e \frac{e_m}{\|\mathbf{e}_\tau\|_e}, \quad (12.76)$$

with

$$\|\mathbf{e}_\tau\|_e \leq e_m = \eta_{tol} [(\|\mathbf{e}_\tau\|^2 + \|\boldsymbol{\tau}_h\|^2)/n_{el}]^{\frac{1}{2}}. \quad (12.77)$$

In contact problems with friction the history variables are present, such as the plastic slip (e.g. see the algorithm in Box 13) which have to be transferred to the new mesh. Thus we have to use special strategies within a nonlinear adaptive process.

If history independent problems are considered, the transfer can be done without problems by projecting the deformation to the new mesh and iterating the nodal values to equilibrium. If history dependent problems, like frictional contact or inelasticity, are considered, internal variables have to be transferred. Such a process, described in Section 12.7.2, cannot be performed without loss of accuracy.

To avoid such a new source of error due to the transfer of state variables, a second strategy can be proposed in which no transfer of variables is needed. Here, for every re-meshing the computation starts again from the first load step. The density of the elements for a new mesh can be defined by the smallest local element size which is obtained during computation over all load steps  $k = 1, \dots, k_{max}$ :

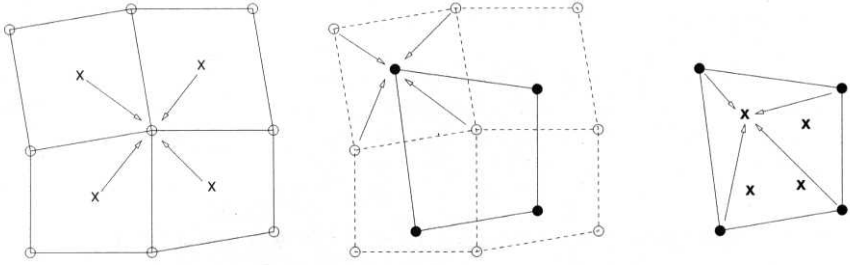
$$h^e = \min_{k=1 \dots k_{max}} \{h_k^e\}. \quad (12.78)$$

The procedure ends if the finite element solution for the adapted mesh is below the prescribed tolerance at all load steps regarding  $\eta_k \leq \eta_{tol}, \forall k$ . This refinement strategy ensures that the local element size is smaller every load step than the element size defined by the average error of a single load step  $e_m$ . Obviously, (12.78) does not lead to optimal meshes, in the sense that the local element size is as small for an actual load step within the complete process.

### 12.7.2 Transfer of history variables

If an adaptive procedure of evolving meshes is considered, the variables of the finite element formulation have to be transmitted to the new mesh. These are the nodal  $\mathbf{u}$  and internal variables in the solid, like plastic strains, or at the contact surface, like frictional slip,  $\mathbf{i} = (\mathbf{E}^p, \mathbf{g}_T^s)$  which have to be updated in each load step. Note that the transfer of the available data  $\mathbf{u}, \mathbf{E}^p, \mathbf{g}_T^s, \dots$  leads in general to an inconsistent dependence of these data on each other in the new mesh.

Several procedures for the history data transfer are possible, e.g. see Ortiz and Quigley (1991) and Lee and Bathe (1994). The procedure which is applied here for the transfer of the inner variables is described in Figure 12.21. The internal variables are projected via the procedures discussed in Section 12.3 onto the nodes in the old



**Fig. 12.21** Transfer of history variables.

mesh (left picture in Figure 12.21). Nodal and internal variables are then transmitted to the nodes in the new mesh (middle picture in Figure 12.21). Finally, the internal variables are interpolated to the GAUSS points (see right picture in Figure 12.21). This procedure is also applied at the contact interface. In that case we have to interpolate data related to a curve in two-dimensional, and related a surface in three-dimensional, problems. The displacements at the contact interface can be obtained by interpolating the displacements from the element nodes. This is usually achieved by isoparametric interpolations of the adjacent finite elements. However, an improved interpolation can be obtained using higher order polynomials, e.g. HERMITE polynomials as discussed in Section 8.5.1. In contact formulations, like the node-to-segment elements, the constitutive data related to the contact interface are stored as nodal values. In that case, the first step in Figure 12.21 can be omitted within the history transfer.

Independent of the choice of transfer procedure, an error in the transferred data cannot be avoided. This error stems from two sources: the transfer error itself, and from the fact that the data in the old mesh are already affected by the discretization error. Therefore, after transfer the discrete weak form of equilibrium is in general violated  $\mathbf{G}^{h+1}(\mathbf{u}^{trans}, \mathbf{i}^{trans}) \neq \mathbf{0}$ . By iterating the nodal variables  $\mathbf{u}$  to equilibrium with the NEWTON-RAPHSON scheme for a fixed set of active contact constraints,

$$\mathbf{G}^{h+1}(\mathbf{u}^{h+1}, \mathbf{i}^{trans}) = \mathbf{0}, \quad (12.79)$$

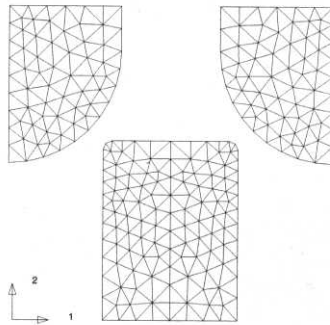
the current state in the new mesh is obtained. It should be mentioned here that the tangential stiffness matrix  $\mathbf{K}_T$  computed with these transferred values can be badly conditioned, especially in the case of large strains. In addition, we may also have a large residuum, hence the solution of (12.79) is not easy to achieve, and may require considerably more iterations than usual. Particularly for sheet metal forming simulations, this can be a severe problem. In this context, the transfer of the current state makes the iteration to equilibrium even more difficult.

In the numerical equilibrium state of the new mesh (12.79), the strains differ in general from those of the old mesh, whereas the internal variables  $\mathbf{i}^{trans}$  remain fixed after transformation to the new mesh. As a consequence, smaller strains can occur in the new mesh, and hence the stresses can also be reduced considerably.

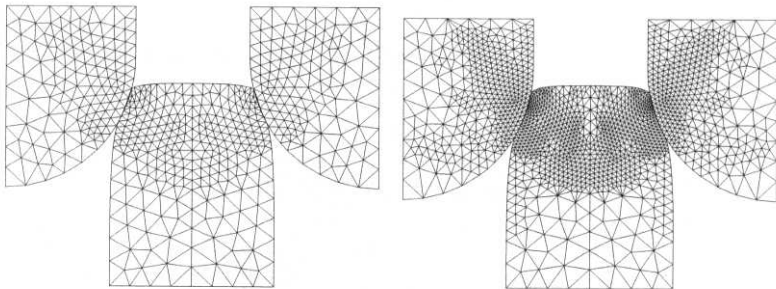
### 12.7.3 Numerical examples

In this section two different problems are discussed. The first is associated with frictional contact, and the second considers finite inelastic strains with contact. Both computations were performed using a version of FEAP which included adaptive finite element analysis.

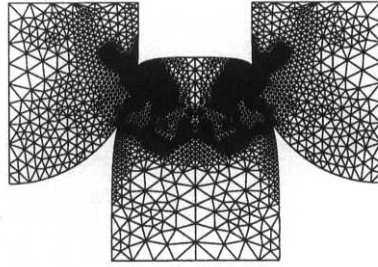
**Rubber sealing.** We consider the problem of a sealing which is in frictional contact with a rigid part. The sealing consists of an elastic block which is pressed into two almost rigid parts, fixed at the straight edges. The block is driven into the rigid parts by a prescribed displacement at the bottom. The problem and its initial mesh are depicted in Figure 12.22. The error due to the time history is not considered here. Thus, error indication is based on the error measure given in (12.68) for stationary problems. Error computation is only performed at specific load steps in the complete load history, which are chosen heuristically from engineering observations of the



**Fig. 12.22** Problem with initial mesh.



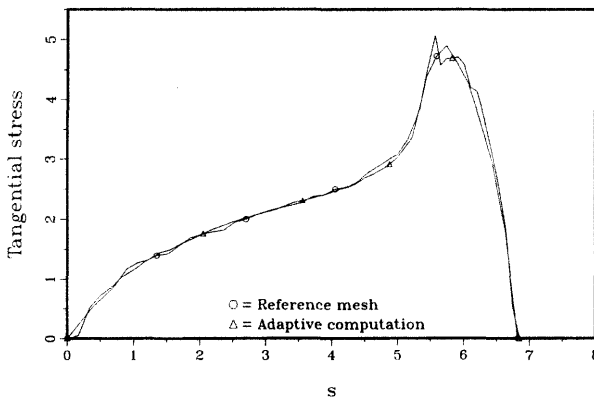
**Fig. 12.23** (a) First refined mesh, (b) Second refined mesh.



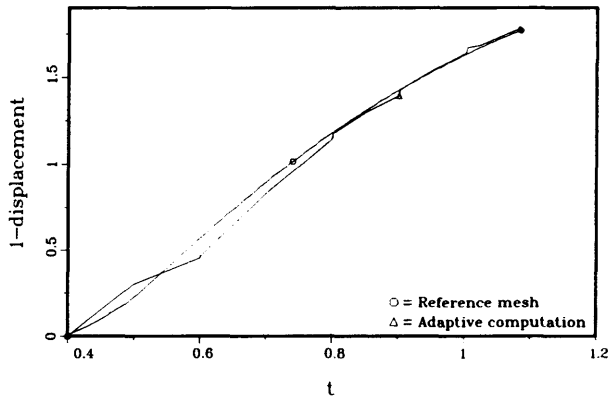
**Fig. 12.24** Third refined mesh.

stress fields. As can be seen from Figure 12.26 by the jumps of the displacements in the 1-direction, the mesh is refined at times  $t = 0.5$ ,  $t = 0.6$ ,  $t = 0.8$ ,  $t = 0.9$  and  $t = 1.0$ . Refined meshes which are obtained at  $t = 0.9$  are depicted in Figures 12.23 and 12.24. These show clearly that the main refinement takes place close to the contact interface, hence the last term in (12.68) governs the error indication.

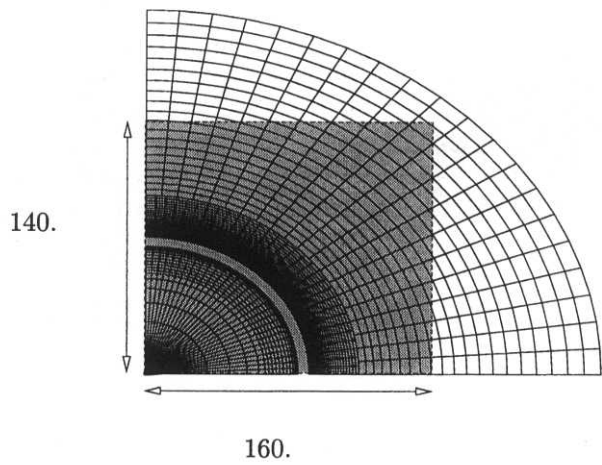
Figure 12.25 shows the contribution of the tangential stress. This has been confirmed by a reference solution using a uniformly refined mesh with over 16,000 degrees of freedom. We see that the adaptive method is more efficient, since it needs only about 5000 degrees of freedom, and hence fewer unknowns, to arrive at the same result. This is also confirmed by Figure 12.26, which shows the evolution of a characteristic displacement in the contact area. Note that the displacements depict jumps from one mesh to the next, which move the new displacements close to the reference



**Fig. 12.25** Tangential stress in contact zone,



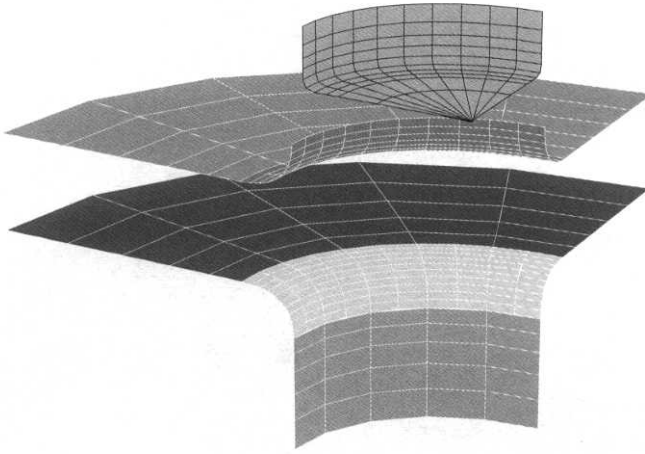
**Fig. 12.26** Evolution of vertical displacement.



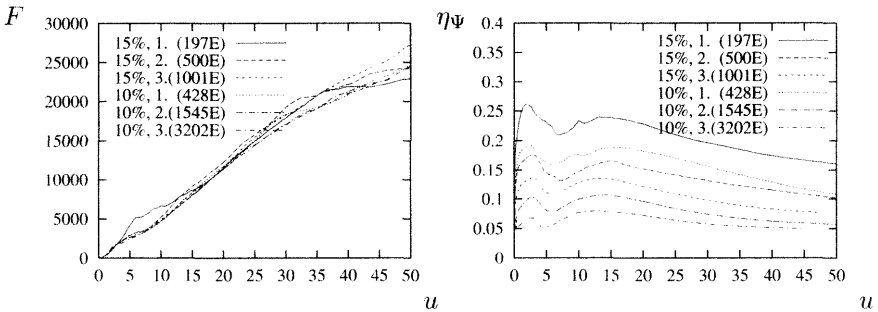
**Fig. 12.27** Deep drawing of a sheet.

solution. Such behaviour is not observed when finite deformations of elasto-plastic solids are treated.

**Deep drawing of a sheet.** The process of deep drawing of a sheet is simulated as illustrated in Figure 12.27, where the grey sheet is located between a punch and a die. The sheet is discretized by shell elements for finite strain plasticity as described in

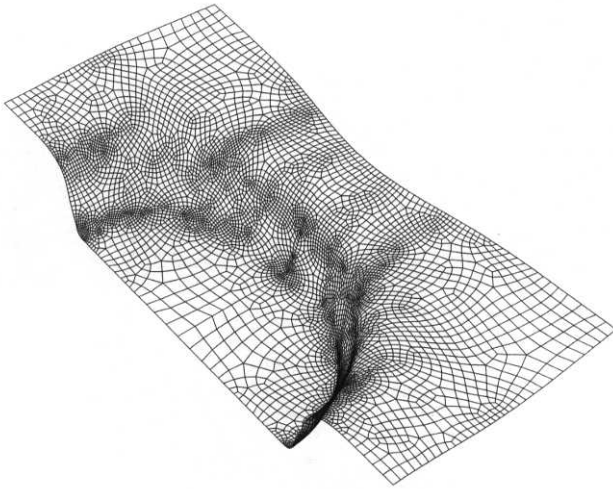


**Fig. 12.28** NURBS-description of the tools.



**Fig. 12.29** Vertical applied contact force and relative error  $\eta_\Psi$  of the deep drawing of the sheet.

Wriggers et al. (1996) and Eberlein and Wriggers (1999). The following geometrical data describe the geometry of the problem (see Figure 12.27): the thickness of the sheet which is  $t = 0.7$ , and the length and height of the sheet which are assumed to be  $L = 160$  and  $H = 140$ . The material data of the sheet are YOUNG'S modulus



**Fig. 12.30** Last mesh in the final configuration with  $TOL = 10\%$ .

$E = 110000$ , POISSON'S  $\nu = 0.3$ , yield stress  $\sigma_Y = 200$  and linear hardening parameter  $h = 1200$ .

The rigid tools, punch, blank holder and die are described with special smooth BÉZIER interpolation, leading to *NURBS*-surfaces, see Figure 12.28. The contact is formulated using a frictionless penalty method. The discretization was derived in Section 8.3.1.

In the simulation, the die and blank holder are positioned near each other with a distance of  $d = 0.1$ . The punch is pushed after the first touch towards the sheet in intervals of  $\Delta u = 0.05$ , until it has reached the final position of  $u = 50.0$ .

The adaptive procedure for this type of contact problem was developed in Han and Wriggers (2002). Large plastic strains occur in the deformation process, and hence large errors due to transfer of the current state are expected, see Section 12.7.2. Therefore, the strategy without transfer is preferable, which computed the errors within a complete computation of the entire load history based on a fixed mesh. After that, the whole process is recalculated using a refined mesh based on the previous error computation.

The adaptive simulations are performed such that the error is limited by 15% and 10% for the relative measure given in (12.75). Again, meshes are chosen within the adaptive procedure which fulfil the tolerance in the first load. The sum over the vertical contact forces of the die, and the relative error of the simulations with respect to  $\eta$  are plotted against the total punch displacement  $u$  in Figure 12.29. Two adaptive re-meshing steps for the sheet were needed to limit the error within the prescribed tolerances of 15% and 10%, respectively. In Figure 12.30 the last mesh of the final configuration which fulfils the prescribed tolerance of 10% is shown.

# 13

---

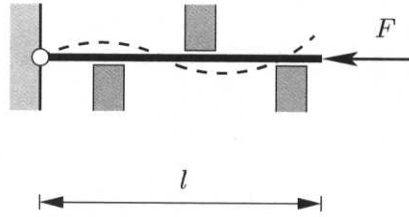
## *Computation of Critical Points with Contact Constraints*

The postcritical behaviour of contact problems cannot be studied in general by using standard path-following methods like the arc-length schemes given by Riks (1972) and others. These methods work well for applications where pure snap-through appears, even if contact takes place during the deformation process, e.g. see Simo et al. (1986). Besides the snap-through behaviour of structures, bifurcation can dominate the response, especially for shell structures. In the case of bifurcation, special algorithms like branch-switching or related techniques have to be employed if secondary branches have to be computed, e.g. see Riks (1984) or Wagner and Wriggers (1988).

An example for bifurcation within a contact problem is given by the buckling of a EULER beam which can come into contact with rigid obstacles, see Figure 13.1. In this case, path-following algorithms with branch-switching and contact formulations have to be modified since several post-critical solutions are possible due to different active contact constraints. These lead to solutions which are associated with different branches. In general, constraints associated with contact are inequalities. Thus, special techniques have to be used to incorporate contact constraints into the formulation.

An interesting aspect is the computation of critical points of a structure when contact constraints are present. In Simo et al. (1986) results of a buckling analysis with unilateral contact using arc-length procedures are published. Due to the emerging inequality constraints, the underlying mathematics is rather complex. Related literature is restricted mainly to unilateral contact problems with rigid obstacles, see Huy and Werner (1986), Endo et al. (1984) or Klarbring and Björkman (1992). In those papers, the mathematics of path following and critical points with inequality constraints originating from unilateral contact problems is studied. Klarbring and





**Fig. 13.1** Buckling of a beam under the presence of contact constraints.

Björkman (1992) describe the occurrence of end points in these cases, where the equilibrium path ends. Those points are not real critical points, in the sense that the stiffness matrix is not singular.

