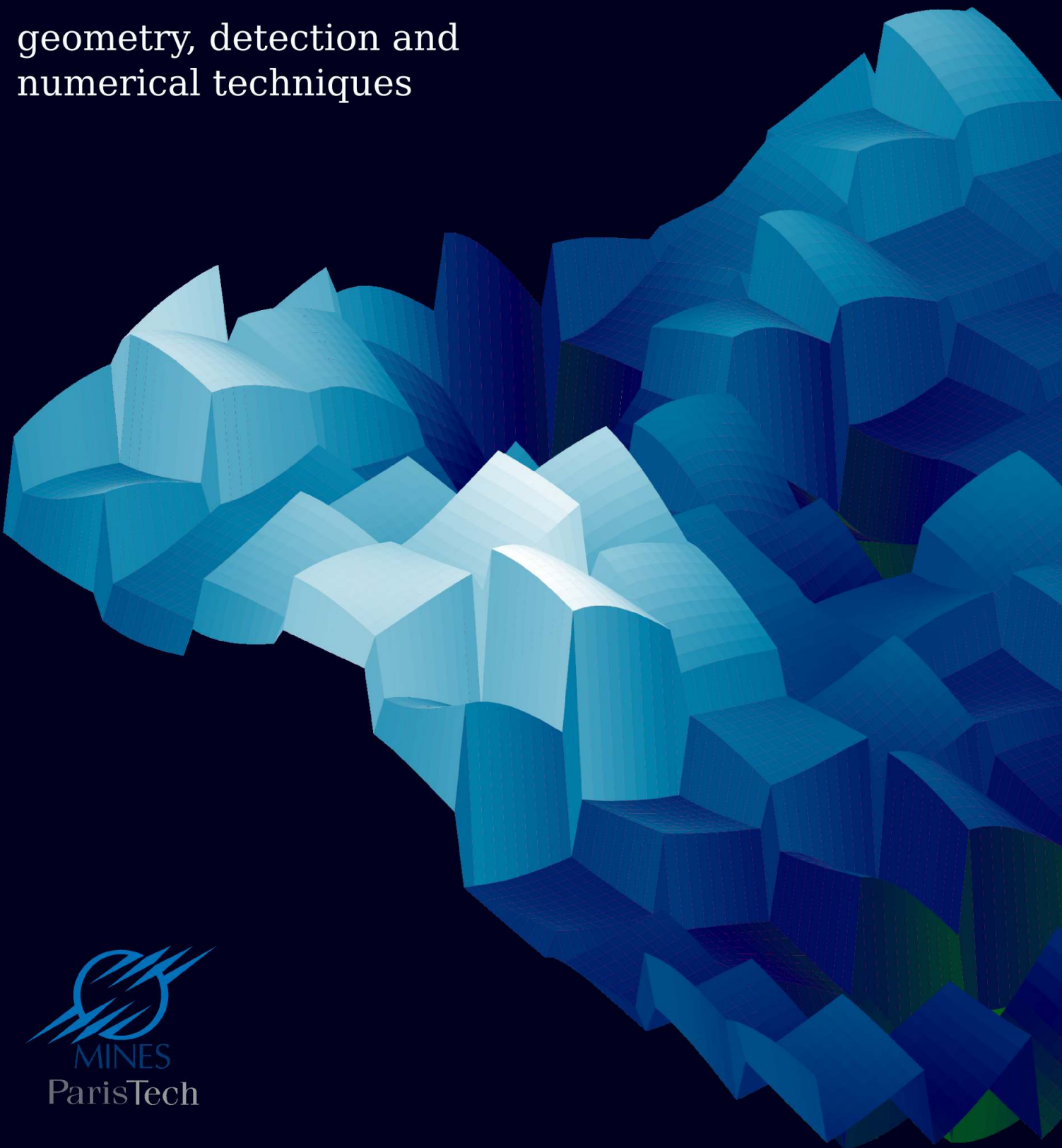


Ph.D. dissertation  
by V.A. Yastrebov

# Computational Contact Mechanics

geometry, detection and  
numerical techniques





École doctorale n°432 : Sciences des Métiers de l'Ingénieur

**Doctorat ParisTech**

**T H È S E**

pour obtenir le grade de docteur délivré par

**l'École nationale supérieure des mines de Paris**

**Spécialité « Mécanique »**

*présentée et soutenue publiquement par*

**Vladislav A. YASTREBOV**

le 25 mars 2011

**Computational contact mechanics:  
geometry, detection and numerical techniques**

**Mécanique numérique du contact :  
géométrie, détection et techniques de résolution**

Directeurs de thèse : **Georges CAILLETAUD,**  
**Frédéric FEYEL**

**Jury**

<b>M. Peter WRIGGERS</b> , Professeur, Leibniz Universität Hannover	Président
<b>M. Pierre ALART</b> , Professeur, Université de Montpellier	Rapporteur
<b>M. Daniel NÉLIAS</b> , Professeur, Université de Lyon, INSA-Lyon	Rapporteur
<b>M. Georges CAILLETAUD</b> , Professeur, Centre des Matériaux, Mines ParisTech	Directeur
<b>M. Frédéric FEYEL</b> , Professeur, Onera	Directeur
<b>M. Boris E. MELNIKOV</b> , Professeur, Université Polytechnique de St. Pétersbourg	Examineur
<b>M. Jean-François MOLINARI</b> , Professeur, École Polytechnique Fédérale de Lausanne	Examineur
<b>M. François COMTE</b> , Docteur, Snecma	Examineur





## Ph.D. dissertation:

Author: **Vladislav A. Yastrebov**  
Centre des Matériaux, MINES ParisTech  
CNRS UMR 7633, France

Title: **Computational contact mechanics:  
geometry, detection and numerical techniques**

Content: 383 pages, 151 figures

Specialization: Computational Mechanics  
Defense date: 25 March 2011  
Defense place: Mines ParisTech, 60 bvd Saint Michel, Paris, France

## Scientific advisers:

Prof. **Georges Cailletaud**  
Centre des Matériaux, MINES ParisTech, CNRS UMR 7633, France

Prof. **Frédéric Feyel**  
ONERA, The French Aerospace Lab, France

## Ph.D. defense committee:

President Prof. Dr.-Ing. **Peter Wriggers**  
Leibniz Universität Hannover, Germany

Reviewer Prof. **Pierre Alart**  
Université de Montpellier, France

Reviewer Prof. **Daniel Nélías**  
Université de Lyon, INSA-Lyon, France

Member Prof. **Boris E. Melnikov**  
St. Petersburg State Polytechnical University, Russia

Member Prof. **Jean-François Molinari**  
École Polytechnique Fédérale de Lausanne, Switzerland

Member Ph.D. **François Comte**  
Snecma, France





*To my teachers*



# Abstract

The goal of this work is to derive a consistent framework for the treatment of contact problems within the Finite Element Method using the Node-to-Segment discretization. Three main components of the computational contact have been considered: geometry, detection and resolution techniques. For the sake of completeness, the mechanical aspects of contact as well as numerous numerical algorithms and methods have been discussed. A new mathematical formalism called “s-structures” has been employed through the entire dissertation. It results in a comprehensive coordinate-free notations and provides an elegant apparatus, available for other mechanical and physical applications. Several original ideas and extensions of standard techniques have been proposed and implemented in the finite element software ZéBuLoN (Z-set). Numerical case studies, presented in the dissertation, demonstrate the performance and robustness of the employed detection and resolution schemes.

Le but de ce travail était de fournir un cadre cohérent pour le traitement des problèmes de contact en utilisant une discrétisation de type nœud à segment. Trois aspects principaux de la mécanique numérique du contact ont été particulièrement considérés : la description de la géométrie, le problème de détection de contact et les techniques de résolution. Le manuscrit contient cependant une présentation complète de la mécanique du contact et des algorithmes numériques qui lui sont attachés. Un nouveau formalisme mathématique – les s-structures – est employé dans l’ensemble de la thèse. Il fournit un cadre de formulation intrinsèque qui permet d’exprimer de façon compacte un grand nombre de problèmes de mécanique et de physique. La thèse propose plusieurs idées originales et des extensions des techniques classiques, qui ont toutes été mises en œuvre dans le code de calcul par éléments finis ZéBuLoN (Z-set). Plusieurs études de cas, présentées dans la thèse, viennent démontrer les performances et la robustesse des méthodes numériques utilisées pour la détection et la résolution.



# Acknowledgements

The three years of my PhD thesis at the Centre des Matériaux of the École des Mines de Paris were a fruitful period of my personal and professional development. The wideness of the subject and the freedom I had in its investigation allowed me to develop many aspects of the computational contact mechanics under a wise guidance of my scientific advisers Georges Cailletaud and Frédéric Feyel, always helpful and open to the new ideas. I highly appreciate their deep knowledge and top professionalism as well as the kind and respectful attitude towards me. For the possibility to work with them I am obliged to the recommendation of Boris Evgenievich Melnikov, whose permanent help and attention I greatly acknowledge.

I thank a lot my dear friends Nikolay Osipov and Djamel Missoum-Benziane for their constant support, friendship, help and the warm atmosphere they created. I am also very much obliged to Djamel for my rapid improvements in French. Further, I am grateful to Bahram Sarbandi who helped me a lot to set up in France and to Yoann Guilhem, Julian Frachon and Guillaume Abrivard for their open-mindedness and the “accueil chaleureux en France”. I was glad to spend the 3 years of my thesis in the same office with my colleague and friend, Sophie Cartel.

I would like to thank Lingtao Sun, Laurent Maze, Florine Maes, Guillaume Martin from the new generations of PhD students for their friendliness. I deeply appreciated the collaboration and friendly interaction with Julian Durand and Michael Fischlschweiger. I thank my colleagues Ozgur Aslan, Guruprasad Padubidri, Saber El-Arem, Jarmila Savkova, Prajwal Sabnis, Johann Rannou, Jean-Didier Garaud, Vincent Chiaruttini, Ibrahima Gueye, Eva Hérigné, Henry Proudhon, Matthieu Maziere and many others for their kind attitude. I am also grateful to my elder colleagues and advisers André Pineau, Jean-Louis Chaboche, Esteban Busso, David Ryckelynck, Samuel Forest, Stephane Quilici, Farida Azzous, Françoise di Rienzo, Anne-Francoise Gourgues and many others for their personal example, support and encouragements.

A special thanks is reserved to Liliane Locicero, Konaly Sar, Odile Adam, Isabelle Olzenski and Anne Piant for their permanent help and moreover, for their empathy and friendly attention. I thank my french teacher Cécile Brossaud for interesting lessons as well as for her patience and creativity.

I thank Siarhei Dubouski and Konstantin Kuzmenkov for bringing the “Russian spirit” to the laboratory’s life. Also I would like to thank Olga Trubienko, now my friend, for our fruitful collaboration on some aspects of my thesis.

A special acknowledgment goes to my professors from Saint-Petersburg State Polytechnical University who demonstrated to me the beauty and depths of science: Vladimir Alexandrovich Palmov, Boris Alexandrovich Smolnikov, Victor Alexeevich Pupyrev, Pavel Andreevich Zhilin, Victor Nilovich Naumov, Anton Miroslavovich Krivtsov, Artem Semenovitch Semenov and Sergey Nikolaevich Kolgatin. I thank my school teacher, Mikhail Sergeevich Zhitomirsky, who gave me a solid base in Mathematics.

I appreciated that Peter Wriggers, Boris E. Melnikov, Jean-François Molinari and François Comte found the time to participate in the committee of my PhD defense. I am very grateful to Pierre Alart and Daniel Nélías for their attentive review of my thesis.

The financial support of the CNRS-SNECMA (grant #47900) is greatly acknowledged as well as the efficient collaboration with François Comte.

Very sincere thanks are addressed to my dear parents, who gave me the life, love and everything they could to make my life happy. I greet my dear brother Igor and thank him for his friendship, attention and brotherhood. And I thank my beloved wife Alexandra for her love, patience, support in my decisions, wise advices and for giving birth to our son Andrey, whose smile makes every day of my life sunny.





# Notations

## Vectors and tensors:

- **Scalar** (zero-order tensor) – small latin and greek letters:

$$a, \alpha, b, \dots$$

- **Vector** (first-order tensor) – underlined small bold latin and greek letters:

$$\underline{c}, \underline{\beta}, \underline{d}, \dots$$

- **Second-order tensor** – capital bold latin letters underlined twice:

$$\underline{\underline{E}}, \underline{\underline{F}}, \dots$$

- **Higher order tensor** – capital bold latin letters underlined twice with upper left index of order:

$${}^3\underline{\underline{G}}, {}^4\underline{\underline{H}}, \dots$$

## V-Vectors and V-tensors:

- **V-scalar** (“vector of scalars”) – small latin and greek letters underlined by a wave:

$$\underline{\sim} i, \underline{\sim} \gamma, \dots \in {}^m_1\mathbf{S}_0^n$$

- **V-vector** (“vector of vectors”) – small latin and greek letters underlined by a line and a wave:

$$\underline{\sim} j, \underline{\sim} \varepsilon, \dots \in {}^m_1\mathbf{S}_1^n$$

- **V-tensor** (“vector of tensors”) – capital bold latin letters underlined by a double line and a wave:

$$\underline{\underline{\sim}} K, \underline{\underline{\sim}} L, \dots \in {}^m_1\mathbf{S}_2^n$$

## T-Vectors and T-tensors:

- **T-scalar** (“tensor of scalars”) – capital bold latin letter underlined by a double wave:

$$\underline{\underline{\sim}} M, \underline{\underline{\sim}} N, \dots \in {}^m_2\mathbf{S}_0^n$$

- **T-vector** (“tensor of vectors”) – small latin and greek letters underlined by a line and a double wave:

$$\underline{\underline{\sim}} o, \underline{\underline{\sim}} \eta, \dots \in {}^m_2\mathbf{S}_1^n$$

- **T-tensor** (“tensor of tensors”) – capital bold latin letters underlined by a double line and a double wave:

$$\underline{\underline{P}}, \underline{\underline{Q}}, \dots \in {}^m_2\mathbb{S}_2^n$$

### Vector and tensor operations:

- $\|\underline{a}\|$  – euclidean norm of a vector;
- $\det \underline{\underline{A}}$  – determinant of a tensor;
- $\underline{\underline{I}}$  – unit tensor;
- $\underline{\underline{I}}$  – unit t-scalar;
- $\text{tr} \underline{\underline{A}}$  – trace of a tensor;
- $\underline{\underline{A}}^{-1}$  – inverse of tensor;
- $\underline{\underline{A}}^T$  – transpose of tensor;
- $\underline{\underline{A}} \cdot \underline{\underline{B}} = \underline{\underline{C}}^{i+j-2}$  – scalar or dot product;
- $\underline{\underline{A}} \times \underline{\underline{B}} = \underline{\underline{C}}^{i+j-1}$  – vector or cross product;
- $\underline{\underline{A}} \otimes \underline{\underline{B}} = \underline{\underline{A}} \underline{\underline{B}} = \underline{\underline{C}}^{i+j}$  – tensor product;
- $\underline{\underline{A}} \cdot \underline{\underline{B}} = \underline{\underline{C}}^{i+j-4}$  – tensor contraction.

### Other operations:

- $(\bullet)' = \frac{d\bullet}{dt}$  – full time derivative;
- $\delta(\bullet), \Delta(\bullet)$  – first variations;
- $\bar{\delta}(\bullet), \bar{\Delta}(\bullet)$  – full first variations;
- $\Delta\delta(\bullet)$  – second variation;
- $\bar{\Delta}\bar{\delta}(\bullet)$  – full second variation;
- $\nabla \otimes (\bullet)$  – gradient;
- $\nabla \cdot (\bullet)$  – divergence;
- $\nabla \times (\bullet)$  – rotor.

### Miscellaneous:

- $\delta_i^j$  – Kronecker's delta  $\delta_i^j = 1$ , if  $i = j$  else  $\delta_i^j = 0$ ;
- $\langle x \rangle = \frac{1}{2}(x + |x|)$  – Macaulay brackets;
- $[\bullet, \bullet]$ ;  $(\bullet, \bullet)$ ;  $(\bullet, \bullet]$  – closed, open, open-closed intervals;
- $\forall, \exists, \exists!, \exists!!$ ,  $\nexists$  – for all, exists, exists only one, exists infinitely many, does not exist;
- $\Rightarrow, \Leftarrow, \Leftrightarrow$  – sufficient, necessary, sufficient and necessary conditions;
- $\min, \max, \text{ext}, \sup, \inf$  – minimum, maximum, extremum, supremum, infimum;
- $\widetilde{\min}, \widetilde{\max}$  – global minimum, global maximum;
- $i = 1, n$  –  $i$  changes from 1 to  $n$ .

### Abbreviations:

- PM, LMM, ALM – penalty, Lagrange multiplier, augmented Lagrangian methods;
- FEM, FEA – Finite Element Method, Finite Element Analysis;
- CAD – Computer-Aided Design;
- NTN, NTS – Node-to-Node, Node-to-Segment discretizations;
- MPC – Multi-Point Constraints;
- PDN – Partial Dirichlet-Neumann;
- SDMR, MDMR – Single Detection Multiple Resolution, Multiple Detection Multiple Resolution.



#### **Remark on Macaulay brackets, $\text{dist}(\cdot, \cdot)$ and $\theta(\cdot)$ functions.**

*Throughout the dissertation we use the notation of Macaulay brackets*

$$\langle x \rangle = \begin{cases} x, & x \geq 0, \\ 0, & x < 0 \end{cases}; \quad \langle -x \rangle = \begin{cases} -x, & x \leq 0, \\ 0, & x > 0 \end{cases}$$

*The  $\theta$  function is a similar notation widely used in both engineering and mathematical literature*

$$\theta(x) = \max(x, 0) = \begin{cases} x, & x \geq 0, \\ 0, & x < 0 \end{cases}; \quad \theta(-x) = \min(x, 0) = \begin{cases} -x, & x \leq 0, \\ 0, & x > 0 \end{cases}$$

or a more general  $\text{dist}(\cdot)$  function

$$\text{dist}(x, \Omega) = \begin{cases} \text{dist}(x, \partial\Omega), & x \notin \Omega \\ 0, & x \in \Omega, \end{cases}$$

where  $\text{dist}(x, \partial\Omega)$  is a somehow defined distance from point  $x$  to the closure of the set  $\Omega$ . For example, in the simplest case  $\Omega = \mathbb{R}_-$ ,  $x \in \mathbb{R}$ , then  $\partial\mathbb{R}_- = 0$

$$\text{dist}(x, \mathbb{R}_-) = \begin{cases} x, & x \geq 0, \\ 0, & x < 0 \end{cases}; \quad \text{dist}(x, \mathbb{R}_+) = \begin{cases} -x, & x \leq 0, \\ 0, & x > 0. \end{cases}$$

All these functions are equivalent for the considered case and interchangeable, so the reader is invited to interpret the Macaulay brackets as one of above mentioned functions to which he is more accustomed.

---

# Contents

<b>1</b>	<b>Introduction to contact mechanics</b>	<b>5</b>
1.1	Historical remark . . . . .	8
1.1.1	The augmented Lagrangian method . . . . .	10
1.2	Basics of the numerical treatment . . . . .	12
1.2.1	Contact detection . . . . .	13
1.2.2	Contact discretization . . . . .	13
1.2.3	Contact resolution . . . . .	16
1.3	Motivation and overview . . . . .	18
<b>2</b>	<b>Geometry in contact mechanics</b>	<b>21</b>
2.1	Introduction . . . . .	22
2.2	Interaction between contacting surfaces . . . . .	27
2.2.1	Normal gap . . . . .	28
2.2.2	The closest point . . . . .	33
2.2.3	Aircraft's shadow projection method . . . . .	39
2.2.4	Tangential relative sliding . . . . .	42
2.2.5	First order variations . . . . .	45
2.2.6	Second order variations . . . . .	52
2.3	Numerical validation . . . . .	61
2.4	Discretized geometry . . . . .	64
2.4.1	Shape functions and finite elements . . . . .	64
2.4.2	Geometry of contact elements . . . . .	67
2.5	Enrichment of contact geometry . . . . .	74
2.5.1	Example of enrichment . . . . .	87
<b>3</b>	<b>Contact detection</b>	<b>91</b>
3.1	Introduction . . . . .	92
3.2	All-to-all detection . . . . .	96
3.2.1	Preliminary phase . . . . .	96
3.2.2	Detection phase . . . . .	100
3.3	Bucket sort detection . . . . .	105
3.3.1	Preliminary phase . . . . .	106
3.3.2	Numerical tests . . . . .	107
3.3.3	Detection phase . . . . .	110
3.3.4	Multi-face contact elements . . . . .	111
3.3.5	Improvements . . . . .	112

3.4	Validation and performance . . . . .	113
3.5	Case of unknown master-slave discretizations . . . . .	115
3.6	Parallel contact detection . . . . .	120
3.6.1	General presentation . . . . .	120
3.6.2	Single Detection, Multiple Resolution approach . . . . .	120
3.6.3	Multiple Detection, Multiple Resolution approach . . . . .	121
3.6.4	Scalability test . . . . .	123
3.7	Conclusion . . . . .	125
<b>4</b>	<b>Formulation of contact problems and resolution methods</b>	<b>127</b>
4.1	Unilateral contact with a rigid plane . . . . .	128
4.1.1	Interpretation of contact conditions . . . . .	131
4.1.2	Friction . . . . .	133
4.1.3	Interpretation of frictional conditions . . . . .	137
4.1.4	Non-classical friction and adhesion laws . . . . .	140
4.2	Unilateral contact with an arbitrary rigid surface . . . . .	146
4.2.1	Non-penetration condition . . . . .	147
4.2.2	Hertz-Signorini-Moreau's contact conditions . . . . .	150
4.2.3	Interpretation of contact conditions . . . . .	151
4.2.4	Frictional conditions and their interpretation . . . . .	152
4.3	Contact between deformable solids . . . . .	153
4.3.1	General formulation and variational inequality . . . . .	153
4.3.2	Remarks on Coulomb's frictional law . . . . .	159
4.4	Variational equality and resolution methods . . . . .	162
4.5	Penalty method . . . . .	162
4.5.1	Frictionless case . . . . .	163
4.5.2	Example . . . . .	165
4.5.3	Nonlinear penalty functions . . . . .	168
4.5.4	Frictional case . . . . .	170
4.6	Method of Lagrange multipliers . . . . .	173
4.6.1	Frictionless case . . . . .	174
4.6.2	Frictional case . . . . .	175
4.6.3	Example . . . . .	178
4.7	Augmented Lagrangian Method . . . . .	185
4.7.1	Introduction . . . . .	185
4.7.2	Application to contact problems . . . . .	188
4.7.3	Example . . . . .	195
<b>5</b>	<b>Numerical procedures</b>	<b>203</b>
5.1	Newton's method . . . . .	204
5.1.1	One-dimensional Newton's method . . . . .	204
5.1.2	Multidimensional Newton's method . . . . .	207
5.1.3	Application to nondifferentiable functions . . . . .	208
5.1.4	Subdifferentials and subgradients . . . . .	209
5.1.5	Generalized Newton method . . . . .	212
5.2	Return mapping algorithm . . . . .	216
5.3	Finite Element Method . . . . .	223



5.3.1	Introduction . . . . .	223
5.3.2	Contact elements . . . . .	227
5.3.3	Discretization of the contact interface . . . . .	230
5.3.4	Virtual work for discretized contact interface . . . . .	232
5.3.5	Linearization of equations . . . . .	234
5.3.6	Example . . . . .	236
5.4	Residual vectors and tangent matrices . . . . .	237
5.4.1	Penalty method: frictionless case . . . . .	237
5.4.2	Penalty method: frictional case . . . . .	239
5.4.3	Augmented Lagrangian method: frictionless case . . . . .	247
5.4.4	Augmented Lagrangian method: frictional case . . . . .	249
5.5	Method of partial Dirichlet-Neumann boundary conditions . . . . .	256
5.5.1	Description of the numerical technique . . . . .	256
5.5.2	Frictionless case . . . . .	257
5.5.3	Frictional case . . . . .	262
5.5.4	Remarks . . . . .	262
5.6	Technical details . . . . .	264
5.6.1	Rigid master surface . . . . .	264
5.6.2	Multi-face contact elements and smoothing techniques . . . . .	265
5.6.3	Heterogeneous friction . . . . .	267
5.6.4	Short remarks . . . . .	269
<b>6</b>	<b>Numerical examples</b>	<b>273</b>
6.1	Two dimensional problems . . . . .	274
6.1.1	Indentation by a rigid flat punch . . . . .	274
6.1.2	Elastic disk embedded in an elastic bored plane . . . . .	279
6.1.3	Indentation of an elastic rectangle by a circular indenter . . . . .	283
6.1.4	Axisymmetric deep cup drawing . . . . .	285
6.1.5	Shallow ironing . . . . .	289
6.1.6	Axisymmetric post buckling of a thin-walled cylinder . . . . .	294
6.2	Three dimensional problems . . . . .	298
6.2.1	Accordion post buckling folding of a thin-walled tube . . . . .	298
6.2.2	Hydrostatic extrusion of a square plate through a circular hole . . . . .	302
6.2.3	Frictional sliding of a cube on a rigid plane . . . . .	309
<b>7</b>	<b>Conclusions and perspectives</b>	<b>315</b>
7.1	Original contributions . . . . .	316
7.1.1	Geometry . . . . .	316
7.1.2	Detection . . . . .	317
7.1.3	Resolution . . . . .	318
7.2	Intermediate results and perspectives . . . . .	319
7.2.1	Normal contact of rough surface . . . . .	319
7.2.2	Aeronautical applications . . . . .	321

<b>A</b>	<b>Vectors, tensors and s-structures</b>	<b>325</b>
A.1	Fundamentals . . . . .	325
A.2	Vector space basis . . . . .	330
	A.2.1 Transformation matrices, covariant and contravariant objects . . . . .	332
	A.2.2 Gradient operator or Hamilton's operator . . . . .	334
A.3	Sub-basis, vector function of v-scalar argument . . . . .	336
A.4	Tensors . . . . .	339
A.5	Tensor as a linear operator on vector space . . . . .	346
A.6	S-structures . . . . .	347
	A.6.1 Formal definition, notations and types . . . . .	350
	A.6.2 Simple operations . . . . .	353
	A.6.3 Invariant s-structures . . . . .	354
	A.6.4 Scalar products of V-vectors . . . . .	356
	A.6.5 Inverse v-vector . . . . .	360
	A.6.6 Isomorphism of s-space and tensor space . . . . .	361
	A.6.7 Tensor product of v-vectors . . . . .	364
A.7	Reduced form of s-structures . . . . .	365

# Chapter 1

## Introduction to contact mechanics

### **Résumé de Chapitre 1 «Introduction à la mécanique de contact»**

*Ce chapitre présente une brève histoire des développements de la mécanique du contact, de sa naissance jusqu'à nos jours. On insiste en particulier sur les aspects numériques et on présente une formulation mathématique rigoureuse des problèmes associés. La littérature concernant la méthode du Lagrangien augmenté est exposée en détail.*

*De plus, ce chapitre introduit les notions de base qu'on exploite au cours du manuscrit. Pour donner une vue globale sur le traitement numérique des problèmes de contact, on éclaire toutes les étapes de l'algorithme : la détection du contact, la discrétisation et la résolution. On présente la revue des possibilités existantes pour chaque de ces étapes et on argumente les choix qui seront effectués par la suite : méthode de «bucket sort» modifiée pour la détection, méthode du Lagrangien augmenté et méthode de pénalisation pour la résolution et la discrétisation «Nœud-à-Segment».*

*On expose également les éléments qui ont motivé ce travail et on présente le plan du manuscrit.*

From a mechanical point of view, at macroscale, *contact* is a notion for all types of interactions between separate bodies coming in touch. Direct contact between solids allows to transfer a load, a heat and an electric charge from one body to another. The physics of the contact interaction is particularly rich and complicated, due to the multiscale and multiphysical nature of the phenomenon. The branch of mechanical engineering studying this interaction is called *tribology* – a science of relative motion of interacting surfaces in a comprehensive framework combining mechanical, physical and chemical effects at different scales. This dissertation presents the mathematical description and modeling of the *mechanical aspects* of this interaction.

Contact problems in mechanics of deformable solids can be singled out in a particular class. There are several reasons for that. Contact occurs in

the interface formed by two separate continuous surfaces. However, the contact constraints arising in this interface cannot be replaced by ordinary boundary conditions imposed on both contacting surfaces. At the same time, the contact interface itself cannot be simply considered as an internal surface. In an idealized case, the contact interface is a zero thickness layer, which sustains only compressive stress in the direction orthogonal to the contact interface (Fig. 1.1,a.), any stretching leads to vanishing of the contact interface (Fig. 1.1,b.). In case of frictionless contact, the contact interface contrary to an ordinary internal surface, does not sustain any tangential efforts, which allows two surfaces slide relatively to each other (Fig. 1.1,a.). In case of frictional contact, tangential resistance of the contact interface is similar to the resistance of an elasto-plastic material with saturation. For example, in case of the classic Coulomb's friction law in stick state, the contact interface represents an internal surface – no separation, no tangential sliding – locally both surfaces remain glued to each other (Fig. 1.1,c.). If a critical shear stress is reached, the surfaces start to slip relatively to each other, however the nonzero shear stress remains equilibrated (Fig. 1.1,d.). It follows from this simple representation that the contribution of the contact interface to the energy of the system is always zero except in the case of frictional slip.

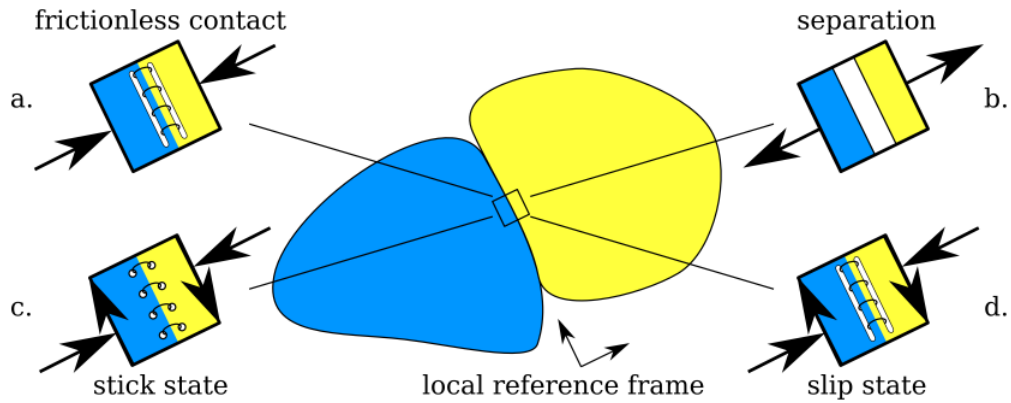


Figure 1.1: Analogy between contact interface and internal interface: **a** – frictionless contact sustains compressive stress in the local reference frame, **b** – any stretching leads to vanishing of contact interface, **c** – frictional contact interface can transfer shear stress; **d** – in Coulomb's friction law in stick state there is no relative sliding up to reaching a critical shear stress.

Mechanical problems are classically formulated as boundary value problems, where a governing differential equations should be fulfilled within the domain  $\Omega$  and ordinary boundary conditions are imposed on the domain's closure  $\partial\Omega$ . The balance of virtual work yields a weak (integral) form of this boundary value problem, which presents a basis on which the structural Finite Element Method is constructed. Contact constraints are formulated as sets of inequalities. Such a formulation is not usual for boundary value problems. The rigorous construction of a variational principle leads to a variational inequality

instead of a classic variational equality. Such a new mathematical structure requires new solution approaches. The problem becomes even more complex when a frictional effect is assumed at the interface. Coulomb's friction law states that tangential resistance depends upon the normal contact pressure, but the latter is known only if the solution is known. Roughly speaking, the boundary conditions are solution dependent, which naturally leads to difficulties in the formulation of the frictional contact problem. Moreover, the nature of Coulomb's friction law yields a nonsmooth energy functional resulting in even more difficulties from a numerical point of view. As pointed out in the book of Kikuchi and Oden [Kikuchi 88] "Frictional contact problem between continuous deformable solids involves formidable mathematical difficulties".

Another mathematical difficulty in contact mechanics arises from a rigorous description of continuous interacting surfaces. First, contacting bodies may penetrate each other or be separated. In both cases, a bijection between points of the contacting surfaces does not always exist. Second, the finite element discretization results only in piecewise smooth contacting surfaces, which leads to mathematical and numerical difficulties. Third, a considerable effort has to be undertaken to derive a rigorous linearization of the variational principle, which in turn requires second order variations of the normal gap and the tangential sliding, which is not an easy task. Basic knowledges of differential geometry is needed to obtain the relevant quantities.

The discretization of the contact interface is a third challenge in computational contact mechanics. A simple and stable discretization for conforming meshes, i.e. each node on one contacting surface has a corresponding node on the other surface, can be established only in case of small deformations and infinitely small relative sliding. Such a discretization is called *Node-to-Node*. A less simple but multipurpose discretization implies the creation of contact pairs consisting of a node of one surface and a corresponding segment of the other surface. This approach is known as *Node-to-Segment discretization*. However, this discretizations does not fulfill the so called Babuška-Brezzi conditions and leads to an unstable discretization. Recently, new techniques based on *segment-to-segment discretizations* – *Nitsche* and *mortar methods* – have been successfully introduced in computational contact mechanics, however, the computer implementation of these methods for a general case presents a real challenge both from algorithmic and technical points of view. Seeking for a stable and relatively simple discretization of the contact interface is still in progress.

All forementioned difficulties are related to the resolution phase of the contact algorithm. It follows the detection phase, which determines the contacting pairs on discretized surfaces. At first glance, the detection phase is a standalone task, but in reality it appears to be strongly connected with the discretization type of the contact interface, with the definition of the gap function and with the type of contact (e.g., simple contact or self-contact). The detection phase may present a bottleneck for an efficient treatment of contact problems, both for rapidity and robustness. The contact detection becomes one of the most crucial points for an efficient parallelization of the whole

resolution scheme. Elaboration and implementation of an efficient contact detection algorithm is an absolute necessity for a robust and fast Finite Element Analysis of large contact problems.

In this introduction, after a brief historical review, the main notions of contact mechanics and related applications will be given, followed by a short presentation of the general concept of contact treatment in the framework of the Finite Element Method and implicit integration. The questions of detection, discretization and resolution will then be addressed in first approximation. Further, the main physical aspects of frictional contact will be introduced. Finally the contents of the dissertation will be presented.

## 1.1 Historical remark

The modern contact mechanics is about 130 years old. It started in 1882 with the publication of **Hertz's** famous paper *On the contact of elastic solids* [Hertz 82], which gives the solution for frictionless contact between two ellipsoidal bodies. This problem had arisen from the problem of the optical interference between glass lenses. Further developments in the contact theory appear only in the beginning of XX<sup>th</sup> century in application to railways, to reduction gears and to rolling contact bearing industry. Progress in contact mechanics was associated with removing the restrictions of the Hertz theory, such as pure elasticity, frictionless and small deformations. A large contribution has been made by the Russian school of mechanicians, starting from **Galin** [Galin 53], [Galin 76] and **Muskhelishvili** [Muskhelishvili 66]. A synthesis of analytical solutions and approaches for contact problems can be found in monographs [Lurie 70], [Alexandrov 83], [Johnson 94], [Goryacheva 98], [Goryacheva 01], [Vorovich 01], etc.

Since the analytical solution is achievable only for a few simple geometries, boundary conditions, and mostly for linear materials, only rough approximations based on these solutions can be established for complicated frictional contact problems. These problems come from industrial needs and are usually coupled with complex geometries, boundary conditions and non-linear materials. For that reason, with approaching computer age, more and more numerically based semi-analytical solutions for contact problems appear. But it is still not sufficient to answer the industrial demand for a fast and accurate resolution of contact problems, which may include friction, wear, adhesion, large deformations, large sliding and non-linear material.

Since 1965 (NASTRAN) the Finite Element Method (FEM) becomes one of the most usable and efficient tools for the treatment of problems in structural mechanics. In order to fulfill industrial demands related to contact problems, the scientific society worked out a rigorous mathematical framework valid for incorporation of the contact in the Finite Element Method. This task required formidable efforts from the mathematico-mechanical community. First, the frictionless Signorini's problem (unilateral contact between a deformable body and a rigid foundation) has been treated, further the developed approaches have been extended to the case of unilateral frictional contact in small and

large deformations and finally to bilateral<sup>1</sup> or multibody contact. At the same time, the engineering practice tested the solution schemes and proposed new challenging tasks. The work on a stable approach for treatment of large sliding frictional contact is still in progress.

The history of the computational contact began in 1933 with the works of **Signorini** who was the first who formulated the general problem of the equilibrium of a linearly elastic body in frictionless contact with a rigid foundation [Signorini 33], [Signorini 59]. The works of **Fichera** represents the first treatment of questions of existence and uniqueness of the *variational inequalities* arising from the minimization of functionals on convex subsets of Banach spaces, which yields from his rigorous analysis of a class of Signorini's problems [Fichera 63], [Fichera 64], [Fichera 72]. Variational inequality is a new structure in the field of the optimization theory; new approaches are required to make use of such formulations for practical problems of physics and mechanics. "Inequalities in mechanics and physics" by **Duvaut** and **Lions** (first published in French and rapidly translated in English [Duvaut 76]) was a real scientific breakthrough in this direction, the authors investigated the solution of frictional contact problems and large deformation contact. Among the early relevant contributions related to contact problems, the following can be enumerated **Cocu** [Cocu 84], **Panagiotopoulos** [Panagiotopoulos 85], **Rabier et al** [Rabier 86]. A consistent description of the variational inequality approach to contact problems is given in the book by **Kikuchi and Oden** [Kikuchi 88], where among other important results the existence and uniqueness of the solution of Signorini's problem is proven. Stability questions of contact problem solution have been discussed by **Klarbring** [Klarbring 88]; examples of non-uniqueness or non-existence were demonstrated by **Klarbring** [Klarbring 90] and **Martins et al.** [Martins 94]. The existence and uniqueness results for dynamic contact problems can be found in **Martins and Oden** [Martins 87], **Jarusek and Eck** [Jarusek 99] and others.

The *frictionless* contact problem formulated as a variational inequality presents a special type of minimization problems with inequality constraints, which can be efficiently treated in a standard manner (penalty method, Lagrange multiplier method, augmented Lagrangian method, etc.). Unfortunately, there is no associated minimization principle for the *frictional* contact problem [Kikuchi 88], [Mijar 00]. Such a problem is rather complicated and unusual for optimization theory since the energy of the system (objective function) depends on the frictional status which depends on the normal contact pressure, which in turn depends on the displacements, i.e. on the solution of the problem which again depends on the energy of the system. Since there is no smooth energy functional associated with the frictional contact problem, its formulation and resolution present real challenges.

The assumption of a known a priori contact interface on the current computational step results in a reformulation of the variational inequality into a variational equality problem with a special contact term; the form of this term

---

<sup>1</sup>*bilateral* - contact between two or more deformable solids, in contrast to *unilateral* contact - contact between a deformable and a rigid solid.



depends on the method chosen to enforce the contact constraints. Among the well-known and widely used methods there are: barrier and penalty methods, Lagrange multiplier methods and their combinations. Another branch of methods makes use of different techniques from mathematical programming: application of the simplex method to contact problems can be found in [Chand 76], parametric quadratic programming method is employed in [Klarbring 86], [Zhong 88]. Separately from these two branches, there is a group of direct methods, which treats the contact problem independently from the structural one: the flexibility method proposed by Francavilla and Zienkiewicz [Francavilla 75], modified and improved by Jean [Jean 95], rarely mentioned in the scientific literature, in practice this method demonstrates a higher robustness and rapidity in comparison to ordinary methods if the number of nodes in contact remains moderate. But this method is not applicable for large contact problems and its parallelization is hardly possible. A detailed description of the method and its application can be found in [Wronski 94]. A complete list of methods used for the numerical treatment of contact problems can be found in [Wriggers 06] and [Laursen 02].

### 1.1.1 The augmented Lagrangian method

As mentioned in the previous section, the assumption of a known a priori contact surface allows to replace the variational inequality by a variational equality with an additional contact term. The form of this contact term depends upon the choice of the optimization method; the most usable in contact mechanics are the Lagrange multiplier method, the linear penalty method and an augmented Lagrangian method, the two latter methods are implemented in leading modern finite element analysis softwares: ANSYS [Bhashyam 02], [Oatis 07], [ANS 05], ABAQUS [ABA 07], COMSOL [COM 10] and others. In this dissertation all forementioned methods are considered, but a particular attention is paid to the augmented Lagrangian method, possessing several advantages in comparison to other methods.

Within the framework of classical *Lagrange multiplier method* (LMM), contact conditions are exactly satisfied by the introduction of extra degrees of freedom called Lagrange multipliers. The constrained minimization problem converts into an unconstrained saddle point problem often called min-max problem. Due to inequality constraints this formulation has to be considered in combination with an active set strategy [Luenberger 03], [Murty 88], i.e. a check and update of active and passive constraints should be integrated in the convergence loop. Moreover, the additional degrees of freedom of the LMM introduce supplementary computational efforts. *Penalty method* (PM) is simple to implement and to interpret from the physical point of view, but, on the other hand, the contact conditions are fulfilled exactly only in case of the infinite penalty parameter which results in ill-conditioning of the numerical problem. The *augmented Lagrangian method* (ALM) is a sort of Lagrange multiplier formulation regularized by penalty functions. It yields a smooth energy functional and fully unconstrained problem, resulting in exact fulfillment of contact constraints with a finite value of the penalty parameter.



In this section a few historical remarks concerning the augmented Lagrangian method are given. For a more detailed background the reader is referred to the articles and books cited below.

The augmented Lagrangian method has been proposed in the first raw approximation by **Arrow and Solow** in 1958 [Arrow 58b]. Further a more elaborated version of the ALM method for optimization problems subjected to equality constraints has been independently proposed by **Hestenes** [Hestenes 69] and **Powell** [Powell 69] in 1969. As mentioned by **Pietrzak** [Pietrzak 97] it was proposed "rather in an intuitive way" and a lot of questions have not been considered. The way to apply the ALM method to optimization problems with inequality constraints has been developed by **Rockafellar** [Rockafellar 70], [Rockafellar 73b] and **Wierzbicki** [Wierzbicki 71].

Using the augmented Lagrangian method as well as the Lagrange multiplier method leads to the saddle point problem, i.e. the objective function is to be minimized by "ordinary" primal variables (*e.g.*, displacement degrees of freedom (dof) in the displacement based FEM) and is to be maximized by dual variables - Lagrange multipliers which represent contact stresses. All forementioned authors approach this min-max (saddle point) problem by an independent consecutive updating of the primal and dual degrees of freedom. An algebraic formula is used to update the Lagrange multipliers at each iteration step and consequently a standard minimization procedure is used to update the primal degrees of freedom. This idea has been worked out by **Powell** [Powell 69]. Nowadays such an approach is employed under the name of Uzawa's algorithm and the full method is referred as a nested augmented Lagrangian algorithm. Another approach has been developed by **Fletcher** [Fletcher 70]. It consists in a continuous minimization of the resulting saddle problem with a simultaneous update of both primal and dual variables.

One of the first applications of the augmented Lagrangian method to frictionless contact problem can be found in **Glowinski and Le Tallec** [Glowinski 89] and **Wriggers, Simo and Taylor** [Middleton 85]. The first application of the augmented Lagrangian method with Uzawa's algorithm to frictional problems has been reported by **Simo and Laursen** [Simo 92]. The first successful attempt to apply the coupled augmented Lagrangian method to frictional contact problems has been undertaken by **Alart** [Alart 88], and **Alart and Curnier** [Alart 91]. The augmented Lagrangian approach has been elaborated by developing the perturbation approach to convex minimization as proposed in [Rockafellar 70] and first applied by Fortin [Fortin 76] to viscoplastic flow problems (rather similar to frictional contact problems).

Further developments of the ALM method to large deformations, large sliding and nonlinear materials can be found in [Heegaard 93], [Mijar 04a], [Mijar 04b], *etc.* A comprehensive investigation on the implementation of the ALM method in the framework of the Finite Element Method to large deformation frictional contact problems has been carried out by **Pietrzak and Curnier** [Pietrzak 97], [Pietrzak 99]. The attempts to work out a technique for penalty parameter updating are worth mentioning, since it became a crucial factor for convergence of the ALM. A direction was proposed in early works [Hestenes 69] and [Powell 69]. The need was mentioned

by Rockafellar [Rockafellar 73b], discussed in [Alart 97] and an approach has been proposed by Mijar and Arora [Mijar 04a], [Mijar 04b]; another phenomenological approach has been proposed in [Bussetta 09]. An early attempt to parallelize the ALM has been undertaken by Barboteu and Alart [Barboteu 99] for particular structures.

The augmented Lagrangian method combines advantages of both methods LMM and PM and avoids their drawbacks, precisely it converges to the exact solution for a finite value of the penalty coefficient and if a nested update of dual variables is used, there is almost no additional computational efforts. Following Pietrzak, we would like to emphasize the smoothing effect of the ALM which is not the only advantage over ordinary LMM. Even in case of a smooth objective function the ALM method shows its superiority. The ordinary LMM does not fully reduce the optimization problem with inequality constraints to an unconstrained problem, since the condition of positivity of the Lagrange multipliers  $\lambda \geq 0$  has to be satisfied. The ALM method does not have this restriction and therefore is better for practical use. An elaborated presentation of the method will be given in Section 4.7.

## 1.2 Basics of the numerical treatment of contact problems

The part of the implicit Finite Element code aimed at the treatment of contact problems consists in the following steps: contact detection, construction of “contact elements”, incorporation of these elements with associated residual vectors and tangential matrices in the general nonlinear problem and finally resolution of the resulting problem. Here we give the main ideas and a general view of these steps, which will be presented in details further in the corresponding chapters.

*Contact elements* are a kind of “bridge elements” between locally separated but potentially interacting surfaces. Each contact element contains components (nodes, edges, segments or their parts) of both surfaces; the composition of these components depends upon the choice of the *contact discretization* method. Each contact element has its own vector of unknowns, residual vector and tangential matrix, which are assembled with unknowns, residual vectors and matrices of ordinary structural elements. The set of unknowns and the structure of the residual vector and the tangential matrix are determined by the *resolution method*. For example, in addition to primal unknowns (e.g. displacement) contact elements may contain dual unknowns (Lagrange multipliers) representing contact stresses.

The *Contact detection* is a step preceding all others. The aim of this step is to create contact elements containing the **proximal** components of both surfaces which may contact on the current solution step. As a consequence the detection algorithm is based on a search for the closest components and presents a particular algorithmic task. The criterion of proximity is either provided by a user or is chosen automatically based on boundary conditions and/or discretization of contacting surfaces. In order to incorporate contact elements in the resolution cycle, they should be created before a contact occurs

and if needed should be removed and recreated at each solution step. Contrary to this scheme, in case of explicit integration, the searching step consists in the detection of penetration, which has already occurred.

In order to treat contact problems, from the programmer's point of view, a standard finite element code has to be complemented by

1. a class governing contact;
2. a contact detection algorithm;
3. a class of contact elements;
4. the corresponding residual vectors and tangential matrices.

### 1.2.1 Contact detection

The development of numerical methods and the increasing demands on complexity (large deformation, large sliding, self-contact) and size of problems in computational contact mechanics entailed the development of contact detection techniques. As previously mentioned the contact detection presents a purely algorithmic task and is strongly connected with the discretization of the contact interface. For example, in the case of Node-to-Node discretization, the contact detection consists simply in establishing close pairs of nodes: nodes from one surface form pairs with their closest opponents from another surface. Since the Node-to-Node discretization is limited to small deformation and infinitely small slidings, once created contact pairs do not change during the solution steps. Node-to-segment discretization requires a more elaborated detection procedure: for nodes of one surface (slave) the closest point on the other surface (master) has to be found, the master segment possessing this point complemented by the slave node forms a Node-to-Segment contact element.

This simple detection procedure generates several difficulties. First, the detection of the closest point on the master segments may fail if the slave node is not sufficiently close to the master surface or if the latter is not smooth, which is the case in case of finite element discretization of the surface. The numerical scheme of the closest point detection is based on the seeking for a minimum of the distance function, but on the one hand this minimum does not always exist, and on the other hand there may be several or infinitely many equivalent minimum points. Second, the detection has to be organized in a smart way. Large contact problems imply a large number of contacting nodes on both surfaces, that is why a simple detection technique, based on a comparison of distances from each slave node to all components of the master surface, leads to an excessively time-consuming algorithm, especially if contact elements must be frequently updated.

Segment-to-segment discretization requires totally different detection algorithms based on surface topologies. Since we confine ourselves to consideration of the Node-to-Segment contact discretization, the questions of detection for other discretizations will be omitted. The geometrical questions of the closest point definition will be discussed in Chapter 2 and the detection algorithms will be presented in Chapter 3.

### 1.2.2 Contact discretization

As already mentioned, the contact discretization predetermines the structure of contact elements transferring efforts from one contacting surface to another. Three main types of discretizations may be distinguished:

- Node-to-Node, NTN
- Node-to-Segment, NTS
- Segment-to-Segment, STS

The simplest and the oldest *Node-to-Node* discretization [Francavilla 75] (Fig. 1.2) does not allow any finite sliding or large deformations and introduces restrictions on mesh generation. On the other hand it passes the contact patch test – uniform pressure is transferred correctly through the conforming contact interface. The NTN discretization is applicable for linear and quadratic elements in two dimensional case and only to linear elements in three dimensional case. The NTN technique smoothes the asymmetry between contacting surfaces. However, the normal vector for each pair of nodes is usually determined according to one of the surfaces. Different possibilities of normal definition are presented in Remark 3.2 in Section 3.5.

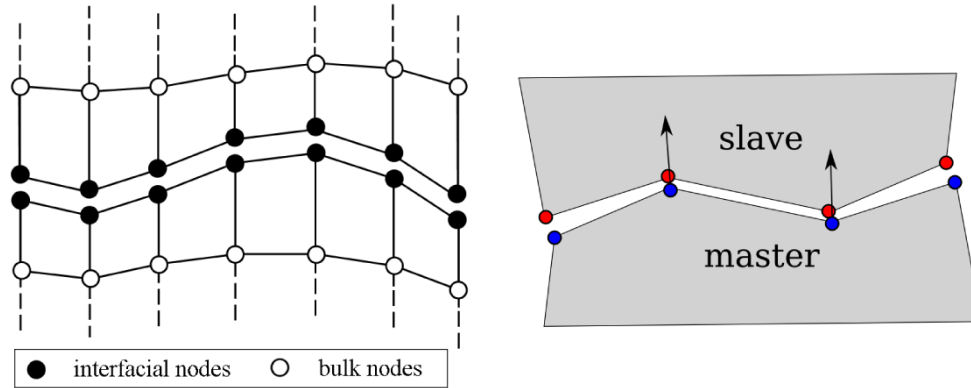


Figure 1.2: Graphical representation of the Node-to-Node discretization, associated pairs of nodes and corresponding normals constructed on the master surface.

*Node-to-segment* (Fig. 1.3) is a multipurpose discretization technique [Hughes 77], valid for non-conforming meshes, large deformation and large sliding. But this discretization is not stable and does not pass the contact patch test proposed by Taylor and Papadopoulos [Taylor 91] for non-conforming meshes – a uniform contact pressure cannot be transferred correctly through the contact interface (see. Fig. 1.4). However, this discretization technique passes this patch test in “double pass” for LMM, which means that at each solution step the problem is solved twice: on the first step one assignment of master and slave surfaces is assumed and on the second step the master and slave surfaces are exchanged. A comprehensive discussion of

the contact patch test for NTS discretizations can be found in [Crisfield 00c], where a new approach combining linear and quadratic shape functions is also suggested. Recently, a modification of the NTS discretization has been proposed [Zavarise 09a], which passes the patch test if the PM is used. Besides the drawbacks of this discretization, it is quite simple and robust, that is why it is often implemented in commercial Finite Element Analysis packages. Contact detection and resolution techniques presented in this dissertation are suitable for the NTS discretization.

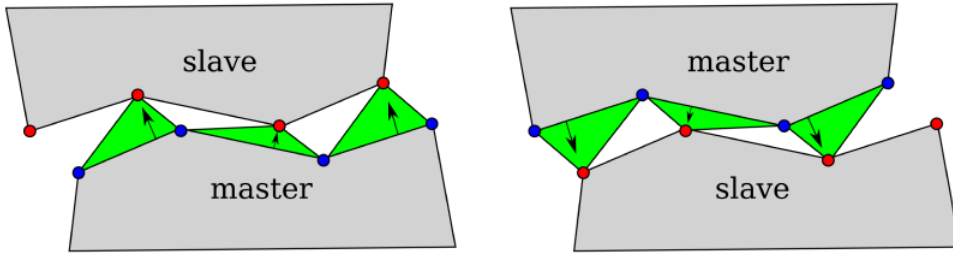


Figure 1.3: Graphical representation of the Node-to-Segment discretization for different choices of master and slave.

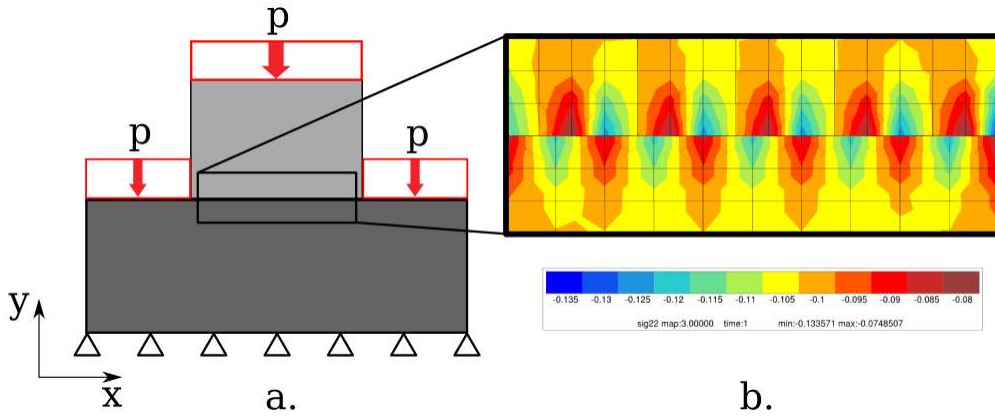


Figure 1.4: **a** – Scheme of the Taylor contact patch test; **b** – the resulting nonuniform distribution of the stress component  $\sigma_{yy}$  in case of the NTS discretization.

Recently, another technique based on a symmetric Node-to-Segment discretization, the *Contact Domain Method*, has been proposed in [Oliver 09], [Hartmann 09]. The discretization of the contact interface is based on a full triangulation of the zone between contacting surfaces based on surface nodes (Fig. 1.5). This formulation seems to be rather stable and passes the patch test, but its three dimensional implementation reported in [Oliver 10] is not applicable for arbitrary discretizations of the contacting surfaces.

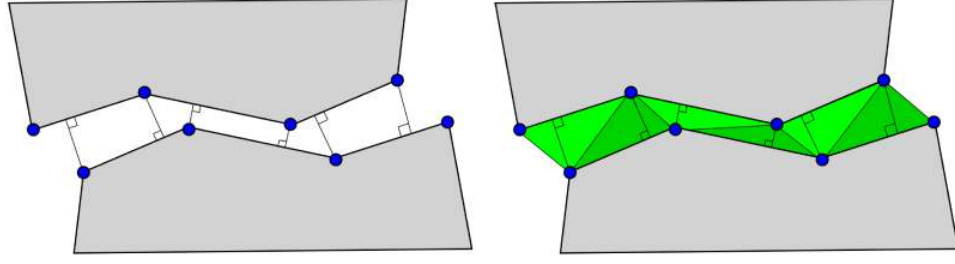


Figure 1.5: Graphical representation of the discretization in the Contact Domain Method, including a triangulation of the contact interface.

*Segment-to-Segment* discretization (Fig. 1.6) has been first proposed by Simo et al. [Simo 85] for the two dimensional case (see also [Zavarise 98]). Recently such a discretization has been efficiently applied to two and three dimensional problems coupled with the *mortar method* for nonconforming meshes, inspired by the domain decomposition methods [Wohlmuth 01]. This technique is stable and passes the patch test but its implementation for a general case presents a great challenge, “a nightmare”, according to Tod A. Laursen, one of the authors of the mortar method’s implementation for two and three dimensional both structural and contact problems [Puso 03], [Puso 04], [Yang 05], [Yang 08b], [McDevitt 00].

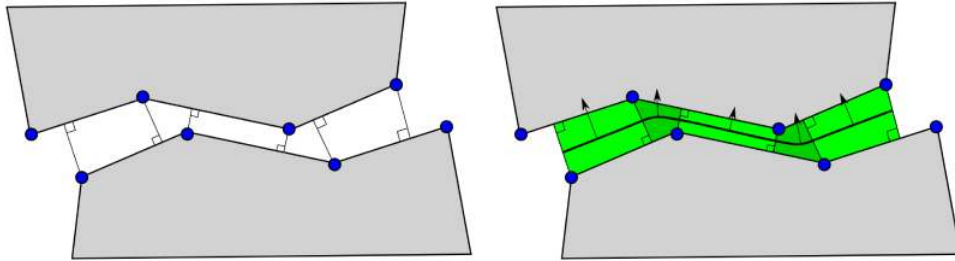


Figure 1.6: Graphical representation of the Segment-to-Segment discretization, contact elements and an intermediate surface.

A standalone discretization technique is needed for *Nitsche* method [Becker 03], [Wriggers 08], Gauss points of one surface play the role of slave nodes. The comparison of Nitsche and mortar techniques can be found in [Fritz 04].

The basic idea of the mortar method appeared in the second half of the 80s and in the beginning of 90s for domain decomposition techniques between non-conforming subdomains, see, for example, [Bernardi 90]. In 1998 Belgacem [Belgacem 98] has adapted the mortar method for the multibody or bilateral frictionless contact problem. Further in the beginning of 2000s the rigorous formulation adapted to frictional contact problem subjected to



large deformations and large slidings has been established by **McDevitt and Laursen** [McDevitt 00]. The mortar method consists either in introducing an intermediate contact surface where contact pressure is defined or in using as mortar surface one of the contacting surfaces, for details see [Wriggers 06]. The mortar based formulation leads to a consistent formulation of the frictional contact problem for large sliding and large deformations. It allows to pass the contact patch test for nonconforming meshes and does not suffer from spurious penetrations like the NTS (see Fig. 1.7 from [Zavarise 98]).

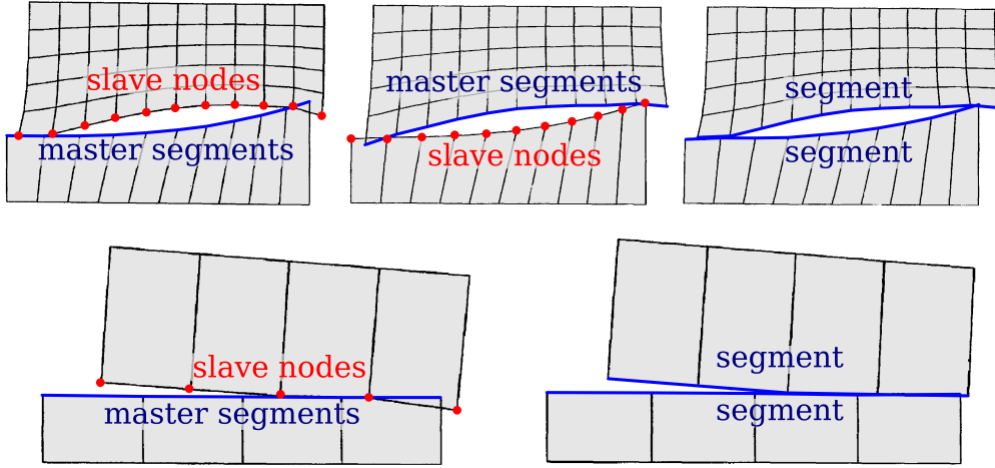


Figure 1.7: Example of spurious penetrations of NTS discretization and accurate treatment in the framework of Segment-to-Segment (adapted from Zavarise and Wriggers [Zavarise 98]).

### 1.2.3 Contact resolution

As mentioned above, the rigorous formulation of a variational principle for contact problems, results in a variational inequality subjected to geometrical constraints [Kikuchi 88]. These constraints can be brought as additional terms in the objective energy functional by means of penalty, Lagrange multiplier or other methods [Bertsekas 84], [Bertsekas 03], [Luenberger 03], [Bonnans 06], etc. Such an operation converts the constrained optimization, where constraints are given as inequalities, into an unconstrained or partly unconstrained one. If one supposes the active contact zone to be known, then the variational inequality can be replaced by a variational equality, which finally results in an unconstrained problem written in a standard form of variational equality [Wriggers 06]. This problem can be treated as a standard nonlinear minimization problem by means available in the given finite element code. A solver for systems of linear equations and a method for the treatment of nonlinear problems are needed. Note that since the contact constraints are given as inequalities, a special attention has to be paid to the definition of the active contact zone. For penalty and augmented Lagrangian methods, this

task is trivial. For the Lagrange multiplier method, an active set strategy should be employed (for the definition of the active set strategy see, for example [Luenberger 03]).

The resulting unconstrained minimization problem is not sufficiently smooth, which may result in slow convergence of the employed iteration scheme or even in divergence. The stability of the numerical scheme depends on the discretization and on the solution parameters. Note that the frictional contact renders the tangent matrix nonsymmetric, which presents a problem for several solvers (like conjugate gradient method) and for the parallel treatment of the problem: the Schwarz theory for nonsymmetric problems is less satisfactory than for positive definite symmetric problems [Toselli 05]. The way out has been proposed in [Laursen 92], [Laursen 93] for augmented Lagrangian method with Uzawa's algorithm – governing equations of Coulomb's friction have been linearized by the operator splitting technique, first recognized in [Glowinski 89], i.e. the entire problem is recast in two subproblems, which are solved once at each solution step. Augmented Lagrangian and Lagrange multiplier methods derive a non-symmetric tangent matrices only for the slip state. This is due to the nonassociativity of Coulomb's friction law, i.e. slip occurs in the plane of the constant contact pressure. The penalty method suffers from a non-symmetry both in stick and slip states. A solution has been proposed in [Wriggers 06], it consists in a similar treatment of all the deviations from the stick state, i.e. no difference between normal and tangential deviations from the stick are made. Another approach yielding a symmetric tangent matrix in stick state has been proposed in [Konyukhov 05], based on a rigorous covariant description of the contact geometry. The same authors proposed a symmetrization of different friction models based on the augmented Lagrangian method coupled with Uzawa's algorithm [Konyukhov 07b].

### 1.3 Motivation and overview

The principal motivation of this dissertation is implementation of a robust and fast sequential and parallel contact algorithm in the implicit Finite Element software **Z-set (ZéBuLoN)** [Besson 97]. Since there is no specific predefined application, the algorithm should be multipurpose. The principle requirement is the efficient treatment of large contact problems within sequential or parallel framework on parallel computers with distributed memory and within the nonoverlapping domain decomposition methods. Another aim predefined the size of the manuscript is to provide a reader with a consistent theoretical and methodological foundation of the computational contact mechanics.

The three principal parts of the thesis are: geometry of the contact, contact detection and resolution techniques. All parts are interdependent, but we tried to render each chapter more or less self-sufficient. So the sequence order of chapters is rather arbitrary. It is worth mentioning that in all chapters, except the chapter devoted to contact detection, a new algebra is used. It arises from a generalization of the tensor algebra and operates with abstract s-structures, elements of s-spaces. A comprehensive presentation of s-structures in a general



context is given in Appendix A. It follows the presentation of vectors and higher-order tensors which is given in a form conforming with s-structures.

Chapter 2 gives an elaborated presentation of the geometrically precise theory of contact. The relevant geometrical quantities and notions are introduced: closest points, gap function, tangential relative sliding. Several original ideas related to the geometrically precise theory of contact are presented.

In Chapter 3, different contact detection schemes for Node-to-Segment discretization are elaborated in minute detail. Sequential and parallel implementation are discussed both for known and unknown a priori master-slave discretizations. Some numerical examples are given.

The main governing equations are given in Chapter 4. The presentation starts from the primitive case of a unilateral contact with a rigid plane. Based on this simple case an interpretation of frictional and contact constraints is given by ordinary Dirichlet and Neumann boundary conditions. Next, Signorini's problem is presented and finally a bilateral (multibody) framework is given. The resulting variational inequality is recast in a variational equality using penalty, Lagrange multiplier and augmented Lagrangian methods. Related weak forms for frictionless and frictional contact for each method are presented and the resulting algorithm is illustrated by a simple example of unilateral contact.

Chapter 5 provides the reader with a minimal knowledge of the numerical schemes used in computational contact mechanics. Newton's method and its generalization for the case of nonsmooth functions and the return mapping algorithm are presented. Next, a short introduction in the standard formalism of the Finite Element Method is given, followed by the derivation of the expressions needed for implementing of the penalty and augmented Lagrangian method in a finite element code. Linearized forms adapted for the Newton's method are deduced. Finally, the details of implementation of the partial Dirichlet-Neumann approach considered in Chapter 4 are discussed, followed by several technical remarks on the implementation of the contact algorithms in a Finite Element environment.

Numerical examples are brought together in Chapter 6. Contact problems with known analytical solution and examples demonstrating the performance of the implemented methods for highly nonlinear problems are presented. Finally, in Chapter 7 the main contributions and the short term perspectives of the dissertation are summarized.



## Chapter 2

# Geometry in contact mechanics

### **Résumé de Chapitre 2 «Géométrie en mécanique de contact»**

*Le deuxième chapitre présente un travail fondamental et original portant sur la définition d'une théorie précise de la géométrie du contact. Après avoir démontré l'importance de la description géométrique et les ambiguïtés qui peuvent être liées à des situations pathologiques, on introduit quelques définitions de base telles que la distance de séparation, le point le plus proche et la vitesse tangentielle relative. On discute en détail les subtilités (asymétrie, non-unicité, existence) de la définition du point le plus proche et la différence entre le minimum et infimum dans cette définition. Puis on donne la forme qui définit de façon rigoureuse la distance de séparation normale ; cette forme est adaptée aux surfaces lisses par morceaux et est bien adaptée aux méthodes numériques de détection.*

*Cependant la séparation normale et la méthode de projection associée ont quelques inconvénients : en particulier, le point de projection n'est pas une fonction continue de la position du point esclave. La distance de glissement est associée directement à deux positions consécutives de la projection et l'énergie dissipée due au frottement est liée à la distance de glissement. Ainsi la forme faible du système devient-elle non continue, si bien que la convergence de la résolution n'est pas assurée. Ceci est un argument pour construire une nouvelle procédure de projection, dite «ombre portée». Le point de projection de l'esclave sur la surface maître est construit en fonction de l'ombre projetée du point esclave, la lumière étant émise par un point imaginaire, situé à une distance finie, ou à l'infini. Cette procédure permet de retrouver une projection continue et, en conséquence une forme faible mieux adaptée à la résolution numérique.*

*Puis on dérive les première et deuxième variations des variables cinétiques (la séparation normale et le paramètre local) pour la géométrie continue. Tous les calculs ont été faits en exploitant un nouveau formalisme, dit «algèbre de S-structure», qui est une généralisation des opérations sur les tenseurs à des opérations sur des champs des tenseurs. On obtient toutes les expressions nécessaires pour intégrer une géométrie arbitraire et non-linéaire dans un code de calcul par élément finis pour la projection normale et la projection d'ombre portée. Les expressions complètes sont comparées d'une part avec les «vraies» variations obtenues numériquement, et avec les expressions analytiques simplifiées souvent utilisées dans le calcul numérique. L'étude statistique démontre l'avantage et la convergence des formes rigoureuses.*

*En fin de ce chapitre, on propose une nouvelle méthode d'enrichissement de la géométrie de contact, inspirée par la méthode X-FEM. Le but est de prendre en compte une géométrie complexe de la surface sans la discrétiser. Cette méthode est utile pour la simulation du frottement anisotrope et dans le cas où la géométrie change localement en raison d'un changement d'état de déformation et de contrainte (usure, mécanismes d'intrusion–extrusion, etc. . . ).*

## 2.1 Introduction

Contact phenomenon takes place at the interface between solids. This fact implies a strong connection of the contact problem with a rigorous description of the geometry of contacting surfaces. The first continuum based description of the contact problem was given by Simo and Laursen [Laursen 93] and Laursen [Laursen 94]. Such a geometrical description still presents an interesting topic for research in computational contact mechanics, see e.g. recent articles by Konyukhov [Konyukhov 06b], [Konyukhov 06a], [Konyukhov 09]. The mathematical formulation of frictionless contact conditions leads to equations connecting the *normal contact pressure*  $\sigma_n$  with the *mutual penetration* of bodies, expressed by a signed *gap function*, frequently the *normal gap function*,  $g_n$  is used. The formulation of frictional contact leads to the connection between *shear or tangential contact stress vector*  $\underline{\sigma}_t$  and the *relative tangential sliding velocity*  $\underline{\dot{g}}_t$ . The contact stress has to be integrated over the *contact surface*  $\Gamma_c^i$  of each solid, where  $i \in [0; N_c]$  and  $N_c$  the total number of contacting surfaces.

Let us show how important the geometry is for contact mechanics and how the geometrical description may predetermine methods and approaches which are used in numerical treatment of contact. As is known, there is an ambiguity in the definition of the normal gap  $g_n$  between contacting surfaces. At first glance, it seems easy to determine the normal gap for each point of a contacting surface as a distance to the closest point of the second surface: for a point  $M$  of the first surface  $M \in \partial A$  one seeks for the closest point  $N$  on the other surface  $N \in \partial B$ . Three problems arise from such a definition:

### P1 Asymmetry of surfaces (Fig. 2.1,a.)

If instead of seeking for the closest point  $N \in \partial B$  to the point  $M \in \partial A$ , one inverts the problem and searches for the point  $M' \in \partial A$  closest to

the  $N \in \partial B$ , then the points  $M$  and  $M'$  do not coincide as soon as the contacting surfaces are not parallel at least locally. It means that there is no one-to-one equitable correspondence between surface points. It implies an asymmetry in the gap function and consequently in the entire geometrical description.

**P2** *Non-uniqueness of the closest point (Fig. 2.1,b.)*

For example, the center of a circle does not have a single closest point on the circle, but all the points on the circle are equally close to its center. All other points have a unique closest point on the circle. The uniqueness of the closest point refers to the curvature of the considered curve or surface and has been discussed in detail by Heergaard and Curnier [Heergaard 96], Pietrzak [Pietrzak 97] and Konyukhov [Konyukhov 08]. The limit case of the infinite curvature corresponds to the third problem.

**P3** *Requirement of smoothness (Fig. 2.1,c.)*

Smoothness of at least one of the contacting surfaces (master) is not sufficient but necessary condition for existence of the normal projection<sup>1</sup> point. The smooth surface allows a rigorous mathematical description of contact, a robust detection procedure and a reliable convergence of numerical schemes. However, surfaces in the Finite Element framework are only piecewise smooth due to the discretization. The non-smoothness represents another source for existence of multiple closest points and generates blind angles in normal projection domains, discontinuous normal vector field and related problems - oscillations and possible divergence of the numerical solution.

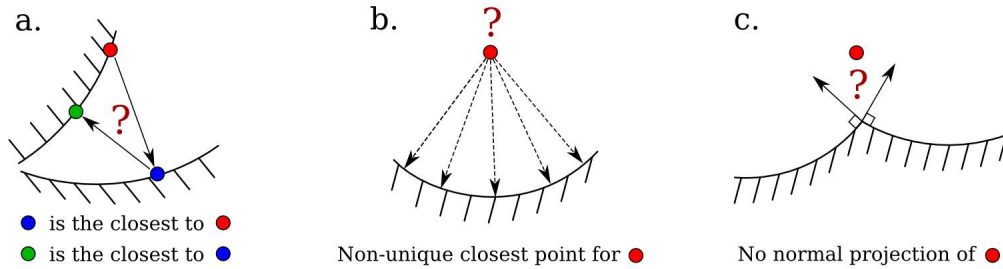


Figure 2.1: Geometry related problems: **a** – asymmetry of the closest point definition; **b** – non-uniqueness of the closest point; **c** – nonexistence of the normal projection point.

All these difficulties affect the geometrical description of the contact. The asymmetry of the closest point detection (**P1**) results in an asymmetric

<sup>1</sup>by normal projection of a slave we imply such a point on the master surface that the vector connecting the slave and its projection is collinear with the normal vector constructed on the master surface

treatment of contact surfaces, it leads to the so-called “master-slave” approach (also called “target-impactor” or “target-contactor”). For each point of the slave surface  $\underline{r}_i \in \Gamma_s$  the closest point on the master surface  $\underline{\rho}_i \in \Gamma_m$  has to be determined, i.e.

$$\underline{\rho}_i = \left\{ \underline{\rho}_i \in \Gamma_m \mid \forall \underline{\rho} \in \Gamma_m : \|\underline{\rho}_i - \underline{r}_i\| \leq \|\underline{\rho} - \underline{r}_i\| \right\}.$$

Due to the non-uniqueness of the closest point (P2), a different technique fulfilling additional conditions on the uniqueness and continuity of the projection, can be elaborated (see section 2.2.3). That allows to improve the convergence and to avoid nonphysical discontinuities in sliding path, which is crucial for a rigorous description of frictional contact.

The non-smoothness of the surface (P3) arises from the discontinuity of local bases across boundaries of adjacent segments or faces in the finite element discretization and produces convergence problems and oscillations in the finite element framework. The main remedy consists in smoothing the master surface over several segments [Pietrzak 97], [Wriggers 01], [Krstulović-Opara 02].

As one can see, the definition of the gap function and the closest point brings out a series of difficulties. This short preface allows to realize the importance of a well founded geometrical approach needed to deal with contact problems. Especially, it worth mentioning that classical contact detection techniques are strongly connected with the closest point definition. All results concerning the definition of the closest point will be used in Chapter 3 devoted to the development of a reliable contact detection procedure.

The aim of this chapter is an elaborated analysis of geometry related questions in computational contact mechanics in the framework of the FEM and the Node-to-Segment discretization. According to the high standards which were set up in computational contact mechanics by Simo and Laursen [Laursen 93] and in order to provide a multipurpose and discretization independent framework, we start from the continuous description of the contact geometry. Such an approach is valid both for classical Node-to-Segment discretization for any type of finite elements and for special cases, namely unilateral contact with a rigid surface and smoothed master surfaces:

- *Unilateral contact with a rigid surface:* contact between a deformable solid and a rigid surface, the latter can be described by an analytical function or a CAD model, see [Hansson 90] for frictionless and [Heege 96] for frictional cases; among engineering problems subjected to this case there are metal forming and metal processing, rubber-metal and tire-road contact, etc.
- *Smoothed master surface:* for many reasons it is advantageous to replace a piecewise smooth master surface by a  $C^1$  smooth surface (NURBS, Bézier, Gregory patches, etc) based on information from several adjacent master segments; this procedure ensures a continuous projection on the master surface and leads to a better convergence [Pietrzak 97], [Padmanabhan 01], [Puso 02], [Wriggers 01], [Krstulović-Opara 02] and

[Wriggers 06], [Chamoret 04], etc; all smoothing procedures enrich the geometry and require an adapted geometrical description, which can be easily derived from the continuous description of the contact geometry.

The geometrical description is confined to the interaction between a point (slave) and a surface (master). Such a description is adapted to the classic NTS contact discretization, which as known introduces some important drawbacks. Nevertheless it is widely used, rather robust and simple. We make a start from a short description of a point-surface geometry. Furthermore, a rigorous closest point definition will be given and two key geometrical quantities will be introduced, the normal gap  $g_n$  and the tangential sliding velocity  $\underline{\dot{g}}_t$ . Next, a new projection technique (“shadow” projection) will be proposed, providing a continuous and unique projection of a slave point onto  $C^0$ ,  $C^1$  and  $C^2$  continuous surfaces.

The next part will be devoted to the incorporation of contact terms into the weak form arising from the application of the principle of virtual works in the Finite Element Method. The weak form consists of the total virtual work of the internal and external forces over volumes and surfaces of contacting solids. Variations of geometrical quantities – normal gap  $g_n$  and tangential relative slip  $\underline{\dot{g}}_t$ :

$$\delta g_n, \quad \delta \underline{\dot{g}}_t$$

are needed to derive the contribution of the contact interface to the balance of virtual works. The resulting equation is nonlinear, and its linearized form is required to provide a solution within the implicit integration scheme in the Finite Element Method. For that purpose, the second variations of these geometrical quantities are needed:

$$\Delta \delta g_n, \quad \Delta \delta \underline{\dot{g}}_t$$

Initially, the geometrical variations were obtained for particular finite element discretizations in [Parisich 89]; they can also be found in a recent book [Wriggers 06]. Laursen and Simo [Laursen 93] established a new standard in computational contact mechanics by deriving all equations in continuous form, which can then be easily applied to any discretization of the contact interface. Different forms of these expressions can be found, e.g. in [Pietrzak 97], [Pietrzak 99]. Simplified expressions using the assumption of a zero normal gap can be found in [Konyukhov 05]. Here, all required variations will be obtained without any assumptions concerning the value of the normal gap. All computations have been carried out independently from other authors using a new mathematical formalism (see Appendix A). Resulting expressions have been validated numerically (Sec. 2.3) and a closed “ready-to-implement” form, adapted for the Finite Element Method, is given in Sec. 2.4. In our computation, we do not use any assumption concerning the size of the normal gap. This differs from the approach, where the normal gap is assumed to be negligibly small  $g_n = 0$ , which makes the computation of necessary variations quite simple, see [Konyukhov 05], [Wriggers 06], but leads to the lost of the

## 2.1 Introduction

---

quadratic convergence of Newton's method. In Section 2.3 both formulations are compared.

All needed expressions will be obtained for two different projection types: the classical normal projection and an original technique – the *shadow projection method*, which smooths over some drawbacks of the normal projection (non-uniqueness or nonexistence of the projection). Another original contribution of this chapter is the introduction of enriched master geometry, see Section 2.5.



## 2.2 Interaction between contacting surfaces

### Notation

- Time:  $t$
- Radius vector of the slave point:  $\underline{r}_s = \underline{r}_s(t) \in \Gamma_s$
- Master surface parametrization by v-scalar:

$$\underline{\xi} = \underline{\xi}(t) = \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \end{bmatrix}$$

- Surface coordinates of the slave point's projection onto the master surface:

$$\underline{\xi}_\pi = \underline{\xi}_\pi(t)$$

- Radius vector of the slave point projection onto the master surface:

$$\underline{\rho} = \underline{\rho}(t, \underline{\xi}_\pi) \in \Gamma_m$$

- V-vector of covariant tangential vectors on the surface (covariant surface basis at the projection point):

$$\nabla_{\underline{\xi}} \underline{\rho}(t, \underline{\xi}_\pi) = \left. \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right|_{\underline{\xi}_\pi} \sim \begin{bmatrix} \frac{\partial \underline{\rho}}{\partial \xi_1} \\ \frac{\partial \underline{\rho}}{\partial \xi_2} \end{bmatrix}$$

- V-vector of contravariant tangential basis vectors on the surface (contravariant surface basis at the projection point):

$$\overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} = \bar{\underline{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}}$$

- First covariant fundamental surface metric matrix (t-scalar):

$$\underline{\underline{A}} = \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T$$

- First contravariant fundamental surface metric matrix (t-scalar):

$$\bar{\underline{\underline{A}}} = \underline{\underline{A}}^{-1} = \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} \cdot \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}}^T$$

## 2.2 Interaction between contacting surfaces

- Second covariant fundamental surface matrix (t-scalar):

$$\underline{\underline{H}} = \underline{\underline{n}} \cdot \frac{\partial^2 \underline{\underline{p}}}{\partial \underline{\underline{\xi}}^2}$$

- Unit vector normal to the master surface at the projection point:

$$\underline{\underline{n}} = \underline{\underline{n}}(t, \underline{\underline{\xi}}_\pi) = \frac{\underline{\underline{\rho}}_1 \times \underline{\underline{\rho}}_2}{\|\underline{\underline{\rho}}_1 \times \underline{\underline{\rho}}_2\|}$$

- Normal gap:

$$g_n = g_n(t) = (\underline{\underline{r}}_s - \underline{\underline{\rho}}) \cdot \underline{\underline{n}} \quad (2.1)$$

- Tangential sliding velocity:  $\underline{\underline{\dot{g}}}_t = \frac{\partial \underline{\underline{p}}}{\partial \underline{\underline{\xi}}} \underline{\underline{\xi}}$



### Remark 2.1 on s-structures (v- and t- notations).

At first sight, introducing s-structures seems cumbersome. However, the description of spatial interaction between surfaces requires both 3D quantities associated with the space and 2D quantities associated with the surface; consequently, the use of indices and Einstein summation is ambiguous, because for some quantities index changes from 1 to 3, for other - from 1 to 2. In differential geometry this ambiguity is avoided by the use of Greek and Roman alphabets to distinguish summation limits. However, I wished to get rid of indices and to express everything in shorter and more transparent notations of the direct tensor language. Moreover, a consistent formulation of s-structures over the space of vectors and tensors leads to an improvement and simplification of the tensor apparatus in many cases as shown in Appendix A. The new formalism of s-structures (set-structures) is introduced and explained in details in Appendix A.6. In this chapter a reduced form of s-structures is employed (see Section A.7).

### 2.2.1 Normal gap

#### Problem definition

As mentioned in the introduction, the definition of the gap<sup>2</sup> between close surfaces and the closest point determination are not trivial tasks. These two quantities are of a big importance for computational contact mechanics. The closest point projection is used during a local contact detection procedure and its change describes the tangential sliding between surfaces. The normal gap strongly connected with the closest point plays a crucial role in description of

<sup>2</sup>By “gap” between contacting surface we will understand any scalar function of the slave point and the master surface, which is positive if there is no contact, zero if there is a contact and negative if a penetration takes place. By “normal gap” we will understand the shortest distance between the slave node and the master surface with the same sign rules.

the normal contact.

Formally, the absolute value of the normal gap  $|g_n|$  is defined as the closest distance from the slave node to the master surface. It can be interpreted as the minimal radius  $R \in [0; \infty]$  of a sphere (circle in 2D) with its center placed in the slave point  $\underline{r}_s$  and touching (but not intersecting) the master surface. Remark that it does not matter if this sphere touches the master surface in one, several or infinitely many points. The gap is positive if the slave point is outside of the solid enveloped by the master surface, else it is negative. The sign of the normal gap means that

- $g_n > 0$  – solids are locally separated;
- $g_n = 0$  – solids are locally in contact;
- $g_n < 0$  – solids locally penetrate each other.

The normal gap can be also defined by using the closest point concept: the absolute value of the normal gap is then the distance between the slave node and the closest point on the master surface. It is considered positive  $g_n > 0$  if the dot product of the vector connecting the slave node and the closest point  $\underline{\rho}$  with a vector normal to the surface is positive  $(\underline{r}_s - \underline{\rho}) \cdot \underline{n}(\underline{\rho}) > 0$  and negative otherwise. The normal at the master surface in every point is oriented outwards the bulk. Remark that for shell structures each surface has two opposite normals. In the latter case to define the sign of the gap one needs to know the history of the master-slave interaction. Formally, the point  $\underline{\rho}$  on the master surface is called the closest point to the slave point  $\underline{r}_s$  if all other points of the master surface are not closer than  $\|\underline{\rho} - \underline{r}_s\|$ . According to this definition for any closed master surface, the closest point always exists, but it is not always unique. Several closest points (or even an infinite number) may satisfy this definition.

An attempt to summarize different approaches and to overcome the difficulties related to the definition of the normal gap and the closest point has been undertaken by Konyukhov et al. [Konyukhov 08], where authors introduce a *continuous projection domain* and the *generalized closest point procedure*. However, the proposed approach does not cover all cases and requires proximity of the slave node to the master surface. Here, we develop a new approach to the normal gap and the closest point definition valid for piecewise  $C^1$ -smooth master surface and for any position of the slave point. This new approach is based on the rigorous definition of the normal gap. Some techniques to overcome non-uniqueness of the closest point will be also proposed.

**Definition**

**Definition. Normal gap.**

An absolute value of the **normal gap**  $|g_n|$  between the slave point  $\underline{r}_s$  and the master surface  $\Gamma_m$  is related to the infimum of the distance functional  $F(\underline{\rho})$  on the closed set of the master surface  $\underline{\rho} \in \Gamma_m$  in the following way:

$$F(\underline{\rho}) = \frac{1}{2} (\underline{r}_s - \underline{\rho})^2, |g_n| = \sqrt{2 \inf_{\underline{\rho}' \in \Gamma_m} \{F(\underline{\rho}')\}} \quad (2.2)$$

The **normal gap** is positive if the slave point is outside of the solid enveloped by the master surface, otherwise the normal gap is negative.



**Remark 2.2 on the difference between infimum and minimum .**

Frequently, for the definition of the normal gap a “minimum” of the distance function is used instead of an “infimum”, moreover the distance functional is assumed to be smooth enough. This imprecision often leads to a wrong or even non-existing normal gap (see Fig. below). Rigorously speaking, a minimum of a function can be studied only on an open set  $\underline{\rho} \in \Gamma_m \setminus \partial\Gamma_m$  and it is defined for any continuous function (see below). A function can have infinitely many minima or does not have them at all, infimum always exists and is always unique. However, on an open set, minimum point  $x^*$  always belongs to the set, infimum  $M$  does not always belong to the image of the set  $M \notin \text{Im}(F)$ , it does when a closed set is considered. For scalar functions the following equations define minima and infimum:

- Minimum of function  $F(x) \in C^0_{(x_0; x_1)}$

$$F(x^*) = \min_{x \in (x_0; x_1)} [F(x)] \Leftrightarrow \exists \varepsilon, \forall x \in (x_0; x_1) \text{ and } 0 < \|x - x^*\| < \varepsilon : F(x) \geq F(x^*)$$

- Minimum of function  $F(x) \in C^1_{(x_0; x_1)}$

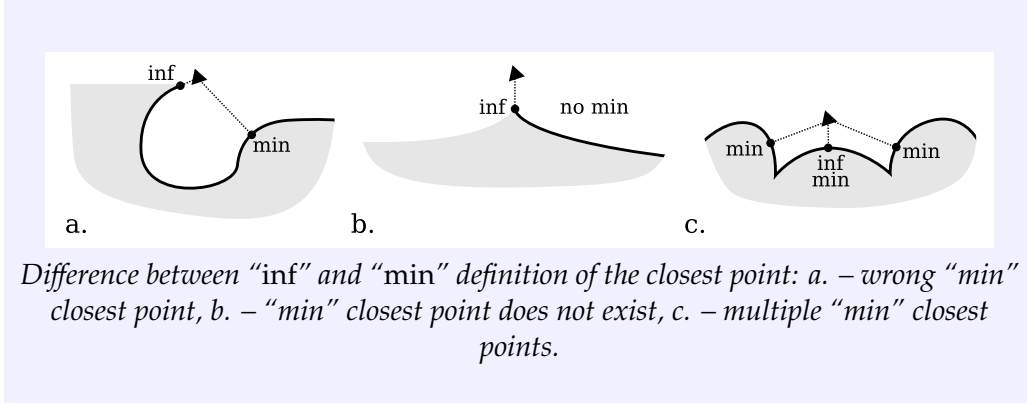
$$F(x^*) = \min_{x \in (x_0; x_1)} [F(x)] \Leftrightarrow \left. \frac{\partial F}{\partial x} \right|_{x^*} = 0 \text{ and } \left. \frac{\partial^2 F}{\partial x^2} \right|_{x^* -} > 0 \text{ and } \left. \frac{\partial^2 F}{\partial x^2} \right|_{x^* +} > 0$$

- Minimum of function  $F(x) \in C^2_{(x_0; x_1)}$

$$F(x^*) = \min_{x \in (x_0; x_1)} [F(x)] \Leftrightarrow \left. \frac{\partial F}{\partial x} \right|_{x^*} = 0 \text{ and } \left. \frac{\partial^2 F}{\partial x^2} \right|_{x^*} > 0$$

- Infimum of function  $F(x) \in C^0_{(x_0; x_1)}$

$$M = \inf_{x \in (x_0; x_1)} [F(x)] \Leftrightarrow \begin{cases} \forall x \in (x_0; x_1) : F(x) \geq M \\ \forall M' > M, \exists x \in (x_0; x_1) : F(x) \leq M' \end{cases}$$



According to this definition, the normal gap always exists. It follows from the remark that the infimum and the global minimum  $\widehat{\min}$  are equivalent for a closed set  $\Gamma_m$  except edges (see Fig. in Remark 2.3 and Fig. 2.2). Moreover, according to the definition a smoothness of the distance functional is not required.

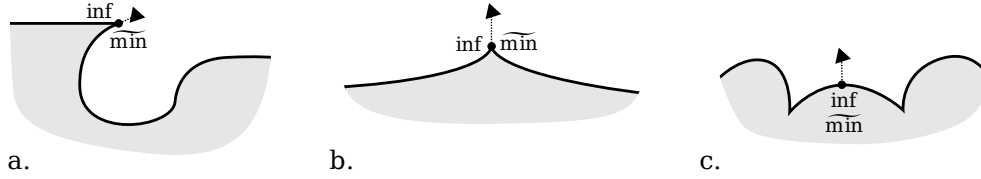


Figure 2.2: Equivalence between  $\inf$  and  $\inf [\min_i]$  closest point definition for  $C^0$  function far from edges.

From a numerical point of view, it is not convenient to solve directly Eq. (2.2) for a  $C^0$  surface. It is better to split it into several smooth open sets and points of nondifferentiability. So let us consider the master  $\Gamma_m$  as the union of parametrized surfaces  $\Gamma_s^i$ , parametrized edges  $\Gamma_e^j$  and vertices  $\Gamma_v^k$  (Fig. 2.3)

$$\Gamma_m = \left\{ \bigcup_{i=1, N_s} \Gamma_s^i \right\} \cup \left\{ \bigcup_{j=1, N_e} \Gamma_e^j \right\} \cup \left\{ \bigcup_{k=1, N_v} \Gamma_v^k \right\}.$$

The parametrization of surfaces (segments) is defined by a scalar  $\xi \in I_\xi$  in 2D or by v-scalar  $\tilde{\xi} \in I_\xi$  in 3D, so  $\underline{\rho} = \underline{\rho}(\xi)$ ,  $\tilde{\xi} \in I_\xi$  where  $I_\xi$  is an open set. The parametrization of edges in 3D is defined by scalar parameter  $\zeta \in I_\zeta$ , where  $I_\zeta$  is also an open set. Finally, in 3D, a master surface consists of:

- $N_s$  connected parametrized open and  $C^2$ -smooth surfaces (segments)

$$\underline{\rho}^i(\xi) \in \Gamma_s^i, i \in [1, N_s], \xi \in I_\xi, \frac{\partial^3 \underline{\rho}}{\partial \xi^3} < \infty$$

## 2.2 Interaction between contacting surfaces

- $N_e$   $C^2$ -smooth open parametrized edges

$$\underline{\rho}^j(\zeta) \in \Gamma_e^j, j \in [1; N_e], \zeta \in I_\zeta, \frac{\partial^3 \underline{\rho}}{\partial \zeta^3} < \infty$$

- $N_v$  vertices

$$\underline{\rho}_k \in \Gamma_v^k, k \in [1; N_v].$$

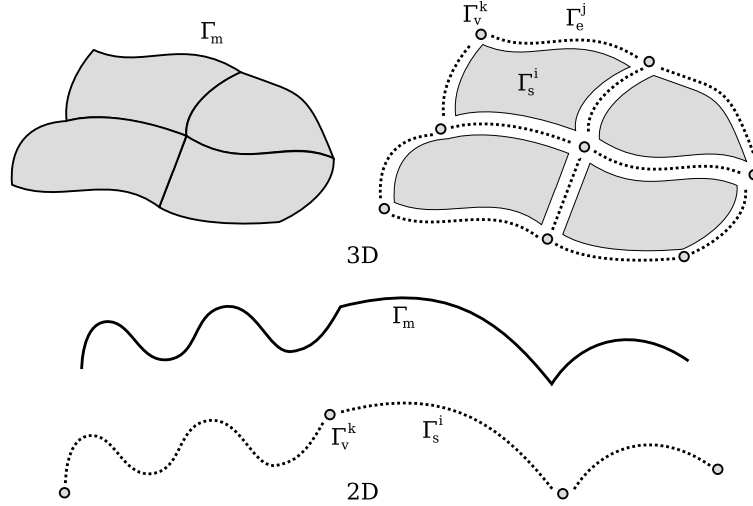


Figure 2.3: Split of the 3D master surface into sets of surfaces, edges, vertices and 2D master curves into sets of segments and vertices.

The absolute value of the normal gap can be expressed in a more appropriate form from a numerical point of view:

$$|g_n| = \sqrt{2 \inf_{\underline{\rho} \in \Gamma_m} \{F(\underline{\rho})\}}$$

$$\inf_{\underline{\rho} \in \Gamma_m} \{F(\underline{\rho})\} = \begin{cases} \text{3D : } \inf \left\{ \begin{aligned} &\inf_{\underline{\rho}^i \in \Gamma_s^i, i \in [1; N_s]} \left\{ \text{ext } F(\underline{\rho}^i(\xi)) \right\} && \text{term 3D.1} \\ &\inf_{\underline{\rho}^j \in \Gamma_e^j, j \in [1; N_e]} \left\{ \text{ext } F(\underline{\rho}^j(\zeta)) \right\} && \text{term 3D.2} \\ &\inf_{\underline{\rho}^k \in \Gamma_v^k, k \in [1; N_v]} \{F(\underline{\rho}^k)\} && \text{term 3D.3} \end{aligned} \right\} \\ \text{2D : } \inf \left\{ \begin{aligned} &\inf_{\underline{\rho}^i \in \Gamma_s^i, i \in [1; N_s]} \left\{ \text{ext } F(\underline{\rho}^i(\xi)) \right\} && \text{term 2D.1} \\ &\inf_{\underline{\rho}^k \in \Gamma_v^k, k \in [1; N_v]} \{F(\underline{\rho}^k)\} && \text{term 2D.2} \end{aligned} \right\} \end{cases} \quad (2.3)$$

Remark that there is neither local nor global minimum in the definition, they have been replaced by extremum (ext). It does not change the result, but may facilitate the numerical treatment, because the investigation of the sign of second or higher order derivatives of the functional  $F$  is no longer needed. A sufficient and necessary condition for extremum of a smooth function is the equality of the first derivative to zero:

$$F(x) \in C^2(x), F(x^*) = \text{ext } F(x) \Leftrightarrow \left. \frac{\partial F}{\partial x} \right|_{x^*} = 0.$$

According to Eq. (2.3) the infimum can correspond to one, several or infinitely many points on the master surface  $\Gamma_m$ . One of these points can be defined as the closest point.

### 2.2.2 The closest point

**The closest point on 3D surface and 2D curve** Let us analyze Eq. (2.3) term by term; the extremum of the distance function on the  $i$ -th smooth and open surface  $\Gamma_s^i$  (Fig. 2.4, Eq. 2.3 term 3D.1)

$$\underline{\rho}' \in \Gamma_m^i, F(\underline{\xi}) \in C^2(I_\xi), \text{ext}_{\underline{\xi} \in I_\xi} [F(\underline{\rho}'(\underline{\xi}))] \Leftrightarrow$$

gives the following equation for the closest projection point:

$$\Leftrightarrow \frac{\partial F}{\partial \underline{\xi}} = 0 \Leftrightarrow (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = 0 \quad (2.4)$$

Since  $F(\underline{\rho})$  is  $C^\infty$  smooth by  $\underline{\rho}$ , then the requirement of smoothness of  $F(\underline{\xi})$  by  $\underline{\xi}$  is equivalent to the smoothness of  $\underline{\rho}(\underline{\xi})$  by  $\underline{\xi}$ . The last equality in Eq. (2.4) means that the vectors of the local basis in an extremum point  $\underline{\rho}(\underline{\xi}^j)$  are orthogonal to the vector connecting this point with the slave point. The unit normal to the master surface in the extremum point can be evaluated as a cross product of the local basis vectors

$$\underline{n} = \frac{\frac{\partial \underline{\rho}}{\partial \xi_1} \times \frac{\partial \underline{\rho}}{\partial \xi_2}}{\left\| \frac{\partial \underline{\rho}}{\partial \xi_1} \times \frac{\partial \underline{\rho}}{\partial \xi_2} \right\|} \quad (2.5)$$

All vectors which are orthogonal to both vectors of the local basis are collinear with the normal vector, consequently from (2.4) and (2.5)  $\underline{r}_s - \underline{\rho}(\underline{\xi}^j) = \alpha \underline{n}(\underline{\xi}^j)$ , where  $\alpha$  is nothing but the distance between the slave point and its projection taken with an appropriate sign

$$\underline{r}_s - \underline{\rho}(\underline{\xi}^j) = \pm \left\| \underline{r}_s - \underline{\rho}(\underline{\xi}^j) \right\| \underline{n}(\underline{\xi}^j). \quad (2.6)$$

If  $\underline{\xi}^j$  is a point of the global minimum on the current surface  $\Gamma_s^j$ , then the norm appearing in (2.6) can be replaced by the local normal gap  $g_n^j$

$$\underline{r}_s - \underline{\rho}(\underline{\xi}^j) = g_n^j \underline{n}(\underline{\xi}^j). \quad (2.7)$$

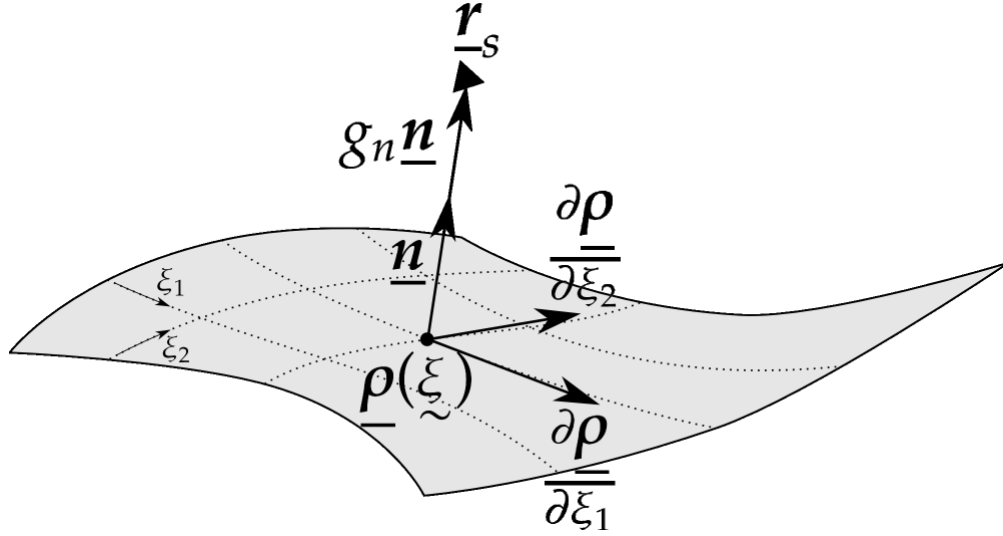


Figure 2.4: Geometry of a smooth master surface with a close slave point.