### 13.1 INEQUALITY CONSTRAINTS FOR CONTACT

Let us assume that the structure is already discretized. In this case the constraint equations are given by nodal values. In this context (see Chapter 3) the gap or penetration  $g_s$  associated with a typical slave node  $s$  is indicated by the inequality

$$g_s = (\mathbf{x}_s^2 - \bar{\mathbf{x}}_1^1) \cdot \bar{\mathbf{n}}^1 \geq 0, \quad \forall s \in \mathcal{J}_C. \quad (13.1)$$

$\bar{\mathbf{n}}^1$  denotes the normal to a master segment as defined in Chapter 7,  $\mathbf{x}_s^2$  defines the current position of the slave node, and  $\bar{\mathbf{x}}_1^1$  defines the projection of the slave node  $\mathbf{x}_s^2$  onto the current position of the master segment. Inequality (13.1) has to be checked for all candidate contact nodes  $s$  which are contained in the finite set  $\mathcal{J}_C$  of possible contact nodes. In general, this set of nodes is given by all nodes lying on the surface of the contacting bodies. For  $g_s \leq 0$  the constraint equation for a node  $s$  becomes active ( $s \in \mathcal{J}_A$ ), otherwise the constraint is inactive ( $s \in \mathcal{J}_I$ ) with  $\mathcal{J}_C = \mathcal{J}_A \cup \mathcal{J}_I$  and  $\mathcal{J}_A \cap \mathcal{J}_I = \emptyset$ .

It should be noted that equation (13.1) is valid for the general contact of two or more bodies; see Section 3.1 for a detailed description. Here and in the following, only the case of frictionless contact is considered. The standard penalty method can be applied to enforce constraint (13.1), and the following term has to be appended to the discretized form of total potential energy functional  $\Pi$  (e.g. see Section 5.3.2 equation (5.32)):

$$\Pi_c(\mathbf{v}) = \frac{\epsilon}{2} \sum_{s=1}^{n_c} g_s^2, \quad \forall s \in \mathcal{J}_A, \quad (13.2)$$

with the penalty parameter  $\epsilon \rightarrow \infty$  and  $n_c$  being the number of slave nodes in contact. It should be noted that only the active constraints are incorporated in (13.2).

The variation of (13.2) yields an additional term to the residual vector  $\mathbf{R}$

$$\mathbf{R}_c = \epsilon \sum_{s=1}^{n_c} g_s \mathbf{N}_s, \quad \forall s \in \mathcal{J}_A. \quad (13.3)$$

Here,  $\mathbf{N}_s$  denotes the distribution of the reaction forces in the master element; see Chapter 7 for details. Thus the replacement of inequality (13.1) by the penalty term (13.3) yields the discretized weak form for frictionless contact:

$$\mathbf{G}_c(\mathbf{u}, \lambda) = \mathbf{R}(\mathbf{u}) + \mathbf{R}_c - \lambda \mathbf{P} = \mathbf{0}, \quad (13.4)$$

which is basis for the following study of stability problems with contact.

## 13.2 CALCULATION OF STABILITY POINTS

In this chapter special algorithms for the detection of bifurcation or limit points are summarized. A criterion for stability is that the second derivative of the potential energy  $\Pi$  becomes zero. This is equivalent to saying  $\det \mathbf{K}_{T_i} = 0$ , where in the discrete case  $\mathbf{K}_{T_i} = \left. \frac{\partial \mathbf{G}}{\partial \mathbf{u}} \right|_i$  is the HESSIAN or tangent matrix at a known state  $\mathbf{u}_i$ . For  $\det \mathbf{K}_{T_i} > 0$  the equilibrium path is stable, and for  $\det \mathbf{K}_{T_i} < 0$  it is unstable.

From the mathematical standpoint, different possibilities exist to calculate singular points  $\mathbf{G}(\mathbf{v}, \lambda) = \mathbf{0}$ . One can, for example, use the system

$$\mathbf{G}(\mathbf{u}, \lambda, \phi) = \begin{Bmatrix} \mathbf{G}(\mathbf{u}, \lambda) \\ \mathbf{K}_T(\mathbf{v}, \lambda) \phi \\ l(\phi) - 1 \end{Bmatrix} = \mathbf{0}. \quad (13.5)$$

This extended system has been introduced by Werner and Spence (1984) for the calculation of limit and symmetrical bifurcation points, for instance. In equation (13.5),  $l$  denotes some normalizing functional which prevents the trivial solution  $\phi = \mathbf{0}$ . Extended systems like (13.5) are associated with  $2n$  unknowns, which seem to considerably increase the numerical effort.

The treatment of bifurcation problems, especially the calculation of secondary branches, requires additional considerations. Near stability points the associated eigenvalue problem has to be solved in order to calculate the number of existing branches.

Since the contact constraints are represented by inequalities, we have to solve an inequality eigenvalue problem, which here is represented by the linearized eigenvalue problem contained in (13.5),

$$\mathbf{K}_T(\mathbf{u}) \phi \geq \omega \phi \quad \forall \mathbf{u}, \phi \in \mathcal{V}, \quad (13.6)$$

where  $\mathcal{V} = \{ \mathbf{u} | g_s(\mathbf{u}) \geq 0, \quad \forall s \in \mathcal{J}_C \}$ . To solve this linearized eigenvalue problem we make use of a result obtained by Huy and Werner (1986) for linear variational eigenvalue inequalities.

Each eigensolution  $(\mathbf{u}, \omega)$  of the eigenvalue inequality (13.6) is a solution of the eigenvalue equality

$$(\mathbf{K}_T(\mathbf{u}) - \omega \mathbf{I}) \phi = \mathbf{0} \quad \forall \mathbf{u}, \phi \in \mathcal{W}, \quad (13.7)$$

where  $\mathcal{W} = \{\mathbf{u} | g_s(\mathbf{u}) = 0, \quad \forall s \in \mathcal{J}_A\}$  is a subspace of  $\mathcal{V}$  being defined by the active constraints.

### 13.3 EXTENDED SYSTEM WITH CONTACT CONSTRAINTS

To compute singular points in the presence of contact constraints when frictionless contact is assumed, one can apply algorithms which are based on the idea of the extended system. These are described in detail in Wriggers et al. (1988) and Wriggers and Simo (1990), for example. Applying the extended system procedure for the system defined in (13.5) yields, with the tangent matrix  $\mathbf{K}_{Tc} = \mathbf{K}_T + \mathbf{K}_c$ :

$$\begin{bmatrix} \mathbf{K}_{Tc} & \mathbf{0} & -\mathbf{P} \\ [\mathbf{K}_{Tc} \phi]_{,\nu} & \mathbf{K}_{Tc} & [\mathbf{K}_{Tc} \phi]_{,\lambda} \\ \mathbf{0}^T & \frac{\phi^T}{\|\phi\|} & 0 \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{u} \\ \Delta \phi \\ \Delta \lambda \end{Bmatrix} = - \begin{Bmatrix} \mathbf{G}_S(\mathbf{u}) + \mathbf{G}_C(\mathbf{u}) \\ \mathbf{K}_{Tc}(\mathbf{u}) \phi \\ \|\phi\| - 1 \end{Bmatrix}. \quad (13.8)$$

Here  $\mathbf{K}_c$  is the matrix notation for the derivatives of the contact terms  $\mathbf{G}_c$  with respect to the displacements ( $\mathbf{K}_c = \frac{\partial \mathbf{G}_c}{\partial \mathbf{v}}$ ). The penalty method was chosen to incorporate the contact inequality constraints in the equation system  $\mathbf{G}$ . The vector  $\mathbf{G}_c$  consists of penalty terms that are subjected to the changes in the active set as described in the previous sections. It can be thought of as two strategies for the choice of the active set. The first is a constant verification and reorganization of the active set after each iteration step. The second possibility is a change of the active set only after convergence has been achieved. The first time a constraint becomes active the active set is formed and held fixed until the iteration converges. Thereafter, the active set is verified and changes are made accordingly. The whole process is repeated until the final convergence, where no changes in the active set are necessary. Compared to the first strategy, this means a higher computational effort caused by the additional iteration loop. On the other hand, the iteration process might become more stable by this second approach as changes in the active set occur less frequently.

The partitioning algorithm can be used without major difficulties for the solution of equation system (13.8); see Box 14. The numerical derivative that is used to compute the vectors  $\mathbf{h}_1$  and  $\mathbf{h}_2$  deserves special attention:

$$\begin{aligned} \mathbf{h}_1 &\approx \frac{1}{\epsilon} [\mathbf{K}_{Tc}(\mathbf{u} + \epsilon \phi) \Delta \mathbf{u}_P - \mathbf{P}], \\ \mathbf{h}_2 &\approx \frac{1}{\epsilon} [\mathbf{K}_{Tc}(\mathbf{u} + \epsilon \phi) \Delta \mathbf{u}_G + \mathbf{G}_S(\mathbf{u}) + \mathbf{G}_c(\mathbf{u})]. \end{aligned} \quad (13.9)$$

The tangent matrices  $\mathbf{K}_c(\mathbf{u})$  and  $\mathbf{K}_c(\mathbf{u} + \epsilon \phi)$  in (13.9) can differ substantially if the displacement values  $\mathbf{u}$  and  $\mathbf{u} + \epsilon \phi$  cause a change in the active set. The result is that

Solve:

$$[\mathbf{K}_{Tc}(\mathbf{v}^i)] \Delta \mathbf{u}_P^{i+1} = \mathbf{P}$$

$$[\mathbf{K}_{Tc}(\mathbf{u}^i)] \Delta \mathbf{u}_G^{i+1} = -\mathbf{G}_S(\mathbf{u}^i) - \mathbf{G}_c(\mathbf{u}^i)(\mathbf{u}^i)$$

Compute:

$$\mathbf{h}_1^i = \frac{1}{\epsilon} [(\mathbf{K}_{Tc}(\mathbf{u}^i + \epsilon \phi^i) \Delta \mathbf{u}_P^{i+1} + -\mathbf{P}]$$

$$\mathbf{h}_2^i = \frac{1}{\epsilon} [(\mathbf{K}_{Tc}(\mathbf{u}^i + \epsilon \phi^i) \Delta \mathbf{u}_G^{i+1} + \mathbf{G}_S(\mathbf{u}^i) + \mathbf{G}_c(\mathbf{u}^i)]$$

Solve:

$$[\mathbf{K}_{Tc}(\mathbf{v}^i)] \Delta \phi_1^{i+1} = -\mathbf{h}_1^i$$

$$[\mathbf{K}_{Tc}(\mathbf{u}^i)] \Delta \phi_2^{i+1} = -\mathbf{h}_2^i$$

Compute increments:

$$\Delta \lambda^{i+1} = \frac{-\phi^{iT} \Delta \phi_2^{i+1} + \|\phi^i\|}{\phi^{iT} \Delta \phi_1^{i+1}}$$

$$\Delta \mathbf{u}^{i+1} = \Delta \lambda^{i+1} \Delta \mathbf{u}_P^{i+1} + \Delta \mathbf{u}_G^{i+1}$$

Update:

$$\lambda^{i+1} = \lambda^i + \Delta \lambda^{i+1}$$

$$\mathbf{u}^{i+1} = \mathbf{u}^i + \Delta \mathbf{u}^{i+1},$$

$$\phi^{i+1} = \Delta \lambda^i \Delta \phi_1^{i+1} + \Delta \phi_2^{i+1}$$

Box 14 Partitioning algorithm for the extended system with contact.

the structure of the matrix  $\mathbf{K}_c$  is no longer the same. To prevent this, the assembly of  $\mathbf{K}_c(\mathbf{u} + \epsilon \phi)$  is made based on the active set used for  $\mathbf{K}_c(\mathbf{u})$ . In other words, when the numerical derivative is computed, no new search for the closest master segment of each slave node is performed. The active set is not changed either. The disadvantage is that penetration can occur, or that adhesive forces are applied for a degree of freedom in the active set whose gap was closed but now becomes open. But taking into account the order of magnitude of  $\epsilon \phi$  with  $\epsilon = 10^{-7}$ , these effects are negligible.

### 13.4 EXAMPLES

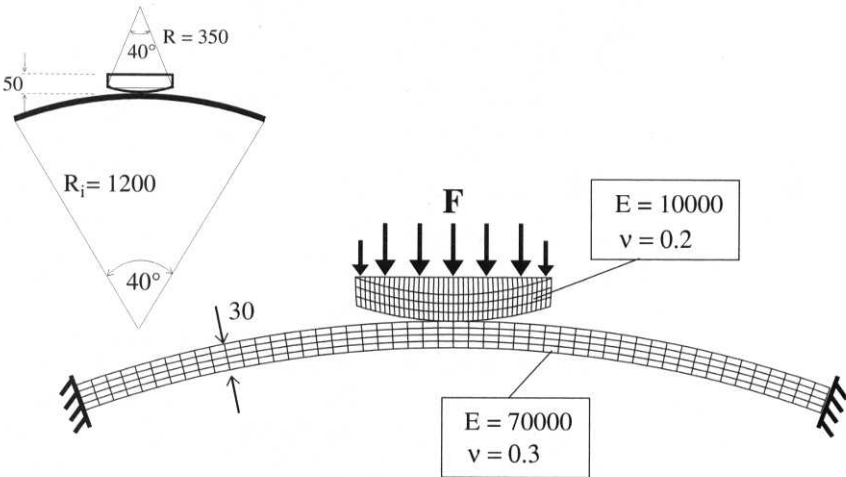
In this section only the extended system will be studied for various continuum mechanical contact problems. All the examples of this section were computed with the finite element program FEAP, where the contact algorithm of Section 9.3.3 and the extended system procedure of Section 13.3 were implemented.

#### 13.4.1 Block pressing on arch

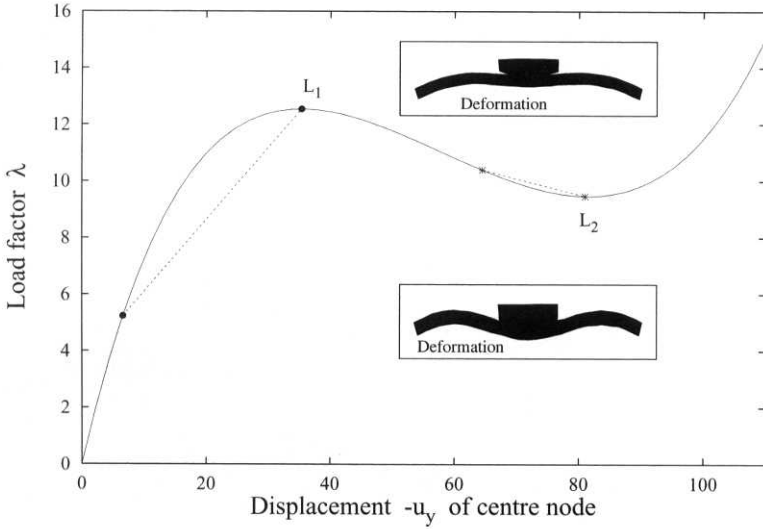
A block pressing on an arch is the next example with contact. The geometrical and material data are to be seen in Figure 13.2. The block is loaded with a unit load in all the nodes, except for the corner nodes, where half of the load is specified. The block is located on top of the arch. The discretization of the arch is 20 layers with 400 elements each, and the block has 11 layers with 300 elements each. Besides the arch in Figure 13.2 that is clamped, the example was also computed with a hinged arch. A sideways movement of the block is prevented by fixating the  $x$ -displacements of the upper corner nodes of the block.

Figure 13.3 shows the equilibrium path and the typical behaviour of a clamped arch with two limit points. The extended system yields for the given discretization the exact coordinates with  $L_1 : (u_y; \lambda) = (-35.35; 12.54)$  and  $L_2 : (u_y; \lambda) = (-80.95; 9.46)$ . The symmetric deformation of the arch is given in the smaller pictures.

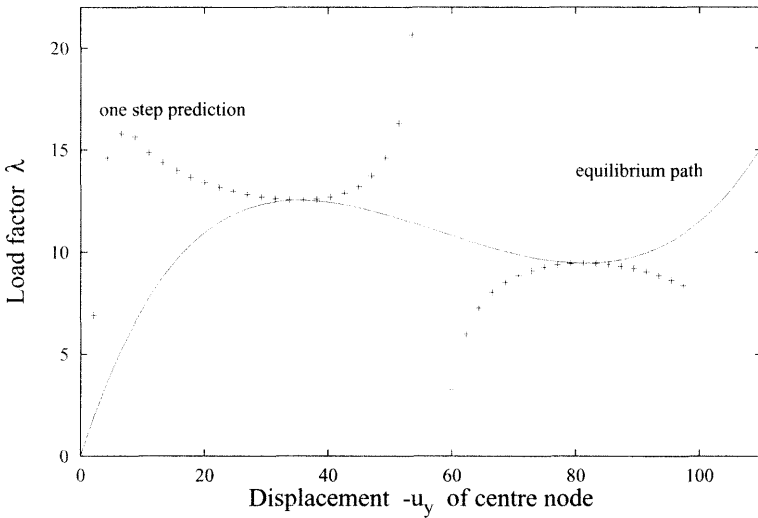
The one step predictions in Figure 13.4 are quite good; the curve goes through both points. A comparison of one step prediction with Figure 13.4 confirms the convergence radius of the extended system computations.



**Fig. 13.2** Block pressing on clamped arch.



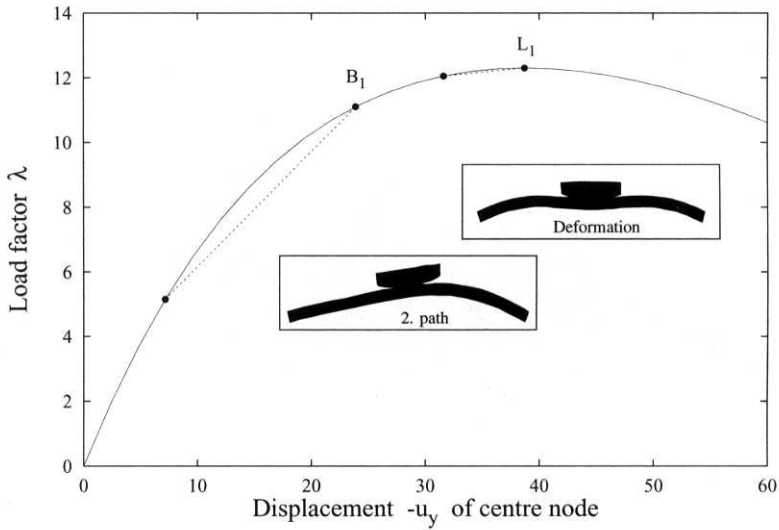
**Fig. 13.3** Equilibrium path for the block pressing on a clamped arch.