However, Eq. (2.4) is not simple, different cases are possible. Let us analyze them in detail. To employ Newton's method to solve it, one has to linearize this equation, at point  $\underline{\xi}$  it gives

$$\left[ (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right]_{\underline{\xi} + \delta \underline{\xi}} = \left[ (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right]_{\underline{\xi}} + \left[ -\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} + (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \right]_{\underline{\xi}} \cdot \delta \underline{\xi} = 0$$

The increment of the surface parameter is then:

$$\delta \underline{\xi} = \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} - (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \right]^{-1} \cdot \left[ (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right]. \quad (2.8)$$

Remark that we cannot substitute  $(\underline{r}_s - \underline{\rho})$  in (2.8) according to (2.7), since the latter is correct only in the extremum point. It can be shown that the inversed t-scalar in (2.8) is always defined for all  $\underline{\xi}$  which are not solutions of Eq. (2.4). So if one makes a start from a non extremum point, at least the first step of the Newton's method (Eq. (2.8)) is ensured. If the starting point is an extremum point, then the solution has been achieved; Newton's method cannot provide us with another extremum. To find all solutions of Equation (2.4) by a simple Newton's method, the master surface  $\Gamma_m^i$  has to be divided into regions containing only one extremum or a more advanced method has to be used. In case of a 2D geometry, the term 2D.1 from Eq. (2.3) gives the following equation for the increment of  $\xi$ :

$$\delta \xi = \frac{(\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}}{\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} - (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2}}. \quad (2.9)$$



It can be also shown that the denominator in (2.9) is not zero for non extremum points.

Different simple and non-trivial cases of extremum location are presented in Figure 2.5 for 2D and in Figure 2.6 for 3D geometries. Within the chosen numerical scheme it is complicated to analyze cases when the functional  $F$  has several or infinitely many extrema (Fig. 2.5: b,c,d and Fig. 2.6: b,c,d).

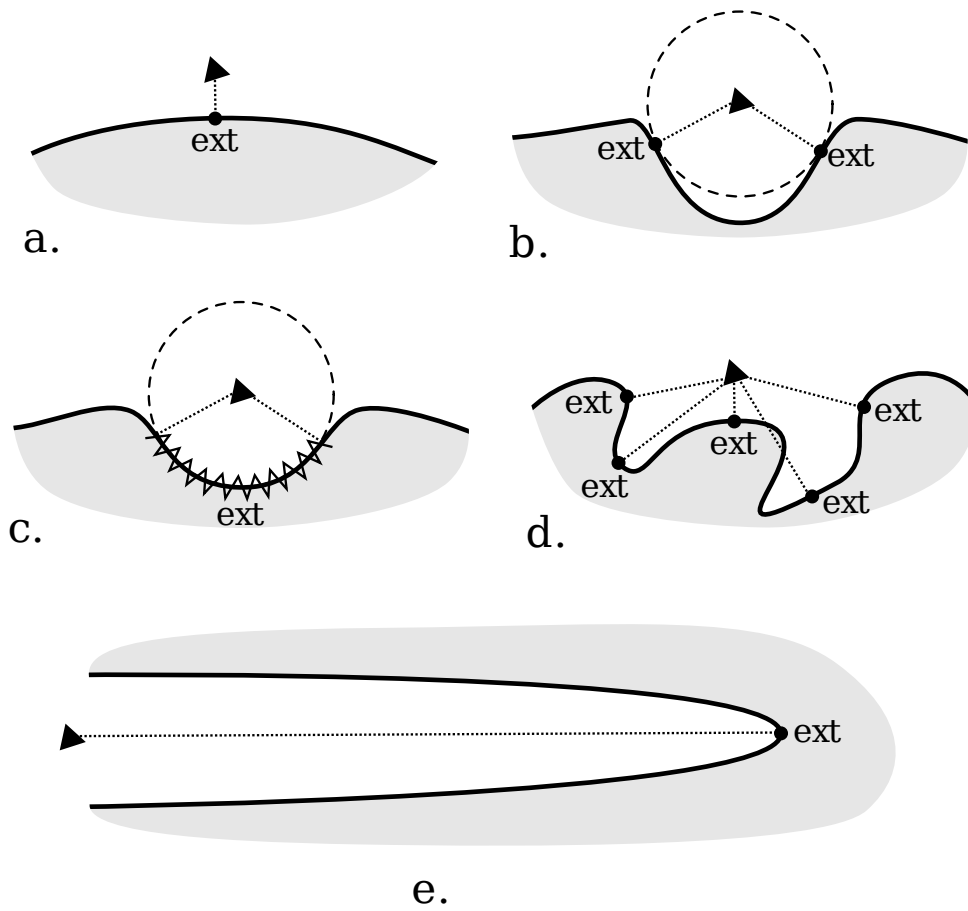


Figure 2.5: Different cases of extremum location in 2D: a. – unique extremum (global minimum), b. – several extrema located at the same distance from the slave point, c. – case of infinite number of extrema, slave point is situated in the local center of curvature of the master surface, d. – case of several different extrema, e. – unique extremum (global maximum).

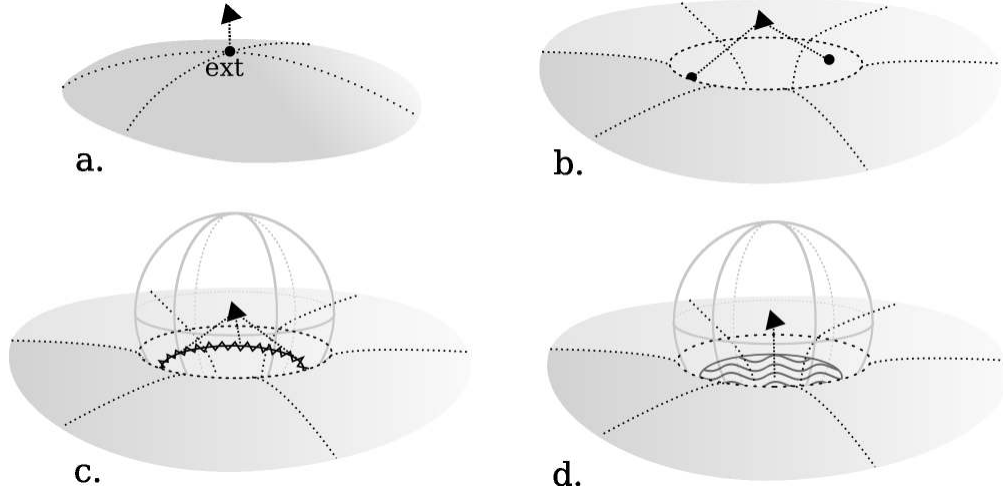


Figure 2.6: Different cases of extremum location in 3D: a. – unique extremum (global minimum), b. – several extrema located at the same distance from the slave point, c. – case of infinite number of extrema (curve), c. – case of infinite number of extrema (surface).

**The closest point on a 3D curve** An extremum point on the  $j$ -th smooth curve  $\Gamma_e^j$  (term 3D.2) can be determined (Fig. 2.7) just like as for the surface:

$$\begin{aligned} \underline{\rho}' \in \Gamma_e^j, F(\zeta) \in C^2(I_\zeta), \text{ext}_{\zeta \in I_\zeta} [F(\underline{\rho}'(\zeta))] &\Leftrightarrow \\ \Leftrightarrow \frac{dF}{d\zeta} = 0 \Leftrightarrow (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \zeta} = 0 \Leftrightarrow (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial s(\zeta)} = 0, \end{aligned} \quad (2.10)$$

where  $s$  is a classical parametrization of the curve and it denotes the curve's length

$$ds = \sqrt{\frac{\partial \underline{\rho}}{\partial \zeta} \cdot \frac{\partial \underline{\rho}}{\partial \zeta}} d\zeta.$$

Three normalized basis vectors can be associated with the curve:

- a unit tangential vector

$$\underline{\tau} = \frac{\partial \underline{\rho}}{\partial s};$$

- a unit normal vector pointing to a center of curvature of the curve

$$\underline{\nu} = \frac{1}{\kappa} \frac{\partial^2 \underline{\rho}}{\partial s^2};$$

- a unit binormal vector defined as the cross product of the first two

$$\underline{\beta} = \underline{\tau} \times \underline{\nu};$$

where  $\kappa$  is the curve's curvature defined as

$$\kappa(\zeta) = \frac{\left\| \frac{\partial \underline{\rho}}{\partial \zeta} \times \frac{\partial^2 \underline{\rho}}{\partial \zeta^2} \right\|}{\left\| \frac{\partial \underline{\rho}}{\partial \zeta} \right\|^3}.$$

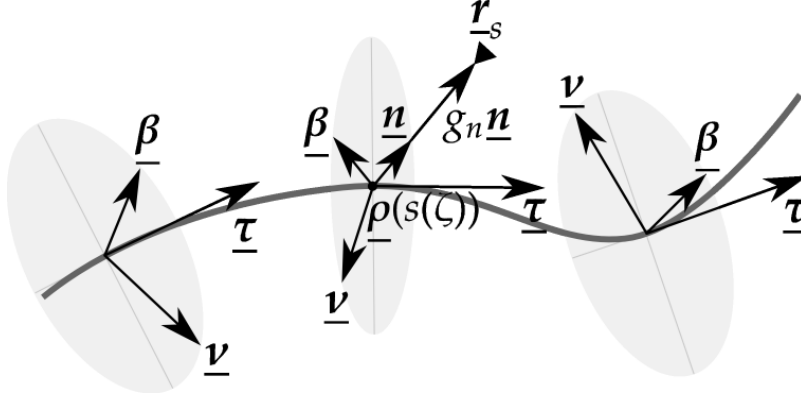


Figure 2.7: Point-curve interaction in three dimensional space.

The numerical scheme to find an extremum point is derived just like for a surface or a curve in 2D. The linearization of Eq. (2.10) gives:

$$\left[ (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \zeta} \right]_{\zeta^0 + \delta \zeta} = \left[ (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial \underline{\rho}}{\partial \zeta} \right]_{\zeta^0} + \left[ -\frac{\partial \underline{\rho}}{\partial \zeta} \cdot \frac{\partial \underline{\rho}}{\partial \zeta} + (\underline{r}_s - \underline{\rho}) \cdot \frac{\partial^2 \underline{\rho}}{\partial \zeta^2} \right]_{\zeta^0} \cdot \delta \zeta = 0$$

So the increment of the parameter  $\delta \zeta$  is always defined if the last term in square brackets is not zero, i.e. if

$$(\underline{r}_s - \underline{\rho}(\zeta)) \cdot \underline{t} \neq 0,$$

where  $\underline{t} = \frac{\partial \underline{\rho}}{\partial \zeta}$ , then  $\zeta$  is not solution of (2.10), then a finite increment for the curve parameter is given by:

$$\delta \zeta = \frac{(\underline{r}_s - \underline{\rho}) \cdot \underline{t}}{\underline{t} \cdot \underline{t} - \kappa \underline{v} \cdot (\underline{r}_s - \underline{\rho})} \quad (2.11)$$

The last terms (3D.3) and (2D.2) in (2.3) are trivial: the distance from the slave point has to be measured to each vertex of the master surface, no iterations are needed. Finally, Eq. (2.8), (2.11) in 3D and (2.9) in 2D allow to determine an extremum on surfaces  $\Gamma_m^i$  and edges  $\Gamma_e^j$  in 3D and on curves  $\Gamma_e^j$  in 2D respectively. Next, the infimum over all extrema has to be determined. However, as already mentioned sometimes the closest point is not unique, or even the number of these points is uncountable. A remedy for that is a selection procedure between several or an infinite number of closest points: for example, one can choose the

new closest point according to the previous position of the closest point. The new point has to be chosen in agreement with the relative motion.

Let us demonstrate the closest point projection procedure for a special master geometry and slave point's path (Fig. 2.8,a). The master surface  $ABCD$  consists of a straight segment  $AB$ , a circle arc  $BC$  and another straight segment  $CD$ . The slave point follows the path 1-6-10-1. In Fig. 2.8,b, the following projection zones are presented:

- all points in "K" zone are projected on the arc  $BC$ ;
- "L" - on the segment  $AB$ ;
- "M" - on the vertex  $B$ ;
- "N" - on the vertex  $C$ ;
- "P" - on the segment  $CD$ .

Some projection zones are separated by red dashed lines, they correspond to discontinuities in the closest point projection, i.e. if a slave point crosses such a line, its projection (closest point on the master surface) jumps from one segment to another. For example, on path 1 – 2, the slave point moves in the projection zone "K" and has a projection on the segment  $BC$ . At point 2, as shown in Fig. 2.8,c, the slave point has two projections: point  $P$  on the arc  $BC$  and point  $Q$  on the segment  $AB$ . Next, on path 2 – 3 – 4 – 5 – 6 there are no discontinuities in the closest point projection. At point 6, the closest point jumps from vertex  $B$  to  $C$ . Point 7 is a special point on the path, all points of the segment  $BC$  are closest points; so in compliance with forementioned any starting point in Newton's procedure will be chosen as the closest one, consequently sometimes the continuity of the projection point can be retained if the starting point is chosen appropriately, however in the presented case there is no way for that, so there is a discontinuity in point 7. On path 7 – 8 – 9 – 10 the projection is continuous. At point 10, there are two closest points on different segments. Path 10 – 1 is continuous.

From a numerical point of view, it is not crucial that on red dashed lines several closest points exist, because the probability to find a slave node exactly on these lines is zero. What is important is that, when such a line is crossed, the projection point jumps relatively far from its position at the previous step. Decreasing the time step to zero does not decrease the jump size. Another observation can be made: in domains  $M$  and  $N$ , any change of position of the slave point does not result in the change of position of the closest point.

A small change in position of the slave point does not always result in a small change in the closest point projection, i.e. the closest point motion over the master surface is not continuous. As seen from the example, this difficulty is sometimes related to the fact that the slave point is relatively far from the master surface (points 6 and 7 in Fig. 2.8), but it may be also crucial even for sufficiently close slave nodes (point 10 in Fig. 2.8). In the case of far slave point the jump can be very big even if the master surface is  $C^2$ -smooth. The energy dissipation within one computational increment in frictional contact problems

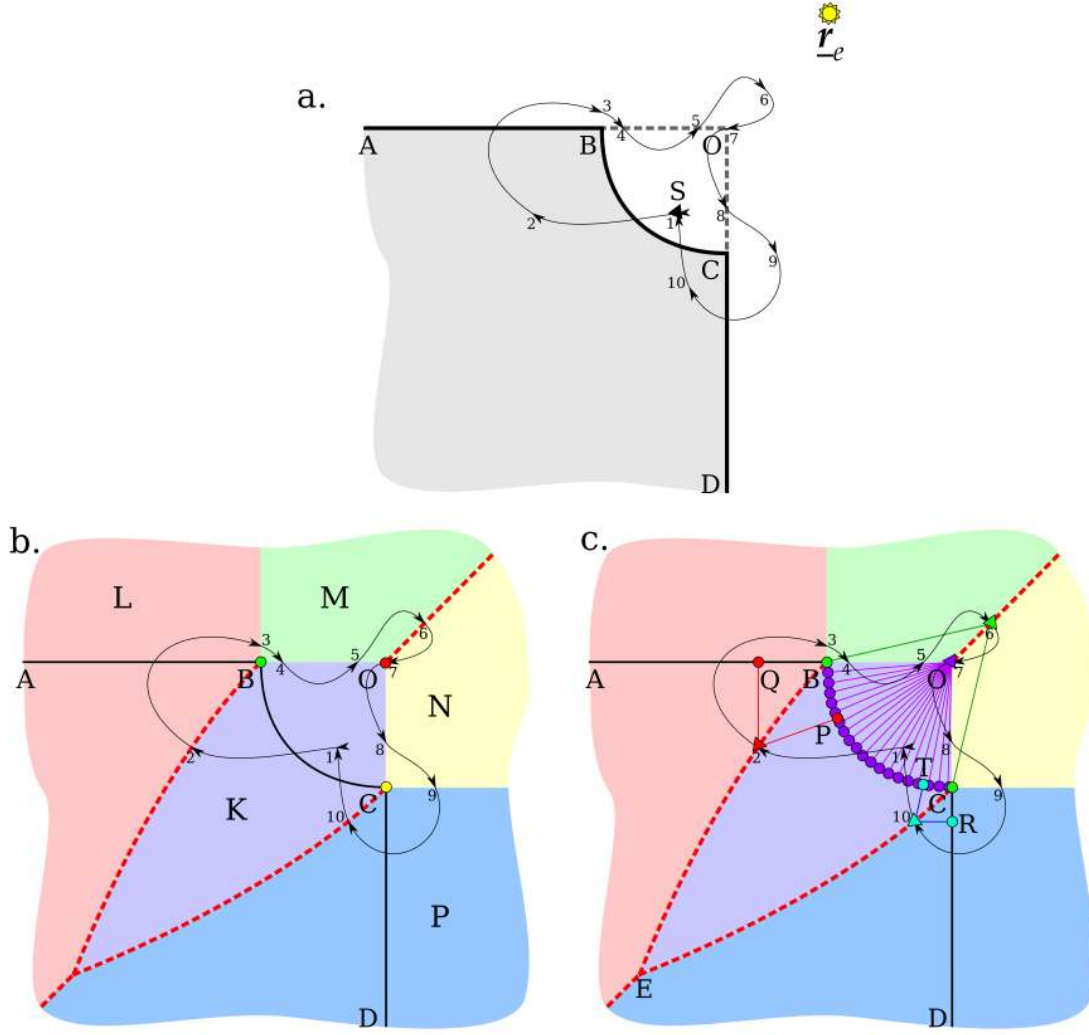


Figure 2.8: Path of a slave point close to a piecewise smooth master surface, projection zones and related problems of the closest point projection.

is related to the sliding distance. If a small change in geometry results in a jump of the projection point, so the sliding distance changes abruptly, then the corresponding dissipated energy will also experience a jump, i.e. the virtual work of frictional forces becomes non-continuous. In this case the convergence of Newton's method cannot be ensured. These consequences can not be avoided in the classical approach, when the gap and sliding distance are measured with respect to the *closest projection point* on the master surface. That is why a new approach will be proposed in the next section.

### 2.2.3 Aircraft's shadow projection method



### Idea 2.1 Aircraft's shadow projection method.

As an alternative to the closest point projection, we propose to determine the gap and the relative sliding according to a new projection technique – aircraft's shadow projection or simply shadow projection. It establishes the relation between a slave point and its shadow or "back-shadow" on the master surface. This shadow is cast by an imaginary light-emitting-point (emitter) placed at some distance from the slave-master system. This technique has been inspired from unique and continuous aircraft's shadow cast by the Sun on the Earth's surface. Also this shadow projection point can be used as a starting point for Newton's procedure to detect the closest projection point for complex surfaces.

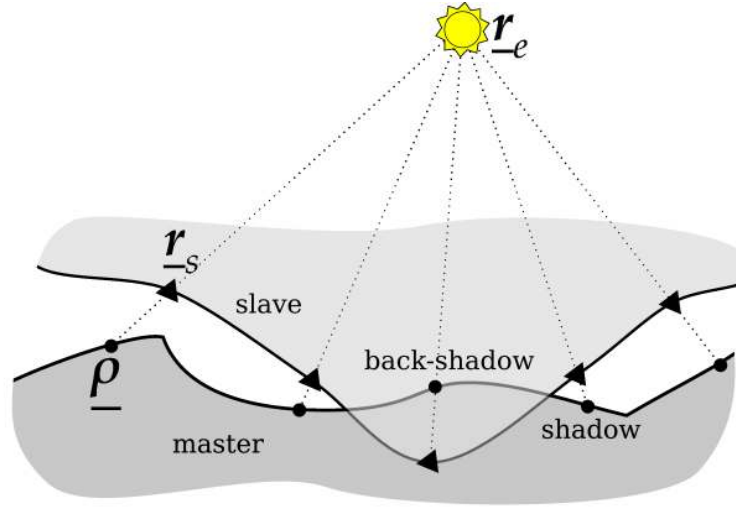


Figure 2.9: An imaginary light-emitting-point (emitter)  $\underline{r}_e$  produces shadows and back-shadows  $\underline{\rho}$  of slave nodes  $\underline{r}_s$  onto the nonsmooth master surface.

If the shadow exists, then it is unique, if the master surface does not cast "shadows" on its own. To find a shadow or a "back-shadow"  $\underline{\rho}$  (see Fig. 2.9) of a slave point  $\underline{r}_s$  due to an imaginary point  $\underline{r}_e$  emitting light, one has to solve a nonlinear equation

$$\underline{F}(\underline{\xi}, \alpha) = \underline{\rho}(\underline{\xi}) - \underline{r}_e - \alpha(\underline{r}_s - \underline{r}_e) = 0, \quad (2.12)$$

where  $\alpha > 0$  is an unknown coefficient. To solve Eq. (2.12) by Newton's method, it should be linearized

$$\underline{F}(\underline{\xi}, \alpha) + \begin{bmatrix} \frac{\partial \underline{F}}{\partial \underline{\xi}} \\ \frac{\partial \underline{F}}{\partial \alpha} \end{bmatrix}^T \begin{bmatrix} \delta \underline{\xi} \\ \delta \alpha \end{bmatrix} = 0 \quad \Leftrightarrow \quad \underline{\rho} - \underline{r}_e - \alpha(\underline{r}_s - \underline{r}_e) + \begin{bmatrix} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \\ -(\underline{r}_s - \underline{r}_e) \end{bmatrix}^T \begin{bmatrix} \delta \underline{\xi} \\ \delta \alpha \end{bmatrix} = 0$$

We take the dot product of the latter expression with the v-vector  $\begin{bmatrix} \frac{\partial F}{\partial \xi} & \frac{\partial F}{\partial \alpha} \end{bmatrix}^T$  and extract the increments  $\delta \xi^i$  and  $\delta \alpha^i$

$$\begin{bmatrix} \delta \xi^i \\ \delta \alpha^i \end{bmatrix} = - \left[ \begin{array}{cc} \frac{\partial \underline{\rho}}{\partial \xi} \cdot \frac{\partial \underline{\rho}}{\partial \xi} & -(\underline{\mathbf{r}}_s - \underline{\mathbf{r}}_e) \cdot \frac{\partial \underline{\rho}}{\partial \xi} \\ -(\underline{\mathbf{r}}_s - \underline{\mathbf{r}}_e) \cdot \frac{\partial \underline{\rho}}{\partial \xi} & (\underline{\mathbf{r}}_s - \underline{\mathbf{r}}_e)^2 \end{array} \right]_{\xi^i, \alpha^i}^{-1} \left[ \begin{array}{c} \frac{\partial \underline{\rho}}{\partial \xi} \cdot (\underline{\rho} - \underline{\mathbf{r}}_e - \alpha(\underline{\mathbf{r}}_s - \underline{\mathbf{r}}_e)) \\ -(\underline{\mathbf{r}}_s - \underline{\mathbf{r}}_e) \cdot (\underline{\rho} - \underline{\mathbf{r}}_e - \alpha(\underline{\mathbf{r}}_s - \underline{\mathbf{r}}_e)) \end{array} \right]_{\xi^i, \alpha^i} \quad (2.13)$$

It can be easily shown that the inverse matrix always exists, i.e. the determinant is nonzero. The unknowns at  $i$ th increment become

$$\xi^{i+1} = \xi^i + \delta \xi^i, \quad \alpha^{i+1} = \alpha^i + \delta \alpha^i$$

and  $\{\xi^0; \alpha^0\}$  is the starting point. In this method, there is no need to split the master surface into faces/segments, edges and vertices as it has been done for the closest point projection. The projection point is assumed to be always located on one of the master faces/segments  $\underline{\rho} \in \Gamma_m^i$ . This is due to the fact that the probability to get a shadow projection on an edge or a vertex is zero. Since the projection zones of edges and vertices are simply surfaces and lines respectively. Their measures are zero in comparison to the measure of any volume (corresponding to projection zones of faces) in three-dimensional space, so rigorously the shadow projection is situated on one of faces of the master for *almost all slave points* if it exists.

The proposed method has some drawbacks:

- it is complicated to check that for a given light emitter there are no self-shadows (see Fig. 2.10,c), i.e. there is no master points for which Eq. (2.12) has more than one solution;
- it is not always easy to find a proper position for the emitter satisfying the non self-shadow conditions, moreover such an emitter may not exist.

However there are also several advantages:

- If the master surface allows to find an emitter which does not create self-shadows, it is rather simple to find the shadow projection point and moreover such a point is unique.
- The projection of the master surface is continuous and does not depend on smoothness and curvatures of the master surface; One can check the continuity of the shadow projection by carrying out an imaginary test with the slave point's path and the master geometry given in Fig. 2.8,a for the emitter  $\underline{\mathbf{r}}_e$ . This statement is valid if slave nodes remain closer to the master surface than the emitter, which is always the case if the emitter is chosen to be infinitely remote from the master.

## 2.2 Interaction between contacting surfaces

- If the emitting point is placed infinitely far from the master surface, which corresponds to the Sun shadow of an aircraft representing a slave point; in this case Eq. (2.12) becomes simpler:

$$\underline{F}(\underline{\xi}, \alpha) = \underline{\rho}(\underline{\xi}) + \alpha \underline{e} - \underline{r}_s = 0, \quad (2.14)$$

where  $\underline{e}$  is a constant unit vector pointing in the direction of the infinitely remote emitter - a pointer to the emitter. The resulting equation for increments  $\delta \underline{\xi}, \delta \alpha$  in Newton's method can be constructed as in Eq. 2.13

$$\begin{bmatrix} \delta \underline{\xi}^i \\ \delta \alpha^i \end{bmatrix} = - \begin{bmatrix} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} & \underline{e} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \\ \underline{e} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot (\underline{\rho} + \alpha \underline{e} - \underline{r}_s) \\ \underline{e} \cdot (\underline{\rho} + \alpha \underline{e} - \underline{r}_s) \end{bmatrix} \bigg|_{\underline{\xi}^i, \alpha^i} \quad (2.15)$$

where at the solution point, the coefficient  $\alpha$  is nothing, but a new "shadow" gap  $g_s$  (different from the normal gap  $g_n$ )

$$\underline{r}_s = \underline{\rho} + g_s \underline{e}.$$

As it will be demonstrated later, it is much easier to work with the latter expression than with the similar one for the normal projection

$$\underline{r}_s = \underline{\rho} + g_n \underline{n}.$$

In Fig. 2.10, different master surfaces and different positions of the emitter are represented.

### 2.2.4 Tangential relative sliding

Consideration of frictional contact requires tracking of relative motion of two surfaces both in normal and tangential directions. As shown later, the variation of the tangential relative velocity  $\dot{\underline{g}}_t$  enters in the main equations governing frictional contact (see Chapter 4). To obtain an accurate expression for  $\dot{\underline{g}}_t$  one has to take into account two independent motions, those of the slave point and of the master surface and to introduce the deformation of the latter.

A simple 1D example depicted in Fig. 2.11 demonstrates the relative motion of a point  $A$  over a straight segment  $BC$ . Absolute velocities  $\underline{v}_A, \underline{v}_B, \underline{v}_C$  correspond to points  $A, B$  and  $C$  respectively. Let us express the relative tangential velocity  $\dot{\underline{g}}_t$  of the point  $A'$  (projection of the point  $A$ ) on the segment  $BC$ . If the segment  $BC$  is parametrized with  $\zeta \in [0; 1]$ , then the point  $A'$  can be expressed as

$$\underline{\rho}(A') = \zeta \underline{\rho}(C) + (1 - \zeta) \underline{\rho}(B). \quad (2.16)$$

Since the problem is one-dimensional, then

$$\underline{\rho}(A) = \underline{\rho}(A') \quad (2.17)$$



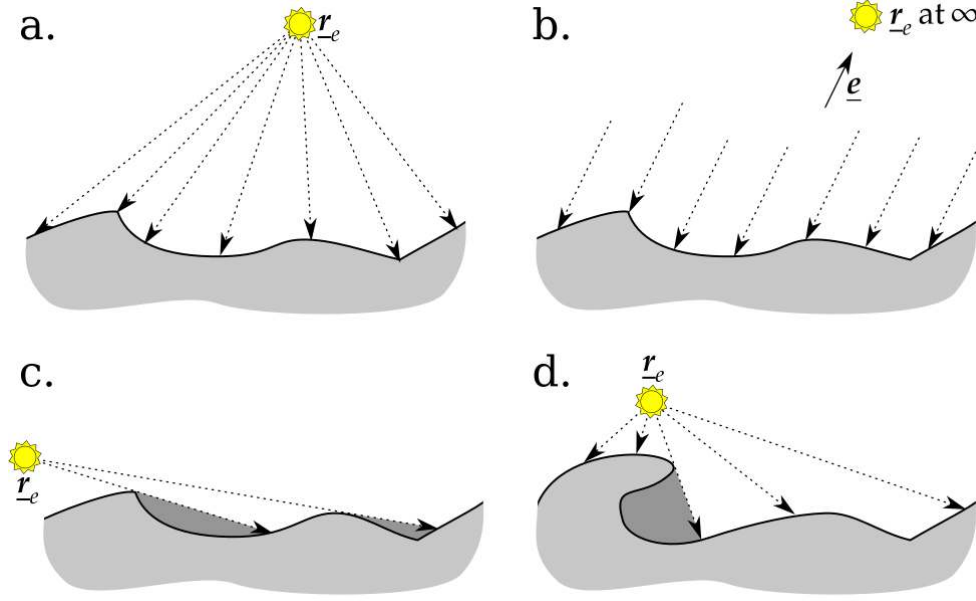


Figure 2.10: Possible emitter-master configuration: unique projection of slave nodes possessing a shadow: **a** – close emitter; **b** – infinitely remote emitter; invalid configurations: **c** – incorrectly chosen emitter (there are self-shadows on the master surface); **d** – the emitter deriving a unique projection for the given master surface does not exist (self shadows are always present).

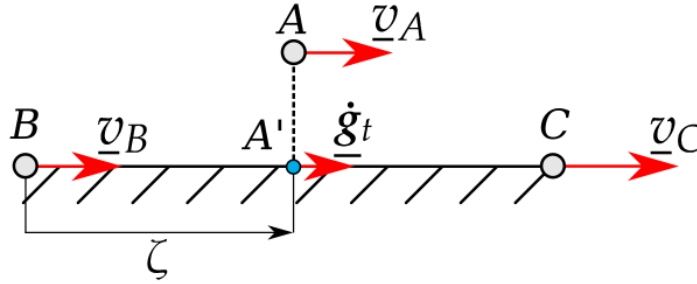


Figure 2.11: Scheme for definition of the relative tangential velocity in 1D.

and the parameter  $\zeta$  can be expressed as

$$\zeta = \frac{(\underline{\rho}(A) - \underline{\rho}(B)) \cdot (\underline{\rho}(C) - \underline{\rho}(B))}{\|\underline{\rho}(C) - \underline{\rho}(B)\|^2}$$

Then the absolute velocity of the projection point can be calculated as

$$\underline{v}_{A'} = \underbrace{\zeta \underline{v}_C + (1 - \zeta) \underline{v}_B}_{= \frac{\partial \underline{\rho}(A')}{\partial t}} + \underbrace{[\underline{\rho}(C) - \underline{\rho}(B)] \dot{\zeta}}_{= \frac{\partial \underline{\rho}(A')}{\partial \zeta} \dot{\zeta}}$$

## 2.2 Interaction between contacting surfaces

Finally, to get the relative tangential velocity one has to subtract from this expression the velocity of the projection point  $\frac{\partial \underline{\rho}(A')}{\partial t} = \underline{v}_{A'}(\zeta)$  for a fixed  $\zeta$

$$\underline{\dot{g}}_t = \underline{v}_{A'} - \frac{\partial \underline{\rho}(A')}{\partial t} = [\underline{\rho}(C) - \underline{\rho}(B)]\dot{\zeta} + \underbrace{\zeta \underline{v}_C + (1 - \zeta) \underline{v}_B - \frac{\partial \underline{\rho}(A')}{\partial t}}_{=0}$$

The expression for the relative tangential sliding is then

$$\underline{\dot{g}}_t = \frac{\partial \underline{\rho}(A')}{\partial \zeta} \dot{\zeta}. \quad (2.18)$$

The velocity of the parameter  $\dot{\zeta}$  can be computed by taking the derivative of Eq. (2.16), substituting (2.17) and evaluating the dot product with vector  $\underline{t} = \underline{\rho}(C) - \underline{\rho}(B)$

$$\dot{\zeta} = \frac{[\underline{v}_A - \zeta \underline{v}_C - (1 - \zeta) \underline{v}_B] \cdot \underline{t}}{\|\underline{t}\|^2}$$

Eq. (2.18) is nothing but the derivative of the projection point vector  $\underline{\rho}(A')$  along the changing vector field  $\underline{\rho}$  of the master segment  $BC$ .

In the general case the tangential relative sliding velocity is given as

$$\underline{\dot{g}}_t = \frac{\partial \underline{\rho}}{\partial \tilde{\xi}}^T \tilde{\xi}, \quad (2.19)$$

where  $\frac{\partial \underline{\rho}}{\partial \tilde{\xi}}$  is a v-vector of the local basis and  $\tilde{\xi}$  is a convective parameter of the master surface. Note that sometimes instead of classical time derivative notation  $\underline{\dot{g}}_t$ , a  $\underline{\dot{g}}_t$  notation is employed (see [Pietrzak 97], [Pietrzak 99]) to emphasize the fact that there is no vector, the time derivative of which is equal to the tangential relative sliding velocity between two deformable bodies. The required velocity vector is obtained as the Lie derivative of the vector connecting slave point  $\underline{r}_s$  and its projection  $\underline{\rho}_\pi$ . The Lie derivative evaluates the change of a tensor field<sup>3</sup> along the change of a vector field. The objective expression for the relative tangential sliding velocity has been obtained by Laursen and Simo [Laursen 93], [Laursen 94] and following by Curnier, He and Klarbring [Curnier 95]. A critical discussion of different forms for the tangential relative velocity can be found in [Pietrzak 97]. In the following we will make use of the variation of the tangential relative sliding, which is stated as

$$\delta \underline{g}_t = \frac{\partial \underline{\rho}}{\partial \tilde{\xi}}^T \delta \tilde{\xi}, \quad (2.20)$$

<sup>3</sup>by *tensor field* here we mean any order tensor field.

The length of the sliding path is computed as the following integral

$$\underline{g}_t = \int_{t_0}^t \|\dot{\underline{g}}_t\| dt' = \int_{t_0}^t \left\| \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \underline{\xi} \right\| dt' = \int_{t_0}^t \sqrt{\underline{\xi}^T \underline{\mathbb{A}} \underline{\xi}} dt',$$

where  $t_0$  is the time of switch to the slip state,  $t$  is the actual time, and  $\underline{\mathbb{A}}$  is the first fundamental t-scalar of the surface.

### 2.2.5 First order variations

As already mentioned, the expressions of the first variations of the geometrical quantities are needed to incorporate contact in the weak form. Now, the first order variations of the normal  $\delta g_n$  and “shadow”  $\delta g_s$  gaps as well as the variations of projection point coordinates  $\delta \underline{\xi}$  for normal and shadow projections will be derived for 3D surfaces and 2D curves. We will make use of s-structures presented in detail in Appendix A. Results presented in this section for normal projections have been obtained in collaboration with **Olga Trubienko**, the master student at the Department of Mechanics and Mathematics, Moscow State University, in Feb-Mar 2010.

#### Normal projection case

The slave point vector is represented by the sum of the projection vector and the normal vector multiplied by the normal gap

$$\underline{r}_s = \underline{\rho} + g_n \underline{n}. \quad (2.21)$$

The full variation of Eq. (2.21) gives

$$\bar{\delta} \underline{r}_s = \bar{\delta} \underline{\rho} + \delta g_n \underline{n} + g_n \bar{\delta} \underline{n}. \quad (2.22)$$

It can be shown that  $\bar{\delta} g_n = \delta g_n$ ; to show that, one can take a partial derivative of expression (2.22) with respect to  $\underline{\xi}$ ;  $g_n$  can be considered as the third independent coordinate (the first two are given by  $\underline{\xi}$ ) of the slave point in the local master basis. Expanding the full variations in (2.21) leads to

$$\delta \underline{r}_s = \delta \underline{\rho} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} + \delta g_n \underline{n} + g_n \left( \delta \underline{n} + \frac{\partial \underline{n}}{\partial \underline{\xi}}^T \delta \underline{\xi} \right). \quad (2.23)$$

Since  $\underline{r}_s$ ,  $\underline{\xi}$ ,  $g_n$  depend only on time, their full variations  $\bar{\delta}$  coincide with the variation  $\delta$ .


**Remark 2.3 on variations.**

By variation of the function  $f(t, x(t))$  at point  $\{t_0, x(t_0)\}$  we mean the following

$$\delta f(t_0, x(t_0)) = \lim_{\delta t \rightarrow 0} (f(t_0 + \delta t, x(t_0)) - f(t_0, x(t_0))) = \left. \frac{\partial f}{\partial t} \right|_{t_0} \delta t.$$

The last equality is correct if the function  $f$  has a  $C^1$  smoothness by  $t$  within a certain neighborhood of the point  $t_0$ . In the considered framework, for example, the variation of the projection point  $\underline{\rho}$  is

$$\delta \underline{\rho}(t_0) = \lim_{\delta t \rightarrow 0} (\underline{\rho}(t_0 + \delta t, \underline{\xi}(t_0)) - \underline{\rho}(t_0, \underline{\xi}(t_0))).$$

For the full variation of the function  $f(t, x(t))$  we get

$$\begin{aligned} \bar{\delta} f(t_0, x(t_0)) &= \lim_{\delta t \rightarrow 0} (f(t_0 + \delta t, x(t_0 + \delta t)) - f(t_0, x(t_0))) = \\ &= \left. \frac{\partial f}{\partial t} \right|_{t_0} \delta t + \left. \frac{\partial f}{\partial x} \right|_{t_0} \delta x = \delta f + \left. \frac{\partial f}{\partial x} \right|_{t_0} \delta x, \end{aligned}$$

The second equality is correct if the function  $f$  is  $C^1$  smooth by  $t$  and  $x$  within a certain neighborhood of the point  $\{t_0, x(t_0)\}$ .

The scalar product of the expression (2.23) with the normal vector  $\underline{n}$  gives directly the first variation of the normal gap  $g_n$ :

$$\delta g_n = \underline{n} \cdot (\delta \underline{r}_s - \delta \underline{\rho}), \quad (2.24)$$

as

$$\underline{n} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = 0; \quad \underline{n} \cdot \delta \underline{n} = 0 \quad \text{and} \quad \underline{n} \cdot \frac{\partial \underline{n}}{\partial \underline{\xi}} = 0. \quad (2.25)$$

The dot product of Eq. (2.23) with the covariant basis v-vector gives

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}) = \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} + g_n \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \underline{n} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{n}}{\partial \underline{\xi}}^T \delta \underline{\xi} \right),$$

expressing  $\delta \underline{\xi}$  we get

$$\delta \underline{\xi} = \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T + g_n \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{n}}{\partial \underline{\xi}}^T \right]^{-1} \cdot \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}) - g_n \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \underline{n} \right), \quad (2.26)$$

where the first term in square brackets is the first fundamental surface metric matrix (t-scalar)

$$\underline{\mathbb{A}}_{\approx} = \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \quad (2.27)$$

and the second term should be transformed to get rid of the normal vector's derivative, so

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} = 0 \Rightarrow \frac{\partial}{\partial \underline{\xi}} \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} \right) = 0 \Leftrightarrow \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{n}}{\partial \underline{\xi}} = -\frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \cdot \underline{n} \quad (2.28)$$

which is nothing but the second fundamental surface matrix (t-scalar) with minus sign

$$\underline{H} = \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \cdot \underline{n}. \quad (2.29)$$

The last term in (2.26) should be also transformed in order to avoid the variation of the normal, since it is more convenient to work with variations of the local basis vectors. Similarly to (2.28) we get

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} = 0 \Rightarrow \delta \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} \right) = 0 \Leftrightarrow \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \underline{n} = -\delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n}. \quad (2.30)$$

Substituting (2.27), (2.29) and (2.30) into (2.26) gives the expression for the variation of the local coordinate

$$\delta \underline{\xi} = \left[ \underline{\tilde{A}} - g_n \underline{H} \right]^{-1} \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}) + g_n \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right). \quad (2.31)$$

This is the classical expression originally obtained in [Laursen 92], [Laursen 93]. Often, the normal gap is assumed to be relatively small and is neglected, which gives a simpler expression

$$\delta \underline{\xi} = \underline{\tilde{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}) = \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}), \quad (2.32)$$

These two formulations (2.31) and (2.32) will be compared in Section 2.3.

### Shadow projection case: infinitely remote emitter

In case of the “shadow” projection and an infinitely remote emitter, one makes a start from the expression connecting the slave point  $\underline{r}_s$  and its shadow projection  $\underline{\rho}$  through the shadow gap  $g_s$  and the unit vector  $\underline{e}$  pointing to the emitter

$$\underline{r}_s = \underline{\rho} + g_s \underline{e}$$

the variation of this expression gives

$$\delta \underline{r}_s = \delta \underline{\rho} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta \underline{\xi} + \delta g_s \underline{e}. \quad (2.33)$$

## 2.2 Interaction between contacting surfaces

The scalar product of (2.33) with the vector  $\underline{e}$  yields

$$\delta g_s = \underline{e} \cdot \left[ (\delta \underline{r}_s - \delta \underline{\rho}) - \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \delta \underline{\xi} \right], \quad (2.34)$$

As one can see, to evaluate  $\delta g_s$  one needs an expression for  $\delta \underline{\xi}$ ; to get it, we take a dot product of (2.33) with a v-vector containing two unit orthogonal vectors  $\underline{s}_1, \underline{s}_2$  laying in the plane orthogonal to the vector  $\underline{e}$ :

$$\underline{s} = \begin{bmatrix} \underline{s}_1 \\ \underline{s}_2 \end{bmatrix}, \quad \underline{s}_1 \cdot \underline{s}_2 = 0, \quad \underline{s} \cdot \underline{e} = 0$$

$$\underline{s} \cdot (\delta \underline{r}_s - \delta \underline{\rho}) = \underline{s} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \delta \underline{\xi},$$

from which it follows

$$\delta \underline{\xi} = \underline{\underline{P}} \underline{s} \cdot (\delta \underline{r}_s - \delta \underline{\rho}), \quad (2.35)$$

with

$$\underline{\underline{P}} = \left[ \underline{s} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \right]^{-1}. \quad (2.36)$$

It is worth mentioning that expression (2.35) in comparison to (2.31) does not contain, at least explicitly, any gap  $g_n$  or  $g_s$ .

The product  $\underline{\underline{P}} \underline{s}$  can be represented in the following form

$$\underline{\underline{P}} \underline{s} = \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} \cdot \left[ \underline{I} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] \quad (2.37)$$

To prove it we express the v-vector  $\underline{s}$  through the surface basis vectors

$$\underline{s} = \overline{\underline{a}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} + \underline{b} \underline{n} \quad (2.38)$$

Dot products of (2.37) with the contravariant basis v-vector and with the normal vector give respectively

$$\underline{\underline{a}} = \underline{s} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}}, \quad \underline{b} = \underline{s} \cdot \underline{n}, \quad (2.39)$$

as

$$\overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = \underline{I}, \quad \underline{n} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} = 0 \quad \text{and} \quad \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} \cdot \underline{n} = 0$$

Now Eq. (2.36) can be rewritten using (2.37) and (2.39)

$$\underline{\underline{P}} \underline{s} = \left[ \underline{s} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \right]^{-1} \left[ \underline{s} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \right] \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} + \left[ \underline{s} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \right]^{-1} \left[ \underline{s} \cdot \underline{n} \right] \underline{n}, \quad (2.40)$$

where the first pair of square brackets gives a unit t-scalar  $\underline{\underline{I}}$ , and the second pair of square brackets can be replaced by  $\underline{\underline{\zeta}}$

$$\left( \underline{\underline{s}} \cdot \frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}} \right)^{-1} (\underline{\underline{s}} \cdot \underline{\underline{n}}) = \underline{\underline{\zeta}} \Leftrightarrow \underline{\underline{s}} \cdot \underline{\underline{n}} = \left( \underline{\underline{s}} \cdot \frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}} \right) \underline{\underline{\zeta}} \Leftrightarrow \underline{\underline{s}} \cdot \left( \underline{\underline{n}} - \frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}} \underline{\underline{\zeta}} \right) = 0 \Leftrightarrow \underline{\underline{n}} - \frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}} \underline{\underline{\zeta}} = \beta \underline{\underline{e}}, \quad (2.41)$$

where  $\beta$  is another coefficient. Left dot product of the last term with the covariant basis v-vector gives the expression for  $\underline{\underline{\zeta}}$

$$\underline{\underline{\zeta}} = -\beta \bar{\underline{\underline{A}}} \frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}} \cdot \underline{\underline{e}} = -\beta \overline{\frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}}} \cdot \underline{\underline{e}}. \quad (2.42)$$

Substituting (2.42) in (2.40) yields

$$\underline{\underline{P}} \underline{\underline{s}} = \overline{\frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}}} - \beta \left( \overline{\frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}}} \cdot \underline{\underline{e}} \right) \underline{\underline{n}}.$$

Finally, to evaluate the coefficient  $\beta$ , one takes the dot product of the latter equation with vector  $\underline{\underline{e}}$

$$\left( \overline{\frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}}} \cdot \underline{\underline{e}} \right) (1 - \beta \underline{\underline{n}} \cdot \underline{\underline{e}}) = 0 \Rightarrow \beta = \frac{1}{\underline{\underline{n}} \cdot \underline{\underline{e}}}.$$

So Eq. (2.37) has been proven. Using the latter equalities, the variation of the surface parameter (2.35) can be rewritten as

$$\delta \underline{\underline{\xi}} = \overline{\frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{\underline{e}} \otimes \underline{\underline{n}}}{\underline{\underline{e}} \cdot \underline{\underline{n}}} \right] \cdot (\delta \underline{\underline{r}}_s - \delta \underline{\underline{\rho}}). \quad (2.43)$$

#### Remark 2.4 on the special projection operator.

The terms on the right-hand side appearing in Eq. (2.43) present a special nonsymmetric projection operator. The left dot product with a given vector  $\underline{\underline{a}}$  projects the latter on the plane orthogonal to the normal vector, the right dot product projects it on the plane orthogonal to the pointer vector  $\underline{\underline{e}}$ , as presented in the figure below. If the vector standing on right hand is collinear to the pointer  $\underline{\underline{e}}$ , then the projection operator gives zero, if vector standing on left hand is collinear with the normal  $\underline{\underline{n}}$ , the projection operator gives also zero.

$$\left[ \underline{\underline{I}} - \frac{\underline{\underline{e}} \otimes \underline{\underline{n}}}{\underline{\underline{e}} \cdot \underline{\underline{n}}} \right] \cdot \underline{\underline{a}} = \underline{\underline{b}} \Rightarrow \underline{\underline{b}} \cdot \underline{\underline{n}} = 0 \quad \text{and} \quad \underline{\underline{a}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{\underline{e}} \otimes \underline{\underline{n}}}{\underline{\underline{e}} \cdot \underline{\underline{n}}} \right] = \underline{\underline{c}} \Rightarrow \underline{\underline{c}} \cdot \underline{\underline{e}} = 0$$

$$\text{if } \underline{a} = \alpha \underline{e} \text{ then } \left[ \underline{I} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] \cdot \underline{a} = 0$$

$$\text{if } \underline{a} = \alpha \underline{n} \text{ then } \underline{a} \cdot \left[ \underline{I} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] = 0$$

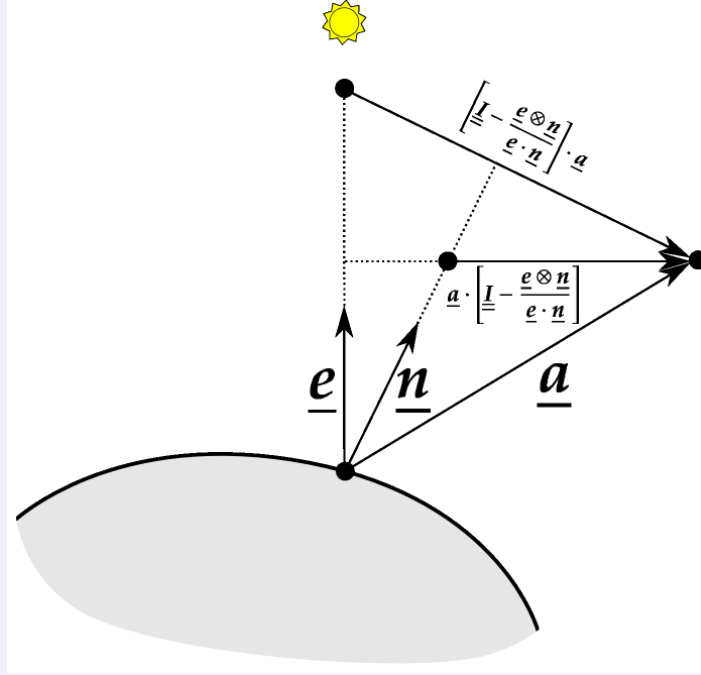


Fig. Representation of the projection operator

Substituting (2.43) into (2.34) yields the expression for the shadow gap's variation

$$\delta g_s = \left\{ \underline{e} - \underline{e} \cdot \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \otimes \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right] \cdot \left( \underline{I} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right) \right\} \cdot (\delta \underline{r}_s - \delta \underline{\rho}). \quad (2.44)$$

It can be shown that the term in square brackets is nothing but a symmetric



projection operator on the basis (tangential) plane:

$$\begin{aligned} \underline{a} \cdot \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \otimes \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right] &= \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \otimes \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right] \cdot \underline{a} = \underline{a} - \underline{a} \cdot \underline{n} \otimes \underline{n} \Leftrightarrow \\ \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \otimes \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right] &= \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \otimes \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right] = \underline{I} - \underline{n} \otimes \underline{n}, \end{aligned} \quad (2.45)$$

carrying this in mind, (2.44) transforms into

$$\delta g_s = \frac{\underline{n}}{\underline{n} \cdot \underline{e}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}). \quad (2.46)$$

### Shadow projection case: close emitter

In case of “shadow” projection from a close emitter, one starts from

$$\underline{r}_s = \underline{\rho} + g_s \underline{e}, \quad (2.47)$$

the difference from the previous case is that vector  $\underline{e}$  is not constant and takes the following form

$$\underline{e} = \frac{\underline{r}_e - \underline{r}_s}{\|\underline{r}_e - \underline{r}_s\|} = \frac{\underline{r}_e - \underline{r}_s}{d_{se}},$$

where  $d_{se} = \|\underline{r}_e - \underline{r}_s\|$  is the distance between the slave point and the emitter, variation of this vector is

$$\delta \underline{e} = -\delta \underline{r}_s^S \quad (2.48)$$

where  $\delta \underline{r}_s^S$  is a normalized projection of the vector  $\delta \underline{r}_s$  on a plane orthogonal to  $\underline{e}$

$$\delta \underline{r}_s^S = (\underline{I} - \underline{e} \otimes \underline{e}) \cdot \frac{\delta \underline{r}_s}{d_{se}} = \underline{\tilde{s}}^T \otimes \underline{\tilde{s}} \cdot \frac{\delta \underline{r}_s}{d_{se}} = \underline{\tilde{s}} \otimes \underline{\tilde{s}} \cdot \frac{\delta \underline{r}_s}{d_{se}}. \quad (2.49)$$

The variation of (2.47) gives

$$\delta \underline{r}_s = \delta \underline{\rho} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} + \delta g_s \underline{e} + g_s \delta \underline{e}. \quad (2.50)$$

Since  $\delta \underline{e} \cdot \underline{e} = 0$ , the dot product of expression (2.50) with vector  $\underline{e}$  gives exactly the same expression as (2.34)

$$\delta g_s = \underline{e} \cdot \left[ (\delta \underline{r}_s - \delta \underline{\rho}) - \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} \right], \quad (2.51)$$

the dot product of (2.50) with v-vector  $\underline{\tilde{s}}$  together with (2.48) gives

$$\underline{\tilde{s}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} = \underline{\tilde{s}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}) + g_s \underline{\tilde{s}} \cdot \delta \underline{r}_s^S, \quad (2.52)$$

## 2.2 Interaction between contacting surfaces

and from (2.36), (2.37) and (2.49) carrying that

$$\underline{\underline{s}} \cdot (\underline{\underline{s}}^T \otimes \underline{\underline{s}}) = \underline{\underline{s}}$$

one has

$$\delta_{\underline{\underline{\xi}}}^{\xi} = \frac{\overline{\partial \rho}}{\partial \underline{\underline{\xi}}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{\underline{e}} \otimes \underline{\underline{n}}}{\underline{\underline{e}} \cdot \underline{\underline{n}}} \right] \cdot \left[ \left( 1 + \frac{g_s}{d_{se}} \right) \delta \underline{\underline{r}}_s - \delta \underline{\underline{\rho}} \right]. \quad (2.53)$$

This expression reduces to (2.43) for an infinitely remote emitter ( $d_{se} \rightarrow \infty$ ) and for the small gap approximation ( $g_s = 0$ ). Substituting the surface parameter variation (2.53) in (2.51) gives the expression for the shadow gap variation

$$\delta g_s = \frac{\underline{\underline{n}}}{\underline{\underline{e}} \cdot \underline{\underline{n}}} \cdot (\delta \underline{\underline{r}}_s - \delta \underline{\underline{\rho}}) - \frac{g_s}{d_{se}} \underline{\underline{e}} \cdot \frac{\partial \rho^T}{\partial \underline{\underline{\xi}}} \otimes \frac{\overline{\partial \rho}}{\partial \underline{\underline{\xi}}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{\underline{e}} \otimes \underline{\underline{n}}}{\underline{\underline{e}} \cdot \underline{\underline{n}}} \right] \cdot \delta \underline{\underline{r}}_s, \quad (2.54)$$

using (2.45) one gets

$$\delta g_s = \frac{\underline{\underline{n}}}{\underline{\underline{e}} \cdot \underline{\underline{n}}} \cdot \left[ \left( 1 + \frac{g_s}{d_{se}} \right) \delta \underline{\underline{r}}_s - \delta \underline{\underline{\rho}} \right] - \frac{g_s}{d_{se}} \underline{\underline{e}} \cdot \delta \underline{\underline{r}}_s. \quad (2.55)$$

Assuming small gaps  $g_s \approx 0$  or infinitely remote emitter also leads to the expressions obtained in the previous section (2.43) and (2.46).

### 2.2.6 Second order variations

#### Normal projection case

To get second order variations :  $\Delta \delta g_n$  and  $\Delta \delta \underline{\underline{\xi}}$ , we take first the full second variation  $\bar{\Delta}$  of the expression (2.23):

$$\begin{aligned} \bar{\Delta} \delta \underline{\underline{r}}_s &= \Delta \delta \underline{\underline{r}}_s = \Delta \delta \underline{\underline{\rho}} + \delta \frac{\partial \rho^T}{\partial \underline{\underline{\xi}}} \Delta \underline{\underline{\xi}} + \Delta \frac{\partial \rho^T}{\partial \underline{\underline{\xi}}} \delta \underline{\underline{\xi}} + \Delta \underline{\underline{\xi}}^T \frac{\partial^2 \rho}{\partial \underline{\underline{\xi}}^2} \delta \underline{\underline{\xi}} + \\ &+ \frac{\partial \rho^T}{\partial \underline{\underline{\xi}}} \Delta \delta \underline{\underline{\xi}} + \underline{\underline{\Delta \delta g_n \underline{\underline{n}}}} + \underline{\underline{\delta g_n \bar{\Delta} \underline{\underline{n}}}} + \underline{\underline{\Delta g_n \delta \underline{\underline{n}}}} + g_n \bar{\Delta} \delta \underline{\underline{n}} \end{aligned} \quad (2.56)$$

The first term is zero by default, because  $\underline{\underline{r}}_s = \underline{\underline{r}}_s(t) \Rightarrow \Delta \delta \underline{\underline{r}}_s = 0$ ; it can be shown also that the second term is zero ( $\Delta \delta \underline{\underline{\rho}} = 0$ ); however, a rather similar term  $\Delta \delta \underline{\underline{n}} \neq 0$ . The underlined terms in (2.56) are orthogonal to the normal vector. The twice underlined term is orthogonal to the tangential plane of the surface basis.

In a dot product of Eq. (2.56) with the normal vector, the underlined terms vanish

$$\Delta \delta g_n = -\underline{\underline{n}} \cdot \delta \frac{\partial \rho^T}{\partial \underline{\underline{\xi}}} \Delta \underline{\underline{\xi}} - \underline{\underline{n}} \cdot \Delta \frac{\partial \rho^T}{\partial \underline{\underline{\xi}}} \delta \underline{\underline{\xi}} - \Delta \underline{\underline{\xi}}^T \underline{\underline{n}} \cdot \frac{\partial^2 \rho}{\partial \underline{\underline{\xi}}^2} \delta \underline{\underline{\xi}} - g_n \underline{\underline{n}} \cdot \bar{\Delta} \delta \underline{\underline{n}}, \quad (2.57)$$

where the third term in the right-hand part contains the second fundamental surface matrix (t-scalar)

$$\underline{n} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi^2} = \underline{H} \quad (2.58)$$

and the last term can be expanded as

$$-\underline{n} \cdot \bar{\Delta} \delta \underline{n} = \bar{\Delta} \underline{n} \cdot \delta \underline{n}, \quad (2.59)$$

since

$$\underline{n} \cdot \delta \underline{n} = 0 \Rightarrow \bar{\Delta}(\underline{n} \cdot \delta \underline{n}) = 0 \Leftrightarrow -\underline{n} \cdot \bar{\Delta} \delta \underline{n} = \bar{\Delta} \underline{n} \cdot \delta \underline{n}.$$

A derivative of a unit vector is orthogonal to the vector itself. It means that the derivative of the normal vector is a combination of in-plane basis vectors, so it can be written in the following form

$$\frac{\partial \underline{n}}{\partial \xi} = \underline{B} \frac{\partial \underline{\rho}}{\partial \xi}. \quad (2.60)$$

The right dot product with the transposed basis v-vector gives

$$\frac{\partial \underline{n}}{\partial \xi} \cdot \frac{\partial \underline{\rho}}{\partial \xi}^T = \underline{B} \frac{\partial \underline{\rho}}{\partial \xi} \cdot \frac{\partial \underline{\rho}}{\partial \xi}^T = \underline{B} \underline{A}, \quad (2.61)$$

where the last term is a product of the t-scalar  $\underline{B}$  with the first fundamental surface metric matrix  $\underline{A}$  and the first term can be transformed

$$\underline{n} \cdot \frac{\partial \underline{\rho}}{\partial \xi} = 0 \Rightarrow \frac{\partial}{\partial \xi} \left( \underline{n} \cdot \frac{\partial \underline{\rho}}{\partial \xi} \right) = 0 \Leftrightarrow \frac{\partial \underline{n}}{\partial \xi} \cdot \frac{\partial \underline{\rho}}{\partial \xi}^T = -\underline{n} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi^2} = -\underline{H}, \quad (2.62)$$

that is nothing but the second fundamental surface matrix. From (2.61) and (2.62) it follows that

$$\underline{B} = -\underline{H} \underline{A}^{-1} = -\underline{H} \bar{\underline{A}},$$

substituting of this expression into (2.60) yields

$$\frac{\partial \underline{n}}{\partial \xi} = -\underline{H} \bar{\underline{A}} \frac{\partial \underline{\rho}}{\partial \xi} \Rightarrow \frac{\partial \underline{n}}{\partial \xi} = -\underline{H} \overline{\frac{\partial \underline{\rho}}{\partial \xi}}. \quad (2.63)$$

The same procedure can be carried out for the variation of the normal vector

$$\begin{aligned} & \left\{ \begin{aligned} \delta \underline{n} = \underline{b}^T \frac{\partial \underline{\rho}}{\partial \xi} \Rightarrow \delta \underline{n} \cdot \frac{\partial \underline{\rho}}{\partial \xi}^T &= \underline{b}^T \underline{A} = \underline{A} \underline{b} \\ \delta \left( \underline{n} \cdot \frac{\partial \underline{\rho}}{\partial \xi} \right) = 0 \Leftrightarrow \delta \underline{n} \cdot \frac{\partial \underline{\rho}}{\partial \xi} &= -\underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} \end{aligned} \right. \Rightarrow \\ & \Rightarrow \underline{b} = -\bar{\underline{A}} \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} \right) \Leftrightarrow \delta \underline{n} = - \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} \right)^T \overline{\frac{\partial \underline{\rho}}{\partial \xi}} \end{aligned} \quad (2.64)$$

Finally

$$\delta \underline{n} = - \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \underline{\xi}} \right)^T \frac{\partial \rho}{\partial \underline{\xi}} \quad \text{and} \quad \bar{\delta} \underline{n} = - \left( \underline{n} \cdot \bar{\delta} \frac{\partial \rho}{\partial \underline{\xi}} \right)^T \frac{\partial \rho}{\partial \underline{\xi}} \quad (2.65)$$

According to Eq. (2.65) one can transform (2.59)

$$\bar{\Delta} \underline{n} \cdot \bar{\delta} \underline{n} = \left( \underline{n} \cdot \bar{\Delta} \frac{\partial \rho}{\partial \underline{\xi}} \right)^T \frac{\partial \rho}{\partial \underline{\xi}} \cdot \frac{\partial \rho}{\partial \underline{\xi}}^T \left( \underline{n} \cdot \bar{\delta} \frac{\partial \rho}{\partial \underline{\xi}} \right) = \left( \underline{n} \cdot \bar{\Delta} \frac{\partial \rho}{\partial \underline{\xi}} \right)^T \bar{\Delta} \left( \underline{n} \cdot \bar{\delta} \frac{\partial \rho}{\partial \underline{\xi}} \right),$$

expanding the full variations and using (2.58) results in

$$\begin{aligned} \bar{\Delta} \underline{n} \cdot \bar{\delta} \underline{n} &= \left( \underline{n} \cdot \Delta \frac{\partial \rho}{\partial \underline{\xi}} + \underline{H} \Delta \underline{\xi} \right)^T \bar{\Delta} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \underline{\xi}} + \underline{H} \delta \underline{\xi} \right) = \\ &= \Delta \underline{\xi}^T \underline{H} \bar{\Delta} \underline{H} \delta \underline{\xi} + \left( \underline{n} \cdot \Delta \frac{\partial \rho}{\partial \underline{\xi}} \right)^T \bar{\Delta} \underline{H} \delta \underline{\xi} + \Delta \underline{\xi}^T \underline{H} \bar{\Delta} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \underline{\xi}} \right) + \left( \underline{n} \cdot \Delta \frac{\partial \rho}{\partial \underline{\xi}} \right)^T \bar{\Delta} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \underline{\xi}} \right) \end{aligned} \quad (2.66)$$

Substituting (2.58), (2.59) and the contracted form of (2.66) in (2.57) gives the second variation of the normal gap

$$\begin{aligned} \Delta \delta g_n &= - \underline{n} \cdot \left( \delta \frac{\partial \rho}{\partial \underline{\xi}} \Delta \underline{\xi} + \Delta \frac{\partial \rho}{\partial \underline{\xi}} \delta \underline{\xi} \right) - \Delta \underline{\xi}^T \underline{H} \delta \underline{\xi} + \\ &+ g_n \left( \Delta \underline{\xi}^T \underline{H} + \underline{n} \cdot \Delta \frac{\partial \rho}{\partial \underline{\xi}} \right) \bar{\Delta} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \underline{\xi}} + \underline{H} \delta \underline{\xi} \right) \end{aligned} \quad (2.67)$$

The same form has been obtained in [Laursen 92]. Often, to accelerate the execution of the program, this expression is simplified based on the assumption that the gap is negligibly small  $g_n \approx 0$ . It provides a simpler expression

$$\Delta \delta g_n = - \underline{n} \cdot \left( \delta \frac{\partial \rho}{\partial \underline{\xi}} \Delta \underline{\xi} + \Delta \frac{\partial \rho}{\partial \underline{\xi}} \delta \underline{\xi} \right) - \Delta \underline{\xi}^T \underline{H} \delta \underline{\xi} \quad (2.68)$$

To derive the second variation of the surface parameter  $\Delta \delta \underline{\xi}$  one computes the dot product of (2.56) with the surface basis v-vector; only the term underlined twice in (2.56) vanishes:

$$\begin{aligned} 0 &= \frac{\partial \rho}{\partial \underline{\xi}} \cdot \delta \frac{\partial \rho}{\partial \underline{\xi}} \Delta \underline{\xi} + \frac{\partial \rho}{\partial \underline{\xi}} \cdot \Delta \frac{\partial \rho}{\partial \underline{\xi}} \delta \underline{\xi} + \Delta \underline{\xi}^T \left[ \frac{\partial \rho}{\partial \underline{\xi}} \cdot \frac{\partial^2 \rho}{\partial \underline{\xi}^2} \right]^* \delta \underline{\xi} + \frac{\partial \rho}{\partial \underline{\xi}} \cdot \frac{\partial \rho}{\partial \underline{\xi}}^T \Delta \delta \underline{\xi} \\ &+ \delta g_n \frac{\partial \rho}{\partial \underline{\xi}} \cdot \bar{\Delta} \underline{n} + \Delta g_n \frac{\partial \rho}{\partial \underline{\xi}} \cdot \bar{\delta} \underline{n} + g_n \frac{\partial \rho}{\partial \underline{\xi}} \cdot \bar{\Delta} \bar{\delta} \underline{n} \end{aligned} \quad (2.69)$$

**Remark 2.5 on simplified form of s-structures.**

As one can see, the boxed term in Eq. (2.69) is quite strange from the point of view of Linear Algebra, since the product of non-consistent matrices and vectors is forbidden. Formally in Linear Algebra one can choose between two possibilities:

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} = \underline{\bullet}^T \quad \text{or} \quad \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = \underline{\bullet}.$$

Here, contrary to the rules of Linear Algebra, this product will imply the third-order  $t$ -scalar  $2 \times 2 \times 2$

$$\frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T = {}^3 \underline{\bullet} \quad \text{and} \quad \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} = {}^3 \underline{\bullet}. \quad (*)$$

This non-consistency appears due to the fact that reduced s-structures, which are used here for simplicity, do not allow to pass to higher order s-structures.

The apparatus developed for full s-structures (see Appendix A) is consistent and rigorous, but notions are more complicated. So in this chapter the reduced form of s-structures is employed and it will be simply supposed that operations (\*) are permitted.

Let us consider the last term in (2.69)

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \bar{\Delta} \delta \underline{n},$$

it can be expressed by taking the second variation of the expression

$$\begin{aligned} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} = 0 &\Rightarrow \bar{\Delta} \delta \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} \right) = 0 \Leftrightarrow \bar{\Delta} \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \underline{n} \right) = 0 \Leftrightarrow \\ \bar{\Delta} \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} + \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \bar{\Delta} \underline{n} + \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \underline{n} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \bar{\Delta} \delta \underline{n} &= 0 \Leftrightarrow \\ \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \bar{\Delta} \delta \underline{n} &= -\bar{\Delta} \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} - \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \bar{\Delta} \underline{n} - \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \underline{n} \end{aligned} \quad (2.70)$$

The variations of the normal vector have to be avoided, so they are replaced by (2.65) which yields

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \bar{\Delta} \delta \underline{n} = -\bar{\Delta} \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} + \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}}^T \left( \underline{n} \cdot \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) + \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}}^T \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right). \quad (2.71)$$

The next two terms in Eq. (2.69) contain variations of the normal vector have also to be replaced

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} = 0 \Rightarrow \delta \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} \right) = 0 \Leftrightarrow \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \underline{n} = -\underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \quad (2.72)$$

## 2.2 Interaction between contacting surfaces

Substituting (2.71), (2.72) in (2.69) and replacing its fourth term by the first fundamental metric matrix gives

$$\begin{aligned}
0 = & \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \Delta \underline{\xi} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} + \Delta \underline{\xi}^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \delta \underline{\xi} + \underline{\mathbb{A}} \Delta \delta \underline{\xi} \\
& - \delta g_n \underline{n} \cdot \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} - \Delta g_n \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} - g_n \left[ \bar{\Delta} \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} + \right. \\
& \left. + g_n \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \left( \underline{n} \cdot \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) + g_n \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \right]
\end{aligned} \tag{2.73}$$

Expanding the boxed term and carrying that

$$\Delta \delta \underline{\rho} = 0 \Rightarrow \Delta \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = 0$$

gives

$$\bar{\Delta} \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} = \underline{n} \cdot \delta \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \Delta \underline{\xi} + \left( \underline{n} \cdot \Delta \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} + \Delta \underline{\xi}^T \underline{n} \cdot \frac{\partial^3 \underline{\rho}}{\partial \underline{\xi}^3} \right) \delta \underline{\xi} + \underline{n} \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \Delta \delta \underline{\xi}, \tag{2.74}$$

the last term contains the second fundamental matrix  $\underline{\mathbb{H}}$ ; remember that

$$\frac{\partial^3 \underline{\rho}}{\partial \underline{\xi}^3} = \overset{3}{\underset{\approx}{\bullet}}$$

substituting Eq. (2.74) in (2.73) allows to group terms containing the second variation of the surface parameter

$$\begin{aligned}
(g_n \underline{\mathbb{H}} - \underline{\mathbb{A}}) \Delta \delta \underline{\xi} = & \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \Delta \underline{\xi} + \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} \right) + \Delta \underline{\xi}^T \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \right) \delta \underline{\xi} \\
& - \underbrace{\underline{n} \cdot \left( \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta g_n + \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta g_n \right)}_{\text{term 1}} - g_n \underline{n} \cdot \left( \delta \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \Delta \underline{\xi} + \Delta \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \delta \underline{\xi} \right) - g_n \Delta \underline{\xi}^T \left( \underline{n} \cdot \frac{\partial^3 \underline{\rho}}{\partial \underline{\xi}^3} \right) \delta \underline{\xi} \\
& + \underbrace{g_n \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \right) \left( \underline{n} \cdot \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right)}_{\text{term 2}} + \underbrace{g_n \left( \bar{\Delta} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \right) \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right)}_{\text{term 3}}
\end{aligned} \tag{2.75}$$

The last step is to expand the full derivatives in the marked terms on the right

hand

$$\begin{aligned}
 \text{term 1 : } \underline{n} \cdot \left( \bar{\Delta} \frac{\partial \rho}{\partial \xi} \delta g_n + \delta \frac{\partial \rho}{\partial \xi} \Delta g_n \right) &= \underline{n} \cdot \left( \Delta \frac{\partial \rho}{\partial \xi} \delta g_n + \frac{\partial^2 \rho}{\partial \xi^2} \Delta \xi \delta g_n + \delta \frac{\partial \rho}{\partial \xi} \Delta g_n + \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \Delta g_n \right) = \\
 &= \underline{n} \cdot \left( \Delta \frac{\partial \rho}{\partial \xi} \delta g_n + \delta \frac{\partial \rho}{\partial \xi} \Delta g_n \right) + \mathbb{H} (\Delta \xi \delta g_n + \delta \xi \Delta g_n) \\
 \text{term 2 : } g_n \left( \delta \frac{\partial \rho}{\partial \xi} \cdot \overline{\frac{\partial \rho}{\partial \xi}}^T \right) \left( \underline{n} \cdot \bar{\Delta} \frac{\partial \rho}{\partial \xi} \right) &= g_n \left( \delta \frac{\partial \rho}{\partial \xi} + \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right) \cdot \left( \bar{\Delta} \frac{\partial \rho}{\partial \xi} \right)^T \left( \underline{n} \cdot \Delta \frac{\partial \rho}{\partial \xi} + \mathbb{H} \Delta \xi \right) \\
 \text{term 3 : } &\text{expands as the term 2.}
 \end{aligned} \tag{2.76}$$

These terms inserted in expression (2.74) give the ultimate expression for the second order variation of the surface parameter

$$\begin{aligned}
 \Delta \delta \xi &= (g_n \mathbb{H} - \mathbb{A})^{-1} \left\{ \frac{\partial \rho}{\partial \xi} \cdot \left( \delta \frac{\partial \rho}{\partial \xi}^T \Delta \xi + \Delta \frac{\partial \rho}{\partial \xi} \delta \xi \right) + \Delta \xi^T \left( \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \rho}{\partial \xi^2} \right) \delta \xi - \right. \\
 &\quad - \underline{n} \cdot \left( \Delta \frac{\partial \rho}{\partial \xi} \delta g_n + \delta \frac{\partial \rho}{\partial \xi} \Delta g_n \right) - \mathbb{H} (\Delta \xi \delta g_n + \delta \xi \Delta g_n) - \\
 &\quad - g_n \underline{n} \cdot \left( \delta \frac{\partial^2 \rho}{\partial \xi^2} \Delta \xi + \Delta \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right) - g_n \Delta \xi^T \left( \underline{n} \cdot \frac{\partial^3 \rho}{\partial \xi^3} \right) \delta \xi + \\
 &\quad + g_n \left( \delta \frac{\partial \rho}{\partial \xi} + \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right) \cdot \frac{\partial \rho}{\partial \xi}^T \bar{\mathbb{A}} \left( \underline{n} \cdot \Delta \frac{\partial \rho}{\partial \xi} + \mathbb{H} \Delta \xi \right) + \\
 &\quad \left. + g_n \left( \Delta \frac{\partial \rho}{\partial \xi} + \frac{\partial^2 \rho}{\partial \xi^2} \Delta \xi \right) \cdot \frac{\partial \rho}{\partial \xi}^T \bar{\mathbb{A}} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} + \mathbb{H} \delta \xi \right) \right\}
 \end{aligned} \tag{2.77}$$

Grouping terms with  $\delta g_n$  and  $\Delta g_n$  gives a shorter expression

$$\begin{aligned}
 \Delta \delta \xi &= (g_n \mathbb{H} - \mathbb{A})^{-1} \left\{ \frac{\partial \rho}{\partial \xi} \cdot \left( \delta \frac{\partial \rho}{\partial \xi}^T \Delta \xi + \Delta \frac{\partial \rho}{\partial \xi} \delta \xi \right) + \Delta \xi^T \left( \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \rho}{\partial \xi^2} \right) \delta \xi - \right. \\
 &\quad - g_n \underline{n} \cdot \left( \delta \frac{\partial^2 \rho}{\partial \xi^2} \Delta \xi + \Delta \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right) - g_n \Delta \xi^T \left( \underline{n} \cdot \frac{\partial^3 \rho}{\partial \xi^3} \right) \delta \xi + \\
 &\quad + \left[ g_n \left( \delta \frac{\partial \rho}{\partial \xi} + \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right) \cdot \frac{\partial \rho}{\partial \xi}^T \bar{\mathbb{A}} - \delta g_n \mathbb{I} \right] \left( \underline{n} \cdot \Delta \frac{\partial \rho}{\partial \xi} + \mathbb{H} \Delta \xi \right) + \\
 &\quad \left. + \left[ g_n \left( \Delta \frac{\partial \rho}{\partial \xi} + \frac{\partial^2 \rho}{\partial \xi^2} \Delta \xi \right) \cdot \frac{\partial \rho}{\partial \xi}^T \bar{\mathbb{A}} - \Delta g_n \mathbb{I} \right] \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} + \mathbb{H} \delta \xi \right) \right\}
 \end{aligned} \tag{2.78}$$

## 2.2 Interaction between contacting surfaces

This expression coincides with one originally obtained in [Laursen 92]. Remark that this expression contains a third derivative of the master surface vector  $\underline{\rho}$ , which imposes a more strict conditions on the smoothness of the master. However, the approximation of this variation for small gaps  $g_n \approx 0$ , leading to a simpler form, does not contain this derivative

$$\begin{aligned} \Delta \delta \underline{\xi} = & \bar{\underline{A}} \left\{ \delta g_n \left( \underline{n} \cdot \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} + \underline{H} \Delta \underline{\xi} \right) + \Delta g_n \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} + \underline{H} \delta \underline{\xi} \right) - \right. \\ & \left. - \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \underline{\xi} + \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta \underline{\xi} \right) - \Delta \underline{\xi}^T \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \right) \delta \underline{\xi} \right\} \end{aligned} \quad (2.79)$$

### Shadow projection case: infinitely remote emitter

It is much easier to compute the second order variations in case of the shadow projection. Variation of Eq. (2.33) gives

$$0 = \bar{\Delta} \delta \underline{\rho} + \Delta \delta g_s \underline{e},$$

expanding the full variations yields

$$\Delta \delta g_s \underline{e} = -\delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \underline{\xi} - \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta \underline{\xi} - \Delta \underline{\xi}^T \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \delta \underline{\xi} - \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \delta \underline{\xi} \quad (2.80)$$

the dot product with vector  $\underline{e}$  gives

$$\Delta \delta g_s = -\underline{e} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \underline{\xi} + \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta \underline{\xi} + \Delta \underline{\xi}^T \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \delta \underline{\xi} \right) - \underline{e} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \delta \underline{\xi}, \quad (2.81)$$

where the last term can be determined by the dot product of (2.80) with v-vector  $\underline{s}$

$$\left( \underline{s} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \Delta \delta \underline{\xi} = -\underline{s} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \underline{\xi} - \underline{s} \cdot \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta \underline{\xi} - \Delta \underline{\xi}^T \underline{s} \cdot \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \delta \underline{\xi}$$

using notation  $\underline{P}$  from (2.36) one has the expression for the second variation of the surface parameter

$$\Delta \delta \underline{\xi} = -\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \left[ \underline{I} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \underline{\xi} + \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta \underline{\xi} + \Delta \underline{\xi}^T \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \delta \underline{\xi} \right). \quad (2.82)$$

where the operator in square brackets is discussed in Remark 2.6. Substituting (2.82) in (2.81) gives

$$\Delta \delta g_s = - \left[ \underline{e} - \left( \underline{e} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \left( \underline{I} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right) \right] \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \underline{\xi} + \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \delta \underline{\xi} + \Delta \underline{\xi}^T \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} \delta \underline{\xi} \right), \quad (2.83)$$



where the expression in square brackets is the same as in Eq. (2.44), so the second variation of the shadow gap can be rewritten as

$$\Delta\delta g_s = -\frac{\underline{n}}{\underline{e} \cdot \underline{n}} \cdot \left( \delta \frac{\partial \rho}{\partial \xi} \Delta \xi + \Delta \frac{\partial \rho}{\partial \xi} \delta \xi + \Delta \xi^T \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right). \quad (2.84)$$

### Shadow projection case: close emitter

In this case we take the variation of expression (2.50)

$$0 = \delta \frac{\partial \rho}{\partial \xi} \Delta \xi + \Delta \frac{\partial \rho}{\partial \xi} \delta \xi + \Delta \xi^T \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi + \frac{\partial \rho}{\partial \xi} \Delta \delta \xi + \Delta \delta g_s \underline{e} + \delta g_s \Delta \underline{e} + \Delta g_s \delta \underline{e} + g_s \Delta \delta \underline{e}. \quad (2.85)$$

The dot product with vector  $\underline{e}$  and account of the following equality  $\delta \underline{e} \cdot \underline{e} = 0$  gives

$$\Delta\delta g_s = -\underline{e} \cdot \left( \delta \frac{\partial \rho}{\partial \xi} \Delta \xi + \Delta \frac{\partial \rho}{\partial \xi} \delta \xi + \Delta \xi^T \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right) - \underline{e} \cdot \frac{\partial \rho}{\partial \xi} \Delta \delta \xi - g_s \underline{e} \cdot \Delta \delta \underline{e}, \quad (2.86)$$

where, due to the orthogonality of a unit vector to its variation and according to (2.48) and (2.49) the last term becomes

$$\underline{e} \cdot \Delta \delta \underline{e} = -\delta \underline{e} \cdot \Delta \underline{e} = -\frac{(\delta \underline{r}_s \cdot \underline{s})(\underline{s} \cdot \Delta \underline{r}_s)}{d_{se}^2} \quad (2.87)$$

The dot product of (2.85) with v-vector  $\underline{s}$  allows to evaluate the second variation of the surface parameter

$$\begin{aligned} \Delta\delta \xi = & -\left( \underline{s} \cdot \frac{\partial \rho}{\partial \xi} \right)^{-1} \underline{s} \cdot \left( \delta \frac{\partial \rho}{\partial \xi} \Delta \xi + \Delta \frac{\partial \rho}{\partial \xi} \delta \xi + \Delta \xi^T \frac{\partial^2 \rho}{\partial \xi^2} \delta \xi \right) - \\ & -\left( \underline{s} \cdot \frac{\partial \rho}{\partial \xi} \right)^{-1} \underline{s} \cdot (\delta g_s \Delta \underline{e} + \Delta g_s \delta \underline{e}) - g_s \left( \underline{s} \cdot \frac{\partial \rho}{\partial \xi} \right)^{-1} \underline{s} \cdot \Delta \delta \underline{e}. \end{aligned} \quad (2.88)$$

Let us consider the right part of the last term using (2.48) and (2.49)

$$\begin{aligned} \underline{s} \cdot \Delta \delta \underline{e} = & -\underline{s} \cdot \Delta \left[ (\underline{I} - \underline{e} \otimes \underline{e}) \cdot \frac{\delta \underline{r}_s}{d_{se}} \right] = \\ & -\underline{s} \cdot \left[ (\Delta \underline{e} \otimes \underline{e} + \underline{e} \otimes \Delta \underline{e}) \cdot \frac{\delta \underline{r}_s}{d_{se}} + (\underline{I} - \underline{e} \otimes \underline{e}) \cdot \delta \underline{r}_s \Delta \frac{1}{d_{se}} \right] = \\ & = \frac{1}{d_{se}^2} \underline{s} \cdot (\Delta \underline{r}_s \otimes \underline{e} \cdot \delta \underline{r}_s + \delta \underline{r}_s \otimes \underline{e} \cdot \Delta \underline{r}_s) \end{aligned} \quad (2.89)$$

in a similar manner, it is easy to show that

$$\underline{s} \cdot \delta \underline{e} = -\underline{s} \cdot \frac{\delta \underline{r}_s}{d_{se}} \quad (2.90)$$

## 2.2 Interaction between contacting surfaces

Substituting (2.89) and (2.90) in (2.88) and carrying (2.37) gives

$$\Delta \delta \tilde{\xi} = -\frac{\overline{\partial \rho}}{\partial \tilde{\xi}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] \cdot \left\{ \delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \Delta \tilde{\xi} + \Delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \delta \tilde{\xi} + \Delta \tilde{\xi}^T \frac{\partial^2 \rho}{\partial \tilde{\xi}^2} \delta \tilde{\xi} - \right. \\ \left. - \frac{1}{d_{se}} (\Delta g_s \delta \underline{r}_s + \delta g_s \Delta \underline{r}_s) + \frac{g_s}{d_{se}^2} (\Delta \underline{r}_s \otimes \underline{e} \cdot \delta \underline{r}_s + \delta \underline{r}_s \otimes \underline{e} \cdot \Delta \underline{r}_s) \right\} \quad (2.91)$$

Assuming small gaps  $g_s \approx 0$  results in a shorter expression

$$\Delta \delta \tilde{\xi} = -\frac{\overline{\partial \rho}}{\partial \tilde{\xi}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] \cdot \left\{ \delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \Delta \tilde{\xi} + \Delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \delta \tilde{\xi} + \Delta \tilde{\xi}^T \frac{\partial^2 \rho}{\partial \tilde{\xi}^2} \delta \tilde{\xi} - \right. \\ \left. - \frac{1}{d_{se}} (\Delta g_s \delta \underline{r}_s + \delta g_s \Delta \underline{r}_s) \right\} \quad (2.92)$$

Substituting (2.91) and (2.87) in (2.86) gives

$$\Delta \delta g_s = -\underline{e} \cdot \left( \delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \Delta \tilde{\xi} + \Delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \delta \tilde{\xi} + \Delta \tilde{\xi}^T \frac{\partial^2 \rho}{\partial \tilde{\xi}^2} \delta \tilde{\xi} \right) + \left\{ \underline{e} \cdot \frac{\partial \rho^T}{\partial \tilde{\xi}} \otimes \frac{\overline{\partial \rho}}{\partial \tilde{\xi}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] \right\} \cdot \\ \cdot \left[ \delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \Delta \tilde{\xi} + \Delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \delta \tilde{\xi} + \Delta \tilde{\xi}^T \frac{\partial^2 \rho}{\partial \tilde{\xi}^2} \delta \tilde{\xi} - \frac{1}{d_{se}} (\Delta g_s \delta \underline{r}_s + \delta g_s \Delta \underline{r}_s) + \right. \\ \left. + \frac{g_s}{d_{se}^2} (\Delta \underline{r}_s \otimes \underline{e} \cdot \delta \underline{r}_s + \delta \underline{r}_s \otimes \underline{e} \cdot \Delta \underline{r}_s) \right] + \frac{g_s}{d_{se}^2} (\delta \underline{r}_s \cdot \underline{s}^T) (\underline{s} \cdot \Delta \underline{r}_s), \quad (2.93)$$

where the expression in braces can be rewritten as

$$\underline{e} \cdot \frac{\partial \rho^T}{\partial \tilde{\xi}} \otimes \frac{\overline{\partial \rho}}{\partial \tilde{\xi}} \cdot \left[ \underline{\underline{I}} - \frac{\underline{e} \otimes \underline{n}}{\underline{e} \cdot \underline{n}} \right] = \underline{e} - \frac{\underline{n}}{\underline{n} \cdot \underline{e}} = -\frac{\underline{n}}{\underline{n} \cdot \underline{e}} \cdot \underline{s}^T \otimes \underline{s}$$

consequently (2.93) transforms into

$$\Delta \delta g_s = -\frac{\underline{n}}{\underline{n} \cdot \underline{e}} \cdot \left( \delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \Delta \tilde{\xi} + \Delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \delta \tilde{\xi} + \Delta \tilde{\xi}^T \frac{\partial^2 \rho}{\partial \tilde{\xi}^2} \delta \tilde{\xi} - \frac{\underline{s}^T \otimes \underline{s}}{d_{se}} \cdot (\Delta g_s \delta \underline{r}_s + \delta g_s \Delta \underline{r}_s) + \right. \\ \left. + \frac{g_s}{d_{se}^2} \underline{s}^T \left( \underline{s} \cdot \Delta \underline{r}_s \underline{e} \cdot \delta \underline{r}_s + \underline{s} \cdot \delta \underline{r}_s \otimes \underline{e} \cdot \Delta \underline{r}_s \right) \right) + \frac{g_s}{d_{se}^2} (\delta \underline{r}_s \cdot \underline{s}^T) (\underline{s} \cdot \Delta \underline{r}_s) \quad (2.94)$$

Assuming small gaps  $g_s \approx 0$  results in a shorter expression

$$\Delta \delta g_s = -\frac{\underline{n}}{\underline{n} \cdot \underline{e}} \cdot \left( \delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \Delta \tilde{\xi} + \Delta \frac{\partial \rho^T}{\partial \tilde{\xi}} \delta \tilde{\xi} + \Delta \tilde{\xi}^T \frac{\partial^2 \rho}{\partial \tilde{\xi}^2} \delta \tilde{\xi} - \frac{\underline{s}^T \otimes \underline{s}}{d_{se}} \cdot (\Delta g_s \delta \underline{r}_s + \delta g_s \Delta \underline{r}_s) \right) \quad (2.95)$$

In the case of an infinitely remote emitter ( $d_{se} \rightarrow \infty$ ) we get the same expressions as in the previous section both for surface parameter (2.43) and the shadow gap (2.46).



**Remark 2.6 on 2D case.**

*In order to get all variations for two dimensional geometries, one has to replace simply all quantities by their two dimensional homologues:*

$$\begin{aligned} \underline{\xi} &\rightarrow \xi, & \frac{\partial^n \underline{\rho}}{\partial \underline{\xi}^n} &\rightarrow \frac{\partial^n \rho}{\partial \xi^n}, & \underline{\mathbb{I}} &= 1, \\ \underline{A} &\rightarrow A = \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial \rho}{\partial \xi}, & \underline{\bar{A}} &\rightarrow \frac{1}{A}, & \underline{H} &\rightarrow H = \underline{n} \cdot \frac{\partial^2 \rho}{\partial \xi^2}, & \overline{\frac{\partial \rho}{\partial \xi}} &\rightarrow \frac{1}{A} \frac{\partial \rho}{\partial \xi} \end{aligned}$$

## 2.3 Numerical validation

The aim of this section is to carry out a numerical validation of the expressions obtained in the previous section. This validation allows us to estimate the errors that we introduce in the incremental solution by approximating the first and the second variations by analytical expressions obtained within the assumption of infinitely small perturbations of the contact geometry. The analytical expressions obtained with and without the assumption of small gaps ( $g_n \approx 0$ ) are compared.

The numerical validation technique can be summarized as follows: for a randomly generated bi-quadratic master surface in 3D  $\underline{\rho}(\underline{\xi})$  and a slave point at a random position  $\underline{r}$ , perturbations  $\underline{\pi}'(\underline{\xi})$  and  $\underline{p}'$  are applied respectively:

$$\underline{\rho} \rightarrow \underline{\rho} + \underline{\pi}', \quad \underline{r} \rightarrow \underline{r} + \underline{p}'.$$

The direction of the perturbation is arbitrary and its value is given in fraction of the master segment dimension (maximal length)

$$\|\underline{\pi}'(\underline{\xi})\| = \varepsilon_p \dim(\underline{\rho}), \quad \|\underline{p}'\| = \varepsilon_p \dim(\underline{\rho}).$$

The initial normal gap  $g_n$  between the slave node and its projection is prescribed, its value is also expressed in fraction of the master segment dimension

$$g_n = \varepsilon_g \dim(\underline{\rho}).$$

Applied perturbations  $\underline{\pi}$  and  $\underline{p}$  result in change of all geometrical quantities:

$$g_n \rightarrow g_n'; \quad \underline{\xi} \rightarrow \underline{\xi}'.$$

### 2.3 Numerical validation

---

The real variation of these quantities are evaluated as the difference between the perturbed state quantities and the initial ones

$$\delta g_n = g'_n - g_n, \quad \delta \xi = \xi' - \xi$$

At the same time, these variations can be estimated according to analytical formulae (2.24) and (2.31) as functions of the initial geometry and its perturbation

$$\delta^a g_n = \delta^a g_n(\underline{\rho}, \underline{r}, \underline{\pi}', \underline{p}'), \quad \delta^a \xi = \delta^a \xi(\underline{\rho}, \underline{r}, \underline{\pi}', \underline{p}')$$

the reduced expression for variations, i.e. without terms containing  $g_n$  will be noted with the  $r$  index  $\delta^r \xi$ ,  $\delta^r \xi = \delta^a \xi$ . Note that initial gap  $g_n$  remains nonzero, it simply does not enter in the expressions for variations. To evaluate the second variations, another perturbation is imposed to the initial geometry  $\underline{\rho}, \underline{r}$ . The value of this perturbation is the same  $\varepsilon_p$ , but direction is different

$$\underline{\rho} \rightarrow \underline{\rho}'', \quad \underline{r} \rightarrow \underline{r}''.$$

The corresponding real change in geometrical quantities is denoted with  $\Delta$

$$\Delta g_n = g''_n - g_n, \quad \Delta \xi = \xi'' - \xi.$$

To evaluate the second variations, the geometry should be perturbed once again. The perturbation is fully determined by two previous perturbations

$$\underline{\rho} \rightarrow \underline{\rho} + \underline{\pi}' + \underline{\pi}'', \quad \underline{r} \rightarrow \underline{r} + \underline{p}' + \underline{p}''.$$

The corresponding change in geometrical quantities will be

$$D g_n = g'''_n - g_n, \quad D \xi = \xi''' - \xi$$

From these three perturbed states, the real second variations of geometrical quantities can be deduced as

$$\Delta \delta g_n = D g_n - \delta g_n - \Delta g_n, \quad \Delta \delta \xi = D \xi - \delta \xi - \Delta \xi.$$

These variations can be estimated by full analytical expressions

$$\Delta \delta^a g_n = \Delta \delta^a g_n(\underline{\rho}, \underline{r}, \underline{\pi}', \underline{p}', \underline{\pi}'', \underline{p}''), \quad \Delta \delta^a \xi = \Delta \delta^a \xi(\underline{\rho}, \underline{r}, \underline{\pi}', \underline{p}', \underline{\pi}'', \underline{p}'').$$

and approximated expressions  $\Delta \delta^r g_n$ ,  $\Delta \delta^r \xi$  which do not contain terms with the normal gap. It remains to compare the full and truncated analytical estimations with the real variations computed for a given negative initial normal gap  $g_n = -\varepsilon_g \dim(\underline{\rho})$  and the given values of the perturbation  $\|\underline{\pi}'\| = \varepsilon_p \dim(\underline{\rho})$ ,  $\|\underline{p}'\| = \varepsilon_p \dim(\underline{\rho})$ . The aim is to demonstrate that when the perturbation of the geometry  $\varepsilon_p$  tends to zero, the analytical expression for the variations tend to the real variations evaluated numerically. Another aim is to show the difference between the truncated expressions (without terms

containing  $g_n$ ) and the accurate ones derived in the previous section. For each value of the initial gap  $g_n$ , to get statistically meaningful data the variations of geometrical quantities have been evaluated 50 000 times.

The relative error between the real first variation of the normal gap  $\delta g_n$  and its analytical estimation is depicted in Fig. 2.12 for different perturbation values  $\varepsilon_p = 10\%, 1\%, 0.1\%, 0.01\%, 0.001\%, 0.0001\%, 0.00001\%$ . Each plot corresponds to a given value of the initial normal gap  $\varepsilon_g = 0.01\%, 0.1\%, 1\%$ .

$$\text{Er}(g_n) = \left| \frac{\delta g_n - \delta^a g_n}{\delta g_n} \right| 100\%.$$

Here and further the left plots depict the general trend and the scatter of the data, whereas the right figures depict the probability density function (not normalized and shifted in vertical direction order to distinguish points corresponding to different perturbations), i.e. number of occurrences for the given precision of the expression. The data distribution is centered at zero, that is why the absolute value is analyzed.

In Fig. 2.13 the relative error between the variation of the surface parameter  $\delta \xi$  and its analytical estimation  $\delta^a \xi$  are presented

$$\text{Er}(\xi) = \frac{\|\delta \xi - \delta^a \xi\|}{\|\delta \xi\|} 100\%$$

as well as the relative error of the analytical estimation obtained with approximation of the zero gap  $g_n = 0$  denoted as “truncated”. As previously, different perturbation values and initial normal gaps are used.

The relative error for the second variation of the normal gap  $\Delta \delta g_n$  is presented in Fig. 2.14

$$\text{Er}(g_n) = \left| \frac{\Delta \delta g_n - \Delta \delta^a g_n}{\Delta \delta g_n} \right| 100\%.$$

and for the second variation of the surface parameter  $\Delta \delta \xi$  is presented in Fig. 2.15

$$\text{Er}(\xi) = \frac{\|\Delta \delta \xi - \Delta \delta^a \xi\|}{\|\Delta \delta \xi\|} 100\%.$$

The analysis of plots 2.12–2.15 shows that the relative error for accurate expressions tends to zero with decreasing perturbation, whereas the error for truncated expressions does not. Note that for second order variations the inverse tendency is observed (increase of error with decreasing perturbation) for very small perturbations, this fact is explained by reaching the limits of the numerical precision.

The obtained results demonstrate that the scatter of the relative error (plots on the left hand) is very large, it is due to possible close to zero values of variations. It is worth noting that the relative error of the first and the second variation of the normal gap have a peak at a certain value

and decay exponentially to zero for higher relative errors. The value of the peak error for the accurate analytical expression decreases with decreasing perturbation, which is not the case for the truncated one. The relative error of the surface parameter variations have a quite large “plateau” rapidly decaying for larger values of errors. This plateau approaches zero error with decreasing perturbation in case of accurate estimation of variations and remains within the same error range for truncated expression. The figures also clearly demonstrate that the accurate results are not affected by the value of the normal gap and always converge to the real value for decreasing perturbation, whereas the truncated expressions are highly affected by the value of the normal gap. Even for relatively small value of the gap  $g_n = 1\%$ , the average relative error of the first and second variation of the surface parameter remains at values of 10% and 100% respectively independently on the value of perturbation. The numerical validation, which we use, may be helpful for verification of any complexity expressions, for example, in case of three-dimensional smoothing of the master surface.