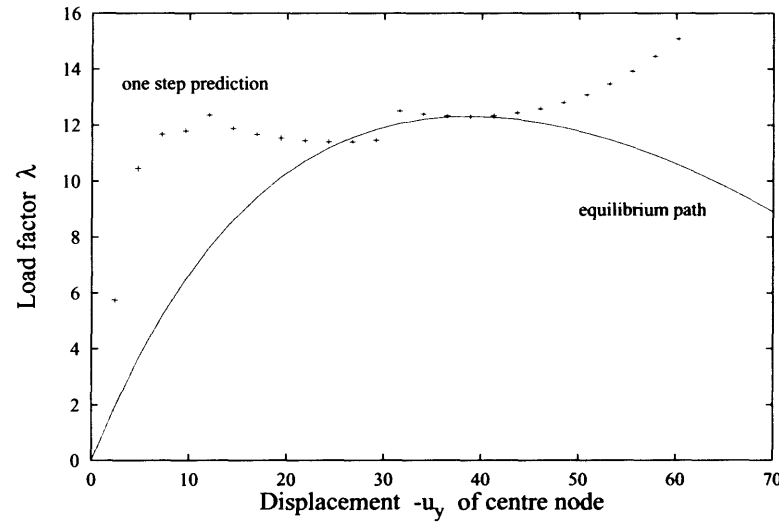
**Fig. 13.4** One step predictions for the block pressing on a clamped arch.

Changing the boundary conditions for the arch in such a way that the ends are hinged enables bifurcation of the equilibrium path, see Figure 13.5. The extended system has no problem in locating the bifurcation point  $B_1 : (u_y; \lambda) = (-23.88; 11.11)$  and the limit point  $L_1 : (u_y; \lambda) = (-38.69; 123.05)$ . The scaled up secondary

path deformation associated with the eigenvector in  $B_1$  can be seen in Figure 13.5. The good results of the one step prediction (Figure 13.6) are in accordance with the convergence radius of the extended system.

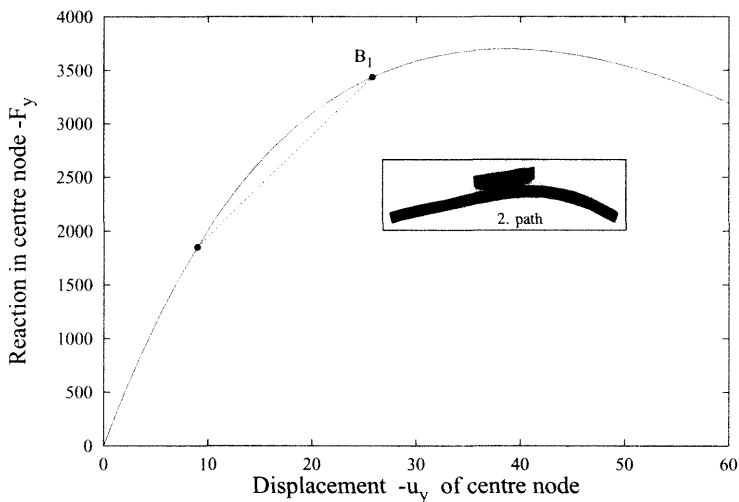


**Fig. 13.5** Equilibrium path for the block pressing on a hinged arch.

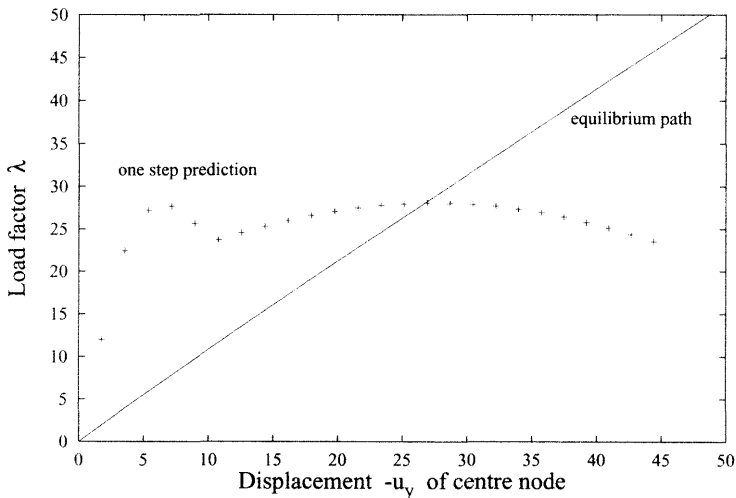


**Fig. 13.6** One step predictions for the block pressing on a hinged arch.

A further interesting load case of this example is the hinged arch with the block on top of it and a unit displacement prescribed in the upper centre node of the block. The equilibrium path in Figure 13.7 has no limit point, only the bifurcation point  $B_1 : (u_y; F_y) = (-25.78; -3435)$  is left.  $F_y$  is the reaction force in the upper centre node of the block. In Figure 13.8, the one step prediction for this load case is given.



**Fig. 13.7** Equilibrium path for the hinged arch with displacement boundary conditions



**Fig. 13.8** One step predictions for the hinged arch with displacement boundary conditions



13.4.2    Two arches

The next example is the structure of Figure 13.9 with two arches. The lower arch is clamped, and the sideways movement ( $x$ -direction) of the upper arch is prevented by the boundary conditions in the corner nodes, where the unit loads are applied.

Since only four node elements with linear shape functions are used, the question arises as to whether the elimination chosen is sufficient for a proper computation of the structure. Therefore, a convergence test with different discretizations is carried out to see how many elements are necessary. Figure 13.10 shows the results for the three discretizations with 1800, 14,000 and 30,000 elements. The curves demonstrate that

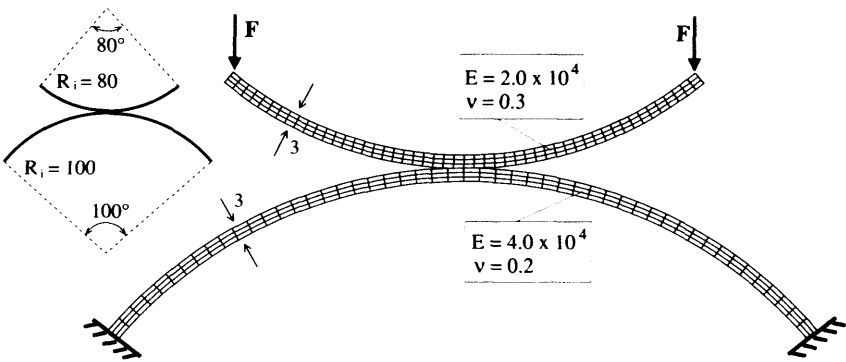


Fig. 13.9    Outline of the two arches.

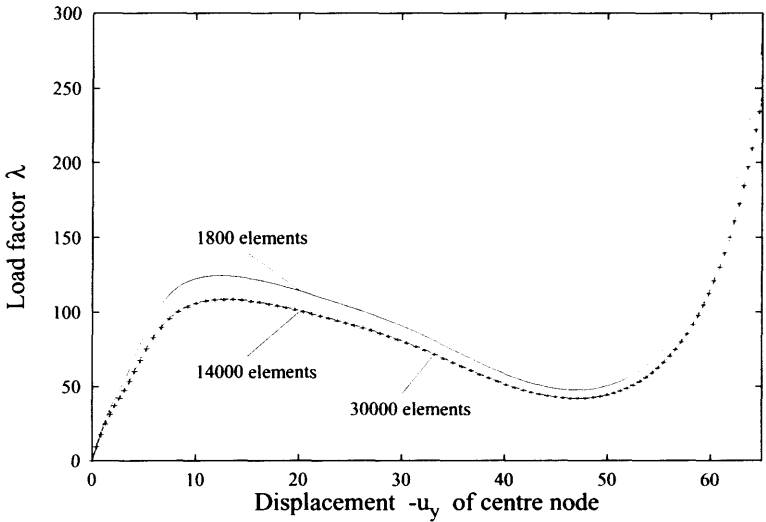
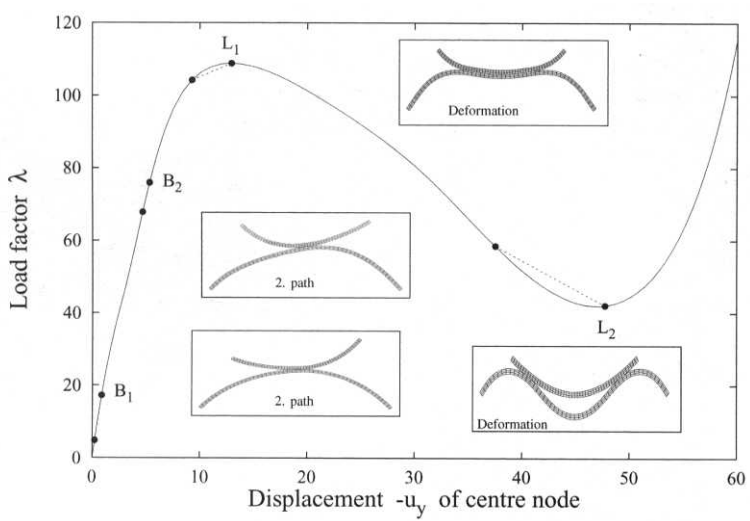
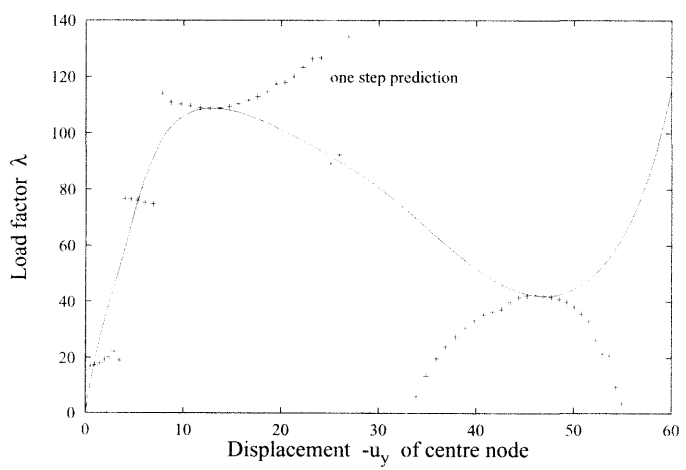


Fig. 13.10    Equilibrium path of the two arches structure for different discretizations.

there is practically no difference between the 14,000 and the 30,000 element curve, so 14,000 elements can be considered as sufficient for this example.



**Fig. 13.11** Equilibrium path with critical points.



**Fig. 13.12** One step prediction for the two arches structure.

Looking in detail at the equilibrium path of Figure 13.11 reveals four critical points. Two bifurcation points with  $B_1 : (u_y; \lambda) = (-0.85; 17.21)$  and  $B_2 : (u_y; \lambda) = (-4.77; 75.94)$ ; and two limit load points with  $L_1 : (u_y; \lambda) = (-12.94; 108.83)$  and  $L_2 : (u_y; \lambda) = (-47.73; 42.14)$  are found. The deformed arches and the secondary path deformation, respectively, are depicted in the small pictures next to each critical point.

A comparison of these results with the one step prediction in Figure 13.12 clearly demonstrates the switching of the prediction from one critical point to another. Again, the similarity of convergence radius of the extended system and the one step prediction curve is apparent.

# *Appendix A*

## *Gauss integration rules*

In finite element computations we always have to evaluate integrals, e.g. the weak form or the tangent matrices. These integrations can be performed on the element level. Since isoparametric elements are usually employed for the discretization, an exact integration is no longer possible. Thus we need numerical integrations. These are usually performed on the reference element  $\Omega_{\square}$ , see Chapter 6. Here we focus especially on contact problems, hence the integration rules are only stated up to two dimensions. These are needed in Chapter 7. For three-dimensional rules applied for three-dimensional solids we refer to Zienkiewicz and Taylor (2000b) or Dhett and Touzot (1985).


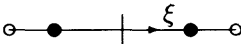

### **A.1 ONE-DIMENSIONAL INTEGRATION**

Since integration is carried out in finite element analysis in the reference configuration,  $\xi \in [-1, +1]$ , all values have to be transformed to this configuration:

$$\int_{(X)} g(X) dX = \int_{-1}^{+1} g(\xi) \frac{dX}{d\xi} d\xi = \int_{-1}^{+1} g(\xi) J_e(\xi) d\xi. \quad (\text{A.1})$$

$g(\xi)$  is the function which has to be integrated and  $J_e$  is the JACOBIAN of the transfor-

**Table A.1** One-dimensional GAUSS integration

$n_p$	$p$	$\xi_p$	$W_p$	
1	1	0	2	
2	1	$1/\sqrt{3}$	1	
	2	$1/\sqrt{3}$	1	
3	1	$-\sqrt{3/5}$	5/9	
	2	0	8/9	
	3	$+\sqrt{3/5}$	5/9	

mation to the reference configuration which can be computed using the isoparametric map defined in section 6.1. The integration will be done numerically, since  $g(\xi) J_e(\xi)$  is in general no longer a polynomial. Hence, the integral (A.1) will be approximated by the sum

$$\int_{-1}^{+1} g(\xi) J_e(\xi) d\xi \approx \sum_{p=1}^{n_p} g(\xi_p) J_e(\xi_p) W_p. \quad (\text{A.2})$$

$W_p$  are weighting factors and  $\xi_p$  denote the coordinates of the evaluation points. The locations  $\xi_p$  and the weighting factors  $W_p$  are stated in Table A.1 up to the order of  $n_p = 3$  for a GAUSS integration.

Polynomials of order  $p = 2n_p - 1$  are integrated exactly by  $n_p$  evaluation points. These rules can be used for two-dimensional contact elements, e.g. see Section 7.3.

## A.2 TWO-DIMENSIONAL INTEGRATION

For evaluation of the weak form in (2.59) which is valid for two-dimensional problems, or for evaluation of the contact element (7.60) or (7.65), we need an integration of the interpolation functions and its derivatives over the element domain  $\Omega_e$ . For this purpose, it is advantageous to transform the integral to the  $\xi$ - $\eta$  coordinate system in the reference element  $\Omega_\square$ :

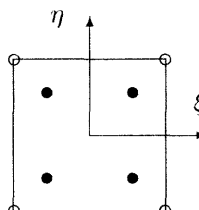
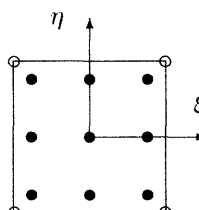
$$\int_{(\Omega_e)} g(\mathbf{X}) dA = \int_{(\Omega_\square)} g(\xi) \det \mathbf{J}_e(\xi) d\xi d\eta = \int_{-1}^{+1} \int_{-1}^{+1} g(\xi, \eta) \det \mathbf{J}_e d\xi d\eta. \quad (\text{A.3})$$

Integration over  $\Omega_{\square}$  is performed by a numerical quadrature formula, since the product  $g(\xi) \det \mathbf{J}_e(\xi)$  does in general not yield a polynomial. Thus, we obtain

$$\int_{-1}^{+1} \int_{-1}^{+1} g(\xi, \eta) \det \mathbf{J}_e d\xi d\eta \approx \sum_{p=1}^{n_p} g(\xi_p, \eta_p) \det \mathbf{J}_e(\xi_p, \eta_p) W_p. \quad (\text{A.4})$$

The weighting factors  $W_p$  and the coordinates of the quadrature points  $\xi_p$  and  $\eta_p$  are contained in Table A.2 for a GAUSS quadrature up to a number of  $n_p = 3 \times 3$  points. These integration rules are exact for polynomials up to the order  $i + k \leq m$ . We note that the integration rules follow from the one-dimensional integration rules via a product formula. Usually, GAUSS rules are applied in finite element computations due to their accuracy. Thus we do not discuss other rules here. More quadrature rules can be found in Dhatt and Touzot (1985), for example.

**Table A.2** Two-dimensional GAUSS quadrature for rectangular elements

$m$	$n_p$	$p$	$\xi_p$	$\eta_p$	$W_p$	Position of points
1	1	1	0	0	4	
3	4	1	$-1/\sqrt{3}$	$-1/\sqrt{3}$	1	
		2	$+1/\sqrt{3}$	$-1/\sqrt{3}$	1	
		3	$-1/\sqrt{3}$	$+1/\sqrt{3}$	1	
		4	$+1/\sqrt{3}$	$+1/\sqrt{3}$	1	
5	9	1	$-\sqrt{3/5}$	$-\sqrt{3/5}$	25/81	
		2	0	$-\sqrt{3/5}$	40/81	
		3	$+\sqrt{3/5}$	$-\sqrt{3/5}$	25/81	
		4	$-\sqrt{3/5}$	0	40/81	
		5	0	0	64/81	
		6	$+\sqrt{3/5}$	0	40/81	
		7	$-\sqrt{3/5}$	$+\sqrt{3/5}$	25/81	
		8	0	$+\sqrt{3/5}$	40/81	
		9	$+\sqrt{3/5}$	$+\sqrt{3/5}$	25/81	

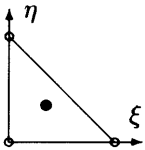
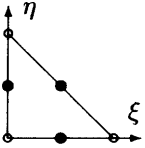
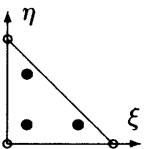
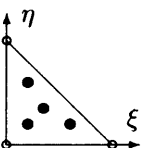
The transformation to the reference element is different for triangular elements. In general, we obtain the following relation:

$$\int_{(\Omega_e)} g(\mathbf{X}) \, dA = \int_0^1 \int_0^{1-\xi} g(\xi, \eta) \, \det \mathbf{J}_e \, d\eta \, d\xi,$$

(A.5)

which again can be evaluated using the quadrature rule (A.4). Table A.3 contains the associated quadrature points and weighting factors for an element with side length 1. The formulas are exact for polynomials  $\xi^k \eta^l$  up to the order  $m$  (with  $m \geq k + l$ ). Again, different quadrature rules with different quadrature points or higher accuracy can be found in Zienkiewicz and Taylor (1989) or Dhatt and Touzot (1985), for example.

**Table A.3** Two-dimensional GAUSS quadrature for triangular elements

$m$	$n_p$	$p$	$\xi_p$	$\eta_p$	$W_p$	Position of points
1	1	1	1/3	1/3	1/2	
2	3	1	1/2	1/2	1/6	
		2	0	1/2	1/6	
		3	1/2	0	1/6	
2	3	1	1/6	1/6	1/6	
		2	2/3	1/6	1/6	
		3	1/6	2/3	1/6	
3	4	1	1/3	1/3	-27/96	
		2	1/5	1/5	25/96	
		3	3/5	1/5	25/96	
		4	1/5	3/5	25/96	

# Appendix B

## Convective Coordinates

Frequently, in computational solid mechanics a special parameterization of the bodies under investigation is performed using convective coordinates. This is also convenient to formulate the contact constraint equations for finite deformations, see Chapter 3. This is because the deformed surface of a body which has to be considered for the set up of contact constraints can be described best by convective coordinates.