## 2.4 Discretized geometry

### 2.4.1 Shape functions and finite elements

The finite element method approximates the real geometry by a so-called *finite element mesh*. Such a mesh consists of *nodes*, which are connected to form *elements*, which are coupled together into a structure. So we deal with coordinates of nodes  $\mathbf{x}$  and their displacements. What about the description of material points which are situated somewhere inside the elements? Their positions are determined by an interpolation of the nodal positions by means of so-called *shape functions*

$$\phi_i(\xi), \quad i = 1, N$$

where  $N$  is the number on nodes which determine the geometry of the given element and  $\xi$  is a vector (v-scalar) of normalized parameters ( $\xi_j \in [-1; 1]$ ), which determine the coordinate in the “parent space” of internal points  $\mathbf{r}$

$$\mathbf{r} = \sum_{i=1}^N \phi_i(\xi) \mathbf{x}_i, \quad (2.96)$$

where  $\mathbf{x}_i$  are nodal coordinates. Shape functions must verify the following equality

$$\phi_i(\xi_j) = \delta_{ij}.$$

The coordinate of the  $j$ -th node in the local frame  $\xi_j$  must verify

Structural finite elements can be classified by the order of shape functions, which depends on the number of nodes which form the element:

- *linear elements* - any internal point is approximated by means of the nodal coordinates and linear shape functions (shape function depends linearly

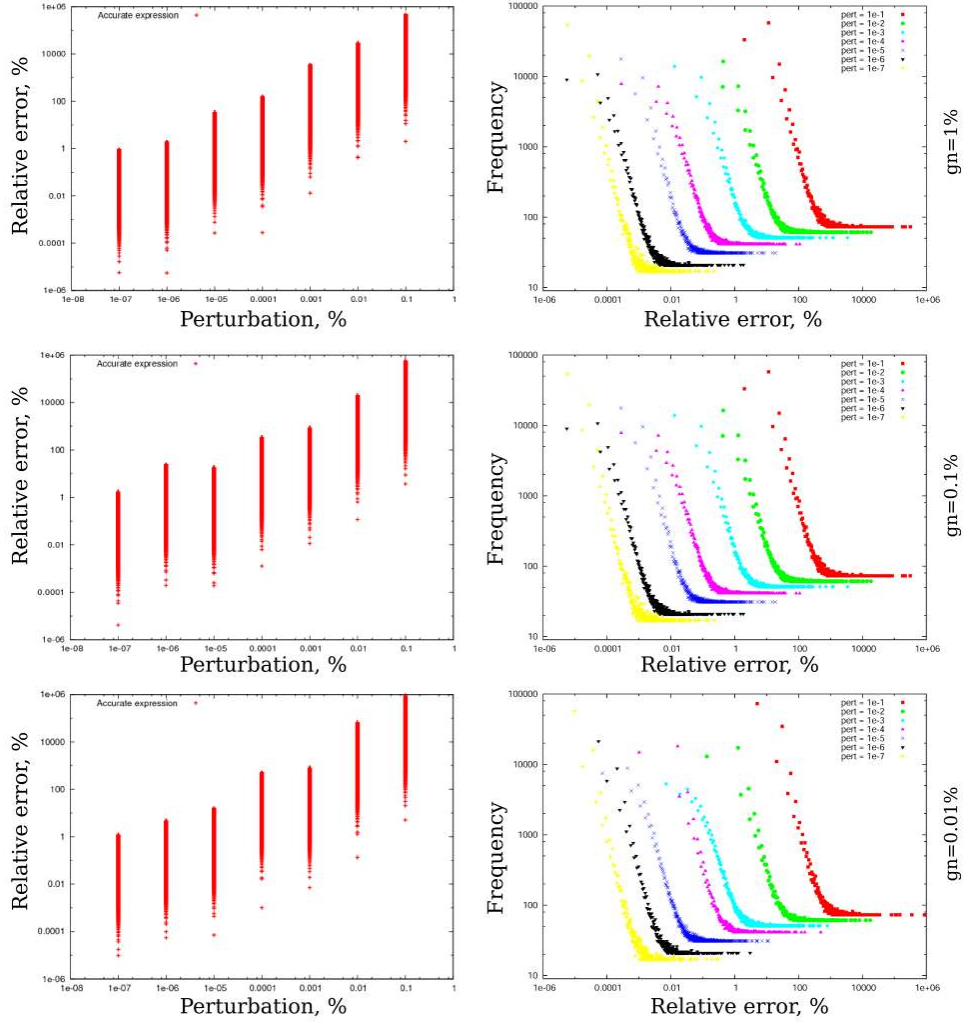


Figure 2.12: Comparison of the real variation of the normal gap and its analytical estimation for different perturbations and initial values of the gap  $g_n$  given in percent of the master segment size.

on each parameter  $\xi_i$ ), i.e.

$$\frac{\partial \phi(\xi)}{\partial \xi_i} \neq 0, \quad \frac{\partial^2 \phi(\xi)}{\partial \xi_i^2} = 0;$$

- *quadratic elements* - any internal point is approximated by nodal coordinates and quadratic shape functions (shape function has second order dependence on each parameter  $\xi_i$ ), i.e.

$$\frac{\partial^2 \phi(\xi)}{\partial \xi_i^2} \neq 0, \quad \frac{\partial^3 \phi(\xi)}{\partial \xi_i^3} = 0;$$

## 2.4 Discretized geometry

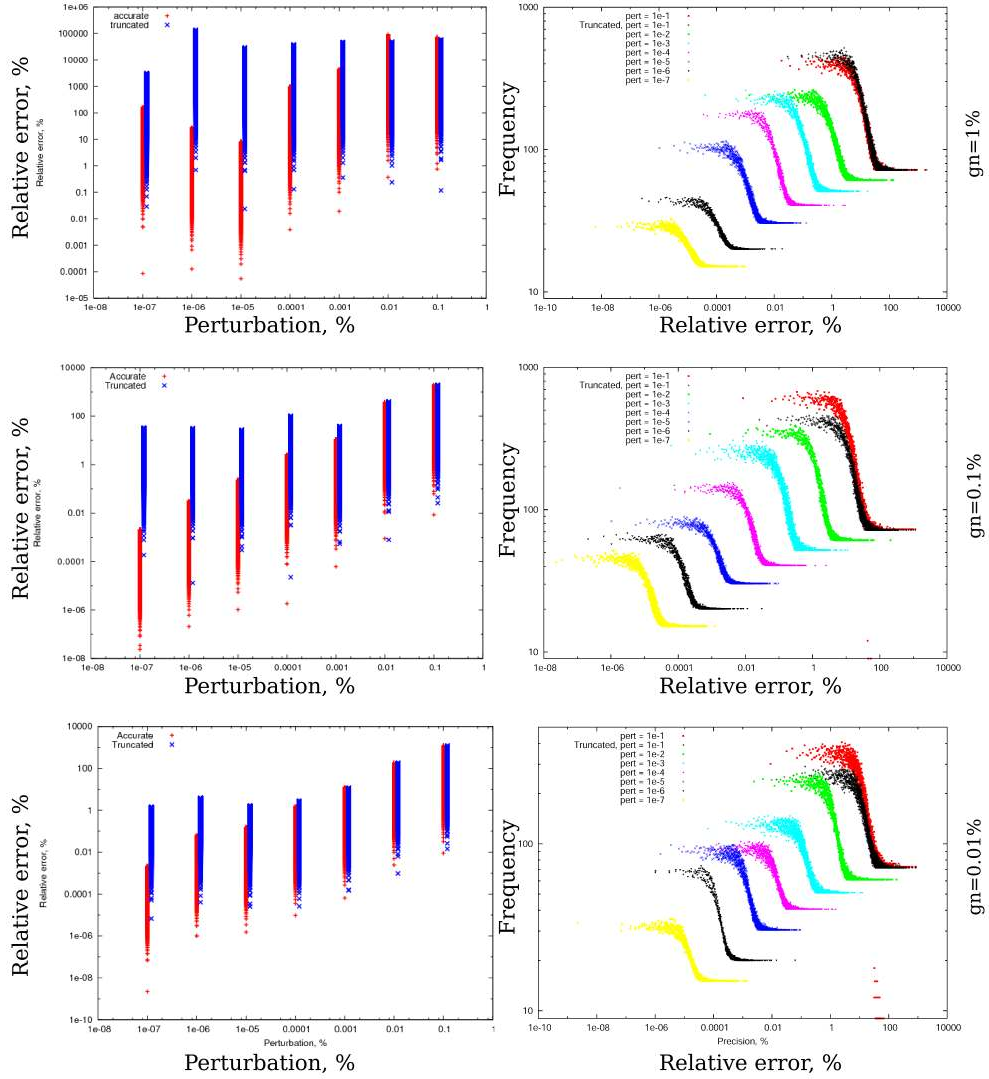


Figure 2.13: Comparison of the real variation of the surface parameter and its analytical estimations (full and truncated) for different perturbations and initial values of the gap  $g_n$  given in percent of the master segment size.

- *higher order elements* will not be considered here.

1D, 2D and 3D elements are available, they may contain different number of nodes, for example, 2D triangle linear element, 3D tetrahedron linear element, 3D prismatic quadratic element, etc. In the following subsections, we will consider an abstract 3D element, which contacting surface is formed by  $N$  nodes with the corresponding order of approximation. A more detailed presentation of the Finite Element formalism will be given in Section 5.3.



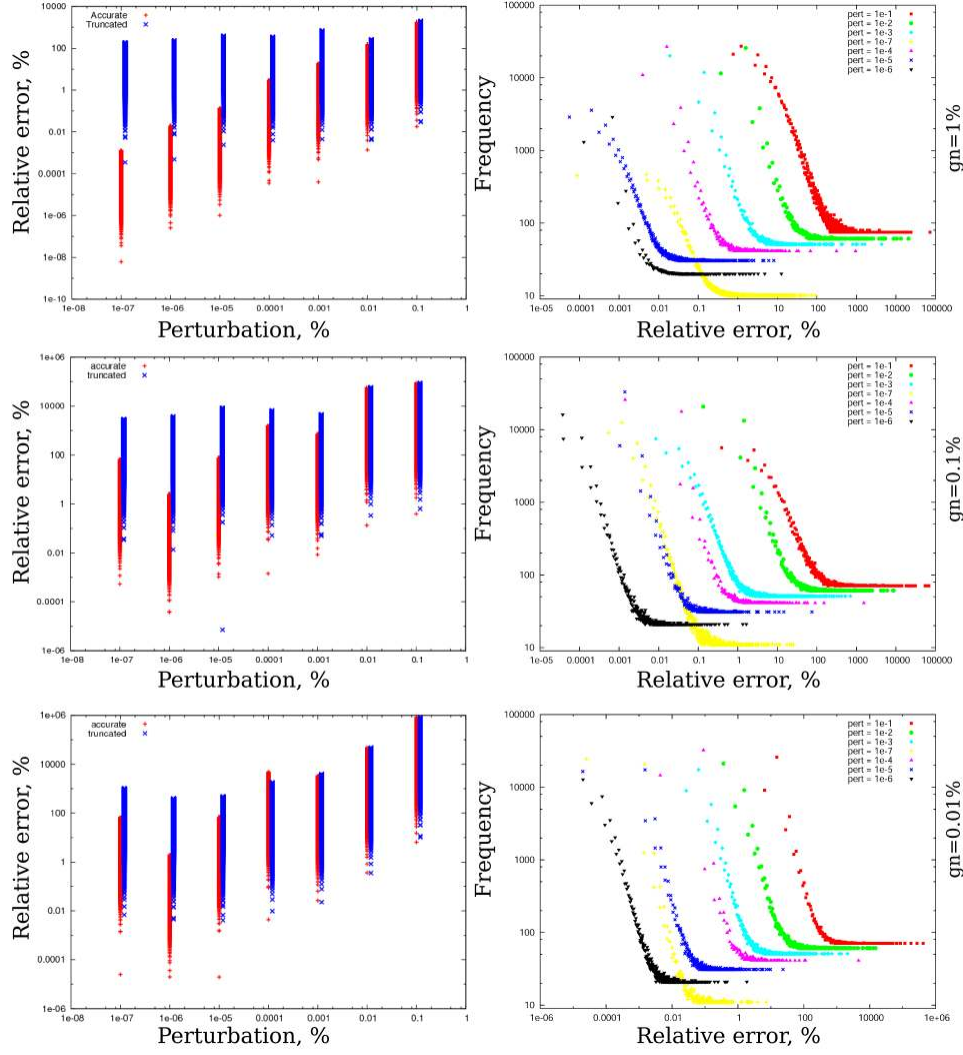


Figure 2.14: Comparison of the real second variation of the normal gap and its analytical estimations (full and truncated) for different perturbations and initial values of the gap  $g_n$  given in percent of the master segment size.

### 2.4.2 Geometry of contact elements

To describe the contact geometry, let us create an abstract element which consists of one slave node  $\underline{x}_0(t)$  and several master nodes attached to one structural element  $\underline{x}_i(t)$ ,  $i = 1, \dots, N$  with a surface spanned on them. The latter is determined by position of nodes and the shape functions of the element. Such an element with some special properties will be called *contact element*. As in the previous sections the master surface is parametrized by  $\underline{\xi}$ <sup>4</sup>. Vector of

<sup>4</sup>the dimension of the surface is equal to the space dimension minus one, master is a curve in 2D and is a surface in 3D case.

## 2.4 Discretized geometry

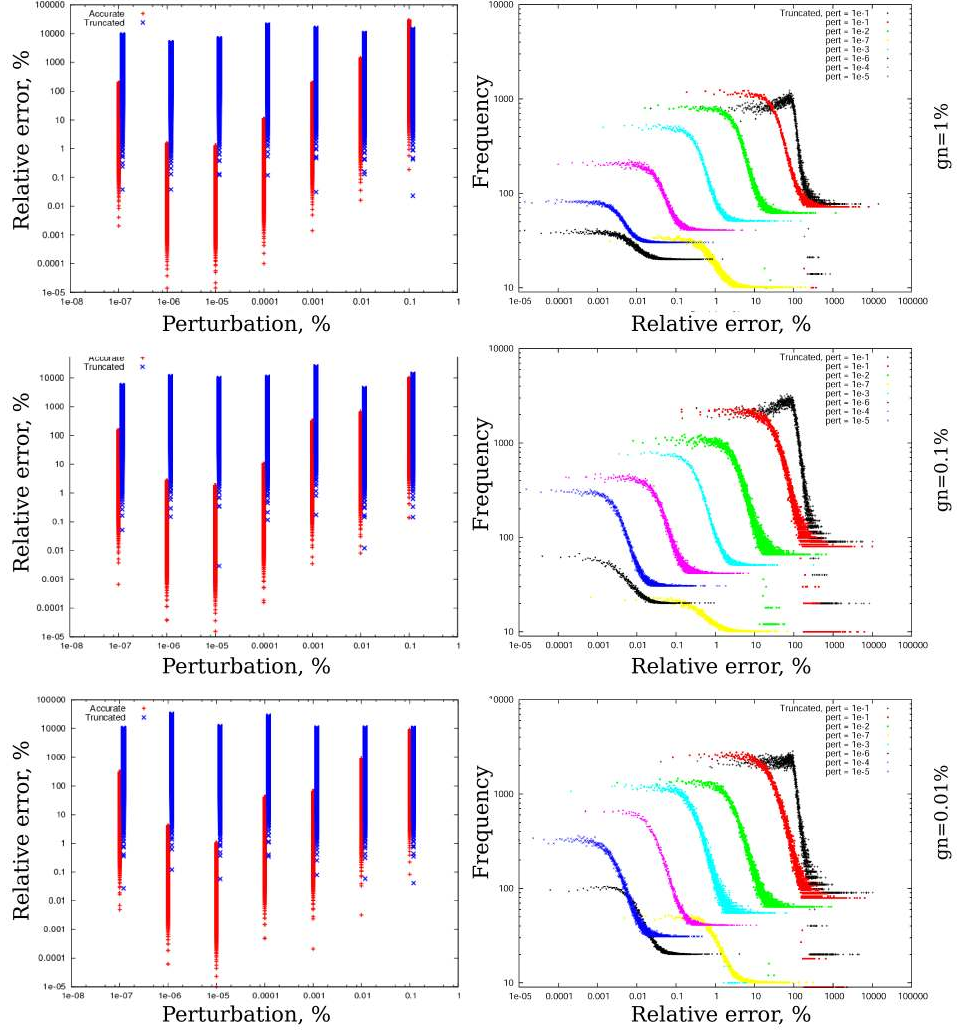


Figure 2.15: Comparison of the real second variation of the surface parameter and its analytical estimations (full and truncated) for different perturbations and initial values of the gap  $g_n$  given in percent of the master segment size.

every point on this surface can be expressed through the nodal positions and shape functions

$$\underline{\rho} = \sum_{i=1}^N \phi_i(\underline{\xi}) \underline{x}_i = \sum_{i=1}^N \phi_i(\xi_1, \xi_2) \underline{x}_i. \quad (2.97)$$

The last equality is valid for 3D geometries, where the master surface can be parametrized by two parameters  $\xi_1$  and  $\xi_2$ , which, as previously, we assemble in the v-scalar  $\underline{\xi} = [\xi_1, \xi_2]^T$ .

Let us derive all the quantities we need to express the first and second order variations in the frame of the contact element. First, we introduce some notations.

- $\left[\underline{\mathbf{X}}\right]_{\text{dim} \times N+1}$  – is an array (v-vector in s-structure notations) containing nodal coordinate vectors of the contact element

$$\underline{\mathbf{X}} \sim \left[\underline{\mathbf{X}}\right] = \left[\underline{\mathbf{X}}(t)\right] = \left[\underline{\mathbf{x}}_0(t), \underline{\mathbf{x}}_1(t), \dots, \underline{\mathbf{x}}_N(t)\right]^T;$$

- $[\Phi]_{1 \times N+1}$  – is an array (v-scalar in s-structure notations) containing zero in the first position and shape functions of the master surface

$$\underline{\phi} \sim [\Phi] = [\Phi(\underline{\xi})] = [0, \phi_1(\underline{\xi}), \dots, \phi_N(\underline{\xi})]^T;$$

- $[\Phi'_i]_{1 \times N+1}$  – is an array (v-scalar in s-structure notations) of derivatives of the shape functions by the surface parameter  $\xi_i$  with zero in the first position

$$\underline{\phi}'_i \sim [\Phi'_i] = \left[\frac{\partial \Phi(\underline{\xi})}{\partial \xi_i}\right] = \left[0, \frac{\partial \phi_1(\underline{\xi})}{\partial \xi_i}, \dots, \frac{\partial \phi_N(\underline{\xi})}{\partial \xi_i}\right]^T = [0, \phi_{1,i}, \dots, \phi_{N,i}]^T;$$

- And so on for the higher order derivations

$$\underline{\phi}''_{ij} \sim [\Phi''_{ij}] = \left[\frac{\partial^2 \Phi(\underline{\xi})}{\partial \xi_i \partial \xi_j}\right] = [0, \phi_{1,ij}, \dots, \phi_{N,ij}]^T,$$

$$\underline{\phi}'''_{ijk} \sim [\Phi'''_{ijk}] = \left[\frac{\partial^3 \Phi(\underline{\xi})}{\partial \xi_i \partial \xi_j \partial \xi_k}\right] = [0, \phi_{1,ijk}, \dots, \phi_{N,ijk}]^T.$$

Now we have to express all the kinematic quantities and their variations from Section 2.2 in a form adapted for the finite element formalism.

- Slave node coordinate vector

$$\boxed{\underline{\mathbf{r}}_s} = \underline{\mathbf{r}}_s(t) = \underline{\mathbf{x}}_0(t) = [S_0]^T [\underline{\mathbf{X}}],$$

where  $[S_0]_{N+1}$  is nothing but a selection vector for the slave node component

$$[S_0] = [1, 0, \dots, 0]^T.$$

- Projection point on the master surface

$$\boxed{\underline{\rho}} = \underline{\rho}(t, \underline{\xi}_\pi) = \sum_{i=1}^N \phi_i(\underline{\xi}_\pi) \underline{\mathbf{x}}_i = [\Phi(\underline{\xi}_\pi)]^T [\underline{\mathbf{X}}(t)] = \boxed{[\Phi]^T [\underline{\mathbf{X}}]} = \underline{\phi}^T \underline{\mathbf{X}}$$

where  $\pi$  in  $\underline{\xi}_\pi$  denotes that the surface parameter corresponds to the projection point;

- Local surface basis at the projection point

$$\left[ \frac{\partial \underline{\rho}}{\partial \xi_i} \right] = \left. \frac{\partial \underline{\rho}(t, \underline{\xi})}{\partial \xi_i} \right|_{\underline{\xi}_\pi} = \left[ \left. \frac{\partial \Phi(\underline{\xi})}{\partial \xi_i} \right|_{\underline{\xi}_\pi} \right]^T [\underline{\mathbf{X}}(t)] = [\Phi'_i]^T [\underline{\mathbf{X}}] = \phi_i'^T \underline{\mathbf{X}};$$

- Unit normal to the master surface at the projection point

$$\underline{\mathbf{n}} = \underline{\mathbf{n}}(t, \underline{\xi}_\pi) = \frac{\frac{\partial \underline{\rho}(t, \xi_\pi)}{\partial \xi_1} \times \frac{\partial \underline{\rho}(t, \xi_\pi)}{\partial \xi_2}}{\left\| \frac{\partial \underline{\rho}(t, \xi_\pi)}{\partial \xi_1} \times \frac{\partial \underline{\rho}(t, \xi_\pi)}{\partial \xi_2} \right\|} = \frac{\left( [\Phi'_1]^T [\underline{\mathbf{X}}] \right) \times \left( [\Phi'_2]^T [\underline{\mathbf{X}}] \right)}{\left\| \left( [\Phi'_1]^T [\underline{\mathbf{X}}] \right) \times \left( [\Phi'_2]^T [\underline{\mathbf{X}}] \right) \right\|},$$

further the  $\underline{\mathbf{n}}$  notation will be used.

Now the variations of the kinematic quantities can be derived in a finite element format. Let us start with the **first order variation of the normal gap**  $\delta g_n$ . From (2.24) using the introduced notation we get

$$\begin{aligned} \delta g_n &= \underline{\mathbf{n}} \cdot (\delta \underline{\mathbf{r}}_s - \delta \underline{\rho}) \Rightarrow \\ \delta g_n &= \underline{\mathbf{n}} \cdot (\delta [\underline{\mathbf{X}}]_0 - [\Phi]^T \delta [\underline{\mathbf{X}}]) = ([S_0]^T - [\Phi]^T) \underline{\mathbf{n}} \cdot \delta [\underline{\mathbf{X}}] \end{aligned} \quad (2.98)$$

or in a component form

$$\delta g_n = \begin{bmatrix} \underline{\mathbf{n}} \\ -\phi_1 \underline{\mathbf{n}} \\ \vdots \\ -\phi_N \underline{\mathbf{n}} \end{bmatrix}^T \cdot \begin{bmatrix} \delta \underline{\mathbf{x}}_0 \\ \delta \underline{\mathbf{x}}_1 \\ \vdots \\ \delta \underline{\mathbf{x}}_N \end{bmatrix} = [\nabla g_n]^T \cdot \delta [\underline{\mathbf{X}}]. \quad (2.99)$$

The **first order variation of the surface parameter**  $\delta \underline{\xi}$  can be expressed from (2.31)

$$\delta \xi_i = (\underline{\mathbf{A}} - g_n \underline{\mathbf{H}})_{ij}^{-1} \left( \frac{\partial \underline{\rho}}{\partial \xi_j} \cdot (\delta \underline{\mathbf{r}}_s - \delta \underline{\rho}) + g_n \delta \frac{\partial \underline{\rho}}{\partial \xi_j} \cdot \underline{\mathbf{n}} \right) \quad (2.100)$$

If we denote  $\underline{\mathbf{C}} = (\underline{\mathbf{A}} - g_n \underline{\mathbf{H}})^{-1}$  we get

$$\delta \xi_i = C_{ij} \left( ([S_0]^T - [\Phi]^T) \frac{\partial \underline{\rho}}{\partial \xi_j} + g_n [\Phi'_j]^T \underline{\mathbf{n}} \right) \cdot \delta [\underline{\mathbf{X}}] \quad (2.101)$$

or in component form

$$\delta \xi_i = C_{ij} \begin{bmatrix} \frac{\partial \underline{\rho}}{\partial \xi_j} \\ -\frac{\partial \underline{\rho}}{\partial \xi_j} \phi_1 + g_n \underline{\mathbf{n}} \phi_{1,j} \\ \vdots \\ -\frac{\partial \underline{\rho}}{\partial \xi_j} \phi_N + g_n \underline{\mathbf{n}} \phi_{N,j} \end{bmatrix}^T \cdot \begin{bmatrix} \delta \underline{\mathbf{x}}_0 \\ \delta \underline{\mathbf{x}}_1 \\ \vdots \\ \delta \underline{\mathbf{x}}_N \end{bmatrix} = [\nabla \xi_i]^T \cdot \delta [\underline{\mathbf{X}}], \quad (2.102)$$

where the basis vectors  $\frac{\partial \rho}{\partial \xi_j}$  may be either inserted explicitly in this expression or expressed through the shape functions

$$\frac{\partial \rho}{\partial \xi_i} = [\Phi'_i]^T [\underline{\mathbf{X}}].$$

The **second order variation of the normal gap**  $\Delta \delta g_n$  can be easily expressed by rewriting (2.67)

$$\begin{aligned} \Delta \delta g_n = & \underbrace{-\underline{\mathbf{n}} \cdot \delta \frac{\partial \rho}{\partial \xi_i} \Delta \xi_i}_{\text{term 1}} - \underbrace{\underline{\mathbf{n}} \cdot \Delta \frac{\partial \rho}{\partial \xi_i} \delta \xi_i - \delta \xi_i (H_{ij} - g_n H_{ik} A^{km} H_{mj}) \Delta \xi_j}_{\text{term 2}} + \\ & + \underbrace{g_n \left( \underline{\mathbf{n}} \cdot \delta \frac{\partial \rho}{\partial \xi_i} \right) A^{ij} \left( \Delta \frac{\partial \rho}{\partial \xi_j} \cdot \underline{\mathbf{n}} \right)}_{\text{term 3}} \end{aligned} \quad (2.103)$$

To express  $\Delta \xi$  and  $\delta \xi$  we use (2.102). The other terms are replaced by their discretized homologues

$$\begin{aligned} \Delta \delta g_n = & \underbrace{-\delta [\underline{\mathbf{X}}]^T \cdot \left\{ \underline{\mathbf{n}} [\Phi'_i] \otimes [\nabla \xi_i]^T + [\nabla \xi_i] \otimes [\Phi'_i]^T \underline{\mathbf{n}} \right\} \cdot \Delta [\underline{\mathbf{X}}]}_{\text{term 1}} \\ & \underbrace{-\delta [\underline{\mathbf{X}}]^T \cdot \left\{ (H_{ij} - g_n H_{ik} A^{km} H_{mj}) [\nabla \xi_i] \otimes [\nabla \xi_j]^T \right\} \cdot \Delta [\underline{\mathbf{X}}]}_{\text{term 2}} \\ & + \underbrace{\delta [\underline{\mathbf{X}}]^T \cdot \left\{ g_n A^{ij} \underline{\mathbf{n}} [\Phi'_i] \otimes [\Phi'_j]^T \underline{\mathbf{n}} \right\} \cdot \Delta [\underline{\mathbf{X}}]}_{\text{term 3}}, \end{aligned} \quad (2.104)$$

Grouping the terms leads to the final expression for the second variation of the normal gap

$$\begin{aligned} \Delta \delta g_n = & \delta [\underline{\mathbf{X}}]^T \cdot \left\{ -\underline{\mathbf{n}} [\Phi'_i] \otimes [\nabla \xi_i]^T - [\nabla \xi_i] \otimes [\Phi'_i]^T \underline{\mathbf{n}} \right. \\ & \left. - (H_{ij} - g_n H_{ik} A^{km} H_{mj}) [\nabla \xi_i] \otimes [\nabla \xi_j]^T + g_n A^{ij} \underline{\mathbf{n}} [\Phi'_i] \otimes [\Phi'_j]^T \underline{\mathbf{n}} \right\} \cdot \Delta [\underline{\mathbf{X}}] = \\ & = \delta [\underline{\mathbf{X}}]^T \cdot [\nabla \nabla g_n] \cdot \Delta [\underline{\mathbf{X}}] \end{aligned} \quad (2.105)$$

where  $[\nabla \xi_i]$  should be substituted according to Eq. 2.102.

The second order variation of the surface parameter  $\Delta \delta \tilde{\xi}$  is derived from its continuous form (2.78).

$$\begin{aligned}
 \Delta \delta \xi_i = & -C_{ij} \left[ \underbrace{\frac{\partial \rho}{\partial \xi_j} \cdot \left( \delta \frac{\partial \rho}{\partial \xi_k} \Delta \xi_k + \Delta \frac{\partial \rho}{\partial \xi_k} \delta \xi_k + \delta \xi_k \frac{\partial^2 \rho}{\partial \xi_k \partial \xi_m} \Delta \xi_m \right)}_{\text{term 1}} \right. \\
 & - \underbrace{g_n \underline{n} \cdot \left( \delta \frac{\partial^2 \rho}{\partial \xi_j \partial \xi_k} \Delta \xi_k + \Delta \frac{\partial^2 \rho}{\partial \xi_j \partial \xi_k} \delta \xi_k + \delta \xi_k \frac{\partial^3 \rho}{\partial \xi_k \partial \xi_j \partial \xi_m} \Delta \xi_m \right)}_{\text{term 2}} \\
 & + \underbrace{\left( -\delta g_n \delta_{kj} + g_n A^{km} \frac{\partial \rho}{\partial \xi_m} \cdot \left\{ \delta \frac{\partial \rho}{\partial \xi_j} + \frac{\partial^2 \rho}{\partial \xi_j \partial \xi_l} \delta \xi_l \right\} \right) \left\{ \Delta \frac{\partial \rho}{\partial \xi_k} \cdot \underline{n} + H_{ks} \Delta \xi_s \right\}}_{\text{term 3}} \\
 & \left. + \underbrace{\left( -\Delta g_n \delta_{kj} + g_n A^{km} \frac{\partial \rho}{\partial \xi_m} \cdot \left\{ \Delta \frac{\partial \rho}{\partial \xi_j} + \frac{\partial^2 \rho}{\partial \xi_j \partial \xi_l} \Delta \xi_l \right\} \right) \left\{ \delta \frac{\partial \rho}{\partial \xi_k} \cdot \underline{n} + H_{ks} \delta \xi_s \right\}}_{\text{term 4}} \right]. \tag{2.106}
 \end{aligned}$$

We recall that  $C_{ij} = (\underline{\underline{A}} - g_n \underline{\underline{H}})_{ij}^{-1}$ . Now we replace all the continuous quantities by their discretized analogues term by term. **Term 1.**

$$\begin{aligned}
 & \frac{\partial \rho}{\partial \xi_j} \cdot \left\{ \delta [\underline{\underline{X}}]^T [\Phi'_k] [\mathbf{V}^{\xi_k}]^T \cdot \Delta [\underline{\underline{X}}] + \left( \delta [\underline{\underline{X}}]^T \cdot [\mathbf{V}^{\xi_k}] \right) [\Phi'_k]^T \Delta [\underline{\underline{X}}] + \right. \\
 & \quad \left. + \frac{\partial^2 \rho}{\partial \xi_k \partial \xi_m} \left( \delta [\underline{\underline{X}}]^T \cdot [\mathbf{V}^{\xi_k}] \right) ([\mathbf{V}^{\xi_m}]^T \cdot \Delta [\underline{\underline{X}}]) \right\} = \\
 & = \delta [\underline{\underline{X}}]^T \cdot \left\{ [\Phi'_k] \frac{\partial \rho}{\partial \xi_j} \otimes [\mathbf{V}^{\xi_k}]^T + [\mathbf{V}^{\xi_k}] \otimes \frac{\partial \rho}{\partial \xi_j} [\Phi'_k]^T + \left( \frac{\partial \rho}{\partial \xi_j} \cdot \frac{\partial^2 \rho}{\partial \xi_k \partial \xi_m} \right) [\mathbf{V}^{\xi_k}] \otimes [\mathbf{V}^{\xi_m}]^T \right\} \cdot \Delta [\underline{\underline{X}}] \tag{2.107}
 \end{aligned}$$

**Term 2.**

$$\begin{aligned}
 & g_n \left\{ \delta [\underline{\underline{X}}]^T \cdot \underline{n} [\Phi''_{jk}] [\mathbf{V}^{\xi_k}]^T \cdot \Delta [\underline{\underline{X}}] + \delta [\underline{\underline{X}}]^T \cdot [\mathbf{V}^{\xi_k}] [\Phi''_{jk}]^T \underline{n} \cdot \Delta [\underline{\underline{X}}] \right. \\
 & \quad \left. + \delta [\underline{\underline{X}}]^T \cdot [\mathbf{V}^{\xi_k}] \left( \underline{n} \cdot \frac{\partial^3 \rho}{\partial \xi_k \partial \xi_j \partial \xi_m} \right) [\mathbf{V}^{\xi_m}]^T \cdot \Delta [\underline{\underline{X}}] \right\} = \\
 & = \delta [\underline{\underline{X}}]^T \cdot \left\{ g_n [\Phi''_{jk}] \underline{n} \otimes [\mathbf{V}^{\xi_k}]^T + g_n [\mathbf{V}^{\xi_k}] \otimes \underline{n} [\Phi''_{jk}]^T \right. \\
 & \quad \left. + g_n \left( \underline{n} \cdot \frac{\partial^3 \rho}{\partial \xi_k \partial \xi_j \partial \xi_m} \right) [\mathbf{V}^{\xi_k}] \otimes [\mathbf{V}^{\xi_m}]^T + \right\} \cdot \Delta [\underline{\underline{X}}] \tag{2.108}
 \end{aligned}$$

**Term 3.**

$$\begin{aligned}
 & \left( -[\nabla g_n]^T \cdot \delta [\underline{\mathbf{X}}] \delta_{kj} + g_n A^{km} \frac{\partial \underline{\rho}}{\partial \xi_m} \cdot \left\{ [\Phi'_j]^T \delta [\underline{\mathbf{X}}] + \frac{\partial^2 \underline{\rho}}{\partial \xi_j \partial \xi_l} [\nabla \xi_l]^T \cdot \delta [\underline{\mathbf{X}}] \right\} \right) \otimes \\
 & \quad \otimes \left\{ [\Phi'_k]^T \Delta [\underline{\mathbf{X}}] \cdot \underline{\mathbf{n}} + H_{ks} [\nabla \xi_s]^T \cdot \Delta [\underline{\mathbf{X}}] \right\} = \\
 & - \left( \delta [\underline{\mathbf{X}}]^T \cdot [\nabla g_n] \delta_{kj} \right) \left( [\Phi'_k]^T \underline{\mathbf{n}} \cdot \Delta [\underline{\mathbf{X}}] + H_{ks} [\nabla \xi_s]^T \cdot \Delta [\underline{\mathbf{X}}] \right) \\
 & + \left( \delta [\underline{\mathbf{X}}]^T \cdot \frac{\partial \underline{\rho}}{\partial \xi_m} [\Phi'_j] g_n A^{km} \right) \left( [\Phi'_k]^T \underline{\mathbf{n}} \cdot \Delta [\underline{\mathbf{X}}] + H_{ks} [\nabla \xi_s]^T \cdot \Delta [\underline{\mathbf{X}}] \right) \\
 & + \left( \delta [\underline{\mathbf{X}}]^T \cdot [\nabla \xi_l] \right) \left( g_n A^{km} \frac{\partial \underline{\rho}}{\partial \xi_m} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi_j \partial \xi_l} \right) \left( [\Phi'_k]^T \underline{\mathbf{n}} \cdot \Delta [\underline{\mathbf{X}}] + H_{ks} [\nabla \xi_s]^T \cdot \Delta [\underline{\mathbf{X}}] \right) = \\
 & \delta [\underline{\mathbf{X}}]^T \cdot \left\{ -\delta_{kj} [\nabla g_n] \otimes \left( \underline{\mathbf{n}} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) + g_n A^{km} [\Phi'_j] \frac{\partial \underline{\rho}}{\partial \xi_m} \otimes \left( \underline{\mathbf{n}} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) \right. \\
 & \quad \left. + g_n A^{km} \left( \frac{\partial \underline{\rho}}{\partial \xi_m} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi_j \partial \xi_l} \right) [\nabla \xi_l] \otimes \left( \underline{\mathbf{n}} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) \right\} \cdot \Delta [\underline{\mathbf{X}}]
 \end{aligned} \tag{2.109}$$

**Term 4** is nothing but the term 3 with  $\Delta$  replaced by  $\delta$  and vice versa

$$\begin{aligned}
 & \Delta [\underline{\mathbf{X}}]^T \cdot \left\{ -\delta_{kj} [\nabla g_n] \otimes \left( \underline{\mathbf{n}} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) + g_n A^{km} [\Phi'_j] \frac{\partial \underline{\rho}}{\partial \xi_m} \otimes \left( \underline{\mathbf{n}} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) \right. \\
 & \quad \left. + g_n A^{km} \left( \frac{\partial \underline{\rho}}{\partial \xi_m} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi_j \partial \xi_l} \right) [\nabla \xi_l] \otimes \left( \underline{\mathbf{n}} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) \right\} \cdot \delta [\underline{\mathbf{X}}] = \\
 & \delta [\underline{\mathbf{X}}]^T \cdot \left\{ -\delta_{kj} \left( [\Phi'_k] \underline{\mathbf{n}} + H_{ks} [\nabla \xi_s] \right) \otimes [\nabla g_n]^T + g_n A^{km} \left( \underline{\mathbf{n}} [\Phi'_k] + H_{ks} [\nabla \xi_s] \right) \otimes \frac{\partial \underline{\rho}}{\partial \xi_m} [\Phi'_j]^T \right. \\
 & \quad \left. + g_n A^{km} \left( \frac{\partial \underline{\rho}}{\partial \xi_m} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi_j \partial \xi_l} \right) \left( \underline{\mathbf{n}} [\Phi'_k] + H_{ks} [\nabla \xi_s] \right) \otimes [\nabla \xi_l]^T \right\} \cdot \Delta [\underline{\mathbf{X}}]
 \end{aligned} \tag{2.110}$$

The final expression for the second order variation of the surface parameter is the sum of the obtained terms (2.107)-(2.110) multiplied by t-scalar  $\underline{\underline{\mathbb{C}}}$  with minus sign

$$-\underline{\underline{\mathbb{C}}} = (g_n \underline{\underline{\mathbb{H}}} - \underline{\underline{\mathbb{A}}})^{-1}$$

$$\begin{aligned}
 \Delta \delta \xi_i = \delta [\underline{\mathbf{X}}]^T \cdot & \left\{ -C_{ij} \left[ [\Phi'_k] \frac{\partial \underline{\rho}}{\partial \xi_j} \otimes [\nabla \xi_k]^T + [\nabla \xi_k] \otimes \frac{\partial \underline{\rho}}{\partial \xi_j} [\Phi'_k]^T \right. \right. \\
 & + \left( \frac{\partial \underline{\rho}}{\partial \xi_j} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi_k \partial \xi_m} - g_n \underline{n} \cdot \frac{\partial^3 \underline{\rho}}{\partial \xi_k \partial \xi_j \partial \xi_m} \right) [\nabla \xi_k] \otimes [\nabla \xi_m]^T \\
 & - g_n [\Phi''_{jk}] \underline{n} \otimes [\nabla \xi_k]^T - g_n [\nabla \xi_k] \otimes \underline{n} [\Phi''_{jk}]^T \\
 & \left. - [\nabla g_n] \otimes \left( \underline{n} [\Phi'_j]^T + H_{js} [\nabla \xi_s]^T \right) - \left( [\Phi'_j] \underline{n} + H_{js} [\nabla \xi_s] \right) \otimes [\nabla g_n]^T \right. \\
 & + g_n A^{km} \left( [\Phi'_j] \frac{\partial \underline{\rho}}{\partial \xi_m} \otimes \left( \underline{n} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) + \left( \underline{n} [\Phi'_k] + H_{ks} [\nabla \xi_s] \right) \otimes \frac{\partial \underline{\rho}}{\partial \xi_m} [\Phi'_j]^T \right) \\
 & + g_n A^{km} \left( \frac{\partial \underline{\rho}}{\partial \xi_m} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi_j \partial \xi_l} \right) \left( [\nabla \xi_l] \otimes \left( \underline{n} [\Phi'_k]^T + H_{ks} [\nabla \xi_s]^T \right) + \right. \\
 & \left. \left. + \left( \underline{n} [\Phi'_k] + H_{ks} [\nabla \xi_s] \right) \otimes [\nabla \xi_l]^T \right) \right] \cdot \Delta [\underline{\mathbf{X}}] = \\
 & = \delta [\underline{\mathbf{X}}]^T \cdot [\nabla \nabla \xi_i] \cdot \Delta [\underline{\mathbf{X}}]
 \end{aligned}
 \tag{2.111}$$

where expressions  $[\nabla g_n]$  and  $[\nabla \xi_i]$  should be substituted according to (2.99) and (2.102) respectively. All derivatives of the projection vector may be directly substituted in this expression explicitly (as evaluated constant vectors) or expressed through the shape functions and nodal values

$$\left. \frac{\partial \underline{\rho}}{\partial \xi_i} \right|_{\xi_\pi} = [\Phi'_i]^T [\underline{\mathbf{X}}], \quad \left. \frac{\partial^2 \underline{\rho}}{\partial \xi_i \partial \xi_j} \right|_{\xi_\pi} = [\Phi''_{ij}]^T [\underline{\mathbf{X}}], \quad \left. \frac{\partial^3 \underline{\rho}}{\partial \xi_i \partial \xi_j \partial \xi_k} \right|_{\xi_\pi} = [\Phi'''_{ijk}]^T [\underline{\mathbf{X}}].$$

The resulting matrices (2.105), (2.111) which connect  $\delta [\underline{\mathbf{X}}]$ ,  $\Delta [\underline{\mathbf{X}}]$  with  $\Delta \delta \xi$  and  $\Delta \delta g_n$  are symmetric.

## 2.5 Enrichment of contact geometry

In the second half of the 90s several approaches based on enriching of the element interpolation functions have been proposed for different problems [Heyliger 89], [Babuška 95].

$$\underline{\mathbf{r}}(t, \zeta) = \underline{\phi}^T(\zeta) \underline{\mathbf{x}}(t) \longrightarrow \underline{\mathbf{r}}_e(t, \zeta) = \left( \underline{\phi}^T(\zeta) + \underline{\psi}_e^T(\zeta) \right) \underline{\mathbf{x}}(t),$$

where  $\underline{\mathbf{r}}_e$  denotes the enriched vector and  $\underline{\psi}_e$  – the enriching function. The entire class of enriched or extended finite element methods got the name XFEM (extended finite element method) [Moës 99] or GFEM (generalized finite element method) or PUM (partition of unity method) [Melenk 96]. The method is used for modeling of dislocations, solidification, two-fluid flows, cracks and



cohesive cracks, the last two examples are based on enriching by Heaviside step function

$$\psi(\zeta) = \alpha H(\zeta - \zeta_0).$$



### Idea 2.2 On enrichment of contact geometry.

*Inspired from the XFEM, we propose an enrichment method for the contact problems. The method consists in enriching the master surface (Fig. 2.16) with an arbitrary enriching function  $h_e$*

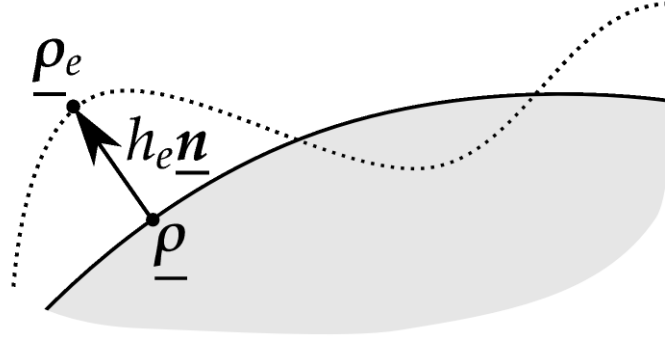
$$\underline{\rho} \longrightarrow \boxed{\underline{\rho}_e = \underline{\rho} + h_e \underline{n}} \Leftrightarrow \underline{\rho}_e(\underline{\xi}) = \underline{\rho}(\underline{\xi}) + h_e(\underline{\xi}, \underline{\varrho}) \underline{n}(\underline{\xi}), \quad (2.112)$$

*where the enriching function  $h_e(\underline{\xi}, \underline{\varrho})$  may depend on the surface parameter  $\underline{\xi}$  and in general may also depend on the local stress-strain state and its history, which can be used, for example, to account for geometry change due to wear, dislocation escape, etc. This dependence is taken into account by means of  $\underline{\varrho}$  variables, which may also include time or loading cycle. Such an enrichment allows to roughly account for submesh features of the contact geometry without an excessive remeshing. The change of the surface geometry due to a complex physics of deformation process can be taken into account within the enriching function as a phenomenological model.*

To keep the formulation consistent, let us impose two conditions on the enriching function  $h_e(\underline{\xi}, \underline{\varrho})$

- $h_e(\underline{\xi}, \underline{\varrho}) \in C^2(\underline{\xi}) \cap C^1(\underline{\varrho})$ , i.e.  $h_e(\underline{\xi}, \underline{\varrho})$  is a  $C^2$ -smooth function by surface parameter and  $C^1$ -smooth by  $\underline{\varrho}$ ;
- in order to avoid self-intersection of the enriched surface, we require that the value of enriching function remains smaller than the minimal local curvature radius of the surface  $|h_e(\underline{\xi})| < \min_i \{1/\kappa_i(\underline{\xi})\}$ .

If the enriching function depends only on the surface parameter (i.e. it does not change in time and does not depend on deformation state) or if the enriching function changes very slowly, then all the variations of the kinematic quantities remain the same (see sections 2.2.5 and 2.2.6) if one replaces all


 Figure 2.16: Enriched geometry  $\underline{\rho}_e$  of the master surface  $\underline{\rho}$ .

quantities related to the master surface by enriched quantities  $\bullet_e$ :

$$\begin{aligned}
 \text{a. } & \boxed{\underline{\rho} \longrightarrow \underline{\rho}_e} : \underline{\rho}_e = \underline{\rho} + h_e \underline{n} \\
 \text{b. } & \boxed{\underline{n} \longrightarrow \underline{n}_e} : \underline{n}_e = \frac{\frac{\partial \underline{\rho}_e}{\partial \xi_1} \times \frac{\partial \underline{\rho}_e}{\partial \xi_2}}{\left\| \frac{\partial \underline{\rho}_e}{\partial \xi_1} \times \frac{\partial \underline{\rho}_e}{\partial \xi_2} \right\|} \\
 \text{c. } & \boxed{\bar{\delta} \underline{\rho} \longrightarrow \bar{\delta} \underline{\rho}_e} : \bar{\delta} \underline{\rho}_e = \bar{\delta} \underline{\rho} + \frac{\partial \underline{\rho}_e}{\partial \xi}^T \delta \xi \\
 \text{d. } & \boxed{\delta \underline{\rho} \longrightarrow \delta \underline{\rho}_e} : \delta \underline{\rho}_e = \delta \underline{\rho} + h_e \delta \underline{n} \\
 \text{e. } & \boxed{\frac{\partial \underline{\rho}}{\partial \xi} \longrightarrow \frac{\partial \underline{\rho}_e}{\partial \xi}} : \frac{\partial \underline{\rho}_e}{\partial \xi} = \frac{\partial \underline{\rho}}{\partial \xi} + \underline{n} \frac{\partial h_e}{\partial \xi} + h_e \frac{\partial \underline{n}}{\partial \xi} \\
 \text{f. } & \boxed{\frac{\partial^2 \underline{\rho}}{\partial \xi^2} \longrightarrow \frac{\partial^2 \underline{\rho}_e}{\partial \xi^2}} : \frac{\partial^2 \underline{\rho}_e}{\partial \xi^2} = \frac{\partial^2 \underline{\rho}}{\partial \xi^2} + \frac{\partial \underline{n}}{\partial \xi} \frac{\partial h_e}{\partial \xi}^T + \frac{\partial h_e}{\partial \xi} \frac{\partial \underline{n}}{\partial \xi}^T + h_e \frac{\partial^2 \underline{n}}{\partial \xi^2} + \underline{n} \frac{\partial^2 h_e}{\partial \xi^2} \\
 \text{g. } & \boxed{\underline{\mathbb{A}} \longrightarrow \underline{\mathbb{A}}_e} : \underline{\mathbb{A}}_e = \frac{\partial \underline{\rho}_e}{\partial \xi} \cdot \frac{\partial \underline{\rho}_e}{\partial \xi}^T \\
 \text{h. } & \boxed{\underline{\mathbb{H}} \longrightarrow \underline{\mathbb{H}}_e} : \underline{\mathbb{H}}_e = \underline{n}_e \cdot \frac{\partial^2 \underline{\rho}_e}{\partial \xi^2}
 \end{aligned} \tag{2.113}$$

Let us consider in detail some of these quantities. The unit normal to the enriched surface can be obtained by substituting (2.113.e) in (2.113.b). Then the numerator takes the form

$$\frac{\partial \underline{\rho}_e}{\partial \xi_1} \times \frac{\partial \underline{\rho}_e}{\partial \xi_2} = \left( \frac{\partial \underline{\rho}}{\partial \xi_1} + \underline{n} \frac{\partial h_e}{\partial \xi_1} + h_e \frac{\partial \underline{n}}{\partial \xi_1} \right) \times \left( \frac{\partial \underline{\rho}}{\partial \xi_2} + \underline{n} \frac{\partial h_e}{\partial \xi_2} + h_e \frac{\partial \underline{n}}{\partial \xi_2} \right), \tag{2.114}$$

where, according to (2.63)<sup>5</sup>

$$\frac{\partial \underline{n}}{\partial \xi} = -\underline{H} \underline{\bar{A}} \frac{\partial \rho}{\partial \xi}, \quad \text{where } \underline{H} \underline{\bar{A}} = \begin{bmatrix} \gamma_1^T \\ \gamma_2^T \end{bmatrix} \Rightarrow \frac{\partial \underline{n}}{\partial \xi} = - \begin{bmatrix} \gamma_1^T \\ \gamma_2^T \end{bmatrix} \frac{\partial \rho}{\partial \xi} = - \begin{bmatrix} \gamma_1^T \frac{\partial \rho}{\partial \xi} \\ \gamma_2^T \frac{\partial \rho}{\partial \xi} \end{bmatrix}. \quad (2.115)$$

Expanding (2.114) and carrying (2.115) gives

$$\begin{aligned} & \left( \frac{\partial \rho}{\partial \xi_1} + \underline{n} \frac{\partial h_e}{\partial \xi_1} - h_e \gamma_1^T \frac{\partial \rho}{\partial \xi} \right) \times \left( \frac{\partial \rho}{\partial \xi_2} + \underline{n} \frac{\partial h_e}{\partial \xi_2} - h_e \gamma_2^T \frac{\partial \rho}{\partial \xi} \right) = \\ & = \frac{\partial \rho}{\partial \xi_1} \times \frac{\partial \rho}{\partial \xi_2} + \frac{\partial \rho}{\partial \xi_1} \times \underline{n} \frac{\partial h_e}{\partial \xi_2} - h_e \frac{\partial \rho}{\partial \xi_1} \times \frac{\partial \rho}{\partial \xi} \gamma_2^T + \underline{n} \times \frac{\partial \rho}{\partial \xi_2} \frac{\partial h_e}{\partial \xi_1} - h_e \underline{n} \times \frac{\partial \rho}{\partial \xi} \gamma_2^T \frac{\partial h_e}{\partial \xi_1} - \\ & - h_e \gamma_1^T \frac{\partial \rho}{\partial \xi} \times \frac{\partial \rho}{\partial \xi_2} + h_e \underline{n} \times \frac{\partial \rho}{\partial \xi} \gamma_1^T \frac{\partial h_e}{\partial \xi_2} + h_e^2 \gamma_1^T \frac{\partial \rho}{\partial \xi} \times \frac{\partial \rho}{\partial \xi} \gamma_2^T, \end{aligned} \quad (2.116)$$

where the cross product of the basis v-vectors can be expressed through a special antisymmetric t-scalar  $\underline{J}$ :

$$\underline{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \underline{J}^T = -\underline{J}, \quad \underline{J} \underline{J}^T = \underline{I}, \quad \underline{J} \underline{J} = -\underline{I}, \quad (2.117)$$

$$\frac{\partial \rho}{\partial \xi} \times \frac{\partial \rho}{\partial \xi} = \underline{\hat{n}} \underline{J}, \quad (2.118)$$

where  $\underline{\hat{n}}$  denotes not normalized vector normal to the initial master surface. Some terms of (2.116) can be simplified in the following manner

$$h_e \frac{\partial \rho}{\partial \xi_1} \times \frac{\partial \rho}{\partial \xi} \gamma_2^T + h_e \gamma_1^T \frac{\partial \rho}{\partial \xi} \times \frac{\partial \rho}{\partial \xi_2} = h_e \underline{\hat{n}} (\gamma_{22} + \gamma_{11}) = h_e \underline{\hat{n}} \text{tr}[\underline{H} \underline{\bar{A}}] \quad (2.119)$$

$$\gamma_1^T \frac{\partial \rho}{\partial \xi} \times \frac{\partial \rho}{\partial \xi} \gamma_2^T = \underline{\hat{n}} \gamma_1^T \underline{J} \gamma_2 = \underline{\hat{n}} (\gamma_{11} \gamma_{22} - \gamma_{12} \gamma_{21}) = \underline{\hat{n}} \det[\underline{H} \underline{\bar{A}}] \quad (2.120)$$

From (2.116), (2.119) and (2.120), the cross product of the basis v-vectors of the enriched surface becomes

$$\begin{cases} \frac{\partial \rho_e}{\partial \xi_1} \times \frac{\partial \rho_e}{\partial \xi_2} = \underline{\hat{n}} (1 - h_e \text{tr}[\underline{H} \underline{\bar{A}}] + h_e^2 \det[\underline{H} \underline{\bar{A}}]) + \\ \quad + \underline{n} \times \left( \frac{\partial \rho}{\partial \xi_2} \frac{\partial h_e}{\partial \xi_1} - \frac{\partial \rho}{\partial \xi_1} \frac{\partial h_e}{\partial \xi_2} \right) - h_e \underline{n} \times \frac{\partial \rho}{\partial \xi} \left( \gamma_2^T \frac{\partial h_e}{\partial \xi_1} - \gamma_1^T \frac{\partial h_e}{\partial \xi_2} \right). \end{cases} \quad (2.121)$$

<sup>5</sup>here it is more convenient to use the components of s-structures

The last two brackets in (2.120) can be transformed using  $\underset{\sim}{J}$  t-scalar

$$\begin{aligned} \frac{\partial \underline{\rho}}{\partial \xi_2} \frac{\partial h_e}{\partial \xi_1} - \frac{\partial \underline{\rho}}{\partial \xi_1} \frac{\partial h_e}{\partial \xi_2} &= -\frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \underset{\sim}{J} \frac{\partial h_e}{\partial \underline{\xi}} \\ \gamma^2 \frac{\partial h_e}{\partial \xi_1} - \gamma^1 \frac{\partial h_e}{\partial \xi_2} &= -\underset{\sim}{H} \underset{\sim}{\bar{A}} \underset{\sim}{J} \frac{\partial h_e}{\partial \underline{\xi}}, \quad \text{as } \begin{bmatrix} \gamma^1 \\ \gamma^2 \end{bmatrix} = \underset{\sim}{H} \underset{\sim}{\bar{A}}. \end{aligned} \quad (2.122)$$

Now it is necessary to express the cross product of the normal vector with the basis v-vector. This cross product lies in the tangent plane to the master surface, moreover

$$\underline{n} \times \frac{\partial \underline{\rho}}{\partial \xi_1} = \alpha_1 \frac{\partial \underline{\rho}}{\partial \xi_2} \quad \text{and} \quad \underline{n} \times \frac{\partial \underline{\rho}}{\partial \xi_2} = \alpha_2 \frac{\partial \underline{\rho}}{\partial \xi_1},$$

which can be rewritten as

$$\underline{n} \times \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = \underline{\alpha} \frac{\partial \underline{\rho}}{\partial \underline{\xi}}.$$

The right dot product of this expression with the covariant basis v-vector gives

$$\left( \underline{n} \times \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} = \underline{\alpha} \underset{\sim}{I} \Leftrightarrow \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \times \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \right) \cdot \underline{n} = \underline{\alpha} \Leftrightarrow \underline{\alpha} = \|\underline{n}\| \underset{\sim}{J} = \hat{n} \underset{\sim}{J},$$

consequently

$$\underline{n} \times \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = \hat{n} \underset{\sim}{J} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \quad \text{and} \quad \underline{n} \times \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} = \hat{n} \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \underset{\sim}{J}^T. \quad (2.123)$$

Substituting (2.122) and (2.123) in the last two terms of (2.121) gives

$$\begin{aligned} \underline{n} \times \left( \frac{\partial \underline{\rho}}{\partial \xi_2} \frac{\partial h_e}{\partial \xi_1} - \frac{\partial \underline{\rho}}{\partial \xi_1} \frac{\partial h_e}{\partial \xi_2} \right) - h_e \underline{n} \times \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \left( \gamma^2 \frac{\partial h_e}{\partial \xi_1} - \gamma^1 \frac{\partial h_e}{\partial \xi_2} \right) &= \\ = -\underline{n} \times \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \left( \underset{\sim}{J} \frac{\partial h_e}{\partial \underline{\xi}} - h_e \underset{\sim}{H} \underset{\sim}{\bar{A}} \underset{\sim}{J} \frac{\partial h_e}{\partial \underline{\xi}} \right) &= -\hat{n} \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} (\underset{\sim}{I} - h_e \underset{\sim}{J}^T \underset{\sim}{H} \underset{\sim}{\bar{A}} \underset{\sim}{J}) \frac{\partial h_e}{\partial \underline{\xi}} \end{aligned} \quad (2.124)$$

Finally, from (2.121) and (2.124) (replacement of the contravariant basis by the covariant) we get the following expression for the numerator of the unit normal vector to the enriched master surface

$$\frac{\partial \underline{\rho}_e}{\partial \xi_1} \times \frac{\partial \underline{\rho}_e}{\partial \xi_2} = \left( 1 - h_e \text{tr}[\underset{\sim}{H} \underset{\sim}{\bar{A}}] + h_e^2 \det[\underset{\sim}{H} \underset{\sim}{\bar{A}}] \right) \hat{n} - \hat{n} \frac{\partial h_e^T}{\partial \underline{\xi}} (\underset{\sim}{I} - h_e \underset{\sim}{J}^T \underset{\sim}{H} \underset{\sim}{\bar{A}} \underset{\sim}{J}) \underset{\sim}{\bar{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}}. \quad (2.125)$$

Note that the t-scalar  $\underset{\sim}{J}^T \underset{\sim}{H} \underset{\sim}{\bar{A}} \underset{\sim}{J} = (\underset{\sim}{J}^T \underset{\sim}{H} \underset{\sim}{\bar{A}} \underset{\sim}{J})^T$  is symmetric.

Now we can write the expression for the normal to the enriched surface

$$\underline{n}_e = \frac{\left( \alpha \underline{n} - \underline{\beta}^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right)}{\sqrt{\alpha^2 + \underline{\beta}^T \underline{\mathbb{A}} \underline{\beta}}}$$

where

$$\alpha = 1 - h_e \operatorname{tr}[\underline{\mathbb{H}} \underline{\bar{\mathbb{A}}}] + h_e^2 \det[\underline{\mathbb{H}} \underline{\bar{\mathbb{A}}}]$$

$$\underline{\beta}^T = \frac{\partial h_e}{\partial \underline{\xi}} (\underline{\mathbb{I}} - h_e \underline{\mathbb{J}}^T \underline{\mathbb{H}} \underline{\bar{\mathbb{A}}} \underline{\mathbb{J}}) \underline{\bar{\mathbb{A}}}$$
(2.126)

It follows naturally from this expression that an adequate enriched geometry (without singularities and self-intersections) is retained if and only if

$$\frac{\partial h_e}{\partial \underline{\xi}} < \infty, \quad \alpha > 0 \text{ and } \underline{\mathbb{I}} - h_e \underline{\mathbb{J}}^T \underline{\mathbb{H}} \underline{\bar{\mathbb{A}}} \underline{\mathbb{J}} > 0.$$

The last two conditions<sup>6</sup> imply that

$$\det(\underline{\mathbb{I}} - h_e \underline{\mathbb{H}} \underline{\bar{\mathbb{A}}}) = 1 - h_e \operatorname{tr}[\underline{\mathbb{H}} \underline{\bar{\mathbb{A}}}] + h_e^2 \det[\underline{\mathbb{H}} \underline{\bar{\mathbb{A}}}] > 0,$$

which is equivalent to the requirement that  $h_e < \min\left\{\frac{1}{\kappa_1}, \frac{1}{\kappa_2}\right\}$ , where  $\kappa_1$  and  $\kappa_2$  are local curvatures of the surface and consequently are the solutions of the following equation (see, e.g. [Konyukhov 08])

$$\kappa^2 - \kappa \operatorname{tr}[\underline{\mathbb{H}} \underline{\bar{\mathbb{A}}}] + \det[\underline{\mathbb{H}} \underline{\bar{\mathbb{A}}}] = 0$$

To compute the explicit forms for two fundamental surface t-scalars ( $\underline{\mathbb{A}}$  and  $\underline{\mathbb{H}}$ ) of the enriched surface; starting from (2.113.g) and using (2.113.e) we get

$$\begin{aligned} \underline{\mathbb{A}}_e &= \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}^T = \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} + \underline{n} \frac{\partial h_e}{\partial \underline{\xi}} + h_e \frac{\partial \underline{n}}{\partial \underline{\xi}} \right) \cdot \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T + \underline{n} \frac{\partial h_e}{\partial \underline{\xi}}^T + h_e \frac{\partial \underline{n}}{\partial \underline{\xi}}^T \right) = \\ &= \underline{\mathbb{A}} + \frac{\partial h_e}{\partial \underline{\xi}} \frac{\partial h_e}{\partial \underline{\xi}}^T + 2h_e \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{n}}{\partial \underline{\xi}}^T + h_e^2 \frac{\partial \underline{n}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{n}}{\partial \underline{\xi}}^T = \\ &= \underline{\mathbb{A}} + \frac{\partial h_e}{\partial \underline{\xi}} \frac{\partial h_e}{\partial \underline{\xi}}^T - 2h_e \underline{\mathbb{H}} + h_e^2 \underline{\mathbb{H}} \underline{\bar{\mathbb{A}}} \underline{\mathbb{H}}, \end{aligned}$$
(2.127)

so

$$\underline{\mathbb{A}}_e = \underline{\mathbb{A}} + \frac{\partial h_e}{\partial \underline{\xi}} \frac{\partial h_e}{\partial \underline{\xi}}^T - 2h_e \underline{\mathbb{H}} + h_e^2 \underline{\mathbb{H}} \underline{\bar{\mathbb{A}}} \underline{\mathbb{H}}$$
(2.128)

<sup>6</sup>It is easy to show that:  $\det(\underline{\mathbb{I}} - h_e \underline{\mathbb{J}}^T \underline{\mathbb{H}} \underline{\bar{\mathbb{A}}} \underline{\mathbb{J}}) = \det(\underline{\mathbb{I}} - h_e \underline{\mathbb{H}} \underline{\bar{\mathbb{A}}})$ .

To derive the second fundamental surface t-scalar, we make use of the equations obtained for the enriched normal vector (2.126) and (2.113.f)

$$\begin{aligned}
 \underline{\underline{H}}_e &= \underline{\underline{n}}_e \cdot \frac{\partial^2 \underline{\underline{p}}_e}{\partial \xi^2} = \frac{\left( \alpha \underline{\underline{n}} + \underline{\underline{\beta}}^T \frac{\partial \underline{\underline{p}}}{\partial \xi} \right)}{\sqrt{\alpha^2 + \underline{\underline{\beta}}^T \underline{\underline{A}} \underline{\underline{\beta}}}} \cdot \left( \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} + \frac{\partial \underline{\underline{n}}}{\partial \xi} \frac{\partial h_e}{\partial \xi}^T + \frac{\partial h_e}{\partial \xi} \frac{\partial \underline{\underline{n}}}{\partial \xi}^T + h_e \frac{\partial^2 \underline{\underline{n}}}{\partial \xi^2} + \underline{\underline{n}} \frac{\partial^2 h_e}{\partial \xi^2} \right) = \\
 &= \frac{1}{\sqrt{\alpha^2 + \underline{\underline{\beta}}^T \underline{\underline{A}} \underline{\underline{\beta}}}} \left( \alpha \underline{\underline{H}} - \alpha h_e \underline{\underline{H}} \underline{\underline{A}} \underline{\underline{H}} + \alpha \frac{\partial^2 h_e}{\partial \xi^2} + \underline{\underline{\beta}}^T \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} - \right. \\
 &\quad \left. - \underline{\underline{\beta}}^T \underline{\underline{H}} \frac{\partial h_e}{\partial \xi}^T - \underline{\underline{\beta}}^T \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial h_e}{\partial \xi} \frac{\partial \underline{\underline{p}}}{\partial \xi}^T \underline{\underline{H}} \underline{\underline{A}} + h_e \underline{\underline{\beta}}^T \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial^2 \underline{\underline{n}}}{\partial \xi^2} \right), \quad (2.129)
 \end{aligned}$$

where the last term can be expressed in the following way (in order to exclude the derivative of the normal vector)

$$\begin{aligned}
 \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \underline{\underline{n}} &= 0 \Rightarrow \frac{\partial}{\partial \xi} \left( \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \underline{\underline{n}} \right) = 0 \Leftrightarrow \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} \cdot \underline{\underline{n}} + \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial \underline{\underline{n}}}{\partial \xi}^T = 0 \Leftrightarrow \\
 \Leftrightarrow \frac{\partial}{\partial \xi} \left( \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} \cdot \underline{\underline{n}} + \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial \underline{\underline{n}}}{\partial \xi}^T \right) &= 0 \Leftrightarrow \frac{\partial^3 \underline{\underline{p}}}{\partial \xi^3} \cdot \underline{\underline{n}} + 2 \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} \cdot \frac{\partial \underline{\underline{n}}}{\partial \xi}^T + \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial^2 \underline{\underline{n}}}{\partial \xi^2} = 0 \\
 \Leftrightarrow \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial^2 \underline{\underline{n}}}{\partial \xi^2} &= - \frac{\partial^3 \underline{\underline{p}}}{\partial \xi^3} \cdot \underline{\underline{n}} + 2 \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} \cdot \frac{\partial \underline{\underline{p}}}{\partial \xi}^T \underline{\underline{H}} \underline{\underline{A}} \quad (2.130)
 \end{aligned}$$

Substituting (2.130) in (2.129) gives the final expression for the second fundamental tensor of the enriched surface

$$\begin{aligned}
 \underline{\underline{H}}_e &= \frac{1}{\sqrt{\alpha^2 + \underline{\underline{\beta}}^T \underline{\underline{A}} \underline{\underline{\beta}}}} \left( \alpha \underline{\underline{H}} - \alpha h_e \underline{\underline{H}} \underline{\underline{A}} \underline{\underline{H}} + \alpha \frac{\partial^2 h_e}{\partial \xi^2} + \underline{\underline{\beta}}^T \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} - \right. \\
 &\quad \left. - \underline{\underline{\beta}}^T \underline{\underline{H}} \frac{\partial h_e}{\partial \xi}^T - \underline{\underline{\beta}}^T \frac{\partial \underline{\underline{p}}}{\partial \xi} \cdot \frac{\partial h_e}{\partial \xi} \frac{\partial \underline{\underline{p}}}{\partial \xi}^T \underline{\underline{H}} \underline{\underline{A}} - h_e \underline{\underline{\beta}}^T \frac{\partial^3 \underline{\underline{p}}}{\partial \xi^3} \cdot \underline{\underline{n}} + 2 h_e \underline{\underline{\beta}}^T \frac{\partial^2 \underline{\underline{p}}}{\partial \xi^2} \cdot \frac{\partial \underline{\underline{p}}}{\partial \xi}^T \underline{\underline{H}} \underline{\underline{A}} \right) \quad (2.131)
 \end{aligned}$$

As one can see, the third derivative of vector  $\underline{\underline{p}}$  appears, which requires that  $\underline{\underline{p}}(\xi) \in C^3(\xi)$ . This condition is satisfied for classic shape functions (linear, quadratic) of the FEM within master faces, but not on their edges.

The form of the expressions of the geometrical variations does not change. The master quantities have to be simply replaced by their enriched homologues, and the variations of the geometrical quantities are expressed by means of the basic variations of the master and slave vectors

$$\delta g_n^e, \delta \xi, \Delta \delta g_n^e, \Delta \delta \xi \longleftarrow \delta \underline{\underline{r}}_s, \delta \underline{\underline{p}}_e, \delta \frac{\partial \underline{\underline{p}}_e}{\partial \xi}, \delta \frac{\partial^2 \underline{\underline{p}}_e}{\partial \xi^2}$$

However from the point of view of the Finite Element Method all variations in the right part except  $\delta \underline{r}_s$  are not basic; some efforts have to be undertaken to convert them in a suitable form for the Finite Element Method. We start from the variation of the projection vector on the enriched surface, according to (2.113.d) and (2.65)

$$\delta \underline{\rho}_e = \delta \underline{\rho} + h_e \delta \underline{n} \Rightarrow \delta \underline{\rho}_e = \delta \underline{\rho} - h_e \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right)^T \underline{\bar{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}}. \quad (2.132)$$

The variation of the covariant basis v-vector  $\delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}$  does not stand alone in expressions of variations, only its dot product with the normal  $\underline{n}_e$  or the basis v-vector  $\frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}$  of the enriched surface appears, since both of them can be expressed through the normal  $\underline{n}$  and the basis v-vector  $\frac{\partial \underline{\rho}}{\partial \underline{\xi}}$  of the original master surface (see (2.113.e) and (2.126))

$$\begin{aligned} \underline{n}_e &= a_1 \underline{n} + a_2^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \\ \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}} &= b_1 \underline{n} + b_2 \frac{\partial \underline{\rho}}{\partial \underline{\xi}}, \end{aligned} \quad (2.133)$$

where

$$\begin{aligned} a_1 &= \frac{\alpha}{\sqrt{\alpha^2 + \underline{\beta}^T \underline{\bar{A}} \underline{\beta}}}; \quad a_2^T = \frac{-\underline{\beta}^T}{\sqrt{\alpha^2 + \underline{\beta}^T \underline{\bar{A}} \underline{\beta}}} \\ \alpha &= 1 - h_e \text{tr}[\underline{\bar{H}} \underline{\bar{A}}] + h_e^2 \det[\underline{\bar{H}} \underline{\bar{A}}]; \quad \underline{\beta} = \frac{\partial h_e}{\partial \underline{\xi}} (\underline{I} - h_e \underline{\bar{J}}^T \underline{\bar{H}} \underline{\bar{A}} \underline{\bar{J}}) \underline{\bar{A}} \\ b_1 &= \frac{\partial h_e}{\partial \underline{\xi}}; \quad b_2 = \underline{I} - h_e \underline{\bar{H}} \underline{\bar{A}} \end{aligned} \quad (2.134)$$

Let us derive the expression for the dot product of the basis vectors  $\underline{n}$  and  $\frac{\partial \underline{\rho}}{\partial \underline{\xi}}$  with variations of the enriched quantities

$$\underline{n} \cdot \delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}} = \underline{n} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} + \delta \underline{n} \frac{\partial h_e}{\partial \underline{\xi}} + h_e \delta \frac{\partial \underline{n}}{\partial \underline{\xi}} \right) = \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} + \underline{n} \cdot \delta \frac{\partial \underline{n}}{\partial \underline{\xi}} h_e, \quad (2.135)$$

where according to (2.63) and (2.65) the last term is

$$\underline{n} \cdot \delta \frac{\partial \underline{n}}{\partial \underline{\xi}} = -\delta \underline{n} \cdot \frac{\partial \underline{n}}{\partial \underline{\xi}} = - \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right)^T \underline{\bar{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \underline{\bar{H}} \underline{\bar{A}} = - \underline{\bar{H}} \underline{\bar{A}} \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right). \quad (2.136)$$

Finally, we get

$$\underline{n} \cdot \delta \frac{\partial \rho_e}{\partial \xi}^T = \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi}^T \left( \underline{I} - h_e \underline{H} \underline{\bar{A}} \right) \quad (2.137)$$

The dot product of the covariant basis v-vector with the first variation of its enriched homologue gives

$$\frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \rho_e}{\partial \xi}^T = \frac{\partial \rho}{\partial \xi} \cdot \left( \delta \frac{\partial \rho}{\partial \xi}^T + \delta \underline{n} \frac{\partial h_e}{\partial \xi}^T + h_e \delta \frac{\partial \underline{n}}{\partial \xi}^T \right) = \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \rho}{\partial \xi}^T - \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} \frac{\partial h_e}{\partial \xi}^T + h_e \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \underline{n}}{\partial \xi}^T, \quad (2.138)$$

where the last term can be derived from

$$\delta \frac{\partial}{\partial \xi} \left( \underline{n} \cdot \frac{\partial \rho}{\partial \xi} \right) = 0 \Leftrightarrow \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \underline{n}}{\partial \xi}^T = -\underline{n} \cdot \delta \frac{\partial^2 \rho}{\partial \xi^2} + \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi}^T \right) \underline{\bar{A}} \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \rho}{\partial \xi^2} + \delta \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial \rho}{\partial \xi}^T \underline{H} \underline{\bar{A}}. \quad (2.139)$$

Substituting the last expression in (2.138) yields a nonsymmetric T-scalar  $\delta \underline{Q} \neq (\delta \underline{Q})^T$

$$\begin{aligned} \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \rho_e}{\partial \xi}^T &= \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \rho}{\partial \xi}^T - \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} \frac{\partial h_e}{\partial \xi}^T - h_e \underline{n} \cdot \delta \frac{\partial^2 \rho}{\partial \xi^2} + \\ &+ h_e \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi}^T \right) \underline{\bar{A}} \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \rho}{\partial \xi^2} + h_e \delta \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial \rho}{\partial \xi}^T \underline{H} \underline{\bar{A}} = \delta \underline{Q} \end{aligned} \quad (2.140)$$

The same procedure has to be carried out with the variation of the derivatives of the basis v-vector. Let us compute  $\underline{n} \cdot \delta \frac{\partial^2 \rho_e}{\partial \xi^2}$  and  $\frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial^2 \rho_e}{\partial \xi^2}$ .

According to (2.113.f)

$$\begin{aligned} \underline{n} \cdot \delta \frac{\partial^2 \rho_e}{\partial \xi^2} &= \underline{n} \cdot \left( \delta \frac{\partial^2 \rho}{\partial \xi^2} + \delta \frac{\partial \underline{n}}{\partial \xi} \frac{\partial h_e}{\partial \xi}^T + \frac{\partial h_e}{\partial \xi} \delta \frac{\partial \underline{n}}{\partial \xi}^T + h_e \delta \frac{\partial^2 \underline{n}}{\partial \xi^2} + \delta \underline{n} \frac{\partial^2 h_e}{\partial \xi^2} \right) = \\ &= \underline{n} \cdot \delta \frac{\partial^2 \rho}{\partial \xi^2} - \delta \underline{n} \cdot \frac{\partial \underline{n}}{\partial \xi} \frac{\partial h_e}{\partial \xi}^T - \frac{\partial h_e}{\partial \xi} \frac{\partial \underline{n}}{\partial \xi}^T \cdot \delta \underline{n} + h_e \underline{n} \cdot \delta \frac{\partial^2 \underline{n}}{\partial \xi^2}, \end{aligned} \quad (2.141)$$

carrying eq.(2.136)

$$\underline{n} \cdot \delta \frac{\partial^2 \rho_e}{\partial \xi^2} = \underline{n} \cdot \delta \frac{\partial^2 \rho}{\partial \xi^2} - \underline{H} \underline{\bar{A}} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} \right) \frac{\partial h_e}{\partial \xi}^T - \frac{\partial h_e}{\partial \xi} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi}^T \right) \underline{H} \underline{\bar{A}} + h_e \underline{n} \cdot \delta \frac{\partial^2 \underline{n}}{\partial \xi^2}, \quad (2.142)$$



where the last term is quite long. First, we evaluate the following expression

$$\begin{aligned} \delta \frac{\partial}{\partial \xi} \left( \underline{n} \cdot \frac{\partial \underline{n}}{\partial \xi} \right) &= 0 \Leftrightarrow \delta \frac{\partial \underline{n}}{\partial \xi} \cdot \frac{\partial \underline{n}^T}{\partial \xi} + \frac{\partial \underline{n}}{\partial \xi} \cdot \delta \frac{\partial \underline{n}^T}{\partial \xi} + \delta \underline{n} \cdot \frac{\partial^2 \underline{n}}{\partial \xi^2} + \underline{n} \cdot \delta \frac{\partial^2 \underline{n}}{\partial \xi^2} = 0 \Leftrightarrow \\ \underline{n} \cdot \delta \frac{\partial^2 \underline{n}}{\partial \xi^2} &= - \underbrace{\delta \frac{\partial \underline{n}}{\partial \xi} \cdot \frac{\partial \underline{n}^T}{\partial \xi}}_{\text{term 1}} - \underbrace{\frac{\partial \underline{n}}{\partial \xi} \cdot \delta \frac{\partial \underline{n}^T}{\partial \xi}}_{\text{term 2}} - \underbrace{\delta \underline{n} \cdot \frac{\partial^2 \underline{n}}{\partial \xi^2}}_{\text{term 3}}, \end{aligned} \quad (2.143)$$

Replacing  $\frac{\partial \underline{n}}{\partial \xi}$  in terms 1 and 2 in (2.143) and carrying of (2.139) yields

$$\begin{aligned} \text{term 1:} \quad \delta \frac{\partial \underline{n}}{\partial \xi} \cdot \frac{\partial \underline{n}^T}{\partial \xi} &= - \delta \frac{\partial \underline{n}}{\partial \xi} \cdot \frac{\partial \rho^T}{\partial \xi} \underline{H} \underline{\bar{A}} = \\ &= \underline{n} \cdot \delta \frac{\partial^2 \rho}{\partial \xi^2} \underline{H} \underline{\bar{A}} - \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} \right) \underline{\bar{A}} \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \rho}{\partial \xi^2} \underline{H} \underline{\bar{A}} - \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \rho^T}{\partial \xi} \underline{H}^2 \underline{\bar{A}}^2 \\ \text{term 2:} \quad \frac{\partial \underline{n}}{\partial \xi} \cdot \delta \frac{\partial \underline{n}^T}{\partial \xi} &= - \underline{H} \underline{\bar{A}} \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \underline{n}^T}{\partial \xi} = \\ &= \underline{n} \cdot \delta \frac{\partial^2 \rho}{\partial \xi^2} \underline{H} \underline{\bar{A}} - \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} \right) \underline{\bar{A}} \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \rho}{\partial \xi^2} \underline{H} \underline{\bar{A}} - \underline{H} \underline{\bar{A}} \delta \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial \rho^T}{\partial \xi} \underline{H} \underline{\bar{A}} \end{aligned} \quad (2.144)$$

The variation of the normal  $\delta \underline{n}$  in the third term in (2.143) has to be replaced by a combination of the basis vectors

$$\delta \underline{n} \cdot \frac{\partial^2 \underline{n}}{\partial \xi^2} = - \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} \right) \underline{\bar{A}} \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \underline{n}}{\partial \xi^2}, \quad (2.145)$$

where the last product can be derived from

$$\frac{\partial^2}{\partial \xi^2} \left( \frac{\partial \rho}{\partial \xi} \cdot \underline{n} \right) = 0 \Leftrightarrow \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \underline{n}}{\partial \xi^2} = - \underline{n} \cdot \frac{\partial \rho_e}{\partial \xi} + 2 \frac{\partial^2 \rho}{\partial \xi^2} \cdot \frac{\partial \rho^T}{\partial \xi} \underline{H} \underline{\bar{A}}. \quad (2.146)$$

Substituting (2.144), (2.145) and (2.146) in (2.143) and consequently in (2.142) gives, after grouping some terms, the final expression

$$\begin{aligned} \underline{n} \cdot \delta \frac{\partial^2 \rho_e}{\partial \xi^2} &= \underline{n} \cdot \delta \frac{\partial^2 \rho}{\partial \xi^2} (\underline{I} - 2h_e \underline{H} \underline{\bar{A}}) - \underline{H} \underline{\bar{A}} \left( \underline{n} \cdot \delta \frac{\partial \rho}{\partial \xi} \right) \frac{\partial h_e^T}{\partial \xi} - \frac{\partial h_e}{\partial \xi} \left( \underline{n} \cdot \delta \frac{\partial \rho^T}{\partial \xi} \right) \underline{H} \underline{\bar{A}} + \\ &\quad + h_e \frac{\partial \rho}{\partial \xi} \cdot \delta \frac{\partial \rho^T}{\partial \xi} \underline{H}^2 \underline{\bar{A}}^2 + h_e \underline{H} \underline{\bar{A}} \delta \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial \rho^T}{\partial \xi} \underline{H} \underline{\bar{A}} + \\ &\quad + h_e \left( \underline{n} \cdot \delta \frac{\partial \rho^T}{\partial \xi} \right) \underline{\bar{A}} \left( - \underline{n} \cdot \frac{\partial \rho_e}{\partial \xi} + 2 \frac{\partial^2 \rho}{\partial \xi^2} \cdot \frac{\partial \rho^T}{\partial \xi} \underline{H} \underline{\bar{A}} + 2 \frac{\partial \rho}{\partial \xi} \cdot \frac{\partial^2 \rho}{\partial \xi^2} \underline{H} \underline{\bar{A}} \right) \end{aligned} \quad (2.147)$$

## 2.5 Enrichment of contact geometry

The dot product of the basis v-vector with the variation of its second derivative can be expressed as

$$\begin{aligned} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \frac{\partial^2 \underline{\rho}_e}{\partial \underline{\xi}^2} &= \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \left( \delta \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} + \delta \frac{\partial \underline{n}}{\partial \underline{\xi}} \frac{\partial h_e}{\partial \underline{\xi}}^T + \frac{\partial h_e}{\partial \underline{\xi}} \delta \frac{\partial \underline{n}}{\partial \underline{\xi}}^T + h_e \delta \frac{\partial^2 \underline{n}}{\partial \underline{\xi}^2} + \delta \underline{n} \frac{\partial^2 h_e}{\partial \underline{\xi}^2} \right) = \\ &= \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \frac{\partial^2 \underline{\rho}}{\partial \underline{\xi}^2} + \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \frac{\partial \underline{n}}{\partial \underline{\xi}} \frac{\partial h_e}{\partial \underline{\xi}}^T + \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial h_e}{\partial \underline{\xi}} \delta \frac{\partial \underline{n}}{\partial \underline{\xi}}^T + h_e \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \frac{\partial^2 \underline{n}}{\partial \underline{\xi}^2} - \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \frac{\partial^2 h_e}{\partial \underline{\xi}^2}, \end{aligned} \quad (2.148)$$

Now we can derive all the variations for the enriched geometry. Although all needed expressions have been just derived without approximation of small gap, in order to keep the text more compact, below we present the variations of the enriched geometrical quantities only within the assumption of small gap. According to (2.24) the variation of the normal gap (2.24) using (2.132) is

$$\delta g_n^e = \underline{n}_e \cdot (\delta \underline{r}_s - \delta \underline{\rho}_e) \rightarrow \delta g_n^e = \underline{n}_e \cdot \left( \delta \underline{r}_s - \delta \underline{\rho} + h_e \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right)^T \underline{\bar{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \quad (2.149)$$

The approximated variation of the surface parameter (2.32) for the enriched surface is

$$\delta \underline{\xi} = \underline{\bar{A}}_e \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}_e) \rightarrow \delta \underline{\xi} = \underline{\bar{A}}_e \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}} \cdot \left( \delta \underline{r}_s - \delta \underline{\rho} + h_e \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right)^T \underline{\bar{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \quad (2.150)$$

The approximated second variation of the normal gap (2.67) is

$$\Delta \delta g_n = -\underline{n}_e \cdot \left( \delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}^T \Delta \underline{\xi} + \Delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}^T \delta \underline{\xi} \right) - \Delta \underline{\xi}^T \underline{\bar{H}}_e \delta \underline{\xi}, \quad (2.151)$$

Replacing of  $\underline{n}_e$  by (2.133) and carrying (2.137) and (2.140), we get

$$\begin{aligned} \Delta \delta g_n &= - \left( \mathbf{a}_1 \underline{n} \cdot \delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}^T + \mathbf{a}_2^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}^T \right) \Delta \underline{\xi} - \\ &\quad - \left( \mathbf{a}_1 \underline{n} \cdot \Delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}^T + \mathbf{a}_2^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \Delta \frac{\partial \underline{\rho}_e}{\partial \underline{\xi}}^T \right) \delta \underline{\xi} - \Delta \underline{\xi}^T \underline{\bar{H}}_e \delta \underline{\xi} \end{aligned} \quad (2.152)$$

$$\Delta \delta g_n = -\mathbf{a}_1 \underline{n} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \underline{\mathbf{b}}_2 \Delta \underline{\xi} + \Delta \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \underline{\mathbf{b}}_2 \delta \underline{\xi} \right) - \mathbf{a}_2^T [\delta \underline{\mathbf{Q}} \Delta \underline{\xi} + \Delta \underline{\mathbf{Q}} \delta \underline{\xi}] - \Delta \underline{\xi}^T \underline{\bar{H}}_e \delta \underline{\xi}, \quad (2.153)$$

where  $\mathbf{a}_1$ ,  $\mathbf{a}_2^T$ ,  $\underline{\mathbf{b}}_2$  can be found in (2.134) and  $\delta \underline{\mathbf{Q}}$ ,  $\Delta \underline{\mathbf{Q}}$  in (2.140).

The expression for the second variation of the surface parameter for the enriched surface can be obtained from (2.79)

$$\begin{aligned} \Delta \delta \xi = \bar{\mathbb{A}}_e \left\{ \delta g_n^e \left( \underline{n}_e \cdot \Delta \frac{\partial \rho_e}{\partial \xi} + \mathbb{H}_e \Delta \xi \right) + \Delta g_n^e \left( \underline{n}_e \cdot \delta \frac{\partial \rho_e}{\partial \xi} + \mathbb{H}_e \delta \xi \right) - \right. \\ \left. - \frac{\partial \rho_e}{\partial \xi} \cdot \left( \delta \frac{\partial \rho_e}{\partial \xi} \Delta \xi + \Delta \frac{\partial \rho_e}{\partial \xi} \delta \xi \right) - \Delta \xi^T \left( \frac{\partial \rho_e}{\partial \xi} \cdot \frac{\partial^2 \rho_e}{\partial \xi^2} \right) \delta \xi \right\}. \end{aligned} \quad (2.154)$$

As previously, we replace  $\underline{n}_e$  and  $\frac{\partial \rho_e}{\partial \xi}$  by (2.133),(2.134) and group terms, next we substitute (2.137), (2.140) and get

$$\begin{aligned} \Delta \delta \xi = \bar{\mathbb{A}}_e \left[ a_1 \mathbb{b}_2 \left( \delta g_n^e \Delta \frac{\partial \rho}{\partial \xi} + \Delta g_n^e \delta \frac{\partial \rho}{\partial \xi} \right) \cdot \underline{n} + \mathbb{H}_e (\delta g_n^e \Delta \xi + \Delta g_n^e \delta \xi) - \right. \\ \left. - \Delta \xi^T \left( \frac{\partial \rho_e}{\partial \xi} \cdot \frac{\partial^2 \rho_e}{\partial \xi^2} \right) \delta \xi + [\delta \mathbb{Q}]^T \mathbb{a}_2 \Delta g_n^e + [\Delta \mathbb{Q}]^T \mathbb{a}_2 \delta g_n^e - \right. \\ \left. - \mathbb{b}_2 \delta \mathbb{Q} \Delta \xi + \mathbb{b}_2 \Delta \mathbb{Q} \delta \xi - \mathbb{b}_1 \underline{n} \cdot \left( \delta \frac{\partial \rho}{\partial \xi} \mathbb{b}_2 \Delta \xi + \Delta \frac{\partial \rho}{\partial \xi} \mathbb{b}_2 \delta \xi \right) \right], \end{aligned} \quad (2.155)$$

where  $a_1$ ,  $\mathbb{a}_2^T$ ,  $\mathbb{b}_1$  and  $\mathbb{b}_2$  have been taken from (2.134) and  $\delta \mathbb{Q}, \Delta \mathbb{Q}$ , from (2.140).

Now it remains to define the enriching functions  $h_e(\xi)$  for the master surfaces and substitute them into (2.113) and consequently in the expressions for the variations. During this operation, one has to take into account that the projection procedure has to be changed (see Fig. 2.17). Briefly the new Newton's procedure used for the definition of the projection point has the same form as (2.8), but obviously all quantities related to the master should be replaced by their enriched homologues

$$\Delta \xi = \left[ \bar{\mathbb{A}}_e - (\underline{r}_s - \underline{\rho}_e) \cdot \frac{\partial^2 \rho_e}{\partial \xi^2} \right]^{-1} \cdot \left[ (\underline{r}_s - \underline{\rho}_e) \cdot \frac{\partial \rho_e}{\partial \xi} \right]. \quad (2.156)$$

The shadow projection procedures undergo the same modifications and are not presented here.

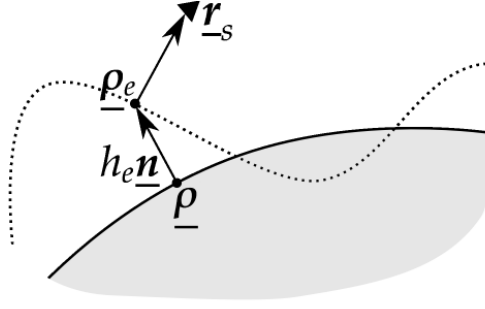


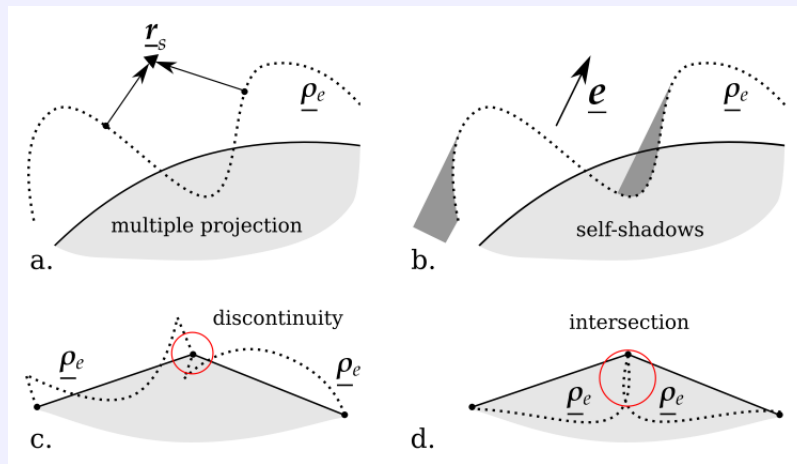
Figure 2.17: Projection procedure for enriched geometry.

### Remark 2.7 on enriched geometry.

As already mentioned, the enriching function  $h_e$  has to be rather smooth  $C^2$  and its value  $|h_e|$  must remain smaller than the minimal local curvature radius. Moreover, if one uses the normal projection, it is necessary to keep in mind that Newton's method allows to determine only one projection point (closest to the starting point). So in case of enrichment of the master surface with a nonlinear function  $h_e(\xi)$ , Newton's method may be insufficient to find the projection point (see Fig. a), a more advanced technique should be used, for example, a dissection method combined with Newton's method.

Contrary to the normal projection, the shadow projection is unique if there is no "self-shadow" from the master surface on its own, so one has to pay attention to avoid shadows due to the enrichment (see Fig. b).

In order to preserve the continuity of the discretized master surface, enriching functions have to be zero at edges of each segment  $\xi_e \in \Gamma_e$ :  $h_e(\xi_e) = 0$  (see Fig. c). It has to be mentioned that there is a possibility of intersection of enriched geometries of adjacent master segments, it has to be also avoided (see Fig. d).



Issues related to the enriching geometry: a – multiple normal projection within one segment (Newton's method is not sufficient to find the closest point); b – possible presence of self-shadows on the master surface; c – discontinuity of enriched master surface; d – self intersection of enriched master surface.

### 2.5.1 Example of enrichment

Below we derive the expressions needed for implementation of a frictionless contact for the case of a linear 2D element enriched by a function  $h_e(\xi)$  (Fig. 2.18), where  $\xi$  is a segment parameter  $\xi \in [0; 1]$ .

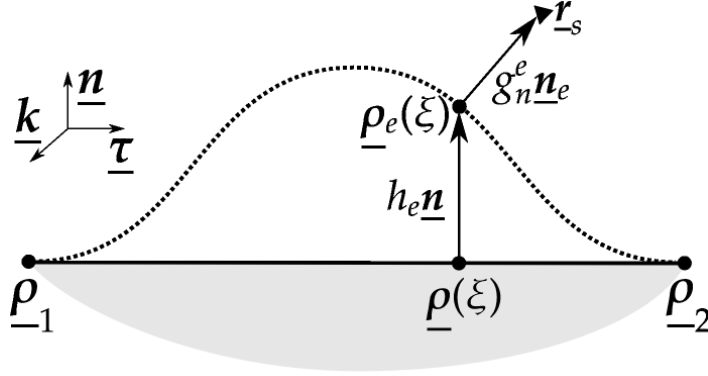


Figure 2.18: Example of linear master segments enriched by cosine wave.

Master segments are assumed to be linear; if one introduces the following notation

$$\underline{t} = \frac{1}{l}(\underline{\rho}_2 - \underline{\rho}_1),$$

where  $l = \|\underline{\rho}_2 - \underline{\rho}_1\|$ , then

$$\begin{aligned} \underline{\rho}(\xi) &= \underline{\rho}_1 + \xi l \underline{t}; & \frac{\partial \underline{\rho}}{\partial \xi} &= l \underline{t}; \\ \underline{n} &= \frac{\underline{k} \times \frac{\partial \underline{\rho}}{\partial \xi}}{\left\| \frac{\partial \underline{\rho}}{\partial \xi} \right\|} = \underline{k} \times \underline{t}, \end{aligned} \quad (2.157)$$

The unit normal vector  $\underline{n}$ , defined in such a way, points outward the solid if the unit vector  $\underline{k}$  point outward the page (to us) and at the same time if master nodes in Fig. 2.18 are enumerated from left to right. The enriched master geometry is defined by

$$\underline{\rho}_e(\xi) = \underline{\rho}(\xi) + h_e(\xi) \underline{n},$$

and consequently, as  $\frac{\partial \underline{n}}{\partial \xi} = 0$

$$\frac{\partial \underline{\rho}_e}{\partial \xi} = l \underline{t} + \frac{\partial h_e}{\partial \xi} \underline{n} \Rightarrow \underline{n}_e = \frac{\underline{k} \times \frac{\partial \underline{\rho}_e}{\partial \xi}}{\left\| \frac{\partial \underline{\rho}_e}{\partial \xi} \right\|} = \frac{l \underline{n} - \frac{\partial h_e}{\partial \xi} \underline{t}}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}}, \quad (2.158)$$

in notations which have been introduced in the previous section

$$\frac{\partial \underline{\rho}_e}{\partial \xi} = b_1 \underline{n} + b_2 l \underline{t}, \quad b_1 = \frac{\partial h_e}{\partial \xi}, \quad b_2 = 1 \quad (2.159)$$

## 2.5 Enrichment of contact geometry

According to (2.149) we get the first variation of the normal gap to the enriched surface

$$\delta g_n^e = \frac{l\underline{n} - \frac{\partial h_e}{\partial \xi} \underline{t}}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \cdot (\delta \underline{r}_s - \delta \underline{\rho}) - \frac{h_e \frac{\partial h_e}{\partial \xi}}{l \sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} \quad (2.160)$$

In a ready-to-implement form, if  $\underline{\rho} = \underline{\rho}^i(t)\phi_i(\xi)$

$$\delta g_n^e = \frac{1}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \left( \left( l\underline{n} - \frac{\partial h_e}{\partial \xi} \underline{t} \right) \begin{bmatrix} 1 \\ -\phi_1 \\ -\phi_2 \end{bmatrix}^T - \frac{h_e \frac{\partial h_e}{\partial \xi}}{l} \underline{n} \begin{bmatrix} 0 \\ \frac{\partial \phi_1}{\partial \xi} \\ \frac{\partial \phi_2}{\partial \xi} \end{bmatrix}^T \right) \cdot \begin{bmatrix} \delta \underline{r}_s \\ \delta \underline{\rho}_1 \\ \delta \underline{\rho}_2 \end{bmatrix} \quad (2.161)$$

From (2.150) adapted to the 2D case, the first variation of the surface parameter is

$$\delta \xi = \frac{1}{l^2} \left( l\underline{t} + \frac{\partial h_e}{\partial \xi} \underline{n} \right) \cdot (\delta \underline{r}_s - \delta \underline{\rho}) + \frac{h_e}{l^2} \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} \quad (2.162)$$

The second variation of the normal gap (2.153) adapted to 2D case for linear elements has the following form

$$\Delta \delta g_n^e = -\frac{l}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \underline{n} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \xi} \Delta \xi + \Delta \frac{\partial \underline{\rho}}{\partial \xi} \delta \xi \right) - \frac{\frac{\partial h_e}{\partial \xi}}{l \sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} (\delta Q \Delta \xi + \Delta Q \delta \xi) - \Delta \xi H_e \delta \xi,$$

where from (2.140)

$$Q = \frac{\partial \underline{\rho}}{\partial \xi} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} - \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} \frac{\partial h_e}{\partial \xi}$$

and from (2.131) we get the expression for  $H_e$

$$H_e = \frac{l \frac{\partial h_e}{\partial \xi}^2}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}}$$

And finally the expression for the second variation of the normal gap for the enriched surface takes the following form

$$\begin{aligned} \Delta \delta g_n^e = & -\frac{l}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \underline{n} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \xi} \Delta \xi + \Delta \frac{\partial \underline{\rho}}{\partial \xi} \delta \xi \right) - \Delta \xi \frac{l \frac{\partial h_e}{\partial \xi}^2}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \delta \xi \\ & - \frac{\frac{\partial h_e}{\partial \xi}}{l \sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \left( \left[ \frac{\partial \underline{\rho}}{\partial \xi} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} - \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} \frac{\partial h_e}{\partial \xi} \right] \Delta \xi + \left[ \frac{\partial \underline{\rho}}{\partial \xi} \cdot \Delta \frac{\partial \underline{\rho}}{\partial \xi} - \underline{n} \cdot \Delta \frac{\partial \underline{\rho}}{\partial \xi} \frac{\partial h_e}{\partial \xi} \right] \delta \xi \right), \end{aligned} \quad (2.163)$$

after grouping terms

$$\Delta \delta g_n^e = \frac{l}{\sqrt{l^2 + \frac{\partial h_e}{\partial \xi}^2}} \left[ \frac{\left\{ \frac{\partial h_e}{\partial \xi}^2 - l^2 \right\} \underline{n} - \frac{\partial h_e}{\partial \xi} l \underline{t}}{l^2} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \xi} \Delta \xi + \Delta \frac{\partial \underline{\rho}}{\partial \xi} \delta \xi \right) - \Delta \xi \frac{\partial h_e}{\partial \xi}^2 \delta \xi \right], \quad (2.164)$$

To get the ready-to-implement expression one needs to use (2.162) for  $\delta\xi$  and  $\Delta\xi$

### Concluding remarks

The enrichment of the contact geometry by an arbitrary function permits:

1. to take into account a complicated geometry within one contact element;
2. to account for a change of the local geometry due to loading conditions.

As mentioned, if the enrichment is chosen to be localized within NTS contact elements, the choice of the enrichment function is limited: its value must be zero at the edges of the master segments. It implies a strong connection between the discretization and the enrichment. A possible application of this approach is the modeling of periodic structures using a regular mesh, Fig. 2.19. Enrichment of thin-walled or beam structure geometries by a constant enriching function seems to be meaningful, since the predominant deformation of such structures does not affect the geometry of the surface (Fig. 2.19,a-b). Moreover, the enrichment technique is the only way to account for the surface topology for shell and beam elements. A possible application is a modeling of contact with grid structures, micro contact with fiber, etc. An anisotropic friction can be simulated implicitly by a special enrichment of the master surface Fig. 2.20, a-b.