The three-dimensional as well as two-dimensional formulation of a surface will be stated here using convective coordinates. These can be viewed as coordinates which are attached to the body, and thus deform with the body, as shown in Figure B.1. In general, the idea is to write the cartesian coordinates  $\{X_A\}$  and  $\{x_i\}$  as functions of the convective coordinates  $\{\Theta^j\}$ ,  $j = 1, 2, 3$ ,

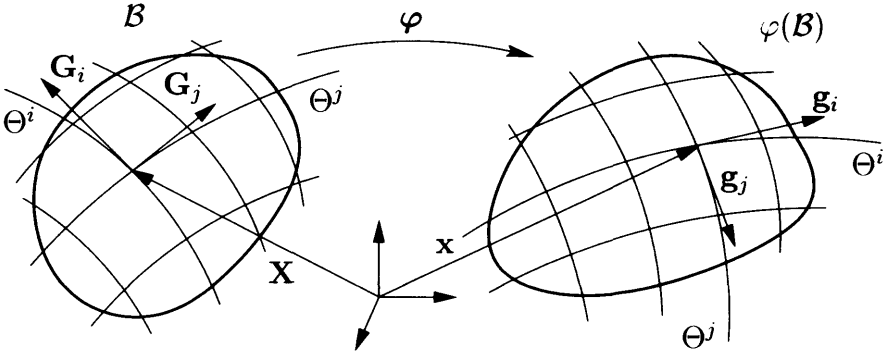
$$X_A = \hat{X}_A(\Theta^1, \Theta^2, \Theta^3), \quad x_i = \hat{x}_i(\Theta^1, \Theta^2, \Theta^3), \quad (\text{B.1})$$

which can then be applied to formulate the equation for a solid in convective coordinates. In compact form we state  $\mathbf{X} = \hat{\mathbf{X}}(\Theta^j)$  and  $\mathbf{x} = \hat{\mathbf{x}}(\Theta^j)$ . If we want to describe a surface using such coordinates, then only two coordinates are necessary, as will be shown later.

A tangent vector to the coordinate lines in a point  $\mathbf{X}$  in  $\mathcal{B}$  is computed via

$$\mathbf{G}_j = \frac{\partial \mathbf{X}}{\partial \Theta^j} = \mathbf{X}_{,j}. \quad (\text{B.2})$$





**Fig. B.1** Convective coordinates in  $B$  and  $\varphi(B)$ .

Assuming sufficient smoothness, the same holds for a point which is described by  $\varphi(\mathbf{X}, t)$  in  $\varphi(B)$ :

$$\mathbf{g}_j = \frac{\partial \varphi(\mathbf{X}, t)}{\partial \Theta^j} = \varphi_{,j}. \quad (\text{B.3})$$

Using the chain rule, together with (B.2) and (B.3),

$$\mathbf{g}_j = \frac{\partial \varphi(\mathbf{X}, t)}{\partial \mathbf{X}} \frac{\partial \mathbf{X}}{\partial \Theta^j} = \mathbf{F} \mathbf{G}_j \quad (\text{B.4})$$

follows. This means that the tangent vectors have the same transformation rule as the line elements  $d\mathbf{x}$  and  $d\mathbf{X}$ , see equation (2.5). With (B.4) we obtain the following form of the deformation gradient:

$$\mathbf{F} = \mathbf{g}_i \otimes \mathbf{G}^i. \quad (\text{B.5})$$

The tangent vectors are covariant vectors which are connected to the contravariant vectors via  $\mathbf{g}_i \cdot \mathbf{g}^k = \delta_i^k$ . Using equation (B.4), we observe

$$\mathbf{F} \mathbf{G}_i \cdot \mathbf{A} \mathbf{G}^k = \delta_i^k \longrightarrow \mathbf{A} = \mathbf{F}^{-T}, \quad (\text{B.6})$$

with the transformation tensor  $\mathbf{A} = \mathbf{F}^{-T}$  for the contravariant base vectors.

Hence, it is relatively easy to apply the *pull back* operation and the *push forward* in the convective setting. The four possible transformations are

$$\begin{aligned} \mathbf{g}_i &= \mathbf{F} \mathbf{G}_i & \mathbf{G}_i &= \mathbf{F}^{-1} \mathbf{g}_i \\ \mathbf{g}^i &= \mathbf{F}^{-T} \mathbf{G}^i & \mathbf{G}^i &= \mathbf{F}^T \mathbf{g}^i. \end{aligned} \quad (\text{B.7})$$

Since covariant and contravariant base vectors denote the base for vectors and tensors, equations (B.7) can be applied for *pull back* and *push forward* operations. As an

example, we consider the *push forward* GREEN-LAGRANGIAN strain tensor, which yields the transformation (2.22)

$$\begin{aligned} \mathbf{E} &= \frac{1}{2} (g_{ik} - G_{ik}) \mathbf{G}^i \otimes \mathbf{G}^k \\ &= \frac{1}{2} (g_{ik} - G_{ik}) \mathbf{F}^T \mathbf{g}^i \otimes \mathbf{F}^T \mathbf{g}^k = \mathbf{F}^T \left[ \frac{1}{2} (g_{ik} - G_{ik}) \mathbf{g}^i \otimes \mathbf{g}^k \right] \mathbf{F}. \end{aligned} \quad (\text{B.8})$$

For contact one has to describe the surface  $S$  of the bodies in contact. This can be done by introducing a position vector of a point on the surface in the current configuration by

$$\varphi_S(\xi^1, \xi^2, t), \quad (\text{B.9})$$

where  $\xi^\alpha$ , the Greek index  $\alpha$  has values of 1 and 2, are the convective coordinates, and  $\varphi_S$  describes the deformed surface of a solid. The same procedure also holds for the initial or reference configuration:

$$\mathbf{X}_S(\xi^1, \xi^2, t). \quad (\text{B.10})$$

The covariant base vectors with regard to the surface description in the initial configuration introduced above are (B.2)

$$\mathbf{A}_\alpha(\xi^1, \xi^2, t) = \frac{\partial \mathbf{X}_S}{\partial \xi^\alpha} = \mathbf{X}_{S,\alpha}, \quad (\text{B.11})$$

where  $\mathbf{A}_\alpha$  are vectors tangent to the surface  $\mathcal{M}$ . These can be used to compute the normal vector

$$\mathbf{N}(\xi^1, \xi^2, t) = \frac{\mathbf{A}_1(\xi^1, \xi^2, t) \times \mathbf{A}_2(\xi^1, \xi^2, t)}{\|\mathbf{A}_1(\xi^1, \xi^2, t) \times \mathbf{A}_2(\xi^1, \xi^2, t)\|}. \quad (\text{B.12})$$

In the current configuration, the tangent vectors follow from (B.9) in the same way as in (B.11):

$$\mathbf{a}_\alpha(\xi^1, \xi^2, t) = \frac{\partial \varphi_S}{\partial \xi^\alpha} = \varphi_{S,\alpha}. \quad (\text{B.13})$$

The normal vector is defined by the cross product

$$\mathbf{n}(\xi^1, \xi^2, t) = \frac{\mathbf{a}_1(\xi^1, \xi^2, t) \times \mathbf{a}_2(\xi^1, \xi^2, t)}{\|\mathbf{a}_1(\xi^1, \xi^2, t) \times \mathbf{a}_2(\xi^1, \xi^2, t)\|}. \quad (\text{B.14})$$

Now one can define the metric tensor  $a_{\alpha\beta}$  and the curvature tensor  $b_{\alpha\beta}$  of the deformed surface as follows:

$$\begin{aligned} a_{\alpha\beta} &= \mathbf{a}_\alpha \cdot \mathbf{a}_\beta, \\ b_{\alpha\beta} &= -\mathbf{a}_\alpha \cdot \mathbf{n}_{,\beta} = \mathbf{a}_{\alpha,\beta} \cdot \mathbf{n}. \end{aligned} \quad (\text{B.15})$$

Using the metric tensor, one can introduce the contravariant base vectors

$$\mathbf{a}^\alpha = a^{\alpha\beta} \mathbf{a}_\beta \quad \text{with} \quad a^{\alpha\beta} a_{\beta\gamma} = \delta_\gamma^\alpha. \quad (\text{B.16})$$

A vector  $\mathbf{v}$  which lies in the tangent plane can be stated either with respect to the covariant or contravariant base,

$$\mathbf{v} = v^\alpha \mathbf{a}_\alpha = v_\alpha \mathbf{a}^\alpha \quad \text{with} \quad v_\alpha a^{\alpha\beta} = v^\beta. \quad (\text{B.17})$$

For an arbitrary three-dimensional vector  $\mathbf{w}$ , one has the representation

$$\mathbf{w} = w^\alpha \mathbf{a}_\alpha + w^3 \mathbf{n} = \mathbf{w}_\alpha \mathbf{a}^\alpha + w_3 \mathbf{n}, \quad (\text{B.18})$$

where  $w^3 = w_3$  and  $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{n}\}$  are the base vectors.

The derivative of a vector  $\mathbf{w}$  with respect to the convective coordinates yields

$$\frac{\partial \mathbf{w}}{\partial \xi^\beta} = \mathbf{w}_{,\beta} = w_{,\beta}^\alpha \mathbf{a}_\alpha + w^\alpha \mathbf{a}_{\alpha,\beta} + w_{,\beta}^3 \mathbf{n} + w^3 \mathbf{n}_{,\beta}. \quad (\text{B.19})$$

The derivatives of the base vectors  $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{n}\}$  with respect to the convective coordinates have to be considered here, since the base vectors are not constant on arbitrary surfaces. The derivation of the base vectors can be expressed using the CHRISTOFFEL tensors,  $\Gamma_{\alpha\beta}^\delta$ , and the metric and curvature tensors from (B.15):

$$\mathbf{a}_{\alpha,\beta} = \Gamma_{\alpha\beta}^\delta \mathbf{a}_\delta + b_{\alpha\beta} \mathbf{n} \quad \text{with} \quad \Gamma_{\alpha\beta}^\delta = \mathbf{a}_{\alpha,\beta} \cdot \mathbf{a}^\delta, \quad (\text{B.20})$$

$$\mathbf{n}_{,\alpha} = -b_\alpha^\beta \mathbf{a}_\beta \quad \text{with} \quad b_\alpha^\beta = a^{\alpha\delta} b_{\delta\beta}. \quad (\text{B.21})$$

However, as we usually approximate the surfaces in this book either by the isoparametric formulation or by other functions (e.g. BÉZIER polynomials; see Section 8.5), we can compute the derivatives in (B.19) directly in the finite element discretization process. Hence, the CHRISTOFFEL tensors  $\Gamma_{\alpha\beta}^\delta$  do not have to be computed explicitly; for details see Eringen (1962).

When using the convective description, the following conditions can be applied to simplify the formulations:

$$\mathbf{a}_\alpha \cdot \mathbf{n} = 0,$$

$$\frac{d}{dt} [\mathbf{n} \cdot \mathbf{n}] = \mathbf{n} \cdot \dot{\mathbf{n}} = 0, \quad (\text{B.22})$$

$$\frac{d}{dt} [\mathbf{a}_\alpha \cdot \mathbf{a}^\beta] = \mathbf{a}_\alpha \cdot \dot{\mathbf{a}}^\beta + \dot{\mathbf{a}}_\alpha \cdot \mathbf{a}^\beta = 0. \quad (\text{B.23})$$

To model the frictional response, one has to introduce the differential of the arc-length  $ds$  of a curve  $\Theta^\alpha(t)$  (the sliding distance) on the surface. It is given by

$$ds = \sqrt{a_{\alpha\beta} \dot{\Theta}^\alpha \dot{\Theta}^\beta} dt. \quad (\text{B.24})$$

The arc-length  $s$  between two parameter values  $t_0$  and  $t_1$  follows from

$$s = \int_{t_0}^{t_1} \sqrt{a_{\alpha\beta} \dot{\Theta}^\alpha \dot{\Theta}^\beta} dt. \quad (\text{B.25})$$

# *Appendix C*

## *Parameter Identification for Friction Materials*

A parameter identification is needed when experimental data of friction tests have to be fitted to an existing constitutive relation for friction.

The identification of the friction material parameters is based on the minimization of a least squares functional. For this purpose, the following optimization problem is regarded:

$$\sum_i (t_{Ti} - \tilde{t}_{Ti})^2 \rightarrow Min, \quad (C.1)$$

with the experimental results  $\tilde{t}_{Ti}$  and the tangential stresses  $t_{Ti}$  evaluated using the new frictional law depending on the material parameters  $a_j$ .

A requirement for a minimum of this functional is

$$\sum_i (t_{Ti} - \tilde{t}_{Ti}) \frac{dt_{Ti}}{da_j} = 0. \quad (C.2)$$

To evaluate (C.2), it is necessary to calculate the frictional stresses as well as their derivatives with respect to the material parameters. Depending on the type of problem, (C.2) leads either to a linear system of equations for the unknown parameter values or to a nonlinear one. In the latter case, a NEWTON type procedure has to be applied

to solve this equation. The necessary ingredients for the case of frictional contact are provided next.

To obtain the frictional stresses and their derivatives, the following frictional law in a general form, including hardening (here only the two-dimensional case is considered), is used:

$$\begin{aligned}
 \text{linear elastic (stick) law: } t_T &= \epsilon_T g_T^e = \epsilon_T (g_T - g_T^p) \\
 \text{slip rule } \dot{g}_T^p &= \lambda \frac{\partial f_s}{\partial t_T} \\
 \text{hardening } \dot{w} &= \lambda r(t_T, w) \\
 \text{Kuhn-Tucker conditions } f_s \leq 0, \quad \lambda \geq 0, \quad f_s \lambda &= 0.
 \end{aligned} \tag{C.3}$$

In this set of equations the slip function  $f_s$  and the hardening function  $r(t_T, w)$  still have to be specified; for examples see Section 4.2. Here, for demonstration purposes, a special choice is made for the slip function:

$$f_s = |t_T| - t_N \left[ \frac{a_1 w}{1 + a_2 w + a_3 w^2} + a_4 \arctan(a_5 w) \right], \tag{C.4}$$

and the hardening function

$$r = 1, \tag{C.5}$$

leading to

$$\begin{aligned}
 \text{slip rule } \dot{g}_T^p &= \lambda \frac{\partial f_s}{\partial t_T} = \lambda \text{sign}(t_T) \\
 \text{hardening } \dot{w} &= \lambda r(t_T, w) = \lambda.
 \end{aligned} \tag{C.6}$$

To solve the parameter identification problem, a time discretization, according to Section 9.5, first has to be carried out:

$$\gamma = \lambda \Delta t \tag{C.7}$$

$$\begin{aligned}
 t_{T\,n+1} &= \epsilon_T (g_{T\,n+1} - g_{T\,n+1}^p) \\
 g_{T\,n+1}^p &= g_{T\,n}^p + \gamma_{n+1} \frac{\partial f_s}{\partial t_T} \\
 w_{n+1} &= w_n + \gamma_{n+1} r(t_T, w).
 \end{aligned} \tag{C.8}$$

This result can be transformed into a nonlinear system of equations with three unknowns  $t_{T\,n+1}$ ,  $w_{n+1}$  and  $\gamma_{n+1}$ :

$$\mathbf{R} = \left\{ \begin{array}{c} f_{s\,n+1} \\ -w_{n+1} + w_n + \gamma_{n+1} r(t_T, w) \\ \frac{t_{T\,n+1}}{\epsilon_T} - g_{T\,n+1} + g_{T\,n}^p + \gamma_{n+1} \frac{\partial f_s}{\partial t_T} \end{array} \right\} = \mathbf{0}. \tag{C.9}$$

For the special choice of the slip function (C.4), this yields

$$\mathbf{R} = \left\{ \begin{array}{c} f_{s\,n+1} \\ -w_{n+1} + w_n + \gamma_{n+1} \\ \frac{t_{T\,n+1}}{\epsilon_T} - g_{T\,n+1} + g_{T\,n}^p + \gamma_{n+1} \text{sign}(t_{T\,n+1}) \end{array} \right\} = \mathbf{0}. \tag{C.10}$$

To obtain the frictional stress  $t_T$ ,  $\mathbf{R} = \mathbf{0}$  is solved with NEWTON's method:

$$\begin{aligned}
 \mathbf{H} \Delta \phi^{i+1} &= -\mathbf{R} \\
 \phi^{i+1} &= \phi^i + \Delta \phi^{i+1}.
 \end{aligned} \tag{C.11}$$

with  $\mathbf{R}$  from (C.9) or (C.10) and

$$\phi = \begin{Bmatrix} t_{T\,n+1} \\ w_{n+1} \\ \gamma_{n+1} \end{Bmatrix}. \quad (\text{C.12})$$

The tangent matrix  $\mathbf{H}$  is given by

$$\mathbf{H} = \begin{bmatrix} \frac{\partial f_s}{\partial t_T} & \frac{\partial f_s}{\partial w} & 0 \\ \gamma \frac{\partial r}{\partial t_T} & -1 + \gamma \frac{\partial r}{\partial w} & r \\ \frac{1}{\epsilon_T} + \gamma \frac{\partial^2 f_s}{\partial t_T^2} & \gamma \frac{\partial^2 f_s}{\partial t_T \partial w} & \frac{\partial f_s}{\partial t_T} \end{bmatrix} \quad (\text{C.13})$$

for  $\mathbf{R}$  from (C.9), and results in

$$\mathbf{H} = \begin{bmatrix} \text{sign } t_T & -t_N \left( \frac{a_1(1-a_3w^2)}{(1+a_2w+a_3w^2)^2} + \frac{a_4a_5}{1+a_5^2w^2} \right) & 0 \\ 0 & -1 & 1 \\ \frac{1}{\epsilon_T} & 0 & \text{sign } t_T \end{bmatrix} \quad (\text{C.14})$$

for  $\mathbf{R}$  from (C.10).

Now the derivatives of the frictional stresses with respect to the material parameters  $a_i$  of the frictional law have to be calculated.

The total differentiation of  $\mathbf{R}$  with respect to the material parameters yields

$$\frac{d\mathbf{R}}{da_j} = \frac{\partial \mathbf{R}}{\partial \phi_{n+1}} \frac{d\phi_{n+1}}{da_j} + \frac{\partial \mathbf{R}}{\partial \phi_n} \frac{d\phi_n}{da_j} + \frac{\partial \mathbf{R}}{\partial a_j}. \quad (\text{C.15})$$

Here  $d\phi / da_j$  are the sensitivities of the variables  $\phi$  (C.12) from the frictional law. Note that the sensitivities of the last time step have to be taken into account in (C.15), because of the history dependence of the frictional law.