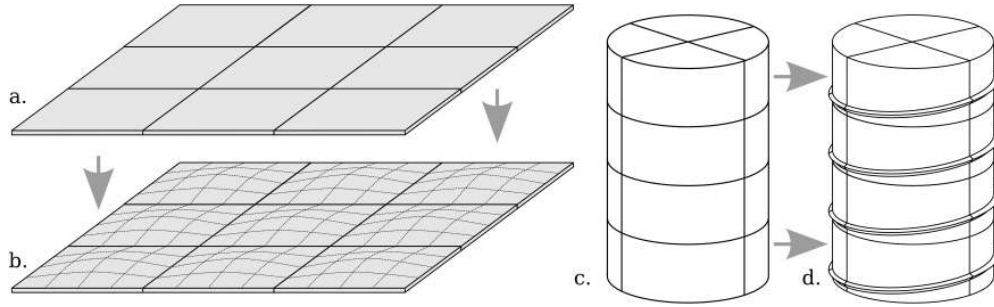


Figure 2.19: Enriched geometry of the master finite element mesh: **a-b** – periodic thin-walled structure, **c-d** – mesh of a screw with 4 turns, represented by enrichment of 16 segments with a screw function  $\underline{\rho}_e$  of the master surface  $\underline{\rho}$ .

If one couples the enrichment technique with a global smoothing procedure (Fig. 2.20, c-d) (for discussion of the latter, see, for example, [Pietrzak 97], [Wriggers 01], [Krstulović-Opara 02]), the mentioned shortcoming of the enrichment approach vanishes. Since the master surface is globally smooth, it is not anymore required that the enriching function is zero along the edges of the master segments/faces. However, it becomes a real challenge to obtain the needed variations of the geometrical quantities for the resulting enriched surface. On the other hand, this coupling makes possible to simulate properly the phenomenon of wear and to enrich the master geometry independently on the mesh.

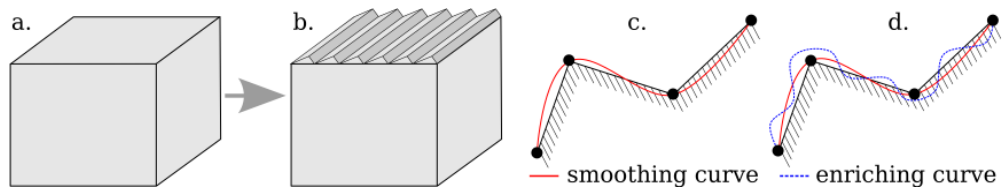


Figure 2.20: **a-b** – an example of enrichment for simulating anisotropic friction; **c-d** – enrichment coupled with smoothing procedure.



## Chapter 3

# Contact detection

### **Résumé de Chapitre 3 «Détection du contact»**

Ce chapitre présente les techniques de détection locale du contact. Cette phase de l'algorithme numérique est responsable de la construction des éléments de contact, qui sont chargés de transférer les efforts d'un corps à l'autre au travers de l'interface discrétisée. Issus de l'application du principe des travaux virtuels, les éléments de contact contribuent avec les éléments structuraux à former le vecteur résidu et la matrice tangente. La phase de détection est importante, car une opération incorrecte à ce niveau conduit à une solution fausse.

Dans le cas d'une discrétisation simple (Nœud-à-Nœud) la procédure de détection est évidente, elle consiste en la détection des paires des nœuds les plus proches. Comme cette discrétisation est applicable seulement dans le cas des petites déformations et des petits glissement, la détection ne doit être exécutée qu'une seule fois. Si on considère le cas des grandes déformations et des grands glissements (par exemple, la discrétisation Nœuds-à-Segment), la détection doit être répétée très souvent, puisque la topologie de l'interface de contact change tout le temps. Pour les problèmes de contact qui possèdent un grand nombre de nœuds à l'interface, la rapidité de détection devient un critère très important pour l'efficacité de l'algorithme : la détection peut en effet prendre autant de temps de calcul que la phase de résolution.

Dans ce chapitre, on présente donc les concepts permettant la détection du contact en calcul par éléments finis implicite. Les notions de base (distance maximale de détection, méthode «bucket sort», détection «tous-à-tous») sont introduites. On propose d'utiliser des «boîtes» découpant le maillage pendant la phase préliminaire afin de réduire la zone de détection. On développe également une méthode pour résoudre le problème des «zones aveugles» et des «nœuds fuyants». En conclusion, les algorithmes de détection classiques basés sur le concept du nœud le plus proche sont critiqués, et quelques contre-exemples sont présentés et discutés en détail.

*On propose une technique de détection basée sur la méthode «bucket sort» et sur la définition rigoureuse du point le plus proche (donnée dans Chapitre 2. Les valeurs optimales des paramètres de la méthode (distance de détection maximale et taille des mailles de détection) sont déduits de façon heuristique sur la base des nombreux essais numériques. On discute également une approche basée sur des éléments de contact à faces multiples ; les questions de détection liées à cette approche sont résolues.*

*La performance de la méthode de détection proposée est vérifiée sur les deux tests: un calcul de contact pneu-asphalte (plus que 300 000 nœuds sur les surfaces de contact) et un calcul de contact entre deux surfaces rugueuses (plus que 2 000 000 nœuds sur les surfaces de contact). Pour certaines configurations, le gain en performance par rapport à la méthode «tous-à-tous» peut atteindre 160 000 (4 secondes contre une semaine).*

*On discute en détail l'adaptation de l'algorithme considéré aux problèmes pour lesquels la discrétisation «esclave-maître» est inconnue a priori. La dernière question abordée est celle de la détection dans le cadre de calcul parallèle. Deux approches (SDMR/MDMR - Détection Singulière/Multiple, Résolution Multiple) sont proposées. On montre pour chacune les détails de l'implémentation dans un code, et également un test de scalabilité.*

Locating the contact points between two surfaces is an important step in the numerical treatment of contact problems. Moreover, this is one of the major computational costs of contact algorithms both in explicit and implicit computations. Fast and accurate detection of contact is not an easy task and has to be considered in detail. Here we present several classical approaches and propose some improvements. We will mainly focus on the so-called local contact detection in case of Node-to-Surface (NTS) discretization for implicit Finite Element Analysis. However, some general remarks are also valid for explicit codes.

First, some notions are introduced, different types and strategies of contact detection are explained. Then simple algorithms and a more elaborated one are discussed in detail both in case of known a priori and unknown master-slave (e.g. self-contact) discretizations. Further some comments are given on parallelization of the discussed algorithms and several examples of contact detection in large contact problems are presented.

### 3.1 Introduction

Roughly, two steps can be distinguished in the contact algorithm: contact detection and resolution. Resolution implies that penetration between contacting solids has to be eliminated by applying repulsive forces to penetrating elements. Consequently the detection phase must determine which elements of the discretized solids are going to penetrate. It is worth mentioning the key difference between contact detection in explicit and implicit resolutions:

- explicit – it is necessary to detect elements that have already penetrated and further apply contact forces;
- implicit – possible penetration has to be known at the beginning of each resolution step in order to include additional degrees of freedom<sup>1</sup> in the problem and to change the residual vector and the stiffness matrix.

In the Finite Element Analysis, contact can occur between discretized deformable bodies or between one discretized deformable body and an analytically defined rigid surface (curve). The penetration can be described in different ways (see Fig. 3.1). The definition of the penetration often introduces an asymmetry in the contact problem and the contacting surfaces have to be treated differently at least locally in space and time. This procedure is strongly connected with the asymmetric geometrical description of contact and the discretization of the contact interface, i.e. contact detection relates to the discretization method, i.e. symmetric discretizations (segment-to-segment type) should use symmetric detection and vice versa. We restrict ourself to node-to-segment discretization, so it is natural to define the penetration as penetration of nodes of one discretized solid (slave) under the segments of the second discretized solid (master) or under an analytical surface. Following we will use these classical notions of master and slave for surfaces and for their components (slave nodes of the slave surface, and master segments and master nodes of the master surface).

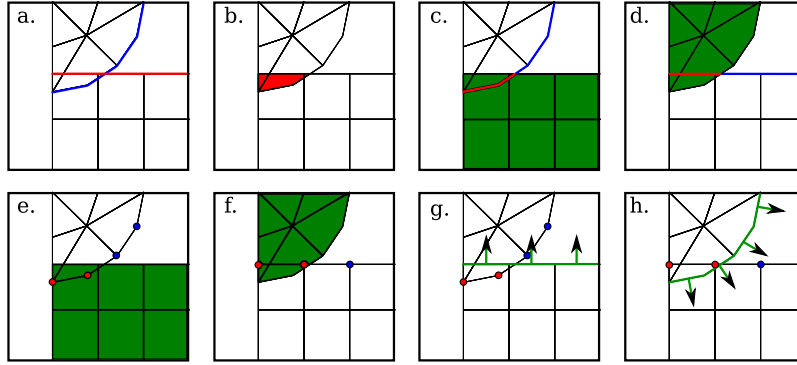


Figure 3.1: Various definitions of the penetration. Symmetric: **a** – segment intersection; **b** – volume intersection. Asymmetric: **c,d** – segment in volume; **e,f** – node in volume; **g,h** – node under surface.

Two contact search phases should be distinguished [Wriggers 06]: spatial search and contact detection. The first notion is used for searching between separate solids coming into contact, i.e. rather between separate geometries than discretizations. Contact spatial search methods are of big importance in multibody systems and discrete element methods where interaction between more or less identical particles such as crashed stone, sand, snow is considered

<sup>1</sup>supplementary degrees of freedom appear in Lagrange multiplier method and the coupled augmented Lagrangian method.

to analyze mud flows, opencast mines, avalanches, etc. It is worth mentioning that previously the particular attention of the scientific community has been paid mostly to this phase of contact search, as local discretization of solids often remained rather moderate and even the simplest all-to-all approach often appeared to be rather efficient and fast enough technique, especially in case of small slip when only one execution of the detection procedure is required.

However, in large deformation and large sliding problems the contact detection was one of the major computational costs, because the contact geometry changes significantly through computation and detection is required almost at each time step. The number of time steps depends on the resolution framework (implicit, explicit) and on the nonlinearity of the problem. Naturally, since the number of time steps is significantly higher in case of explicit calculations the contact detection has been first considered in detail in the framework of explicit codes [Benson 90]. Many simple and efficient techniques have been proposed. They were inspired from hidden line algorithm from computer graphics and based on the bucket sort method. The recent advances in parallel computing make possible extremely large implicit and explicit contact simulations between very finely meshed solids and so it imposes even more severe requirements to the time needed for contact detection. The contact detection algorithms are connected with the contact discretization method and the parallel framework. However, there are not so many publications concerning that phase of the contact algorithm and frequently in moderate size engineering applications the basic ideas of the bucket sort, proposed 20 years ago, is used. Some more recent techniques can be found in [Oldenburg 94], [Bruneel 02], [Fujun 01]. It is worth mentioning that the recently developed mortar based methods require a specific symmetric detection; a method based on the bounding volume trees has been proposed by B. Yang and T.A. Laursen [Yang 08a], [Yang 08b].

The goal of the detection phase in implicit analysis consists in creating “contact elements” – abstract (not structural) elements, which include (in case of a node-to-segment discretization) a slave node and several master nodes united by a master surface segment<sup>2</sup>; normally this master segment should be the closest master element and the slave node has a normal projection on it. The simplest and straightforward method is the all-to-all detection: all slave nodes are projected on all master segments and if for the current slave node one or several projections exist, the closest master segment is chosen. The growth rate of the method is  $O(N_s \times N_m)$ , where  $N_s$  and  $N_m$  are the numbers of slave nodes and master segments respectively. If the master and slave discretization is unknown a priori, the same problem will require four times more time to achieve detection. If one considers a master surface formed by quadrilateral segments<sup>3</sup>, then the determination of the projection requires the

---

<sup>2</sup>in case of Lagrange multiplier or coupled augmented Lagrangian methods, contact elements contain also some additional degrees of freedom (Lagrange multipliers) – one in case of frictionless contact and 2 or 3 in frictional case in 2D and 3D respectively. To keep track of stick-slip-stick behavior it is sometimes necessary to store also an internal history variable.

<sup>3</sup>second order master surface in 2D, or any master surface in 3D except that surface formed from triangles requires an iterative process to determine the projection point.

solution of a nonlinear equation and takes several iterations, however often only one iteration is required to realize that there is no projection.

The first generally accepted simplification of the detection is to start from the detection of the closest node instead of the closest segment. Although the growth rate remains the same  $O(N_s \times N_m)$ , where  $N_m$  now denotes the number of master nodes, the time needed to achieve such a detection is significantly smaller. After detection of the closest master node, the slave node is projected on the master segments adjacent to the determined master node. For each slave node, the closest master segment is then established if it exists. Many sorting algorithms from spatial search can be used for local detection in this framework, based on the search of the closest node.

So there are two strategies: closest-node-and-adjacent-segment and closest-segment. However, both of them are time-consuming and not robust (especially the strategy based on the closest node detection, see remark below). Although detection methods based on the closest node strategy have been used for many years, these methods can not be easily improved to perform a correct contact detection in all cases. That is why we will not base the detection procedure on this strategy. The rigorous formulation of the closest point given in the previous chapter will be exploited.

Let us imagine a set of spatially distributed compact objects (nodes, segments). The problem is to detect for a given object the closest one from this set. Human vision accomplishes this task easily by analyzing just a few objects. It does not need any analysis of the whole set of objects while the simple detection algorithm does, because it is “blind” and needs to “touch” all the objects one by one and compare distances between them (e.g., between their centers of mass). The techniques which have been worked out for contact detection are aimed at reducing the quantity of points to “touch”: bucket sort [Benson 90], [Fujun 00], the heap sort, the Octree method [Williams 99] and others.

Another improvement to reduce the detection time in the master-slave approach consists in considering only those parts of the contact surfaces which are situated in a limited zone (bounding box), where contact can occur in the current solution step. This zone can be confined to a bounding box that will be updated during the computation. It is proposed to determine this detection bounding box as an intersection of bounding boxes spanned on master and slave surfaces. Another improvement, proposed in [Benson 90], consists in a smart update of the closest element : an expensive detection procedure is carried out only in the first time step (for each slave node the closest master element (node, segment) is detected). For the following computational steps, slaves nodes are checked for proximity only with master elements neighboring to the previously determined. However, this method is not very general and sometimes fails. These improvements are not applicable for contact detection in case of unknown a priori master-segment discretization.

As already mentioned, the main difference in detection in explicit and implicit simulations is that for the second case it is necessary to predict possible contact occurrence and to establish contact elements before penetration occurs. This implies that slave nodes approaching master surfaces have to be detected

at a certain distance, the maximal detection distance (MDD), which is an important parameter of the detection procedure. It is good to know that the meaning of the MDD for closest-node and closest-segment based procedures is quite different and will be discussed later.

Further we give a detailed description, analysis and validation of different contact detection techniques. We start from the development of a robust all-to-all detection technique, which is acceptable for a moderate number of nodes in contact. Two strategies are considered based on the closest node and the rigorous closest point definition. Multi-face contact elements, inspired from [Heegaard 93] and [Barboteu 02] are introduced. Some important remarks will be given on the closest node based detection and on the relation between the maximal detection distance and the mesh size. Further, the bucket sort method is considered in details, the optimal bucket size is deduced and validated numerically. An extension of the considered methods to contact in case of unknown a priori master-slave discretization is discussed. Another contribution of this chapter is an extension of the bucket detection method to parallel framework inspired from the so-called Linked Cell Method widely used in molecular dynamic simulations for short-range interactions [Griebel 07]. Some tests of contact detection for very large problems are also presented.

To conclude this introduction, it is worth mentioning that there is a strong correlation between robustness, accuracy of contact detection and the CPU time. This dependence is not always inversely proportional. Sometimes one can sacrifice robustness to keep things simple, a good example is the “closest node” strategy complemented with bucket sort [Benson 90] - a quite simple and rather robust strategy for quadrilateral meshes. However, to preserve both accuracy of the contact detection and simplicity of the algorithm, we propose a new robust and fast detection algorithm based on the rigorous formulation of the closest point and on the bucket sort.

## 3.2 All-to-all detection

All-to-all algorithms are easy and fast to implement but long to execute. Their growth rate is  $O(N_s \times N_m)$ . Their straightforward implementation is not acceptable for large applications, however some simple improvements mentioned above can significantly improve their performance.

### 3.2.1 Preliminary phase

First of all, the key parameter for the contact detection – maximal detection distance (MDD)  $d_{\max}$  – has to be introduced. In case of node-to-segment detection,  $d_{\max}$  determines the following: if a slave node is closer to the master surface than  $d_{\max}$ , then it is supposed that this node can come in contact during the following time step, otherwise not. If one considers a node-to-node detection technique, then the meaning of the maximal detection distance is different. If the distance between a slave  $\underline{r}_i$  node and a master node  $\underline{r}_j$



$d_{ij} = \text{dist}(\underline{r}_i, \underline{r}_j)$ <sup>4</sup> is smaller than the maximal detection distance, then the corresponding slave node  $\underline{r}_j$  and one of the master surfaces containing the mentioned node  $\underline{r}_j$  as its vertex are considered to be potentially in contact during the following time step, otherwise not. This difference naturally results in a limitation on the minimal value of the  $d_{\max}$  for the closest node based detection: MDD has to be greater than one half of the maximal distance between master nodes attached to one segment

$$d_{\max} > \frac{1}{2} \max_{i=1, j=1, k=j+1}^{i=N_m, j=N_n^i-1, k=N_n^i} \text{dist}(\underline{r}_j^i, \underline{r}_k^i), \quad (3.1)$$

where  $N_m$  is a total number of master segments,  $N_n^i$  is a total number of master nodes attached to the  $i$ -th master segment,  $\underline{r}_j^i$  is a coordinate of the  $j$ -th node of the  $i$ -th master segment. If the condition (3.1) is not fulfilled, then some slave nodes coming in contact with the master surface can be lost (see Fig. 3.2)<sup>5</sup>. The value of  $d_{\max}$  can be determined automatically according to the discretization of the master or self-contact surface and to the maximal displacement of nodes on contact interface during one time step. The MDD should be kept as small as possible in order to accelerate the detection procedure and to avoid the creation of non necessary contact elements. For simplicity, it is proposed to keep the MDD unique for the entire contact area. In contrast to closest node based detection, the MDD for the closest point based procedure has no connection with the master surface discretization and can be chosen for all-to-all procedure only according to the maximal displacement of contact nodes. So the following discussions on the optimal choice of the MDD relate to the all-to-all closest node based detection and to all detection strategies in the framework of the bucket sort, which will be considered later.

For a reasonable number of time steps, in a geometrically or physically nonlinear problem, the maximal detection distance can be determined as the dimension of the largest master segment

$$d_{\max} = \max_{i=1, j=1, k=j+1}^{i=N_m, j=N_n^i-1, k=N_n^i} |\underline{r}_j^i - \underline{r}_k^i|. \quad (3.2)$$

Such an estimation is reasonable in case of a regular discretization of the master surface. On the other hand if the distribution of the master nodes is very heterogeneous, i.e. fine surface mesh in one contact region and rough in another, the value of  $d_{\max}$  appears to be highly overestimated for certain regions. This fact decreases the efficiency of the method, but in general for an adequate finite element mesh the increase of the detection time is not so high. The influence of the maximal detection distance on detection time will be discussed later. The possibility of performing an automatic choice of the MDD is of a big practical importance.

In the case of linearly elastic material and frictionless contact, the geometry can change significantly during one time step. So the analysis of the

<sup>4</sup>Here the  $\text{dist}(\underline{r}_i, \underline{r}_j)$  denotes Euclidean metric in the global reference frame  $\text{dist}(\underline{r}_i, \underline{r}_j) = |\underline{r}_i - \underline{r}_j|$ .

<sup>5</sup>Here and further for the sake of simplicity and clarity, almost all figures represent two dimensional cases but can be easily extended to three dimensions.

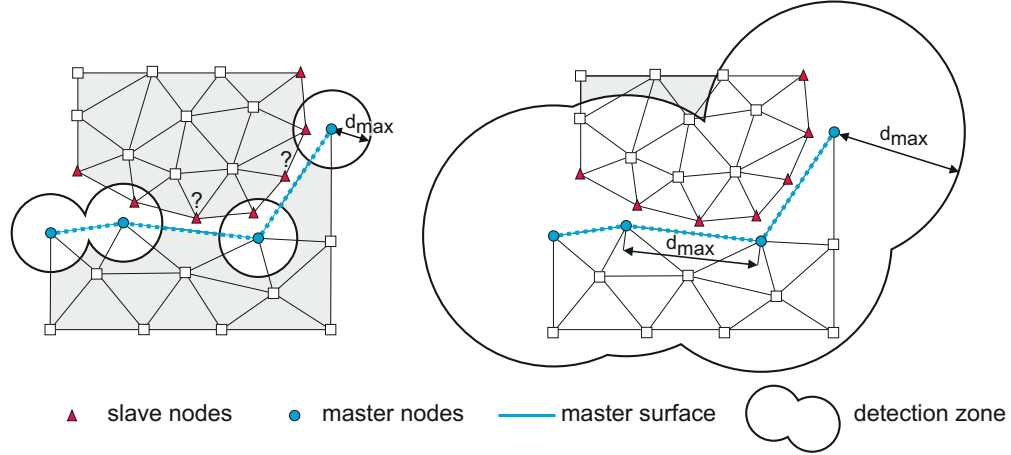


Figure 3.2: Maximal detection distance  $d_{\max}$  for closest node base detection strategies: on the left, not correct; on the right, correct choice.

discretization can only provide a lower bound for  $d_{\max}$  and that is why its value should be augmented manually or automatically according to the deformation and/or displacement rate, for example in the following way

$$d_{\max} = \max \left\{ \max_{\substack{i=N_m, j=N_n^i-1, k=N_n^i \\ i=1, j=1, k=j+1}} |\underline{r}_j^i - \underline{r}_k^i|; 2 \max_{i=1}^{N_c} |\Delta \underline{r}_i| \right\}, \quad (3.3)$$

where  $N_c$  is a total number of slave and master nodes and  $\Delta \underline{r}_i$  is the displacement of the  $i$ -th node. The factor 2 takes care of possible opposite translations of master and slave nodes. In case of remeshing or sufficiently large deformations of the master, the detection parameter  $d_{\max}$  should be recomputed at each remeshing or at each  $N$ -th time step.

To accelerate the procedure before carrying out any detection, the spatial area where contact can take place during the following time step can be limited. It has to contain as few master and slave nodes as possible but obviously it has to include all the nodes potentially coming in contact during the next step. If needed this area has to be frequently updated. We propose to confine this area by a parallelepiped bounding box defined in the global reference frame.

The determination of the bounding box differs for known a priori and unknown master-slave discretizations. In case of unknown master-slave, the bounding box should include all possible contacting surfaces. But frequently the discretization is known a priori even if contact occurs within one body (self-contact). In this case the construction of an optimal bounding box allows to exclude from consideration some nodes which cannot come in contact during the next time step (Fig. 3.3) and consequently it results in an acceleration of the detection procedure. It is worth mentioning that the most general case, where any slave node can potentially come in contact with any master segment during the loading is considered in the discussion. Often, this is not the case and for each slave node the set of possible master segments is limited and partly



predefined. But to take this limitation into account, the detection technique should be tuned for each particular case, which is impractical.

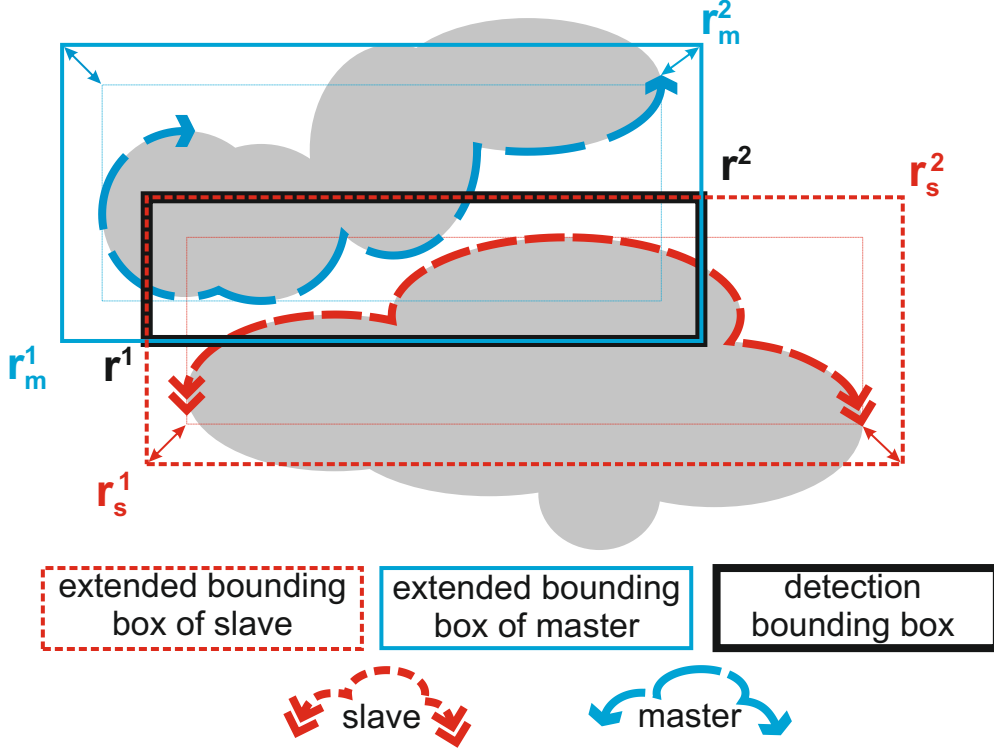


Figure 3.3: Determination of the bounding box for the contact detection procedure in case of a known master-slave discretization.

First of all the dimensions of master and slave surfaces are estimated. It is proposed to construct two independent bounding boxes  $B_s : \{\underline{r}_s^1, \underline{r}_s^2\}$  and  $B_m : \{\underline{r}_m^1, \underline{r}_m^2\}$  containing all slave and master nodes respectively, where  $\underline{r}_s^1$  and  $\underline{r}_m^2$  are the vectors in the global reference frame of two opposite corners determining the bounding boxes. Note that each bounding box confining master and slave nodes includes also a node free margin zone, the size of which is the maximal detection distance at each side:

$$\underline{r}_s^1 : \underline{r}_{\{x,y,z\}}^1 = \min_{i=1}^{N_b} \{\underline{e}_{\{x,y,z\}} \cdot \underline{r}_i\} - d_{\max} \underline{e}_{\{x,y,z\}}, \quad \underline{r}_s^2 : \underline{r}_{\{x,y,z\}}^2 = \max_{i=1}^{N_b} \{\underline{e}_{\{x,y,z\}} \cdot \underline{r}_i\} + d_{\max} \underline{e}_{\{x,y,z\}}, \quad (3.4)$$

where  $N_b$  is a number and  $\underline{r}_i$  is a vector of nodes to be included in the bounding box and  $\underline{e}_{\{x,y,z\}}$  are unit vectors in the global reference frame. The margin of  $\pm d_{\max}$  is introduced to avoid any loss of possible contact elements. Some improvements can be introduced in order to reduce the time needed for the construction of the bounding box. The user can precise that one or several contact surfaces are rigid and do not move, then permanent bounding boxes can be assigned to these surfaces and there is no need to update them. Another possible feature is the prediction by the user that the deformation and

displacement of a contact surface is connected to the displacement of certain nodes. It allows to avoid the verification of all nodes in (3.4). Since the nodal coordinates are stored in memory in the global reference frame, it is much faster to work directly with these coordinates, so no rotation to the bounding boxes must be applied. The resulting bounding box  $B : \{\underline{r}^1, \underline{r}^2\}$  is taken as the intersection of the master and slave bounding boxes  $B = B^m \cap B^s$ . The practice shows that a further contraction of the bounding box does not reduce significantly the detection time. The construction of the bounding box and the verification of the presence of nodes and segments inside the bounding box requires about  $2(N_s + N_m)$  operations.

### 3.2.2 Detection phase

In order to accelerate the detection procedure, one can detect first for each slave node the closest master node (if there are master nodes closer than MDD). It is then enough to find the projection of the slave node on the master segments having this master node as a vertex. The case when only one projection is found is trivial. It remains to create the corresponding contact element spanned on the slave node and the master surface possessing this projection. If several projections are found, the closest one is retained to create a contact element. The case when no projection is found has to be considered in details. There are two possibilities:

1. the slave node is situated in a “blind spot” of the discretized master surface;
2. the slave node does not come in contact but just passes by close to an edge of the master surface.

Let us remind the reader some facts from the previous chapter. Since the finite element method requires only continuity of the discretization ( $\Gamma_c \in C^0$ ) the contacting surface may be not smooth ( $\Gamma_c \notin C^1$ ). Each master segment has its “normal projection” zone (Fig. 3.4), each point in this zone has at least one normal projection onto the master surface. But often in the junction zone of the master segments (at common edges and nodes) the intersections of the “normal projection” zones do not fill the surrounding space entirely. Some gaps are left, in form of prisms and pyramids in 3D or of sectors in 2D. This problem exists not only for linear but for any order master elements. As it has been discussed in the previous chapter, the rigorous definition of the closest point does not possess such a problem: if there is no normal projection on the master segments, the closest point is situated on the edges or on the closest master nodes or it does not exist. However, if one confines to “normal projection” as usual for the NTS discretization, it is necessary to find the closest master segment to establish the contact element. For the sake of generality we will consider the case when only normal projection on master segments is checked, which naturally yields to blind spots.

Three types of blind spots can be distinguished: internal, external or blind spot due to boundary conditions (see Fig. 3.4). If a slave node in a blind spot

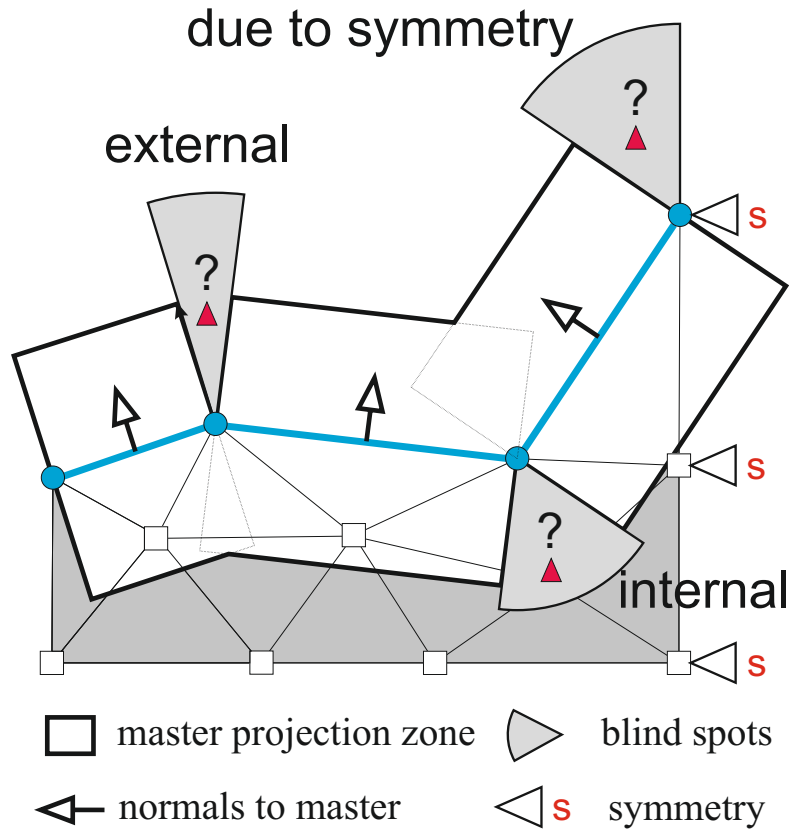


Figure 3.4: Examples of blind spots: external, internal and due to the symmetry boundary conditions.

is overlooked, different consequences depending on the type of blind spot are possible.

- External blind spot. Slave nodes situated in this kind of spot are not detected before they penetrate under the master surface. After such a penetration, it can be detected during the next time step and brought back onto the surface, but the solution has been already slightly changed. In certain cases especially in force driven problems such a penetration can lead to a failure of the solution algorithm.
- Internal blind spot. Contact is predicted correctly, but if a slave node penetrates just a little under the master surface and appears in its internal blind spot, this node will be lost for the contact detection at least during the next time step. Such little penetrations take place if the penalty method for contact resolution is used or just due to the limited precision of the iterative solution.
- Blind spot due to boundary conditions. This type of blind spot is situated at the boundary and can be either internal or external. It appears due

to the presence of symmetric or periodic boundary conditions on the master surface, for example the basic Hertz contact problem with an axisymmetric 2D finite element mesh.

In explicit formulation the internal blind spots are the most dangerous, for implicit – external blind spots.

There are different possibilities to avoid the loss of contact in blind spots:

- Artificial smoothing of master surface for large sliding contact problems [Pietrzak 99], [Wriggers 01], [Krstulović-Opara 02] etc. There are no more gaps in “projection” zones except gaps due to symmetry, i.e. there are almost no more blind spots and the problem of passing by nodes does not exist. However, most of these methods have some inherent drawbacks, e.g. the deformation obtained close to the edge of the active contact zone may be erroneous.
- Master segments can be extended in all directions to cover gaps in the normal projection zone.
- A “proximal volume” can be constructed by an extrusion of the master surface in the normal direction and in the opposite one which fills both projection zones and blind spots. If a slave node is situated in this volume then it is considered as node in contact and the master surface is further detected. “Passing by” nodes can be easily detected as they do not appear in the “proximal volume”.

The first group of methods in general is too “expensive” if one uses them only for the detection purpose. The methods are not applicable for arbitrary meshes. The second group is good and reliable for linear elements. The third method is also quite time consuming.

We use here a rather rough but quite simple and robust treatment of blind spots. If a detected slave node has no projection and is not a passing by node, then the corresponding contact element is constructed with the closest (see [Zavarise 09b]) or randomly chosen master surface attached to the closest master node. For sufficiently small time steps, such an approach is quite reliable. It remains only to determine if the node is passing by or not. One possible technique is represented in Figure 3.5.

First of all, in the preliminary phase, the boundary master nodes surrounding the master contact surface have to be marked. Let us assume that for one of such marked nodes  $\underline{r}_m$ , the closest slave node  $\underline{r}_s$  has been found. If it has no projection onto the master segments attached to the marked master node, two alternatives are possible: either the slave node is situated in a blind spot, or it passes by the master surface. To choose between these two cases, one has to verify if the slave node is located in one of the blind spots attached to the master node, or to check if the slave node is in the local “proximal” volume of the master surface. The second verification seems to be more simple and natural. Note that such a verification is slightly different for locally convex and concave master surface boundaries. The convexity can be known as nodes of

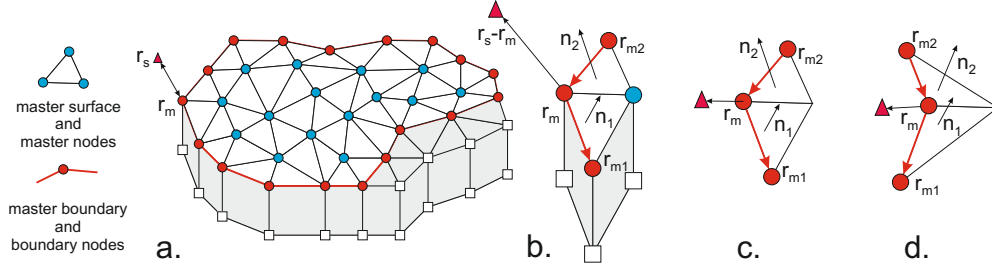


Figure 3.5: Detection of passing by nodes. a – master surface and its boundary; b – zoom on the geometry close to the passing by node; c – convex master boundary; d – concave master boundary.

each master segment are ordered. The condition of convexity is

$$(\mathbf{r}_m - \mathbf{r}_{m2}) \times (\mathbf{r}_{m1} - \mathbf{r}_m) \cdot (\mathbf{n}_1 + \mathbf{n}_2) \geq 0, \quad (3.5)$$

where  $\mathbf{n}_1$  and  $\mathbf{n}_2$  are respectively the average normals to the master segments possessing the edges  $\{\mathbf{r}_m, \mathbf{r}_{m1}\}$  and  $\{\mathbf{r}_m, \mathbf{r}_{m2}\}$ . The criterion of the slave node being in the proximal volume is then

$$\mathbf{n}_2 \times (\mathbf{r}_m - \mathbf{r}_{m2}) \cdot (\mathbf{r}_s - \mathbf{r}_m) \geq 0 \text{ AND } \mathbf{n}_1 \times (\mathbf{r}_{m1} - \mathbf{r}_m) \cdot (\mathbf{r}_s - \mathbf{r}_m) \geq 0 \quad (3.6)$$

If this condition is fulfilled, the slave node is taken into account and the contact element is established with the closest master segment. For the concave surface AND in Eq. (3.5) should be replaced by OR.

As one can see, considering only normal projections on master segments results in difficulties in detection of slave nodes in blind spots and in case of passing by nodes. However, these difficulties are small compared to the main drawback of the contact detection based on the closest node, such a detection is not robust and may fail for nonregular meshes (see Fig. 3.6 and Remark 3.1). So we propose not to start the detection procedure from searching for the closest master node for each slave. According to the rigorous definition, the closest point can be either on the master segments or master edges or master nodes, so for the purpose of robustness it is recommended to search directly the closest point, i.e. to check the projection on all master segments, all master edges except edges surrounding the master contact zone and all master nodes. It allows to avoid problems associated with blind spots and passing by nodes. It increases significantly the robustness of the detection as well.

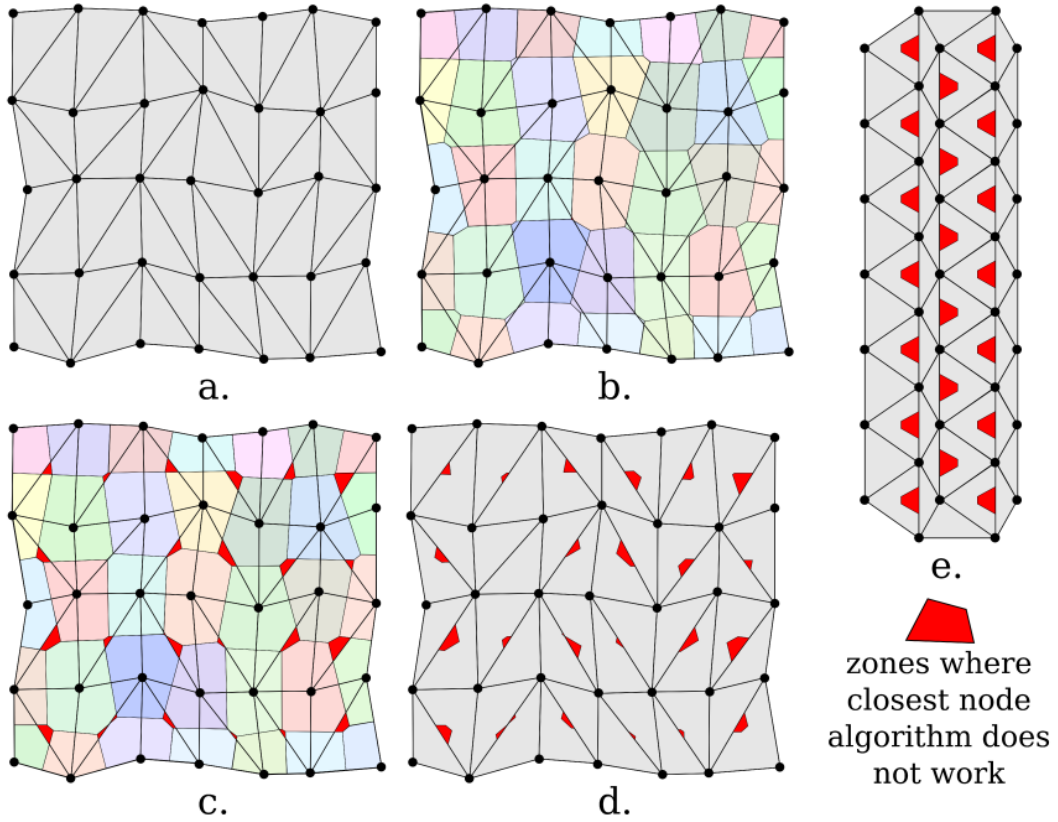


Figure 3.6: Examples of failure of the contact detection algorithm based on the closest node: **a** – flat triangular mesh of master surface; **b** – colors correspond to the proximity zones of the master nodes; **c,d** – if a slave node is situated in a red region, it has its projection on the master segment which is not attached to its closest master node; **e** – on this particular mesh configuration, the fraction of “bad” zones is over 15%.



**Remark 3.1 on the contact detection based on the closest node.**

*The detection based on the closest node takes place in two stages:*

1. *for each slave node, the closest master node is determined;*
2. *if the master node is close enough, then one supposes that the “normal projection” of the slave node onto the master surface is located on one of the master segments that have the master node as one of their vertices; the projection point is then calculated.*

As mentioned in [Benson 90], the fundamental assumption of such an algorithm (the normal projection of the slave node is situated on master segments adjacent to the closest master node) is not always correct. The slave node can penetrate a master segment which is not attached to its closest master node. To demonstrate it the authors give an example of highly distorted quadrilateral surface mesh. However, to keep the detection algorithm fast, they accept the risk. There are several arguments in defense of their choice: they are very experienced users, they used regular quadrilateral mesh and a two pass\* penalty method. These facts reduce the risk almost to zero.

However, for the general case and one pass resolution algorithms, the situation is more dramatic and the risk to overlook penetrations is high, especially for the triangular surface mesh of the master. Even a flat and quite regular surface mesh consisting of triangles may have many zones where slave nodes can penetrate under master segments which are not attached to the closest master nodes (see the figures below and Fig. 3.6). For non-flat meshes, the risk to miss or to create incorrect contact element is significantly higher. That is why we recommend to avoid contact detection methods based on the closest master node.

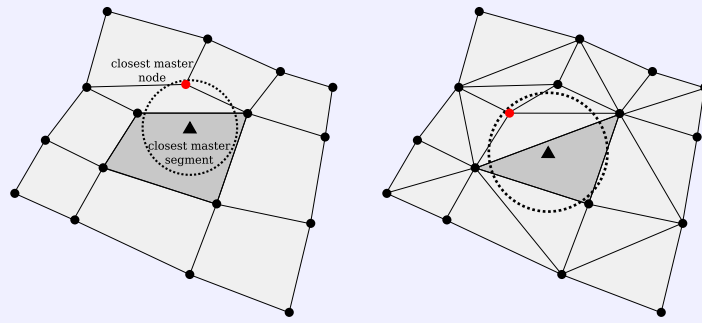


Fig. 3D examples of failure of a detection algorithm based on the closest node search: a flat distorted quadrilateral and triangular master meshes, triangle denotes a slave node for which closest master point is not situated on the master segments having the closest master node as its vertex.

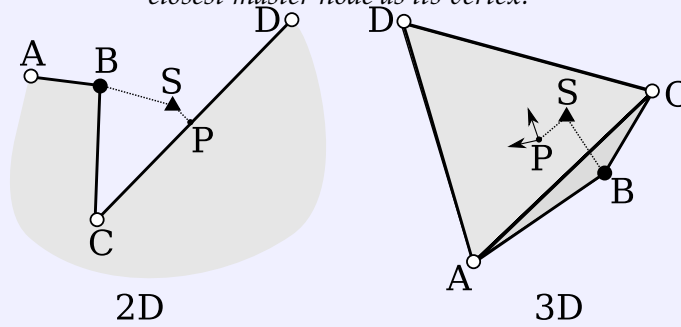


Fig. 2D case: the slave node S has its closest point on the master segment CD not attached to the closest node B; 3D case: the closest point to the slave node S is situated on the segment ACD not attached to the closest node B.

\* one node can be both master and slave during one increment.



### 3.3 Bucket sort detection based on the rigorous definition of the closest point

Bucket sort based on the closest node has been proposed by Benson and Hallquist [Benson 90] for DYNA3D ([Benson 07]). It has been slightly improved in [Fujun 00] and recently revisited with some improvements and several comments concerning parallelization in [Yastrebov 11a]. Here we adapt the bucket approach to account for the rigorous definition of the closest point from the previous section. Many details of the algorithm based on the closest node [Yastrebov 11a] remain unchanged. Nevertheless, the detection of the closest point requires now about 2-3 times more verifications and several times more floating point operations: the slave point is checked against master nodes, edges and elements. On the other hand it, is not necessary to check for passing by nodes and blind spots.

Before discussing particular details, let us derive a short description of the bucket sort detection method. As previously, two phases can be distinguished. In the preliminary phase, the optimal size of the bucket is evaluated, then a potential contact area is determined and split in buckets (cells) by an enumerated regular grid. That allows to reduce locally the area of the closest point search. Finally all slave and master components (nodes, edges, segments) situated in the detection area are distributed in the cells of the grid. In the detection phase, for each slave node, we check for the closest master component in the current cell. If needed, we check one or several neighboring cells for possible proximal master components. Contact element is finally created in a special manner.

#### 3.3.1 Preliminary phase

First of all the maximal detection distance  $d_{\max}$  is chosen equal to the maximal size of the biggest segment of the master surface ( $N_m$  operations). The spatial bounding box is then determined as the intersection of the master and slave bounding boxes exactly as described in the previous section ( $N_m + N_s$  operations). An internal grid should be constructed in a proper way: it should be regular and the cell size  $w$  should be optimum, that is not too large in order to keep the number of slave and master nodes in the cell as small as possible. On the other hand, if one limits the operations to checking only one neighboring layer of cells (9 cells in 2D and 27 in 3D), it can be shown that the cell side must be greater than  $w \geq \sqrt{2}d_{\max}$  for linear elements in order not to overlook segments like  $AB$  for slave node  $S$  in Fig. 3.7. For quadratic elements the limitation depends on the maximal curvature of the master segments, but the rough estimation  $w > \sqrt{2}d_{\max}$  seems to be reliable. If the size of the cell is smaller, more than one layer of neighboring cells has to be checked. It complicates significantly the coding of the algorithm and moreover the growth rate of the maximal number of cells to be checked  $N_c$  is cubical

$$\text{if } w = \frac{d_{\max}}{n}, n > 1 \Rightarrow N_c = (3 + 2n)^3 \quad (3.7)$$



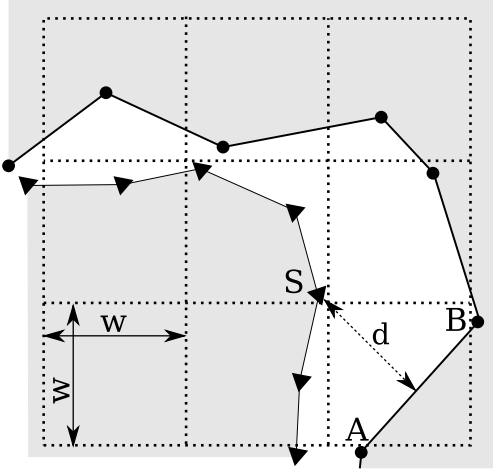


Figure 3.7: Example of 9 buckets (cells) and master-slave solids, triangles - slave nodes, circles - master nodes. Distance between slave node  $S$  and master segment  $AB$  (which is not included in the right bottom bucket) is higher than MDD if cell side  $w > \sqrt{2}d_{\max}$ : there is no risk to overlook possible contact.

The smaller the cell size, the higher the total number of cells and consequently the smaller the number of contact nodes per cell. On the other hand, the small cell size increases the need to carry out the detection in neighboring cells. It can be shown analytically by means of probability methods that for even distribution of master and slave elements both in 2D and 3D cases the minimal detection time is unique and corresponds to the minimal possible cell size. Such a simple analysis predicts a quadratic growth of the detection time in 2D case and cubic in 3D. However, in real simulations, the distribution of the contact elements is not even and so some numerical tests have been performed in order to investigate the dependence of the detection CPU time  $t$  on the cell size  $w$ .

### 3.3.2 Numerical tests

For the purpose of the optimal cell size definition, several artificial finite element meshes have been considered. Slave and master surfaces consist of about 10200 nodes each. Three sets have been considered: proximal meshes with homogeneous (Fig. 3.8, left top) and heterogeneous (Fig. 3.8, left bottom) node distributions and a convex mesh with a heterogeneous node distribution (Fig. 3.8, right). Each set is represented by five different realizations of curved surfaces. By homogeneous node distribution, we mean that the maximal segment dimension does not exceed 200% of the minimal one. In heterogeneous case, this difference reaches 700%.

In Figure 3.9, the dependence of the average detection CPU time and the average number of investigated neighboring cells on the normalized cell size

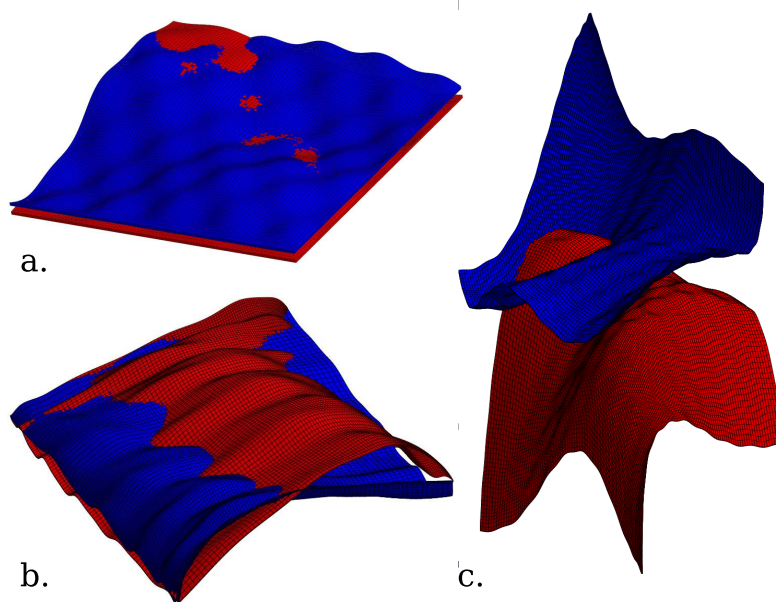


Figure 3.8: Example of finite element meshes used to determine the optimal cell size: proximal meshes with homogeneous (a) and heterogeneous (b) spatial node distribution and convex meshes (c).