(C.15) yields an equation system for the derivatives of the frictional stresses with respect to the material parameters  $d\phi_{n+1} / da_j$ :

$$\mathbf{H} \hat{\phi}_{j\,n+1} = -\frac{\partial \mathbf{R}}{\partial a_j} - \frac{\partial \mathbf{R}}{\partial \phi_n} \frac{d\phi_n}{da_j} \quad (\text{C.16})$$

with

$$\hat{\phi}_{j\,n+1} = \begin{Bmatrix} \frac{dt_{T\,n+1}}{da_j} \\ \frac{dw_{n+1}}{da_j} \\ \frac{d\gamma_{n+1}}{da_j} \end{Bmatrix}. \quad (\text{C.17})$$

**This page intentionally left blank**

# References

- Agelet de Saracibar C. A new frictional time integration algorithm for large slip multi-body frictional contact problems. *Computer Methods in Applied Mechanics and Engineering*, 142:303–334 (1997).
- Agelet de Saracibar C. Numerical analysis of coupled thermomechanical frictional contact problems, computational model and applications. *Archive of Computational Methods in Engineering*, 5:243–302 (1998).
- Agelet de Saracibar C. and Chiumenti M. Numerical analysis of frictional wear contact problems, computational model and applications. *Computer Methods in Applied Mechanics and Engineering*, 177:404–426 (1999).
- Alart P. and Curnier A. A mixed formulation for frictional contact problems prone to Newton like solution methods. *Computer Methods in Applied Mechanics and Engineering*, 92:353–375 (1991).
- Amontons G. On the resistance originating in machines (in French). *Mem. Acad. Roy.*, pages 206–222 (1699).
- Anand L. A constitutive model for interface friction. *Computational Mechanics*, 12:197–213 (1993).
- Archard J.F. Contact and rubbing of flat surfaces. *Journal of Applied Physics*, 24:981–988 (1953).
- Argyris J.H. *Energy Theorems and Structural Analysis*. Butterworth (1960).



- Argyris J.H. and Doltsinis J.S. On the natural formulation and analysis of large deformation coupled thermomechanical problems. *Computer Methods in Applied Mechanics and Engineering*, 25:195–253 (1981).
- Armero F. and Petocz E. Formulation and analysis of conserving algorithms for frictionless dynamic contact/impact problems. *Computer Methods in Applied Mechanics and Engineering*, 158:269–300 (1998).
- Armero F. and Petocz E. A new dissipative time-stepping algorithm for frictional contact problems. *Computer Methods in Applied Mechanics and Engineering*, 179:151–178 (1999).
- Arslan U. *Zur Frage des elastoplastischen Verformungsverhaltens von Sand*. Dissertation, Versuchsanstalt für Bodenmechanik und Grundbau der Technischen Hochschule Darmstadt (1980). Mitteilung 23.
- Attig N. and Esser R. *Molecular Dynamics on Parallel Computers*. World Scientific (1999).
- Avitzur B. and Nakamura Y. Analytical determination of friction resistance as a function of normal load and geometry of surface irregularities. *Wear*, 107:367–383 (1986).
- Axelsson O. *Iterative Solution Methods*. Cambridge University Press (1994).
- Babuska I. and Miller A. A feedback finite element method with a posteriori error estimation: Part i. the finite element method and some basic properties of the a posteriori error estimation. *Computer Methods in Applied Mechanics and Engineering*, 61:1–40 (1987).
- Babuska I. and Rheinboldt W. Error estimates for adaptive finite element computations. *SIAM Journal on Numerical Analysis*, 15:736–754 (1978).
- Babuska I., Strouboulis T., Upadhyay C.S., Gangaraj S.K. and Copps K. Validation of a posteriori error estimators by numerical approach. *International Journal for Numerical Methods in Engineering*, 37:1073–1123 (1994).
- Bandeira A.A., Wriggers P. and de Mattos Pimenta P. Homogenization methods leading to interface laws of contact mechanics. *International Journal for Numerical Methods in Engineering* (2002).
- Bank R.E. PLTMG: A software package for solving elliptic partial differential equations. *Technical Report Vol. 7*, Society for Industrial and Applied Mathematics, Philadelphia (1990).
- Barnsley M. *Fractal Everywhere*. Academic Press, Boston, New York (1988).
- Barnsley M. and Hurd L. *Fractal Image Compression*. AK Peters, Wellesley, Massachusetts (1993).
- Barthold F.J. and Bischoff D. Generalization of Newton type methods to contact problems with friction. *Journal Mec. Theor. Appl.*, pages 97–110 (1988).
- Bartholomew-Biggs M.C. and Hernandez F.G. Using the the kkt-matrix in the augmented Lagrangian SQP method for sparse constraint optimization. *Journal of Optimization Theory and Applications*, 85:201–220 (1995).

- Bathe K.J. *Finite Element Procedures in Engineering Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey (1982).
- Bathe K.J. *Finite Element Procedures*. Prentice-Hall, Englewood Cliffs, New Jersey (1996).
- Bathe K.J. and Bolourchi S. Large displacement analysis of three-dimensional beam structures. *International Journal for Numerical Methods in Engineering*, 14:961–986 (1979).
- Bathe K.J. and Chaudhary A.B. A solution method for planar and axisymmetric contact problems. *International Journal for Numerical Methods in Engineering*, 21:65–88 (1985).
- Bathe K.J., Ramm E. and Wilson E.L. Finite element formulation for large deformation analysis. *International Journal for Numerical Methods in Engineering*, 9:353–386 (1975).
- Bazaraa M.S., Sherali H.D. and Shetty C.M. *Nonlinear Programming, Theory and Algorithms*. Wiley, Chichester, second edition (1993).
- Becker R. and Hansbo P. A finite element method for domain decomposition with non-matching grids. *Technical Report No. 3613*, INRIA Sophia Antipolis (1999).
- Becker R. and Rannacher R. A feed-back approach to error control in finite element methods: Basic analysis and examples. *EAST-WEST J. Numerical Mathematics*, 4:237–264 (1996).
- Belgacem F.B., Hild P. and Laborde P. Approximation of the unilateral contact problem by the mortar finite element method. *C. R. Acad. Sci., Paris, Ser I*, 324:123–127 (1997).
- Belgacem F.B., Hild P. and Laborde P. Extension of the mortar finite element method to a variational inequality modelling unilateral contact. *Mathematical Methods in Applied Sciences*, 9:287–303 (1999).
- Belsky V. A multi-grid method for variational inequalities in contact problems. *Computing*, 51:293–311 (1993).
- Belytschko T., Chiapetta T. and Bartel R.L. Efficient large-scale non-linear transient analysis by finite elements. *International Journal for Numerical Methods in Engineering*, 10:579–596 (1976).
- Belytschko T. and Neal M.O. Contact–impact by the pinball algorithm with penalty and lagrange methods. *International Journal for Numerical Methods in Engineering*, 31:547–572 (1991).
- Benson D.J. and Hallquist J.O. A single surface contact algorithm for the post-buckling analysis of shell structures. *Computer Methods in Applied Mechanics and Engineering*, 78:141–163 (1990).
- Bernadi C., Maday Y. and Patera A. A new nonconforming approach to domain decomposition: the mortar element method. In H. Brezis and J.L. Lions, editors, *Nonlinear Partial Differential Equations and their Application*. Pitman (1994).
- Bertsekas D.P. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York (1984).

- Betsch P. and Steinmann P. Conservation properties of a time finite element method. part i: Time-stepping schemes for n-body problems. *International Journal for Numerical Methods in Engineering*, 49:599–638 (2000).
- Bittencourt E. and Creus G.J. Finite element analysis of three-dimensional contact and impact in large deformation problems. *Computers and Structures*, 69:219–234 (1998).
- Björkman G. The solution of large displacement frictionless contact problems using a sequence of linear complementary problems. *International Journal for Numerical Methods in Engineering*, 31:1553–1566 (1991).
- Björkman G., Klarbring A., Sjödin B., Larsson T. and Rönqvist M. Sequential quadratic programming for non-linear elastic contact problems. *International Journal for Numerical Methods in Engineering*, 38:137–165 (1995).
- Boersma A. and Wriggers P. An algebraic multigrid solver for finite element computations in solid mechanics. *Engineering Computations*, 14:202–215 (1997).
- Boley B.A. and Weiner J.H. *Theory of Thermal Stresses*. Dover Publications, Mineola, New York (1997).
- Bonnet J. and Peraire J. An Alternating Digital Tree (ADT) algorithm for 3d geometric searching and intersection problems. *International Journal for Numerical Methods in Engineering*, 31:1–17 (1991).
- Bowden F.P. and Tabor D. *The Friction and Lubrication of Solids, Part II*. Clarendon Press, Oxford (1964).
- Bufler H. Pressure loaded structures under large deformations. *Zeitschrift für angewandte Mathematik und Mechanik*, 64:287–295 (1984).
- Camacho G.T. and Ortiz M. Computational modelling of impact damage in brittle materials. *International Journal of Solids & Structures*, 33:2899–2938 (1996).
- Carstensen C., Scherf O. and Wriggers P. Adaptive finite elements for elastic bodies in contact. *SIAM Journal of Scientific Computing*, 20:1605–1626 (1999).
- Chadwick P. *Continuum Mechanics, Concise Theory and Problems*. Dover Publications, Mineola (1999).
- Chan S.H. and Tuba I.S. A finite element method for contact problems in solid bodies. *International Journal of Mechanical Sciences*, 13:615–639 (1971).
- Christensen P.W., Klarbring A., Pang J.S. and Strömberg N. Formulation and comparison of algorithms for frictional contact problems. *International Journal for Numerical Methods in Engineering*, 42:145–173 (1998).
- Christensen R.M. A nonlinear theory of viscoelasticity for application to elastomers. *Journal of Applied Mechanics*, 47:762–768 (1980).
- Ciarlet P.G. *Mathematical Elasticity I: Three-dimensional Elasticity*. North-Holland, Amsterdam (1988).

- Cirak F., Ortiz M. and Schröder P. Subdivision surfaces: a new paradigm for thin shell finite-element analysis. *International Journal for Numerical Methods in Engineering*, 47:2039–2072 (2000).
- Cocu M. Existence of solutions of Signorini problems with friction. *International Journal of Engineering Science*, 22:567–575 (1984).
- Conry T.F. and Seireg A. A mathematical programming method for design of elastic bodies in contact. *Journal of Applied Mechanics*, 38:1293–1307 (1971).
- Cooper M.G., Mikic B.B. and Yovanovich M.M. Thermal contact conductance. *International Journal of Heat and Mass Transfer*, 12:279–300 (1969).
- Cottle R.W. and Dantzig G.B. Complementary pivot theory of mathematical programming. *Linear Algebra and its Applications*, 1:103–125 (1968).
- Cottle R.W., Pang J.S. and Stone R.E. *The Linear Complementary Problem*. Academic Press, Boston (1992).
- Coulomb C.A. The theory of simple machines (in french). *Mem. Math. Phys. Acad. Sci.*, 10:161–331 (1785).
- Courtney-Pratt J.S. and Eisner E. The effect of a tangential force on the contact of metallic bodies. *Proceedings of the Royal Society of London*, 238-A:529–550 (1957).
- Crisfield M.A. A fast incremental/iterative solution procedure that handles snap through. *Computers and Structures*, 13:55–62 (1981).
- Crisfield M.A. A consistent co-rotational formulation for non-linear, three-dimensional beam elements. *Computer Methods in Applied Mechanics and Engineering*, 81:131–150 (1990).
- Crisfield M.A. *Non-linear Finite Element Analysis of Solids and Structures*, volume 1. Wiley, Chichester (1991).
- Crisfield M.A. *Non-linear Finite Element Analysis of Solids and Structures*, volume 2. Wiley, Chichester (1997).
- Crisfield M.A. and Shi J. A review of solution procedures and path-following techniques in relation to the non-linear finite element analysis of structures. In P. Wriggers and W. Wagner, editors, *Computational Methods in Nonlinear Mechanics*. Springer, Berlin (1991).
- Curnier A. A theory of friction. *International Journal of Solids & Structures*, 20:637–647 (1984).
- Curnier A. Unilateral contact: Mechanical modelling. In P. Wriggers and P. Panagiotopoulos, editors, *New Developments in Contact Problems*, pages 1–54. Springer, Wien (1999).
- Curnier A. and Alart P. A generalized newton method for contact problems with friction. *J. Mec. Theor. Appl.*, 7:67–82 (1988).
- Curnier A., He Q.C. and Klarbring A. Continuum mechanics modelling of large deformation contact with friction. In M. Raous, M. Jean and J. Moreau, editors, *Contact Mechanics*, pages 145–158. Plenum Press (1995).

- Curnier A., He Q.C. and Telega J.J. Formulation of unilateral contact between two elastic bodies undergoing finite deformation. *C. R. Acad. Sci. Paris*, 314:1–6 (1992).
- Dalrymple T. An analytical overlay technique for describing deformable contact surfaces. In W. Wunderlich, editor, *Proceedings of the European Conference on Computational Mechanics*. Institut für Statik, München (1999).
- Davis T.A. and Duff I.S. A combined unifrontal/multifrontal method for unsymmetric sparse matrices. *ACM Transactions on Mathematical Software*, 25:1–19 (1999).
- De Saxe G. and Feng Z.Q. New inequality and functional for contact with friction: the implicit standard material approach. *Mech. Struct. Mach.*, 19:301–325 (1991).
- Desaguliers J.T. Some experiments concerning the cohesion of lead. *Philosophical Transactions of the Royal Society of London*, 33:345 (1725).
- Desai C.S. and Siriwardane H.J. *Constitutive Laws for Engineering Materials*. Prentice-Hall, Englewood Cliffs, New Jersey (1984).
- Dhatt G. and Touzot G. *The Finite Element Method Displayed*. Wiley, Chichester (1985).
- Diekmann R., Hungershöver J., Lux M., Taenzer L. and Wierum J.M. Efficient contact search for finite element analysis. In E. Onate, editor, *Proceedings of ECCOMAS 2000*. CIMNE, Barcelona (2000).
- Doltsinis I.S. Aspects of modelling and computation in the analysis of metal forming. *Engineering Computations*, 7:2–20 (1990).
- Dostal Z. On preconditioning and penalized matrices. *Numerical Linear Algebra with Applications*, 6:109–114 (1999).
- Dowson D. *History of Tribology*. Longman, New York (1979).
- Duff I.S., Erisman A.M. and Reid J.K. *Direct Methods for Sparse Matrices*. Clarendon Press, Oxford (1989).
- Duvaut G. and Lions J.L. *Inequalities in Mechanics and Physics*. Springer Verlag, Berlin (1976).
- Eberlein R. and Wriggers P. Finite element concepts for finite elastoplastic strains and isotropic stress response in shells: Theoretical and computational analysis. *Computer Methods in Applied Mechanics and Engineering*, 171:243–279 (1999).
- El-Abbasi N. and Bathe K.J. Stability and patch test performance of contact discretizations and a new solution algorithm. *Computers and Structures*, 79:1473–1486 (2001).
- Endo T., Oden J.T., Becker E.B. and Miller T. A numerical analysis of contact and limit-point behaviour in a class of problems of finite elastic deformations. *Computers and Structures*, 18:899–910 (1984).
- Eringen A. *Nonlinear Theory of Continuous Media*. McGraw-Hill, New York, London (1962).
- Eringen A. *Mechanics of Continua*. Wiley, New York, London, Sidney (1967).

- Eterovic A.L. and Bathe K.J. An interface interpolation scheme for quadratic convergence in the finite element analysis of contact problems. In *Computational Methods in Nonlinear Mechanics*, pages 703–715. Springer-Verlag, Berlin, New York (1991).
- Euler L. Sur la diminution de la resistance du frottement. *Mem. Acad. Sci. Berlin*, 4:133–148 (1748a).
- Euler L. Sur le frottement des corps solides. *Mem. Acad. Sci. Berlin*, 4:122–132 (1748b).
- Farin G. *Curves and Surfaces for Computer Aided Geometric Design. A Practical Guide*. Academic Press, Boston, third edition (1993).
- Feder J. *Fractals*. Plenum Press, New York (1988).
- Francavilla A. and Zienkiewicz O.C. A note on numerical computation of elastic contact problems. *International Journal for Numerical Methods in Engineering*, 9:913–924 (1975).
- Fredriksson B. Finite element solution of surface nonlinearities in structural mechanics with special emphasis to contact and fracture mechanics problems. *Computers and Structures*, 6:281–290 (1976).
- Fremont M. Adherence des solides. *Journal des Mechanique Theoretique eet Appliquee*, 6:383–407 (1987).
- Giannakopoulos A.E. The return mapping method for the integration of friction constitutive relations. *Computers and Structures*, 32:157–168 (1989).
- Gill P.E. and Murray W. Numerically stable methods for quadratic programming. *Mathematical Programming*, 14:349–372 (1978).
- Glowinski R. and Le Tallec P. Finite element analysis in nonlinear incompressible elasticity. In *Finite Element, Vol. V: Special Problems in Solid Mechanics*. Prentice-Hall, Englewood Cliffs, New Jersey (1984).
- Goldfarb D. and Idnani A. A numerically stable dual method for solving strictly quadratic programs. *Mathematical Programming*, 27:1–33 (1983).
- Govindjee S. and Mihalic P.A. Viscoelastic constitutive relations for the steady spinning of a cylinder. *Technical Report UCB/SEMM-98/02*, University of California Berkeley (1998).
- Greenwood J.A. and Williamson J.B.P. The contact of normally-flat surfaces. *Proceedings of the Royal Society of London*, 295-A:300–379 (1966).
- Gross D., Hauger W., Schnell W. and Wriggers P. *Technische Mechanik 4*. Springer, Berlin, third edition (1999).
- Gruttmann F., Sauer R. and Wagner W. Theory and numerics of three-dimensional beams with elastoplastic material behaviour. *International Journal for Numerical Methods in Engineering*, 48:1675–1702 (2000).
- Hackbusch W. *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. Teubner Verlag, Stuttgart (1991).

- Hallquist J.O. Nike2d: An implicit, finite-deformation, finite element code for analysing the static and dynamic response of two-dimensional solids. *Technical Report UCRL-52678*, University of California, Lawrence Livermore National Laboratory (1979).
- Hallquist J.O., Goudreau G.L. and Benson D.J. Sliding interfaces with contact-impact in large-scale lagrange computations. *Computer Methods in Applied Mechanics and Engineering*, 51:107–137 (1985).
- Hallquist J.O., Schweizerhof K. and Stillman D. Efficiency refinements of contact strategies and algorithms in explicit fe programming. In D.R.J. Owen, E. Hinton and E. Onate, editors, *Proceedings of COMPLAS III*, pages 359–384. Pineridge Press (1992).
- Han C.S. and Wriggers P. On the error indication of shells in unilateral frictionless contact. *Computational Mechanics*, 28:169–176 (2002).
- Hansson E. and Klarbring A. Rigid contact modelled by CAD surface. *Engineering Computations*, 7:344–348 (1990).
- Haraldsson A. and Wriggers P. A strategy for numerical testing of frictional laws with application to contact between soil and concrete. *Computer Methods in Applied Mechanics and Engineering*, 190:963–978 (2001).
- He Q.C. and Curnier A. Anisotropic dry friction between two orthotropic surfaces undergoing large displacements. *European Journal of Mechanics, A/Solids*, 12:631–666 (1993).
- Heegaard J.H. and Curnier A. An augmented Lagrange method for discrete large-slip contact problems. *International Journal for Numerical Methods in Engineering*, 36:569–593 (1993).
- Heege A. and Alart P. A frictional contact element for strongly curved contact problems. *International Journal for Numerical Methods in Engineering*, 39:165–184 (1996).
- Heinstein M.W. and Laursen T.A. An algorithm for the matrix-free solution of quasistatic frictional contact problems. *International Journal for Numerical Methods in Engineering*, 44:1205–1226 (1999).
- Heinstein M.W., Mello F.J., Attaway S.W. and Laursen T.A. Contact-impact modeling in explicit transient dynamics. *Computer Methods in Applied Mechanics and Engineering*, 187:621–640 (2000).
- Hertz H. Über die Berührung fester elastischer Körper. *Journal für die Reine und Angewandte Mathematik*, 29:156–171 (1882).
- Hilber H., Hughes T.R.J. and Taylor R.L. Improved numerical dissipation for time integration algorithms in structural dynamics. *Earthquake Engineering and Structural Dynamics*, 5:283–292 (1977).
- Hlavacek I., Haslinger J., Necas J. and Lovisek J. *Solution of Variational Inequalities in Mechanics*. Springer-Verlag, Berlin, New York (1988).
- Hogue C. Shape representation and contact detection for discrete element simulations of arbitrary geometries. *Engineering Computations*, 15:374–390 (1998).

- Holm R. *Electric Contacts*. Almqvist and Wiksells, Stockholm (1946).
- Holmberg G. A solution scheme for three-dimensional multi-body contact problems using mathematical programming. *Computers and Structures*, 37:503–514 (1990).
- Hoppe R.H.W. and Kornhuber R. Adaptive multilevel-methods for obstacle problems. *SIAM, Journal of Numerical Analysis*, 31:301–323 (1994).
- Hu G.D., Panagiotopoulos P.D., Panagouli O., Scherf O. and Wriggers P. Adaptive finite element analysis of fractal interfaces in contact problems. *Computer Methods in Applied Mechanics and Engineering*, 182:17–38 (2000).
- Hu G.D. and Wriggers P. On the adaptive finite element method of steady-state rolling contact for hyperelasticity in finite deformations. *Computer Methods in Applied Mechanics and Engineering*, 191:1333–1348 (2002).
- Hughes T.R.J. *The Finite Element Method*. Prentice Hall, Englewood Cliffs, New Jersey (1987).
- Hughes T.R.J., Taylor R.L. and Kanoknukulchai W. A finite element method for large displacement contact and impact problems. In K.J. Bathe, editor, *Formulations and Computational Algorithms in FE Analysis*, pages 468–495. MIT Press, Boston (1977).
- Huy H.D. and Werner B. Linear variational inequalities with application to the buckling problem of the unilateral supported beam. *Numerical Functional Analysis and Optimization*, 8:357–382 (1986).
- Isaacson E. and Keller H.B. *Analysis of Numerical Methods*. Wiley, London (1966).
- Johannsson L. and Klarbring A. Thermoelastic frictional contact problems: Modelling, finite element approximation and numerical realization. *Computer Methods in Applied Mechanics and Engineering*, 105:181–210 (1993).
- Johnson C. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press (1987).
- Johnson C. Adaptive finite element methods for the obstacle problem. *Technical report*, Chalmers University of Technology, Göteborg (1991).
- Johnson C. and Hansbo P. Adaptive finite element methods in computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 101:143–181 (1992).
- Johnson K.L. *Contact Mechanics*. Cambridge University Press (1985).
- Jones R.E. and Papadopoulos P. A yield-limited lagrange multiplier formulation of frictional contact. *International Journal for Numerical Methods in Engineering*, 48:1127–1149 (2000).
- Jourdan F., Alart P. and Jean M. A Gauss-Seidel like algorithm to solve frictional contact problems. *Computer Methods in Applied Mechanics and Engineering*, 155:31–47 (1998).
- Ju W. and Taylor R.L. A perturbed lagrange formulation for the finite element solution of nonlinear frictional contact problems. *Journal of Theoretical and Applied Mechanics*, 7:1–14 (1988).



- Kalker J.J. *Three-dimensional Elastic Bodies in Rolling Contact*. Kluwer Academic, Dordrecht (1990).
- Kane C., Repetto E.A., Ortiz M. and Marsden J.E. Finite element analysis of nonsmooth contact. *Computer Methods in Applied Mechanics and Engineering*, 180:1–26 (1999).
- Keller H.B. Numerical solution of bifurcation and nonlinear eigenvalue problems. In P. Rabinowitz, editor, *Application of Bifurcation Theory*, pages 359–384. Academic Press, New York (1977).
- Kennard E.H. *Kinetic Theory of Gases*. McGraw-Hill, New York, London (1938).
- Khan A.S. and Huang S. *Continuum theory of plasticity*. Wiley, Chichester, New York (1995).
- Khonsari M.M. and Booser E.R. *Applied Tribology: Bearing Design and Lubrication*. Wiley, Chichester (2001).
- Kickinger F. Algebraic multigrid solver for discrete elliptic second order problems. *Technical Report 96-5*, Department of Mathematics, Johannes Kepler University, Linz (1996).
- Kikuchi N. and Oden J.T. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*. SIAM, Philadelphia (1988).
- Klarbring A. A mathematical programming approach to three-dimensional contact problems with friction. *Computer Methods in Applied Mechanics and Engineering*, 58:175–200 (1986).
- Klarbring A. On discrete and discretized nonlinear elastic structures in unilateral contact (stability, uniqueness and variational principles). *International Journal of Solids & Structures*, 24:459–479 (1988).
- Klarbring A. Examples of non-uniqueness and non-existence of solutions to quasistatic contact problems with friction. *Ingenieur Archiv*, 60:529–541 (1990).
- Klarbring A. Contact, friction, discrete mechanical structures and mathematical programming. In P. Wriggers and P. Panagiotopoulos, editors, *New Developments in Contact Problems*, pages 55–100. Springer, Wien (1999).
- Klarbring A. and Björkman G. A mathematical programming approach to contact problems with friction and varying contact surfaces. *Computers and Structures*, 30:1185–1198 (1988).
- Klarbring A. and Björkman G. Solution of large displacement contact problems with friction using newton's method for generalized equations. *International Journal for Numerical Methods in Engineering*, 34:249–269 (1992).
- Klawonn A. An optimal preconditioner for a class of saddle point problems with a penalty term. *SIAM Journal of Scientific Computation*, 19:540–552 (1998).
- Kloosterman G., van Damme R.M.J., van den Boogard A.H. and Huetink J. A geometrical-bases contact algorithm using the barrier method. *International Journal for Numerical Methods in Engineering*, 51:865–882 (2001).

- Klüppel M. and Heinrich G. Rubber friction on self-affine road tracks. *Rubber Chemistry and Technology*, 73:578–606 (2000).
- Knothe K. and Wessels H. *Finite Elemente*. Springer-Verlag, Berlin (1991).
- Knuth D.E. *The Art of Computer Programming*, volume 3, Sorting and Searching. Addison-Wesley, Reading, Mass., USA (1973).
- Korelc J. Automatic generation of finite-element code by simultaneous optimization of expressions. *Theoretical Computer Science*, 187:231–248 (1997).
- Korelc J. Automatic generation of numerical codes with introduction to AceGen 4.0 symbolic code generator. <http://www.fgg.uni-lj.si/Symech> (2000a).
- Korelc J. Computational templates. <http://www.fgg.uni-lj.si/Symech> (2000b).
- Korelc J. and Wriggers P. Symbolic approach in computational mechanics. In D.R.J. Owen, E. Hinton and E. Onate, editors, *Proceedings of COMPLAS 5*. CIMNE, Barcelona (1997).
- Kornhuber R. *Adaptive Monoton Multigrid Methods for Nonlinear Variational Problems*. Teubner, Stuttgart (1997).
- Kornhuber R. and Krause R.H. Adaptive multigrid methods for signorini's problem in linear elasticity. *FU-Berlin, Preprint A-00-15*, pages 1–16 (2000).
- Kragelsky I.V. *Die Entwicklung der Wissenschaft von der Reibung*. Verlag der Akademie der Wissenschaften der UdSSR, Moskau (1956).
- Kragelsky I.V., Dobychin M.N. and Kombalov V.S. *Friction and Wear - Calculation Methods (Translated from The Russian by N. Standen)*. Pergamon Press (1982).
- Krause R.H. and Wohlmuth B.I. A dirichlet-neumann type algorithm for contact problems with friction. *FU-Berlin, Preprint A-01-09*, pages 1–16 (2001).
- Krstulovic-Opara L., Wriggers P. and Korelc J. A c1-continuous formulation for 3d finite deformation frictional contact. *Computational Mechanics* (2002).
- Laursen T.A. and Chawla V. Design of energy conserving algorithms for frictionless dynamic contact problems. *International Journal for Numerical Methods in Engineering*, 40:863–886 (1997).
- Laursen T.A. and Govindjee S. A note on the treatment of frictionless contact between nonsmooth surfaces in fully nonlinear problems. *Communications in Numerical Methods in Engineering*, 10:869–878 (1994).
- Laursen T.A. and Love G.R. Improved implicit integrators for transient impact problems – geometric admissibility within the conserving framework. *International Journal for Numerical Methods in Engineering*, 53:245–274 (2002).
- Laursen T.A. and Simo J.C. Algorithmic symmetrization of coulomb frictional problems using augmented Lagrangians. *Computer Methods in Applied Mechanics and Engineering*, 108:133–146 (1993a).

- Laursen T.A. and Simo J.C. A continuum-based finite element formulation for the implicit solution of multibody, large deformation frictional contact problems. *International Journal for Numerical Methods in Engineering*, 36:3451–3485 (1993b).
- Lee C.Y. and Oden J.T. A-posteriori error estimation of hp-finite element approximation of frictional contact problems. *Computer Methods in Applied Mechanics and Engineering*, 113:11–45 (1994).
- Lee C.Y., Oden J.T. and Ainsworth M. Local a posteriori error estimates and numerical results for contact problems and problems of flow through porous media. In P. Wriggers and W. Wagner, editors, *Computational Methods in Nonlinear Mechanics*, pages 671–689. Springer, Berlin (1991).
- Lee N.S. and Bathe K.J. Error indicators and adaptive remeshing in large deformation finite element analysis. *Finite Elements in Analysis and Design*, 16:99–139 (1994).
- Leibniz G.W. Tentemen de natura et remedlie resistentiarum in machines. *Miscellanea Bero-linensia, Class. mathem.*, 1:307 (1706).
- Lewis R.W. and Schrefler B.A. *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media*. Wiley, Chichester, 2nd edition (2000).
- Litewka P. and Wriggers P. Contact between 3d-beams with rectangular cross section. *International Journal for Numerical Methods in Engineering*, 53:2019–2042 (2002a).
- Litewka P. and Wriggers P. Frictional contact between 3d-beams with rectangular cross section. *Computational Mechanics*, 28:26–39 (2002b).
- Liu W.N., Meschke G. and Mang H.A. A note on the algorithmic stabilization of 2d contact analyses. In L. Gaul and C.A. Brebbia, editors, *Computational Methods in Contact Mechanics IV*, pages 231–240. Wessex Institute, Southampton (1999).
- Lu S.C.H. and Pister K.S. Decomposition of deformation and representation of the free energy function for isotropic thermoelastic solids. *International Journal of Solids & Structures*, 11:927–934 (1975).
- Lubliner J. A model of rubber viscoelasticity. *Mechanics Research Communications*, 12:93–99 (1985).
- Lubliner J. *Plasticity Theory*. MacMillan, London (1990).
- Luenberger D.G. *Linear and Nonlinear Programming*. Addison-Wesley, Reading, Mass., second edition (1984).
- Madhusudana C.V. and Fletcher L.S. Gas conductance contribution to contact heat transfer. *AIAA Journal, Paper 81-1163* (1981).
- Majumdar A. and Bhushan B. Role of fractal geometry in roughness characterization and contact mechanics of surfaces. *Journal of Tribology*, 112:205–216 (1990).
- Malvern L.E. *Introduction to the Mechanics of a Continuous Medium*. Prentice-Hall, Englewood Cliffs, New Jersey (1969).
- Mandelbrot B. *The Fractal Geometry of Nature*. W.H. Freeman and Co., New York (1982).

- Marsden J.E. and Hughes T.J.R. *Mathematical Foundations of Elasticity*. Prentice-Hall, Englewood Cliffs, New Jersey (1983).
- Martins J.A.C., Monteiro Marques M.D.P. and Gastaldi F. On an example of non-existence of solutions to a quasistatic frictional contact problem. *European Journal of Mechanics A/Solids*, 13:113–133 (1994).
- Martins J.A.C. and Oden J.T. Existence and uniqueness results for dynamic contact problems with nonlinear normal and friction interface laws. *Nonlinear Analysis, Theory, Methods and Applications*, 11:407–428 (1987).
- Matthies H. and Strang G. The solution of nonlinear finite element equations. *International Journal for Numerical Methods in Engineering*, 14:1613–1626 (1979).
- Maugin G.A. *The Thermomechanics of Plasticity and Fracture*. Cambridge University Press, Cambridge, New York (1992).
- McDevitt T.W. and Laursen T.A. A mortar-finite element formulation for frictional contact problems. *International Journal for Numerical Methods in Engineering*, 48:1525–1547 (2000).
- Michalowski R. and Mroz Z. Associated and non-associated sliding rules in contact friction problems. *Archives of Mechanics*, 30:259–276 (1978).
- Miehe C. *Zur numerischen Behandlung thermomechanischer Prozesse*. Ph.D. thesis, Institut für Baumechanik und Numerische Mechanik, Universität Hannover, Hannover (1988).
- Mikic B.B. Analytical studies of contact of nominally flat surfaces and effect of previous loading. *Journal of Lubrication Technology*, 93:451–459 (1971).
- Moreau J.J. On unilateral constraints, friction and plasticity. In G. Capriz and G. Stampacchia, editors, *New Variational Techniques in Mathematical Physics*, pages 175–322. CIME, Roma (1974).
- Mroz Z. and Stupkiewicz S. Constitutive modeling of slip and wear in elastic frictional contact. In A. Curnier, editor, *Contact Mechanics*, pages 237–260. Presses Polytechniques et Universitaires Romandes, Lausanne (1992).
- Munjiza A. and Andrews K.R.F. Nbs contact detection algorithm for bodies of similar size. *International Journal for Numerical Methods in Engineering*, 43:131–149 (1998).
- Munjiza A. and Owen D.R.J. Discrete elements in rock blasting. In *Proceedings of the 2nd International Conference on Discret Element Methods (DEM)*, pages 287–300 (1993).
- Munjiza A., Owen D.R.J. and Bicanic N. A combined finite-discrete element method in transient dynamics of fracturing solids. *Engineering Computations*, 12:145–174 (1995).
- Murray W. Sequential quadratic programming methods for large scale problems. *COAP*, 7:127–142 (1997).
- Nackenhorst U. On the finite element analysis of steady state rolling contact. In M.H. Aliabadi and C.A. Brebbia, editors, *Contact Mechanics*, pages 53–60. Computational Mechanics Publications, Southampton (1993).

- Nackenhorst U. Rollkontaktdynamik – Numerische Analyse der Dynamik rollender Körper mit der Finite Elemente Methode. *Technical report*, Institut für Mechanik, Universität der Bundeswehr, Hamburg (2000).
- Needleman A. An analysis of tensile decohesion along an interface. *Journal of the Mechanics and Physics of Solids*, 38:289–324 (1990).
- Neto E.A.D., Hashimoto K., Peric D. and Owen D.R.J. A phenomenological model for frictional contact accounting for wear effects. *Philosophical Transactions of the Royal Society of London*, A 354:373–389 (1996).
- Newmark N.M. A method of computation for structural dynamics. *Proceedings of ASCE, Journal of Engineering Mechanics*, 85:67–94 (1959).
- Nitsche J. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abhandlungen in der Mathematik an der Universität Hamburg*, 36:9–15 (1970).
- Nour-Omid B. and Wriggers P. A note on the optimum choice for penalty parameters. *Communications in Applied Numerical Methods*, 3:581–585 (1987).
- Oden J.T. *Finite Elements of Nonlinear Continua*. McGraw-Hill, New York (1972).
- Oden J.T. Exterior penalty methods for contact problems in elasticity. In W. Wunderlich, E. Stein and K.J. Bathe, editors, *Nonlinear Finite Element Analysis in Structural Mechanics*. Springer, Berlin (1981).
- Oden J.T. and Lin T.L. On the general rolling contact problem for finite deformations of a viscoelastic cylinder. *Computer Methods in Applied Mechanics and Engineering*, 52:297–367 (1986).
- Oden J.T. and Martins J.A.C. Models and computational methods for dynamic friction phenomena. *Computer Methods in Applied Mechanics and Engineering*, 52:527–634 (1986).
- Oden J.T. and Pires E.B. Algorithms and numerical results for finite element approximations of contact problems with non-classical friction laws. *Computers and Structures*, 19:137–147 (1983a).
- Oden J.T. and Pires E.B. Nonlocal and nonlinear friction laws and variational principles for contact problems in elasticity. *Journal of Applied Mechanics*, 50:67–76 (1983b).
- Ogden R.W. *Non-Linear Elastic Deformations*. Ellis Horwood and John Wiley, Chichester (1984).
- Ortega J. and Rheinboldt W. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York (1970).
- Ortiz M. and Quigley J.J. Adaptive mesh refinement in strain localization problems. *Computer Methods in Applied Mechanics and Engineering*, 90:781–804 (1991).
- Paczelt I., Szabo B. and Szabo T. Solution of contact problem using the hp-version of the finite element method. *Computers & Mathematics with Applications*, 38:49–69 (1999).

- Padmanabhan V. and Laursen T. A framework for development of surface smoothing procedures in large deformation frictional contact analysis. *Finite Elements in Analysis and Design*, 37:173–198 (2001).
- Padovan J. and Zeid I. Finite element analysis of steadily moving contact fields. *Computers and Structures*, 2:111–200 (1984).
- Panagiotopoulos P.D. A nonlinear programming approach to the unilateral contact-, and friction-boundary value problem in the theory of elasticity. *Ingenieur-Archiv*, 44:421–432 (1975).
- Panagiotopoulos P.D. *Inequality Problems in Mechanics, Convex and Nonconvex Energy Functions*. Birkhäuser Verlag, Boston, Basel (1985).
- Panagiotopoulos P.D., Panagouli O.K. and Mistakidis E.S. Fractal interfaces in contact problems. theory and numerical applications. In A. Curnier, editor, *Contact Mechanics*, pages 237–260. Presses Polytechniques et Universitaires Romandes, Lausanne (1992).
- Pang J.S. Newton's method for b-differentiable equations. *Math. Oper. Res.*, 15:311–341 (1990).
- Papadopoulos P. and Taylor R.L. A mixed formulation for the finite element solution of contact problems. *Computer Methods in Applied Mechanics and Engineering*, 94:373–389 (1992).
- Parisch H. A consistent tangent stiffness matrix for three-dimensional non-linear contact analysis. *International Journal for Numerical Methods in Engineering*, 28:1803–1812 (1989).
- Parisch H. and Lübbing C. A formulation of arbitrarily shaped surface elements for three-dimensional large deformation contact with friction. *International Journal for Numerical Methods in Engineering*, 40:3359–3383 (1997).
- Peric D. and Owen D.R.J. Computational model for contact problems with friction based on the penalty method. *International Journal for Numerical Methods in Engineering*, 36:1289–1309 (1992).
- Perkins E. and Williams J.R. A fast contact detection algorithm insensitive to object sizes. *Engineering Computations*, 12:185–201 (1995).
- Persson B.N.J. *Sliding Friction, Physical Principles and Applications*. Springer, Berlin, Heidelberg, New York, 2nd edition (2000).
- Pfeiffer F. and Glocker C. *Multibody Dynamics with Unilateral Contacts*. Wiley, New York (1996).
- Pietrzak G. Continuum mechanics modelling and augmented lagrange formulation of large deformation frictional contact problems. *Technical Report 1656*, EPFL, Lausanne (1997).
- Pietrzak G. and Curnier A. Continuum mechanics modeling and augmented lagrange formulation of multibody, large deformation frictional contact problems. In D.R. Owen, E. Hinton and E. Onate, editors, *Proceedings of COMPLAS 5*, pages 878–883. CIMNE, Barcelona (1997).

- Pietrzak G. and Curnier A. Large deformation frictional contact mechanics: continuum formulation and augmented lagrangean treatment. *Computer Methods in Applied Mechanics and Engineering*, 177:351–381 (1999).
- Polzer G. and Meissner F. *Grundlagen zu Reibung und Verschleiss*. VEB Deutscher Verlag für Kunststoffindustrie, Leipzig (1983).
- Rabier P.J., Martins J.A.C., Oden J.T. and Campos L. Existence and local uniqueness of solutions to contact problems in elasticity with nonlinear friction laws. *International Journal of Engineering Science*, 24:1755–1768 (1986).
- Rabinowicz E. *Friction and Wear of Materials*. Wiley, New York, second edition (1995).
- Ramm E. Strategies for tracing the nonlinear response near limit points. In W. Wunderlich, E. Stein and K.J. Bathe, editors, *Nonlinear Finite Element Analysis in Structural Mechanics*. Springer, Berlin, Heidelberg, New York (1981).
- Rank E., Schweingruber M. and Sommer M. Adaptive mesh generation and transformation of triangular to quadrilateral meshes. *Communications in Numerical Methods in Engineering*, 9:121–129 (1993).
- Rannacher R. and Suttmeier F.T. A posteriori error control in finite element methods via duality techniques: Application to perfect plasticity. *Technical Report 97-16*, Institut für Applied Mathematics, SFB 359, University of Heidelberg (1997).
- Raous M. Quasistatic signorini problem with coulomb friction and coupling to adhesion. In P. Wriggers and P. Panagiotopoulos, editors, *CISM Courses and Lectures, No. 384*, pages 101–178. Springer, Wien (1999).
- Raous M., Cangemi L. and Cocu M. A consistent model coupling adhesion, friction and unilateral contact. *Computer Methods in Applied Mechanics and Engineering*, 189 (2000).
- Rebel G., Park K.C. and Felippa C.A. A ccontact-impact formulation based on localised lagrange multipliers. *Technical Report CU-CAS-00-18*, College of Engineering at University of Colorado, Bolder, Colorado 80309 (2000).
- Reese S. Thermomechanische Modellierung gummiartiger Polymerstrukturen. *Technical Report F 01/4*, Institut für Baumechanik und Numerische Mechanik, Universität Hannover (2001).
- Reul O. In-situ-Messungen und numerische Studien zum Tragverhalten der kombinierten Pfahl-Plattengründung. *Technical Report 53*, Institut und Versuchsanstalt für Geotechnik, Technische Universität Darmstadt (2000).
- Rieger H. Experimentelle und theoretische Untersuchungen zur Gummireibung in einem grossen Geschwindigkeits- und Temperaturbereich unter Berücksichtigung der Reibungswärme. *Technical report*, TH Darmstadt (1968).
- Riks E. The application of Newtons method to the problem of elastic stability. *Journal of Applied Mechanics*, 39:1060–1066 (1972).
- Riks E. Some computational aspects of stability analysis of nonlinear structures. *Computer Methods in Applied Mechanics and Engineering*, 47:219–260 (1984).

- Sansour C., Wriggers P. and Sansour J. Nonlinear dynamics of shells: Theory, finite element formulation and integration schemes. *Nonlinear Dynamics*, 13:279–305 (1997).
- Schittkowski K. Solving nonlinear programming problems with very many constraints. *Optimization*, 25:179–196 (1992).
- Schöberl J.. Netgen – an advancing front 2d/3d mesh generator based on abstract rules. *Computer Visualization Science*, 1:41–52 (1997a).
- Schöberl J.. Robust multigrid preconditioning for parameter dependent problems i: the stokes type case. *Technical Report 97-2*, Institut für Mathematik, Johannes Kepler Universität Linz (1997b).
- Schöberl J. Solving the Signorini problem on the basis of domain decomposition techniques. *Computing*, 60:323–344 (1998).
- Scholz C.H. Fractals in geophysics. In B. Mandelbrot, editor, *Fractals*. Birkhäuser Verlag, Boston- Basel (1989).
- Schweizerhof K. and Wriggers P. Consistent linearization for path following methods in nonlinear fe-analysis. *Computer Methods in Applied Mechanics and Engineering*, 59:261–279 (1986).
- Schwetlick H. and Kretschmar H. *Numerische Verfahren für Naturwissenschaftler und Ingenieure*. Fachbuchverlag, Leipzig (1991).
- Sellgren U. and Olofsson U. Application of a constitutive model for micro-slip in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 170:65–77 (1999).
- Sewell M.J. On configuration-dependent loading. *Archives of Rational Mechanics*, 23:321–351 (1967).
- Shai I. and Santo M. Heat transfer with contact resistance. *International Journal of Heat and Mass Transfer*, 24:465–470 (1982).
- Signorini A. Sopra alcune questioni di elastostatica. *Atti della Societa Italiana per il Progresso delle Scienze* (1933).
- Simo J., Taylor R. and Wriggers P. A note on finite element implementation of pressure boundary loading. *Communications in Applied Numerical Methods*, 7:513–525 (1991).
- Simo J., Wriggers P., Schweizerhof K.H. and Taylor R. Postbuckling analysis involving inelasticity and unilateral constraints. *International Journal for Numerical Methods in Engineering*, 23:779–800 (1986).
- Simo J.C. A finite strain beam formulation. The three-dimensional dynamic problem. Part I. *Computer Methods in Applied Mechanics and Engineering*, 49:55–70 (1985).
- Simo J.C. and Hughes T.J.R. *Computational Inelasticity*. Springer, New York, Berlin (1998).
- Simo J.C. and Laursen T.A. An augmented Lagrangian treatment of contact problems involving friction. *Computers and Structures*, 42:97–116 (1992).



- Simo J.C. and Miehe C. Associative coupled thermoplasticity at finite strains: formulation, numerical analysis and implementation. *Computer Methods in Applied Mechanics and Engineering*, 98:41–104 (1992).
- Simo J.C. and Tarnow N. The discrete energy-momentum method. Conserving algorithms for nonlinear elastodynamics. *Zeitschrift für angewandte Mathematik und Physik*, 43:757–792 (1992).
- Simo J.C. and Taylor R.L. Consistent tangent operators for rate-independent elastoplasticity. *Computer Methods in Applied Mechanics and Engineering*, 48:101–118 (1985).
- Simo J.C. and Vu-Quoc L. Three dimensional finite strain rod model. Part II: computational aspects. *Computer Methods in Applied Mechanics and Engineering*, 58:79–116 (1986).
- Simo J.C., Wriggers P. and Taylor R.L. A perturbed Lagrangian formulation for the finite element solution of contact problems. *Computer Methods in Applied Mechanics and Engineering*, 50:163–180 (1985).
- Sloan S.W. A fast algorithm for generating constrained delaunay triangulations. *Computers and Structures*, 47:441–450 (1993).
- Song S. and Yovanovich M.M. Explicit relative contact pressure expression: Dependence upon surface roughness parameters and vickers microhardness coefficients. *AIAA Journal, Paper 87-0152* (1987).
- Spellucci P. Sequential quadratic programming: theory, implementation, problems. *Methods of Operation Research*, 53:183–213 (1985).
- Spellucci P. *Numerische Verfahren der nichtlinearen Optimierung*. Birkhäuser, Basel (1993).
- Stadter J.T. and Weiss R.O. Analysis of contact through finite element gaps. *Computers and Structures*, 10:867–873 (1979).
- Stoer J. On the numerical solution of constrained least squares problems. *SINUMS*, 8:382–411 (1971).
- Strömberg N. A Newton method for three-dimensional fretting problems. *European Journal of Mechanics A/Solids*, 16:573–593 (1997).
- Strömberg N., Johansson L. and Klarbring A. Derivation and analysis of a generalized standard model for contact, friction and wear. *International Journal of Solids & Structures*, 33:1817–1836 (1996).
- Stupkiewicz S. Extension of the node-to-segment contact element for surface-expansion-dependent contact laws. *International Journal for Numerical Methods in Engineering*, 50:739–759 (2001).
- Stupkiewicz S. and Mroz Z. A model of third body abrasive friction and wear in hot metal forming. *Wear*, 231:124–138 (1999).
- Suresh S., Mortensen A. and Needleman A. *Fundamentals of Metal Matrix Composites*. Butterworth-Heinemann, Oxford (1993).

- Tabor D. Friction – The present state of our understanding. *Journal Lubrication Technology*, 103:169–179 (1981).
- Takayasu H. *Fractals in the Physical Sciences*. University Press, Manchester (1990).
- Taliec P.L. and Rahier C. Numerical methods of steady rolling for non-linear viscoelastic structures in finite deformations. *International Journal for Numerical Methods in Engineering*, 37:1159–1186 (1994).
- Taliec P.L. and Sassi T. Domain decomposition with nonmatching grids: augmented lagrangian approach. *Journal of Mathematical Computations*, 647:1367–1396 (1995).
- Taylor R.L. and Wriggers P. Smooth surface discretization for large deformation frictionless contact. *Technical report*, University of California, Berkeley (1999). Report No. UCB/SEMM-99-04.
- Thorpe J.A. *Elementary Topics in Differential Geometry*. Springer-Verlag, New York (1979).
- Timoshenko S.P. and Goodier J.N. *Theory of Elasticity*. McGraw-Hill, New York (1970).
- Truesdell C. and Noll W. The nonlinear field theories of mechanics. In S. Flügge, editor, *Handbuch der Physik III/3*. Springer, Berlin, Heidelberg, Wien (1965).
- Truesdell C. and Toupin R. The classical field theories. In *Handbuch der Physik III/1*. Springer, Berlin, Heidelberg, Wien (1960).
- Turner M.J., Clough R.W., Martin H.C. and Topp L.J. Stiffness and deflection analysis of complex structures. *Journal of the Aeronautical Sciences*, 23:805–823 (1956).
- Tvergaard V. Effect of fibre debonding in a whisker-reinforced metal. *Journal of Material Science in Engineering*, A125:203–213 (1990).
- Vainberg M.M. *Variational Methods for the Study of Nonlinear Operators*. Holden Day, San Francisco (1964).
- Valliappan S., Lee I. and Boonlualohr P. Non-linear analysis of contact problems. In R. Lewis, P. Bettess and E. Hinton, editors, *Numerical Methods in Coupled Problems*, pages 231–253. Wiley, Chichester (1984).
- Verfürth R. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley, Teubner, Chichester, New York, Stuttgart, Leipzig (1996).
- Wagner W. Zur Behandlung von Stabilitätsproblemen mit der Methode der Finiten Elemente. *Technical Report F91/1*, Forschungs- und Seminarberichte aus dem Bereich der Mechanik der Universität Hannover (1991).
- Wagner W. and Wriggers P. A simple method for the calculation of secondary branches. *Engineering Computations*, 5:103–109 (1988).
- Wang S.P. and Nakamachi E. The inside-outside contact search algorithm for finite element analysis. *International Journal for Numerical Methods in Engineering*, 40:3665–3685 (1999).

- Wanheim T., Bay N. and Petersen A.S. A theoretically determined model for friction in metal working processes. *Wear*, 28:251–258 (1974).
- Werner B. and Spence A. The computation of symmetry-breaking bifurcation points. *SIAM Journal on Numerical Analysis*, 21:388–399 (1984).
- Wilkins M.L. Calculation of elasto-plastic flow. In B. Alder, S. Fernbach and M. Rotenberg, editors, *Methods of Computational Physics*, volume 3. Academic Press, Boston, New York (1964).
- Williams J.R. and O'Connor R. A linear complexity intersection algorithm for discrete element simulation of arbitrary geometries. *Engineering Computations*, 12:185–201 (1995).
- Williams J.R. and O'Connor R. Discrete element simulation and the contact problem. *Archives of Computational Methods in Engineering*, 6:279–304 (1999).
- Williams J.R. and Pentland A.P. Superquadrics and modal dynamics for discrete elements in interactive design. *Engineering Computations*, 9:115–127 (1992).
- Willner K. and Gaul L. Contact description by fem based on interface physics. In D.R.J. Owen, E. Hinton and E. Onate, editors, *Proceedings of COMPLAS 5*, volume 1, pages 884–891. CIMNE, Barcelona (1997).
- Wilson E.A. and Parsons B. Finite element analysis of elastic contact problems using differential displacements. *International Journal for Numerical Methods in Engineering*, 2:387–395 (1970).
- Wohlmuth B.I. *Discretization Methods and Iterative Solvers based on Domain Decomposition*. Springer Verlag, Berlin, Heidelberg, New York (2000a).
- Wohlmuth B.I. A mortar finite element method using dual spaces for the lagrange multiplier. *SIAM, Journal of Numerical Analysis*, 38:989–1012 (2000b).
- Wohlmuth B.I. and Krause R.H. Monotone methods on non-matching grids for nonlinear contact problems. *Technical Report 06*, Sonderforschungsbereich 404, Universität Stuttgart (2002).
- Woo K.L. and Thomas T.R. Contact of rough surfaces: A review of experimental works. *Wear*, 58:331–340 (1980).
- Wood W.L. *Practical Time-stepping Schemes*. Clarendon Press, Oxford (1990).
- Wood W.L., Bossak M. and Zienkiewicz O.C. An alpha modification of newmark's method. *International Journal for Numerical Methods in Engineering*, 15:1562–1566 (1981).
- Wriggers P. On consistent tangent matrices for frictional contact problems. In G. Pande and J. Middleton, editors, *Proceedings of NUMETA 87*. M. Nijhoff Publishers, Dordrecht (1987).
- Wriggers P. Finite elements for thermo-mechanical contact and adaptive finite element analysis of contact problems. In P. Wriggers and P. Panagiotopoulos, editors, *New Developments in Contact Problems*, pages 179–246. Springer, Wien (1999).
- Wriggers P. *Nichtlineare Finite Elemente*. Springer, Berlin (2001).

- Wriggers P. and Boersma A. A parallel algebraic multigrid solver for problems in solid mechanics discretized by finite elements. *Computers and Structures*, 69:129–137 (1998).
- Wriggers P., Eberlein R. and Reese S. Comparison between shell and 3d-elements in finite plasticity. *International Journal of Solids & Structures*, 33:3309–3326 (1996).
- Wriggers P. and Imhof M. On the treatment of nonlinear unilateral contact problems. *Ing. Archiv*, 63:116–129 (1993).
- Wriggers P., Krstulovic-Opara L. and Korelc J. Development of 2d smooth polynomial frictional contact element based on a symbolic approach. In W. Wunderlich, E. Stein, E. Ramm and P. Wriggers, editors, *Proceedings of ECCM*. München (1999).
- Wriggers P., Krstulovic-Opara L. and Korelc J. Smooth c1- interpolations for two-dimensional frictional contact problems. *International Journal for Numerical Methods in Engineering*, 51:1469–1495 (2001).
- Wriggers P. and Mieke C. On the treatment of contact constraints within coupled thermomechanical analysis. In D. Besdo and E. Stein, editors, *Proc. of EUROMECH, Finite Inelastic Deformations*. Springer, Berlin (1992).
- Wriggers P. and Mieke C. Contact constraints within coupled thermomechanical analysis - a finite element model. *Computer Methods in Applied Mechanics and Engineering*, 113:301–319 (1994).
- Wriggers P., Mieke C., Kleiber M. and Simo J. A thermomechanical approach to the necking problem. *International Journal for Numerical Methods in Engineering*, 33:869–883 (1992).
- Wriggers P. and Nour-Omid B. Solution methods for contact problems. *Technical Report No. UCB/SESM 84/09*, University of California, Berkeley (1984).
- Wriggers P., Rieger A. and Scherf O. Comparison of different error measures for adaptive finite element techniques applied to contact problems involving large elastic strains. *Computer Methods in Applied Mechanics and Engineering* (2000).
- Wriggers P. and Scherf O. Different a posteriori error estimators and indicators for contact problems. *Mathematics and Computer Modelling*, 28:437–447 (1998).
- Wriggers P., Scherf O. and Carstensen C. Adaptive techniques for the contact of elastic bodies. In J. Hughes, E. Onate and O. Zienkiewicz, editors, *Recent Developments in Finite Element Analysis*, pages 78–86. CIMNE, Barcelona (1994).
- Wriggers P. and Simo J. A note on tangent stiffnesses for fully nonlinear contact problems. *Communications in Applied Numerical Methods*, 1:199–203 (1985).
- Wriggers P. and Simo J. A general procedure for the direct computation of turning and bifurcation points. *International Journal for Numerical Methods in Engineering*, 30:155–176 (1990).
- Wriggers P., Simo J. and Taylor R. Penalty and augmented lagrangian formulations for contact problems. In J. Middleton and G. Pande, editors, *Proceedings of NUMETA Conference*. Balkema, Rotterdam (1985).

- Wriggers P., Van T.V. and Stein E. Finite-element-formulation of large deformation impact-contact -problems with friction. *Computers and Structures*, 37:319–333 (1990).
- Wriggers P., Wagner W. and Miehe C. A quadratically convergent procedure for the calculation of stability points in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 70:329–347 (1988).
- Wriggers P. and Zavarise G. On the application of augmented lagrangian techniques for non-linear constitutive laws in contact interfaces. *Communications in Applied Numerical Methods*, 9:815–824 (1993a).
- Wriggers P. and Zavarise G. Thermomechanical contact – a rigorous but simple numerical approach. *Computers and Structures*, 46:47–53 (1993b).
- Wriggers P. and Zavarise G. On contact between three-dimensional beams undergoing large deflections. *Communications in Applied Numerical Methods*, 13:429–438 (1997).
- Wriggers P., Zavarise G. and Zohdi T. A computational study of interfacial debonding damage in fibrous composites materials. *Computational Materials Science*, 12:39–56 (1998).
- Xu X. and Needleman A. Void nucleation by inclusion debonding in crystal matrix. *Modeling and Simulation in Material Science and Engineering*, 1:111–132 (1993).
- Yovanovich M.M. Thermal contact correlations. *AIAA Journal, Paper 81-1164* (1981).
- Zavarise G. *Problemi termomeccanici di contatto – aspetti fisici e computazionali*. Ph.D. thesis, Ist. di Scienza e Tecnica delle Costruzioni Univ. of Padua, Italy, Padua (1991).
- Zavarise G. and Taylor R.L. Point-to-segment contact geometry revisited. In *Proceedings of joint conference of the Italian group of Computational Mechanics and Ibero-Latin-American Association of Computational Methods in Engineering*, pages 884–891 (1996).
- Zavarise G. and Wriggers P. On the efficiency of new and old strategies for solving contact problems. *Zeitschrift für angewandte Mathematik und Mechanik*, 76:563–564 (1996).
- Zavarise G. and Wriggers P. Contact with friction between beams in 3-d space. *International Journal for Numerical Methods in Engineering*, 49:977–1006 (2000).
- Zavarise G., Wriggers P. and Schrefler B.A. On augmented lagrangian algorithms for thermomechanical contact problems with friction. *International Journal for Numerical Methods in Engineering*, 38:2929–2949 (1995).
- Zavarise G., Wriggers P., Stein E. and Schrefler B. A numerical model for thermomechanical contact based on microscopic interface laws. *Mech. Res. Comm.*, 19:173–182 (1992a).
- Zavarise G., Wriggers P., Stein E. and Schrefler B. Real contact mechanisms and finite element formulation - a coupled thermomechanical approach. *International Journal for Numerical Methods in Engineering*, 35:767–786 (1992b).
- Zhong Z.H. *Finite Element Procedures for Contact-Impact Problems*. Oxford University Press, Oxford (1993).
- Zhong Z.H. and Nilsson L. A contact searching algorithm for general contact problems. *Computers and Structures*, 33:197–209 (1989).

- Zienkiewicz O.C. and Taylor R.L. *The Finite Element Method, 4th Ed.*, volume 1. McGraw Hill, London (1989).
- Zienkiewicz O.C. and Taylor R.L. *The Finite Element Method, 4th Ed.*, volume 2. McGraw Hill, London (1991).
- Zienkiewicz O.C. and Taylor R.L. *The Finite Element Method*, volume 2. Butterworth-Heinemann, Oxford, UK, 5th edition (2000a).
- Zienkiewicz O.C. and Taylor R.L. *The Finite Element Method*, volume 1. Butterworth-Heinemann, Oxford, UK, 5th edition (2000b).
- Zienkiewicz O.C. and Zhu J.Z. A simple error estimator and adaptive procedure for practical engineering analysis. *International Journal for Numerical Methods in Engineering*, 24:337–357 (1987).
- Zienkiewicz O.C. and Zhu J.Z. Adaptivity and mesh generation. *International Journal for Numerical Methods in Engineering*, 32:783–810 (1991).
- Zmitrowicz A. A thermodynamical model of contact, friction and wear: I governing equations. *Wear*, 114:135–168 (1987a).
- Zmitrowicz A. A thermodynamical model of contact, friction and wear: III constitutive equations for friction, wear and frictional heat. *Wear*, 114:199–221 (1987b).
- Zohdi T. and Wriggers P. Modeling and simulating of the decohesion of particulate aggregates in a binding matrix. *Engineering Computations*, 18:79–95 (2001).
- Zohdi T.I. and Wriggers P. A computational model for interfacial damage through microstructural cohesive zone relaxation. *The International Journal of Fracture*, 101:L9–L14 (2000).

**This page intentionally left blank**

# *Index*

- Abrasive wear, 90
- Accelerations, 273
- Accumulated sliding distance, 77
- Active constraints, 390
- Active set strategy, 99
- Adaptive methods, 355
  - dual error estimators, 365
  - friction, 379
  - mesh refinement, 369
  - projection methods, 362
  - residual-based error measures, 358
  - transfer of history variables, 382
- Adhesion, 86, 306
  - algorithm, 306
  - friction, 87
  - normal contact, 86
  - time integration, 306
  - viscosity, 87
- Approach of two bodies, 60
- Arc-length, 408
- Asperities, 58
- Augmented Lagrange method, 109, 288
- Barrier method, 108
- Beam contact, 338
  - tangential gap, 344
  - weak form, 344
  - contact condition, 339
  - contact search, 348
  - Coulomb friction, 345
  - discretization, 346
  - gap function, 338
  - normal vector, 340
  - slip path, 341
  - tangent matrix, 347
  - variation, 342
- Bezier interpolation, 237
- BFGS update, 294
- Bifurcation, 389
- Block Gauss–Seidel method
  - nonlinear, 298
- Body force, 26
- Boundary lubrication, 84
- Bounding box, 263
- Bucket search, 263
- Cauchy stress, 33, 57
- Cauchy theorem, 26
- Central difference scheme, 274
- Christoffel tensor, 408
- Closest point projection, 205, 220, 234
- Coating, 73
- Condition for stick, 50
- Contact-impact, 299
- Contact algorithms
  - active set method, 295
  - Dirichlet–Neumann method, 298
  - dynamic contact, 299
  - Lagrange multiplier method, 286
  - linear complementary problem, 297
  - partitioning method, 290
  - penalty method, 284, 303



- SQP method, 294
- Uszawa method, 288, 304
- Contact conductivity, 316
- Contact constraints, 162
  - continuous, 47
  - discrete, 166, 172, 178, 186, 269, 279
- Contact discretization
  - large deformation, 204
  - small deformation, 162
- Contact element
  - Bezier interpolation, 237
  - Hermitian interpolation, 233
  - node-to-edge, 227
  - node-to-node, 166, 229, 324
  - node-to-segment, 328
  - node-to-surface, 219, 224
  - segment-to-segment, 183
  - smooth interpolation, 231
  - surface-to-surface, 172
- Contact interface
  - fractal surface, 91
  - gas conductance, 317
  - heat conduction, 314
  - micromechanical approach, 58
  - radiation, 319
  - spot conductance, 316
- Contact normal, 47
- Contact search, 261–262
  - beam contact, 348
  - area check, 269
  - bounding box, 263
  - grid cell algorithm, 266
  - heapsort algorithm, 266
  - node-to-segment, 268
  - non-uniqueness, 269
  - octree method, 266
  - spatial search, 264
  - superquadrics, 267
- Contact segments, 183
  - intermediate line, 186
  - residual, 188
  - tangent matrix, 188
- Contact
  - bars, 6
  - bifurcation, 389
  - finite strains, 98
- Continuum element
  - three-dimensional, 146, 155
  - two-dimensional, 152
- Contravariant basis, 407
- Contravariant vectors, 406
- Convective coordinates, 405, 407
- Coulomb–Orowan friction law, 72
- Coulomb friction, 64, 107, 170, 215, 248, 333, 345
  - dynamic law, 64
  - integration, 307
  - regularization, 65
- Courant criterion, 274–275
- Covariant vectors, 406
- Critical points
  - contact, 389
  - extended system, 391
- Cross-constraint method, 111
- Curvature tensor, 52, 407
- Debonding, 87
- Deformation gradient, 21, 146, 312
- Delaunay triangularization, 370
- Density, 26
- Direct constraint elimination, 104, 188
- Direct elimination method, 116
- Directional derivative, 38
  - contravariant basis, 39
  - covariant basis, 39
- Discretized equations
  - dynamics, 150
- Dissipation rate, 90
- Distance function, 47
- Dual formulation, 282
- Dynamic friction coefficient, 64
- Eigenvalue problem, 391
- Elasto-hydrodynamic, 83
- Equation of motions, 273
- Error estimators
  - dual principles, 365
  - residual-based, 358
- Error indicators
  - friction, 379
  - superconvergence, 362, 380
- Euler backward scheme
  - friction, 307
- Existence, 96, 98
- Explicit time integration, 274
- Extended system, 391
- Feasible solution, 295
- Fourier law, 314
- Fractal contact interface, 91, 377
- Fractal graph, 92
- Fractal interpolation function, 91–92
- Friction coefficient, 64, 75, 77
  - material pairings, 64
- Friction
  - augmented Lagrange method, 109
  - anisotropic law, 71
  - bench mark test, 255
  - Coulomb–Orowan law, 72
  - dissipation, 68, 308
  - elastic-plastic approach, 67
  - homogenization, 79
  - metal-forming, 72
  - moving cone, 213, 308
  - path, 51
  - penalty method, 303

- polymer, 73
- rubber, 73
- Shaw law, 72
- slip increment, 307
- soil, 76
- symmetric treatment, 305
- time integration, 302
- time integration for soil, 310
- uniqueness, 96
- Uszawa algorithm, 304
- Frictionless contact, 57, 224, 227, 229, 231, 237, 391
- Frictionless
  - augmented Lagrange method, 109
- Gap, 2, 47, 390
- Gap function, 47, 162, 166, 172, 186, 195, 205, 219
- Gap
  - non penetration condition, 47
  - penetration function, 48
- Gas, 318
- Gauss integration, 195
  - one-dimensional, 402
  - two-dimensional, 403
- Grid cell algorithm, 266
- Heapsort algorithm, 266
- Heat conductance, 326
- Heat conduction
  - gas, 317
  - radiation, 319
  - spots, 316
- Heat flux, 314
- Heat flux vector, 320
- Hertz–Signorini–Moreau conditions, 57
- High contact precision, 58, 288
- History data transfer, 382
- History dependent problems, 382
- Holm–Archard law, 90
- Homogenization, 79
- Hyperelastic material, 32, 129
  - Neo-Hookian, 32
  - rate form, 34
- Ill-conditioning, 288
  - augmented Lagrange method, 110
- Impact, 13
- Implicit time integration, 274
  - friction, 307
- Incremental elasticity tensor, 41
- Incremental hyperelastic constitutive equation, 34
- Initial stress, 41
- Initial value problem
  - thermo-mechanical contact, 320
- Intersection check, 264
- Isoparametric elements, 136
  - computation of gradients, 138
  - deformations, 137
  - one-dimensional shape functions, 139
  - quadrilateral elements, 142
  - three-dimensional interpolation, 143
  - triangular elements, 141
  - two-dimensional shape functions, 141
- Iterated function system, 91
- Jamming, 285
- Kuhn–Tucker–Karush conditions, 3, 57
- Lagrange multiplier method, 100, 114, 163, 280
- Lagrange multiplier
  - algorithm, 286
- Lagrange multipliers, 3
- Lame constants, 33
- Large deformation
  - contact, 204
- Large sliding, 204
- Line search, 295
- Linear Complementary Problem, 297
- Linearization, 118
  - constitutive equations, 40
  - contact terms, 118
  - normal gap, 120
  - normal vector, 120
  - stick case, 124
  - strain measures, 39
  - surface coordinate, 120
  - tangential part, 123
  - variation of gap function, 225
  - weak form, 40, 281
- Load parameter, 271
- Local balance of momentum, 28, 97
- Low contact precision, 56
- Lubrication, 62, 83
- Lubrication
  - viscosity coefficient, 85
- Mass matrix, 148, 163, 273
- Mathematical programming, 99
- Mean plane distance, 60, 112
- Merit function, 295
- Metric tensor, 52, 407
- Micromechanical approach, 58
- Minimization problem, 280
- Minimum distance, 47, 205, 220, 234
- Mortar method, 164, 188, 298
- Moving cone algorithm, 215, 308
- Moving friction cone, 213
- Newmark method, 276
- Newton method, 36, 150, 272, 274, 302, 307, 327, 339, 410
  - linearization, 40
- Nitsche method, 117, 164, 194
- Node-to-edge element, 227
- Node-to-node element, 166, 229, 324
  - friction, 170
  - frictionless, 166
  - Lagrange multipliers, 167

- penalty method, 168
- thermo-mechanical contact, 326
- Node-to-segment element, 204, 213, 231, 298, 328
  - Bezier, 237
  - friction, 207, 242
  - Hermite, 231
  - Lagrange multiplier, 210–211
  - linearization, 210
  - penalty method, 210
  - variation, 205
- Node-to-surface element
  - quad, 219
  - triangle, 225
- Nominal stress, 57
- Non-matching grids, 164, 183
- Non-penetration condition, 47
- Normal contact
  - constitutive equation, 60, 111
  - non penetration condition, 57
  - statistical model, 60
- Normal vector, 47, 407
- Numerical derivative, 392
- Octree method, 266
- Operator split algorithm, 324
- Operator split
  - thermo-mechanical contact, 323
- Parallel computers, 297
- Parameter identification, 78–79, 409
- Penalty method, 4, 102, 115, 166, 225, 232, 239, 280, 284, 390
- Penalty parameter, 101, 285
  - estimation, 285
- Penetration, 47, 390
- Penetration function, 47, 101
- Perturbed Lagrange formulation, 107, 179, 183
- Ploughing, 62
- Polymer
  - friction, 73
  - friction coefficient, 76
- Potential energy, 31
  - contact, 99
- Projection tensor, 57
- Quadratic program, 294–295
- Radiation, 319
- Rate of wear, 90
- Recursive region splitting, 371
- Reference configuration, 20
- Relative error, 382
- Relative tangential velocity, 52, 69, 170
- Representative Volume Element, 79
- Residual
  - current configuration, 157
  - initial configuration, 148
- Return mapping, 213, 261, 307
- Reynolds equation, 85
- Rolling contact, 130
  - creepage, 133
  - gap definition, 131
  - non-penetration condition, 131
  - tangential relative gap velocity, 132
  - weak stick condition, 132
- Rotated frame
  - acceleration, 126
  - Green strains, 128
  - velocity, 126
- Rotating reference frame, 125
- Rubber
  - friction, 73
- Schur complement, 282
- Self-affine surfaces, 74
- Self-contact, 270
- Sequential quadratic programming, 294
- Shaw friction law, 72
- Signorini problem, 95
- Slack variable, 297
- Slip, 50
  - criterion, 68–70, 78, 307
  - linearization, 212
  - moving cone, 215
  - tangential gap, 208
  - velocity, 52
- Smooth contact discretization, 231
  - Bezier, 237
  - friction, 242
  - frictionless, 231
  - Hermite, 231
  - rigid surface, 250
  - three-dimensional, 252
- Snap-through, 389
- Spatial configuration, 20
- Spatial search, 262, 264
- Spot, 316
- SQP method, 294
- Stability, 391
- Static condensation, 282
- Statistical contact laws, 59
- Stick, 50
  - alternative description, 213
  - condition, 63
  - linearization, 211
  - tangential gap, 208
- Strain energy, 30, 32, 99
- Strain tensors
  - Almansi strains, 23, 39
  - Green–Lagrangian strains, 23, 29, 39, 146
  - left Cauchy–Green tensor, 23, 313
  - right Cauchy–Green tensor, 22–23, 40, 312
- Stress divergence, 148
- Stress power, 27–28
- Stress tensor
  - 1. Piola–Kirchhoff stresses, 28

- 2. Piola–Kirchhoff stresses, 28–29, 33, 40–41, 147
  - Cauchy stresses, 26
  - Kirchhoff stresses, 28, 33, 42, 313
- Surface-to-surface element, 172
  - isoparametric interpolation, 173
  - Lagrange multipliers, 174
  - penalty method, 176
  - perturbed Lagrange, 177
- Surface hardness, 60
- Surface loads, 149
- Surface roughness, 58
- Surrounding spheres, 263
- Tangent matrix
  - beam contact, 347
  - node-to-node element, 171
  - contact segments, 188
  - current configuration, 159
  - extended system, 392
  - frictionless, 218, 240
  - initial configuration, 152
  - initial stress matrix, 159
  - initial stresses, 151
  - penalty method, 168
  - slip, 212, 217, 223, 248
  - stick, 211, 214, 248
  - thermo-mechanical contact, 327
- Tangent vector, 47, 405, 407
- Tangential micro displacements, 67
- Tangential relative displacement, 51
- Tangential relative velocity, 52
- Thermal contact resistance, 315
- Thermal expansion coefficient, 313
- Thermo-elasticity
  - Helmholtz energy, 313
  - kinematical split, 312
- Thermo-mechanical contact, 9, 311
  - node-to-node element, 324
  - node-to-segment element, 328
  - gas conductance, 317
  - heat conduction, 314
  - initial value problem, 320
  - multiplicative split, 312
  - operator split, 323
  - radiation, 319
  - spot conductance, 316
  - weak form, 322
- Time integration, 273
  - explicit, 274
  - contact, 299
  - implicit, 276
- Total sliding distance, 53
- Transfer of history variables, 382
- Transfer of state variables, 382
- Tribology, xvii, 61
- Uniqueness, 96
- Unstructured meshes, 183
- Uzawa algorithm, 110, 288, 304
  - symmetric version, 305
  - update formula, 111, 288
- Variation
  - deformation gradient, 29
  - gap, 53
  - gap function, 205, 225
  - Green–Lagrange strain, 29, 147
  - normal gap, 54, 239, 342
  - surface coordinate, 54, 235
  - tangential gap, 54, 206, 344
- Variational inequality, 2, 96, 98, 261, 365
- Velocities, 273
- Viscoelastic material, 129
- Weak form, 40
  - beam contact, 344
  - contact, 99
  - thermo-mechanical contact, 322
  - contact, 98
- Wear, 89
- Wear coefficient, 90
- Wear
  - surface fatigue, 89
  - abrasive, 89
  - adhesive, 89
  - corrosive, 89