$w/d_{\max}$  is represented for different sets. As expected, the detection time for convex meshes is smaller because of the smaller associated bounding boxes. Different discretizations (256x256, 512x512) have been tested. In all the cases, the same dependence takes place. According to the analytical estimation and to the test results, the optimal grid size is the minimal one and equal to the maximal detection distance amplified by  $\sqrt{2}$

$$w = \sqrt{2}d_{\max} \quad (3.8)$$

For such a choice, each grid cell contains the minimal number of elements, but on the other hand it is necessary to carry out the detection procedure in many neighboring cells: in average 12-16 cells from 26 surrounding cells in 3D (Fig. 3.9).

This investigation allows to determine automatically the optimal maximal detection distance and the size of the detection cell depending on the discretization of the master surface. It makes the algorithm user friendly and in most cases accelerates the computation. However, sometimes it is necessary to keep the maximal detection distance significantly smaller than the biggest master element. Such a situation arises in complex self-contact problems with unknown a priori master-slave discretization. In this case, the detection cell size is kept equal to the dimension of the largest master element  $d$  multiplied by  $\sqrt{2}$  and the MDD is set smaller than  $d$ .

When the maximal detection distance is determined and the bounding box is constructed, the internal grid has to be established in the bounding box

-

Figure 3.9: The dependence of the detection time and the average number of neighboring cell investigated during the detection on the normalized cell size.

and further all slave nodes situated in the bounding box as well as all master components (nodes, edges and segments) have to be distributed in the cells of the grid. Since the optimal cell size  $w_{\min} = \sqrt{2}d_{\max}$ , the number of cells in each dimension of the grid is defined as

$$N_{x,y,z} = \max \left\{ \left\lfloor \frac{r_{x,y,z}^2 - r_{x,y,z}^1}{\sqrt{2}d_{\max}} \right\rfloor; 1 \right\}, \quad (3.9)$$

where  $[x]$  stands for the integer part of  $x$ . Such a choice of cell numbers provides the size of the cell  $\Delta x$ ,  $\Delta y$  and  $\Delta z$  not smaller than  $w_{\min}$  at least in case of  $N > 1$

$$\Delta\{x, y, z\} = \frac{r_{x,y,z}^2 - r_{x,y,z}^1}{N_{x,y,z}} \geq \sqrt{2}d_{\max}. \quad (3.10)$$

Each cell of the grid has to be enumerated, the unique integer number  $N \in [0; N_x \times N_y \times N_z - 1]$  is given to each cell with spatial “coordinates”  $i_x$ ,  $i_y$  and  $i_z$ , where  $i_{x,y,z} \in [0; N_{x,y,z} - 1]$

$$N = i_x + i_y N_x + i_z N_x N_y. \quad (3.11)$$

Now the growth rate of the method can be estimated roughly as  $O\left(\frac{N_s N_m}{N_x N_y N_z}\right)$ . If the average number of master and slave nodes per cell is supposed to be

constant<sup>6</sup>  $\varrho = \frac{N}{N_c}$ , where  $N_c = N_x N_y N_z$  and  $N$  is an average number of master and slave nodes, then the growth rate of the method can be rewritten as  $O(N)$ . However, in practice the distribution of nodes is not even and the growth rate is higher. Consequently the clustering of nodes and the nonregularity of the mesh influences significantly the performance of the method.

Slave nodes and master components situated in the bounding box have to be distributed in the cells. For this purpose several arrays  $A^s$  and  $A_n^m, A_e^m, A_s^m$  are created corresponding to slave nodes, master nodes, edges and segments respectively. These arrays contain identification numbers (IDs) of components which are contained in each cell: IDs of slave and master nodes, and edge segments. Another possibility is to construct only one array for master components which will include only master nodes, since edges and segments can be derived from this information; but in this case it is necessary to avoid identical verifications. For example, element  $A^s[i, j]$  keeps the ID of the  $j$ -th slave node in the  $i$ -th cell of the grid,  $i \in [0; N_x N_y N_z - 1]$ ,  $j \in [0; N_i^s]$ ,  $N_i^s$  being the number of slave nodes in the  $i$ -th cell. In average, the number of integer (32 bits) elements in an array does not exceed the number of contact nodes and so even for extremely large problems it makes just a minor contribution in memory requirement. However, the arrays can be replaced by linked-list storages as in [Fujun 00].

For each node with coordinates  $\underline{r} : \{r_x, r_y, r_z\}$  inside the bounding box, the corresponding cell number is easily determined as

$$N_{\text{cell}} = \left\lfloor \frac{r_x - r_x^1}{\Delta x} \right\rfloor + \left\lfloor \frac{r_y - r_y^1}{\Delta y} \right\rfloor N_x + \left\lfloor \frac{r_z - r_z^1}{\Delta z} \right\rfloor N_x N_y, \quad (3.12)$$

master segments and edges are supposed to be in the cell if at least one of their nodes is in the cell. So contrary to nodes, master segments and edges can be associated with several different cells.

#### 3.3.3 Detection phase

All steps described previously represent the preliminary part of the detection algorithm which demands in general 7 to 10 percent of the total detection time. The next steps of the algorithm correspond to the detection of the closest point and the construction of contact elements. For each grid cell  $c_i$  and for each slave node  $\underline{r}_{ij}^s$  in this cell, i.e. for each node with ID  $A^s[i, j]$  we look for the closest point on the master components associated with the current cell, i.e. the closest component among  $A_n^m[i], A_e^m[i], A_s^m[i]$  if they are not empty. Let us suppose that among all master components in the cell the distance to the closest point is  $d_{ij}^s \leq d_{\max}$ . It is obvious that the master components situated in neighboring cells (maximum 8 cells in 2D, 26 in 3D) have to be checked as well. Not all the cells are considered, but only those which boundaries are sufficiently close to the slave node. The criterion of the proximity is the following: if any boundary of the current cell (face, edge or vertex) is closer

<sup>6</sup>the number of master nodes per cell can be considered constant as cell size is proportional to the maximal distance between master nodes.

than the closest master node found previously, i.e. closer than  $d_{ij}^s$ , then the detection procedure has to be carried out in all neighboring cells attached to the boundary one by one (Fig. 3.10).

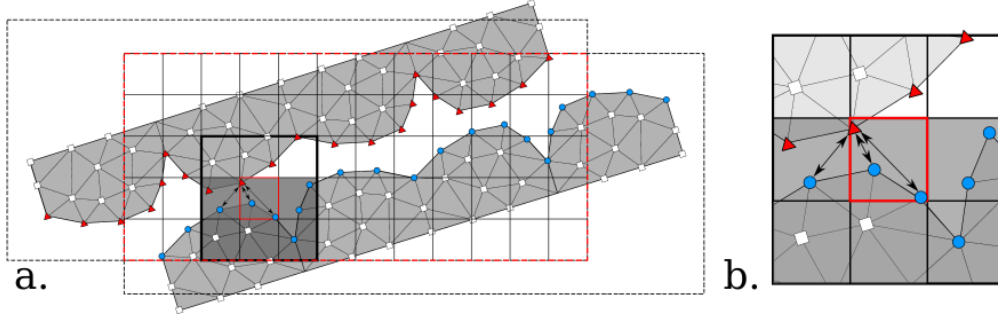


Figure 3.10: Detection of the closest master point: a – two meshes of slave and master solids; corresponding bounding boxes, their intersection and detection grid; the slave node (triangle) in the current cell is checked for proximity against all master components in the current and neighboring cells (in dark gray); b – zoom on the local detection region.

For example, let us consider a vertex of the  $i$ -th cell  $\underline{r}_i^v$ . For instance, after checking all master components in the current cell, we find that the closest component is remote at  $d_{ij}^s$  from the slave. Then if the considered slave node is closer to the vertex than this distance, i.e.  $\underline{r}_i^v: |\underline{r}_{ij}^s - \underline{r}_i^v| < d_{ij}^s$ , then all the master components in one of the neighboring cells attached to the vertex  $\underline{r}_i^v$  have to be considered and consequently  $d_{ij}^s$  has to be decreased or left unchanged if no closer master components is found in these cells. And so on for other cells attached to this corner. In general the same procedure has to be performed for all 8 vertices, 12 edges and 6 faces of the  $i$ -th cell. To get a more optimal algorithm, it is better if such an investigation of neighboring cells starts from the closest faces, further edges and finish the verification with vertices if needed. Note that each verified cell may decrease the  $d_{ij}^s$  and consequently can decrease the number of cells to be checked. In such a manner all possibly proximal slave and master nodes are detected cell by cell. The average number of verified neighboring cells for different meshes is represented in Figure 3.9. This number decreases with increasing normalized grid size  $1/\sqrt{2}w/d_{\max}$  but as the optimal ratio  $w/d_{\max} = \sqrt{2}$  the average number of verified neighboring cells remains quite high, typically 12-16 cells.

### 3.3.4 Multi-face contact elements

In this manner each slave node in the bounding box is checked and for some of them the closest master component has been found. Now let us discuss how to create NTS contact elements. Barboteu and Alart in [Barboteu 02] proposed to use multi-face NTS contact elements. Even earlier, it has been proposed by Heegaard and Curnier for large-slip contact [Heegaard 93]. This

idea is particularly interesting for accelerating the overall computation time, moreover in some cases it renders a more accurate algorithm. It consists in creating contact elements made of one slave node and several master segments, in order

- to avoid frequent updating of the contact elements;
- to treat large sliding (over more than one master segment) during one increment;
- to avoid random choice between two equally close master segments;
- to create a NTS element even when no normal projection has been found.

If the master-slave discretization is not too dense, then NTS contact elements can contain one slave node and all possible master segments, as done in [Alart 04]. Consequently, the values in the global stiffness matrix must be updated, but not its structure, which decreases significantly the total computational time. However, multi-face contact element should be used carefully due to the risk of infinite loop, if a slave node finds itself in a concave region of the master surface.

If a master component has been detected close enough to a given slave node, several segments attached to this master component are included in the contact element (see Fig. 3.11). If a slave node has the closest

- point on a master segment: the current master segment and several neighboring segments are included in contact element;
- point on a master edge: the master segments adjacent to this edge are included in contact;
- master node: the master segments having this node as one of their vertex are included in contact.

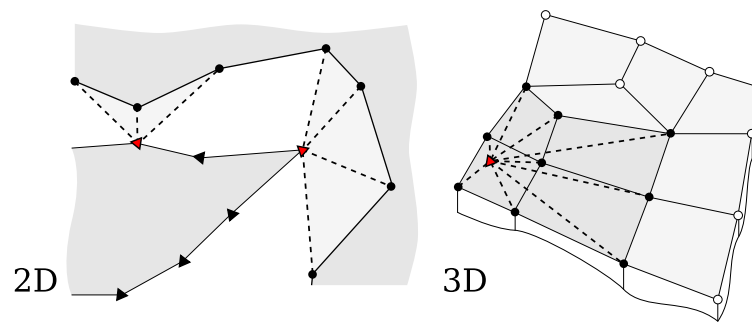


Figure 3.11: Examples of multi-face contact elements in 2D and 3D, dashed lines connect nodes of multi-face contact elements: triangles - slave nodes; circles - master nodes.

Master segments, nodes and edges in such an element can be either active (slave node passes over segment, edge or node) or passive. Only active

components introduce nonzero values to the residual vector and stiffness matrix.

### 3.3.5 Improvements

As proposed in [Benson 90], then used in [Pietrzak 97] and many others, in case of a known master-slave discretization, all slave and master components should be included in the detection procedure only once. Further contact elements which have been already created can be updated in a more efficient manner. If the slave node still has a projection on a master component included in its contact element, this element will not be removed. If not, the detection is performed only in a close vicinity of the former contact element and a new element is created. By close vicinity we mean master components neighboring to the former contact element. Normally it can be rather easily found as the mesh topology is known. If a slave node has no associated contact element on the previous time step, it should be included in the detection procedure. Obviously, the slave bounding box is computed only for such inactive slave nodes, and the master bounding box is still based on all master components. This technique appears to be quite efficient and accelerates significantly the detection phase and the matrix reconstruction.

Another technique consists in increasing the number of buckets, but since this number is connected to the discretization of the master surface, the minimal dimension of buckets is strictly limited. However, there is a way to solve this issue: several artificial nodes can be set on each master segment, which results in an automatic decrease of the limit on the MDD and consequently the number of buckets increases. Two examples are given in Figure 3.12.

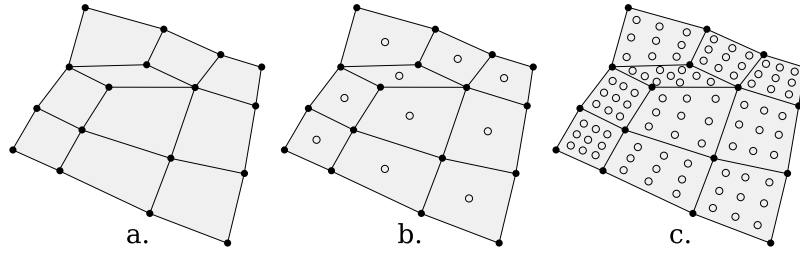


Figure 3.12: Examples of artificial nodes on the master surface: a – original master mesh, b – one supplementary node per segment; c – 9 supplementary nodes per segment.

## 3.4 Validation and performance

The preliminary validation of the proposed detection method based on the bucket sort is easy to carry out on simple meshes. Normally a visual analysis of the constructed contact elements is sufficient. A more secure validation



consists in comparing with the rigorous all-to-all detection which is trivial to implement.

To demonstrate the performance of the implemented detection, we consider a tire-road contact problem. Its simulation can be helpful for example for the improvement of tread patterns (stick increase and noise reduction). We are particularly interested in this problem because contact elements change intensively at each time step and consequently a fast detection procedure is highly desirable, moreover many problems in tire-road contact require transient analysis, i.e. a very large number of computational time steps.

A finely and regularly meshed tire wheel is translated over an artificially rough road surface<sup>7</sup> and its FE mesh is deformed manually accordingly to the road roughness, then the contact detection procedure is used. The finite element mesh of the tire (Fig. 3.13) consists of about 550 000 nodes with a contact zone of about 105 000 nodes. The finite element mesh approximating the road roughness (Fig. 3.14) consists of about 400 000 nodes, the half of them being included in the master contact zone. Established contact elements are shown in Figure 3.14 for different tire-road dispositions and imprint deep. It can be noted that the choice of the bounding box as an intersection of master and slave bounding boxes reduces significantly the number of contact nodes to be considered. The bounding box of the road is kept constant, whereas the bounding box of the tire is updated at each step. The contact detection time at each time step in average is just 1.5-2 seconds on a laptop, i.e. the contact detection time can be neglected in comparison to the system resolution time. The analysis of the detection time shows that the estimation of the maximal detection distance takes about 30% of the time, preliminary stage takes about 20% and the detection procedure requires just 50% of the time.

Another example is an artificial contact between two rough surfaces, each consisting of  $2^{20}$  contact nodes. Rendered surfaces corresponding to the meshes are represented in Figure 3.15. Such a kind of problems requires a longer time for contact detection since the bounding box includes all or almost all contact nodes and there are as many slave nodes as master ones. If one uses the all-to-all method based on the closest node, the reliable estimation of the reference detection time  $T_{\text{ref}}$  exceeds 180 hours (almost 8 days) and  $2^{40}$  distance verifications are needed. The proposed grid detection method requires much less time than the all-to-all method. The time strongly depends on the geometry and discretization, consequently on the constructed bounding box and the number of contact components located in it, for example, for close enough rough surfaces (Fig. 3.15) the detection time is much higher than for convex surfaces and it is almost negligible if both surfaces are not far from each other but not close enough to come in contact. The results are summarized in Table 3.3. Let us note that in both presented computations the quadrilateral master segments are supposed to remain flat for the considered methods.

---

<sup>7</sup>The finite element mesh of the road has been generated by a composition of trigonometrical functions.



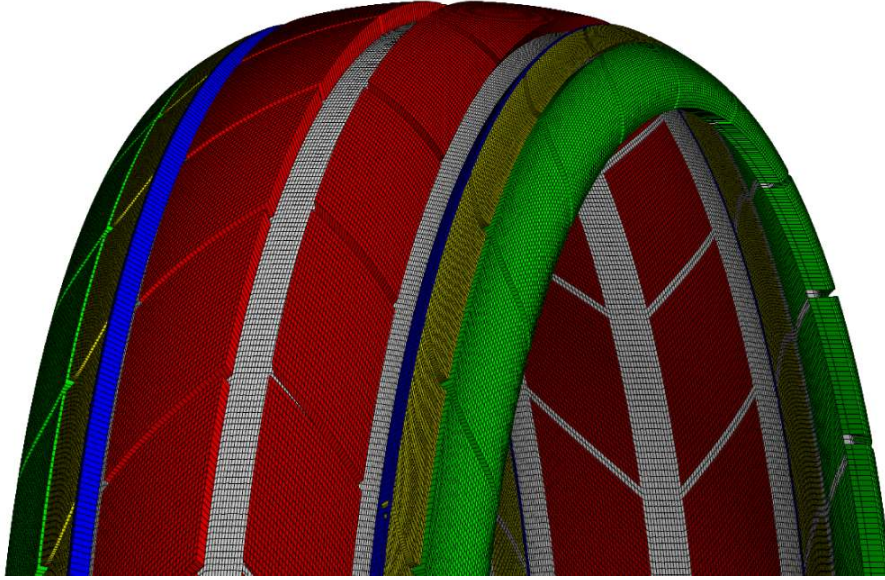


Figure 3.13: A part of the tire finite element mesh (full mesh contains 550 000 nodes, 105 000 slave nodes).

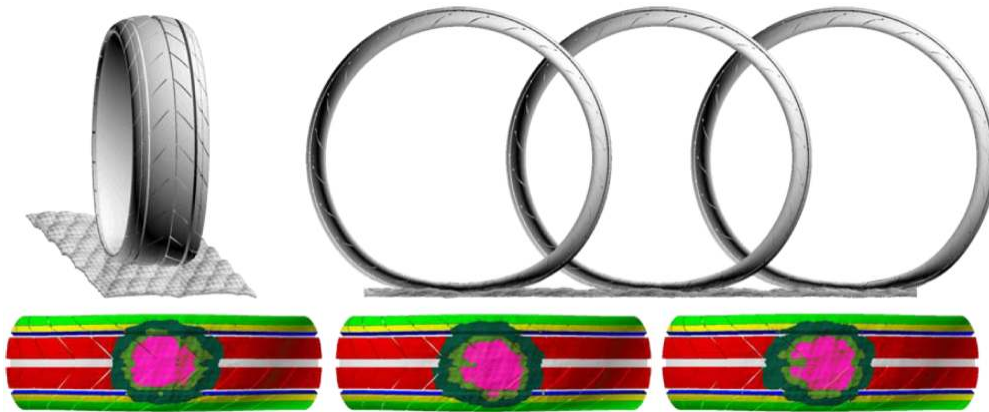


Figure 3.14: Tire-road contact problem: general view, three tire-road dispositions and corresponding contact elements on the bottom of the tire for different imprint configurations.

### 3.5 Case of unknown master-slave discretizations

There are mechanical problems for which the determination of master and slave surfaces presents a big challenge or may be impossible. Among such problems there are multi body systems, problems with complicated geometries (for example highly porous media like metal foams), large deformation problems with non regular discretization and self-contact problems.

This class of contact problems needs a particular contact detection

### 3.5 Case of unknown master-slave discretizations

Table 3.3: Detection of contact between rough surfaces (> 2 millions of master and slave nodes)

Geometry	Nodes in bounding box	Contact elements	Detection time, s	Gain, $T_{\text{ref}}/\text{CPU time}$
Two close surfaces	2 100 000	75 300	2280	>300
Two convex surfaces	340 000	15 800	72	>10 500
Two close but not contacting surfaces	50 000	0	4	>160 000

\*  $T_{\text{ref}} = 180$  hours

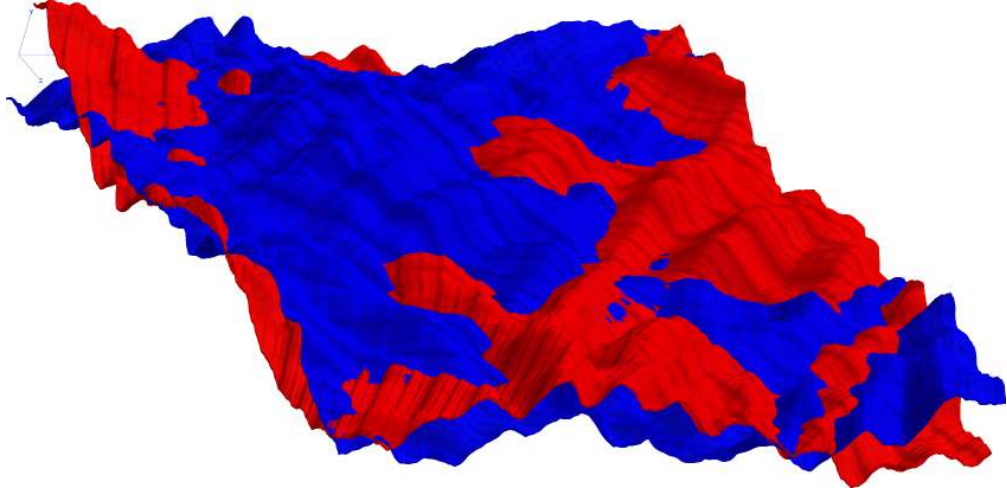


Figure 3.15: Rendered surfaces of two finite element meshes (each contains  $2^{20}$  contact nodes).

procedure. In the past the bucket sort has been successfully applied to contact detection and it is widely known as single surface contact algorithm [Benson 90]. Here a more accurate adaptation of the previously derived formulation for problem with unknown a priori master-slave discretization is proposed. A particular attention is paid to self-contact problems. Some examples will be given in Chapter 6. Such an adaptation demands considerable modifications in all stages of the grid detection procedure. The growth rate of the method is the same as for the case of known a priori master-slave discretization. The method is straightforward and adapted for NTS discretization. In case of a mortar-based formulation, the reader is referred to the recently developed technique proposed in [Yang 08a].

A self-contact is more probable for thin or oblong solids, for which one or two dimensions are much smaller than others, so often shell elements are used to simulate self-contact. But there is a challenge which is illustrated

in Figure 3.16. For a thin solid with one or two sided contact zones, it is complicated to detect the contact with the reverse side (circles in Fig. 3.16) even if in addition to node positions their normal vectors and corresponding surfaces are taken into account. Precisely, an ordinary detection algorithm would suppose that the circles are penetrating under the surface marked with another circle of the same color, consequently associated contact components have to be included in contact. A possible solution is to determine a maximal detection distance smaller than the double minimal thickness of the contacting structure, however very small time steps have to be made. A better solution has been proposed in [Benson 90], where the authors introduce an additional history variable to keep track of the side from which the contact surface has been penetrated. This approach allows the MDD to overpass the thickness of the structure. Remark that this problem is more severe if both surfaces can contact each other.

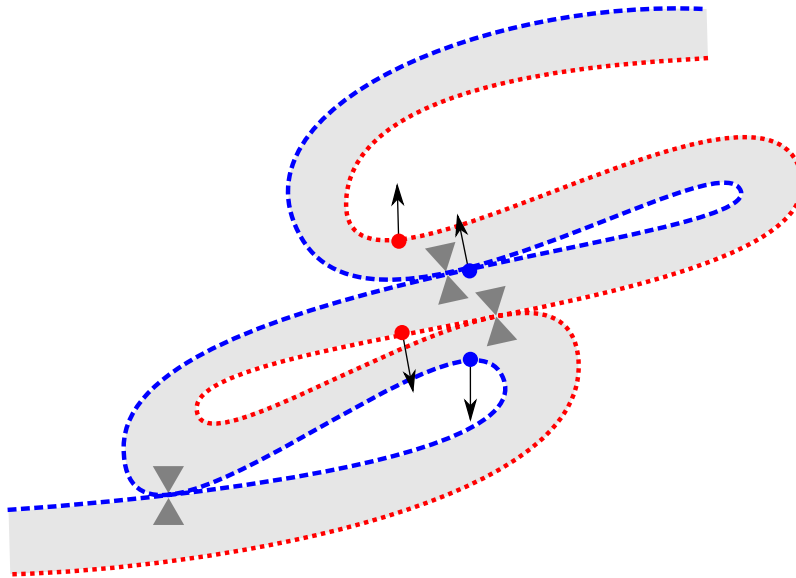


Figure 3.16: Usual case of self-contact in a thin-walled structure: example demonstrates the inherent problem of self-contact detection: triangles mark correctly detected contact; circles mark zones that can be recognized as active contact zones.

Let us enumerate the features of the implementation of the detection method in case of unknown a priori master-slave discretization. The main modification is that not only nodes, edges and segments have to be considered but also associated normals to determine potentially contacting elements.

1. The bounding box has to include all contact components; it can be chosen constant, if we know a priori a sufficiently small area, from where the contact nodes do not escape.
2. Normals have to be assigned to each contact node and contact edge at the beginning of each time step.

3. Only one array  $A^c$  is created and filled with contact components. The process is the same as in case of simple contact.
4. Since we cannot distinguish master and slave, the detection of the closest point has to be carried out for each contact node – future slave node against all contact components in the cell. To be sure that contact components can contact and are not attached to a common element, the normals associated with the slave node and the closest master component should point towards each other  $\underline{n}_i \cdot \underline{n}_j \leq 0$ . Obviously some neighboring cells have to be verified as in case of simple contact.
5. Nodes for which a close enough contact component has been determined should be marked as slaves and should not be included as slave node for another contact element.

Details for self-contact detection based on the closest node are quite similar and are discussed in [Benson 90] and [Yastrebov 11a]. Being adapted for the case of an unknown master-surface, the detection procedure has been verified on the challenging artificial problem of the self-contact within a snail operculum structure containing over 130 000 nodes on the surface, all nodes with attached segments are included in contact (see Fig. 3.17). The detection time is higher than for the contact of the same order with an known a priori master-slave discretization, because the preliminary stage requires the assignation of normals to every node and edge. Moreover, as the main detection stage requires significantly more verifications of distance and normal than in master-slave conception. In practice the difference in detection time between known a priori and unknown master-slave depends significantly on the geometry and its evolution. For example, for the snail operculum problem for the known master-slave the detection time ( $\approx 11$  sec) is only 4 times faster than for unknown a priori master-slave ( $\approx 45$  seconds).

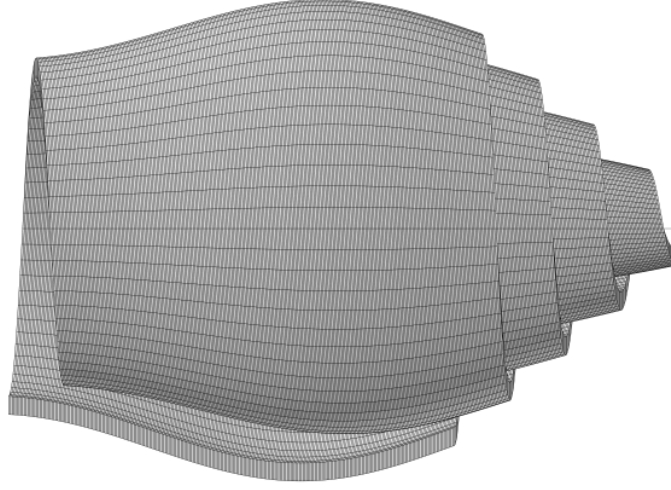


Figure 3.17: Finite element mesh used to test the detection procedure for self-contact problems.



**Remark 3.2 on the definition of the normal vector for contact nodes.**

Normal at node  $\underline{n}$  can be defined in different ways:

- averaged vector of the averaged normals of the adjacent segments

$$\underline{n} = \frac{1}{N} \sum_{i=1,N} \underline{n}_i,$$

where  $N$  is the number of adjacent segments and  $\underline{n}_i$  is the associated average normal of segment  $i$ , normals in the middle of segments or normals in the point where the needed node is situated (Fig.a).

- average between weighted average normals of adjacent segments (Fig.b)

$$\underline{n} = \frac{1}{A} \sum_{i=1,N} \underline{n}_i A_i,$$

where  $A_i$  is area (length) of the  $i$ -th adjacent segment, and  $A$  is the total area (length) of segments  $A = \sum_{i=1,N} A_i$ , and  $n_i$  is the average normal of the  $i$ -th segment.

- If the surface is smooth, then the associated normal is unique (Fig.c).

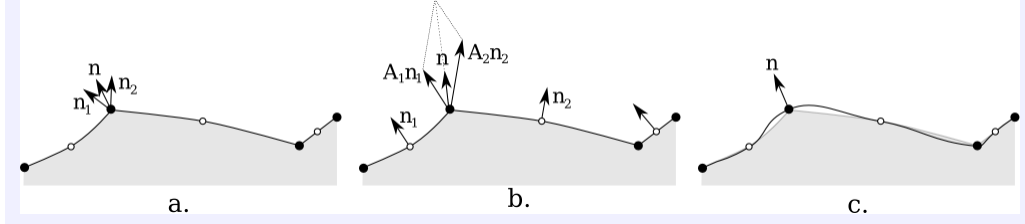


Fig. Different possibilities to define the normal at node/edge: **a** – average between normals of adjacent segments in the node; **b** – average weighted normal; for smoothed surfaces (e.g. with Bézier curve, **c**) the definition of normal is explicit and unique

In conclusion we confirm that the bucket sort method based on a rigorous definition of the closest point can be extended to the case of unknown a priori master-slave discretizations. The required detection time is significantly higher but of the same order of magnitude as the time needed for simple contact detection for the same problem. The availability of such a powerful method for self-contact detection extends significantly the capacities of the finite element analysis of contact problems. Some numerical examples will be given in Chapter 6.

## 3.6 Parallel contact detection

### 3.6.1 General presentation

The sequential treatment of the problems presented above (tire-road, curved contact surfaces, operculum structure) requires either too long computational time or may be just impossible due to the large amount of memory needed. The use of the parallelization paradigm is then a good way out. Many parallelization techniques are available nowadays, the class of non-overlapping domain decomposition or also called iterative substructuring method, is successfully and widely used in computational mechanics ([Farhat 94], [Toselli 05], [Gosselet 06]). It implies a splitting of an entire finite element mesh into subdomains which intersect only at their interfaces. Each subdomain is treated by one or several associated processors, and the global solution is obtained by enforcing displacement continuity and the balance of reactions across subdomains. The use of these techniques with affordable and powerful parallel computers allows to solve very large mechanical problems in a reasonable time. In general the resolution step cannot start before the detection procedure has been finished, so the last one is very important for the efficiency of the parallel computations [Brown 00]. It should not present a bottleneck in the whole process and, if possible, it has to use all the available capacities of parallel computers.

The key point for the contact detection procedure in parallel treatment is the fact that the finite element mesh and possibly the contact surface are divided into some parts associated with different processors and that, in the case of distributed memory, it is not entirely available on a particular processor. Since in principle we need the entire contact surface(s) to perform



the detection procedure, this repartition implies data exchanges between subdomains containing different parts of this surface(s). The amount of data transfer should be kept minimal. This will be our goal in the framework of the contact detection based on the bounding box conception and the bucket sort.

Two strategies for the parallel treatment of contact problems are proposed and analyzed : Single processor Detection, Multiple processor Resolution (SDMR) and Multiple Detection, Multiple Resolution (MDMR). As it is straightforward from the notations, SDMR carries out the contact detection on a single processor whereas MDMR uses all the available resources for the detection procedure. The last implies a parallelization of the detection procedure which will be discussed in details and tested.

### 3.6.2 Single Detection, Multiple Resolution approach

Let us consider the SDMR approach. The main idea is that all necessary information is collected by one processor which carries out the contact detection in the way explained above and distributes consequently the created contact elements among all concerned subdomains. This method can be efficiently applied to any contact problem and is relatively easy to implement. On the other hand, this method does not use efficiently all available resources, i.e., except one, all the processors are idle and inactive during the main detection phase; however all the processors possessing a contact surface are active during the preliminary stage.

At first, the bounding box for the contact detection has to be defined. This task is easily performed in parallel. Each subdomain  $i \in [1; N^c]$  possessing a part of contacting surfaces examines it and derives the corresponding bounding boxes  ${}^m\mathbf{r}_i^1, {}^s\mathbf{r}_i^1, {}^m\mathbf{r}_i^2, {}^s\mathbf{r}_i^2$  and the maximal dimension of the master segment  $d_{\max}^i$ . After data transfer, the global maximal detection distance  $d_{\max} = \max_{i=1, N^c} \{d_{\max}^i\}$  and the master and slave bounding boxes are determined

$${}^{m,s}\mathbf{r}_{\{x,y,z\}}^1 = \min_{i=1}^{N^c} \{{}^{m,s}\mathbf{r}_{i\{x,y,z\}}^1\} - d_{\max}, \quad {}^{m,s}\mathbf{r}_{\{x,y,z\}}^2 = \max_{i=1}^{N^c} \{{}^{m,s}\mathbf{r}_{i\{x,y,z\}}^2\} + d_{\max}. \quad (3.13)$$

Finally the resulting bounding box  $\{\mathbf{r}^1, \mathbf{r}^2\}$  is constructed as the intersection of master and slave bounding boxes, exactly as in the sequential procedure. The data transfer involves at most  $3N^c$  operations but the load is not uniformly distributed between processors, because not all of them contain the contact surface and, for those containing a contact zone, the size of the surface may be quite different. Anyway, this operation is quite fast even for huge meshes.

The next step consists in the union of all the necessary parts of the contact surface in one processor-detector. First, the information about the global bounding box is distributed among the subdomains possessing the contact surface, each of them counts the number of master and slave nodes located in the bounding box, further the subdomain with the maximal number of master and slave nodes is chosen as detector. Another possibility would be that this choice is made in agreement with the processor network topology, to accelerate

the data transfer on the next detection step. At this stage, the data exchange between subdomains remains negligible.

It remains to transfer all master components and slave nodes from the bounding box (global IDs, hosting subdomain ID, coordinates, etc) to the subdomain-detector, to carry out the detection as described in Section 3.3 and to attribute the constructed contact elements to the relevant subdomains. If a contact element is made of a slave node and master nodes from different subdomains, the interface between them has to be created or updated. Duplicated slave nodes have to be formed as well. This step is the most expensive in the terms of data exchange. The technical part of this operation is also quite complicated, because in general it is not possible to exchange directly mesh quantities: nodes, segments, edges. So, the developer has to design an appropriate class structure for geometrical objects, which does not contain any global pointers on the finite element mesh, elements, etc.

#### 3.6.3 Multiple Detection, Multiple Resolution approach

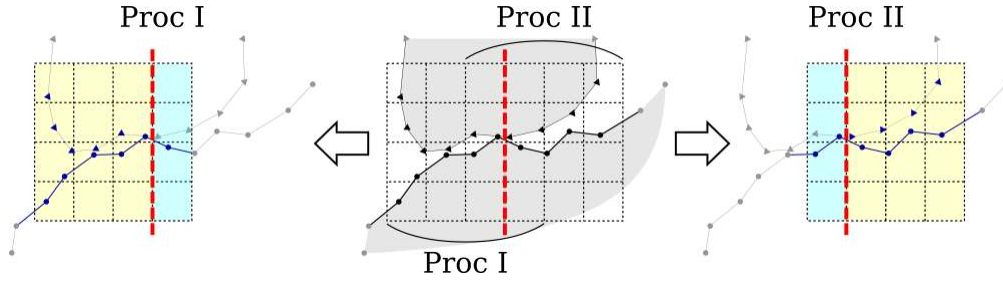
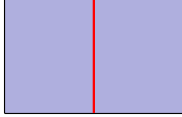
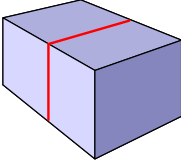
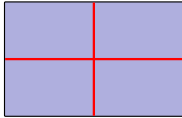
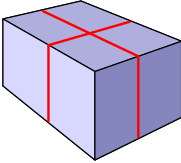
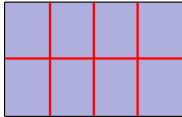
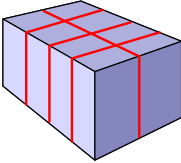
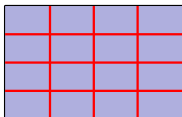
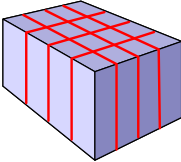


Figure 3.18: Example of cells partition between two processors: each processor gets one half of the total number of cells (with slave nodes and master components – represented by triangles and connected circles respectively) as well as one boundary layer from another half which contain only master segments.

In MDMR the preliminary part of a bounding box construction is exactly the same as in SDMR approach. The key difference between MDMR and SDMR appears in the next step. Instead of transferring all necessary information to the detector, the information is distributed between all subdomains in a special way. As shown above, the grid is constructed in such a way that, for each slave node, only one surrounding layer of neighboring cells has to be checked to find the closest point on the master surface. If self-contact is excluded, we do not care about slave nodes in neighboring cells. That is why the bounding box with associated buckets can be divided into  $N$  non-overlapping parts, each of them consisting of an integer number of buckets. Each part is then extended in all directions (not exceeding the bounding box) by one cell overlapping layer; the extended part is filled only with master components. An example involving two subdomains is presented in Fig. 3.18. It shows the internal cells (non-overlapping with other parts, light yellow color) including both master



Table 3.5: Split of detection buckets for parallel detection (MDMR)

Number of subdomains	split in 2D	split in 3D
2		
4		
8		
16		

components and slave nodes and external cells (shared with neighboring parts, light blue color) including only master segments. Each part is associated with a processor. The nodes and surfaces located in the part (global IDs, hosting subdomain ID, coordinates, etc.) are collected from different subdomains and transferred to the relevant one. The detection can then be carried out independently, i.e. in parallel for each part. No more data exchange is needed, so that the performance and scalability of the MDMR approach are significantly improved. The advantage of the method is that the total number of operations per processor during the main phase of detection does not increase for a given fraction of contacting nodes per number of processors. It is worth mentioning, that during the main detection phase, the number of operations is not homogeneously distributed between processors: there is still a need for a special algorithm to perform a smart split of the detection buckets which would take into account the distribution of contact elements into buckets. Actually we use a simple split taking into account that the number of subdomains is even and that the number of buckets is much smaller than the number of subdomains (see Table 3.5). Since in many applications the contact interface is concentrated in a thin flat zone, the 3D bounding box is not split along its smallest side.

The same parallel procedure can be used for self-contact problems. The first difference is that master and slave nodes are not distinguished and hence all contact components have to be included in the overlapping cells. The

second difference in treating contact components in internal and external (overlapping) cells is that nodes from the latter cannot be assigned as slave nodes. The described method is very similar to the parallelization of the Linked Cell Method widely used in molecular dynamic simulations for short-range interactions [Griebel 07].

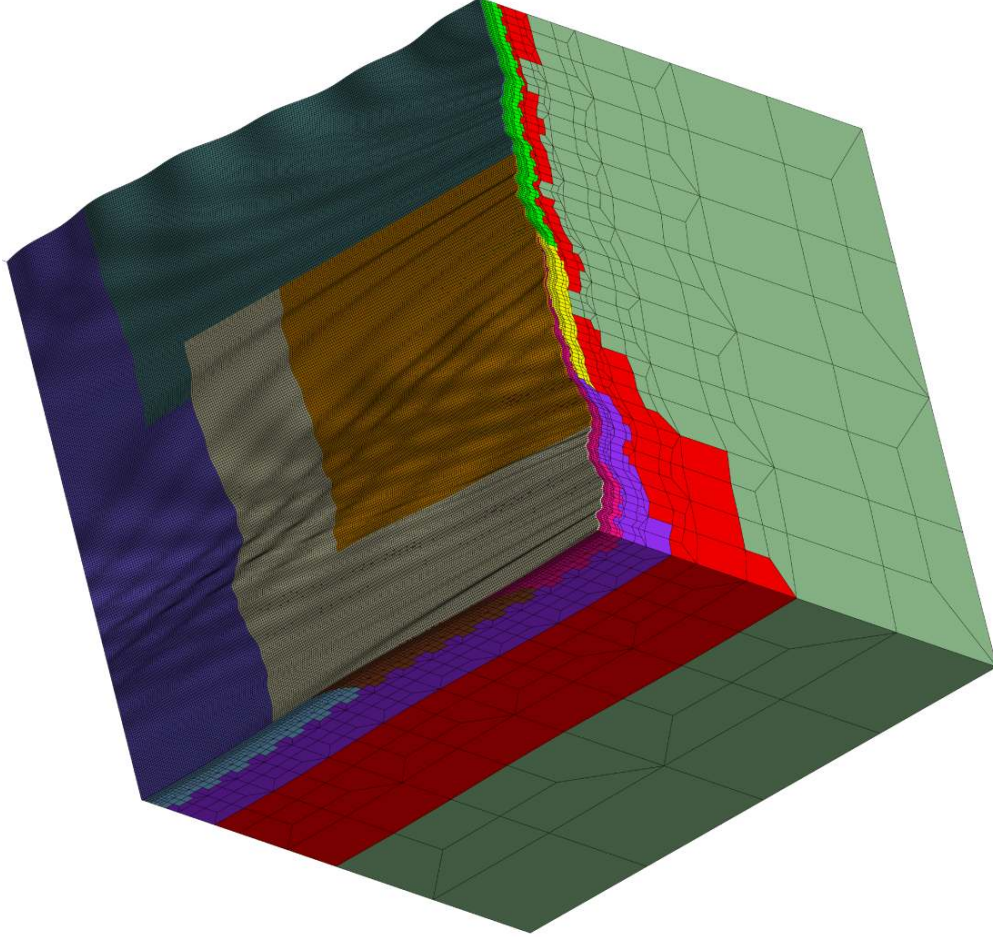


Figure 3.19: Example of the finite element mesh split by the METIS graph partitioning software [Karypis 95],[Karypis 98] into 16 sub-domains for parallel computations, different colors correspond to different subdomains. Remark that only 4 of 16 subdomains have components on the contact surface.

#### 3.6.4 Scalability test

In Figure 3.19, the finite element mesh of a rough surface is presented. Subdomains are highlighted by different colors. The scalability test for MDMR approach has been performed between two such meshes containing over 560 000 nodes and over 66 000 contact nodes each. The results obtained for slightly different surface roughness is represented in Figure 3.20.

“Heterogeneous distribution” of active contact zones means that the parts of bounding box associated with different processors possess a significantly different number of potential contact elements. In the considered case some of subdomains may not have contact elements at all. “Homogeneous distribution” means that this number is similar for different parts ( $\pm \approx 5\%$ ). The gain for a given number of processors is defined as the ratio between the reference CPU time for a single processor to the CPU time of the slowest processor. The average gain is calculated as the ratio of the reference CPU time for a single processor to the average CPU time of all processors. The difference between linear gain and the average gain highlights the time devoted to data exchange between subdomains. The pronounced difference between the gain for heterogeneous and homogeneous active contact zones distributions can be explained by the following observation. If there is no master component in the cell of the slave node and in neighboring cells, the time needed to conclude it is very small. On contrary, if the considered cells are not empty and contain several master components, it takes a much longer time to carry out the closest point detection. Even in these conditions the gain is quite high and its rate does not decrease with increasing number of detecting processors (for a reasonable ratio of contact nodes to number of processors).

The SDMR and MDMR approaches can be efficiently applied to parallel contact treatment. The second approach requires a larger amount of programming but its performance allows to neglect the detection time for large and extremely large contact problems.

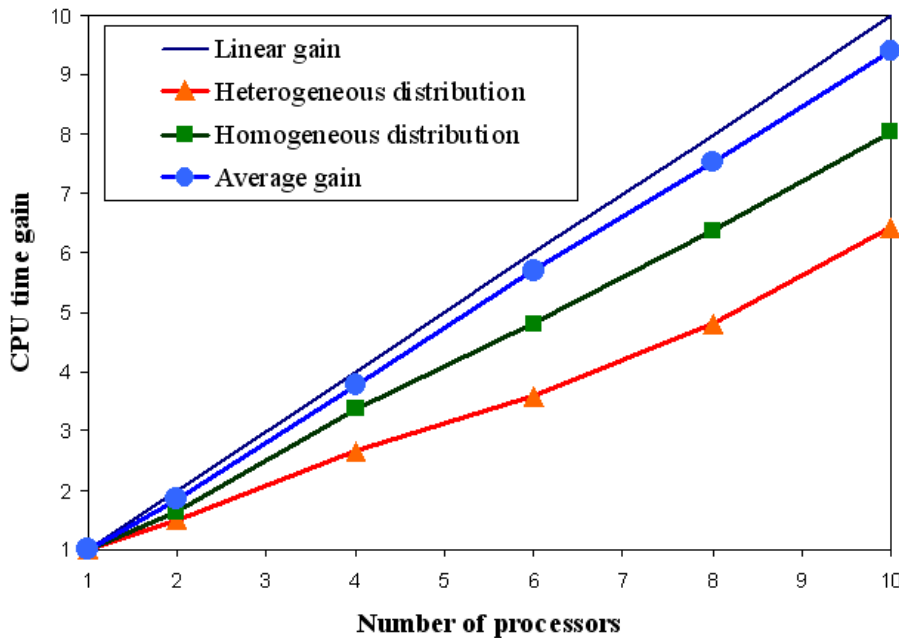


Figure 3.20: Time gain for the parallel contact detection procedure (average gain depicts the gain averaged by processors’ CPU time)..

### 3.7 Conclusion

A very fast local detection method has been elaborated on the base of the bucket method and a rigorous definition of the closest point. Sequential and parallel implementations of the method have been discussed in details for known a priori and unknown master-slave discretizations. Many previously proposed detection techniques starting from the historical article of Benson and Hallquist [Benson 90] are based on the closest node detection. Here we demonstrate that such an approach is not robust. However, the bucket technique can be easily generalized with a rigorous definition of the closest point.

The strong connections between the finite element mesh of the master surface, the maximal detection distance and the optimal dimension of the detection cells are established. The analytical estimation and numerous tests demonstrate that the optimal cell size is equal to the maximal detection distance multiplied by the square root of 2, and that the maximal detection distance can be chosen arbitrary. A reasonable choice for the MDD is the dimension of the biggest master segment if contact geometry changes relatively slowly. A particular attention has been paid to the bounding box construction and to the optimal choice of the neighboring cells to be verified. Techniques based on the closest node strategy and related challenges ("passing by node" and blind spot analysis) have been also discussed. An efficient implementation of the approach on distributed memory parallel computers has been also examined.

The method is very flexible but it is not well adapted for very heterogeneous distributions of the master segment dimensions or for very different mesh densities of the master and slave surfaces. Contrary to the closest node strategy, the dimension of the biggest master segment is not connected with the maximal detection distance but only with cell size. If the master surface has at least one segment which dimension is 10-100 times larger than the dimension of an average segment, the detection time can be rather high, but always less than in all-to-all approach.

The validation of the method has been performed on different contact problems in sequential and parallel cases: contact between rough surfaces with different geometries, tire-road contact, self-contact of a snail operculum and on the extremely large contact problem between two rough meshes including more than 1 000 000 segments on the master surface against 1 000 000 slave nodes. For the latter problem, the detection time changes significantly for different geometries from several seconds to 30-40 minutes in comparison to almost 8 days needed for the all-to-all detection technique.

Two parallel strategies for contact detection have been proposed and elaborated: Single Detection Multiple Resolution (SDMR) and Multiple Detection Multiple Resolution (MDMR). The last one implies the parallelization of the full detection cycle; it can be parallelized in a quite efficient manner, however, for a simple split of the detection zone, the gain depends significantly on the homogeneity of the distribution of the contact elements.

## Chapter 4

# Formulation of contact problems and resolution methods

### **Résumé de Chapitre 4 «Formulation du problème de contact et méthodes de résolution»**

Dans ce chapitre, on introduit les notations de base et on dérive les équations pour les différentes classes de problèmes de contact. On considère d'abord le contact d'un corps déformable avec un plan rigide. Les conditions de Hertz-Signorini-Moreau pour le contact normal et les conditions de contact avec frottement sont présentées. Ce cas simple permet de bien appréhender la spécificité de ces conditions et de les interpréter comme une combinaison de conditions aux limites partielles de Dirichlet et de Neumann. Cette interprétation débouche naturellement sur la méthode de résolution de contact dite «PDN» (Partial Dirichlet-Neumann), ou méthode du statut. Dans la première partie de ce chapitre, on décrit également un grand nombre de lois de frottement différentes.

La deuxième partie est consacrée au contact d'un corps déformable avec une surface rigide de géométrie arbitraire. La notion de séparation normale et son asymétrie sont rappelées. L'interprétation des conditions de contact normal et de contact avec frottement est aussi fournie.

Les équations générales du contact sont formulées en troisième partie. On utilise pour cela une approche qui met en œuvre des inégalités variationnelles. On discute de ses limitations sur quelques cas spécifiques. En raison de ces limitations, notamment pour les problèmes de grands glissements et grandes déformation, on se focalise sur les méthodes pour lesquelles on suppose que la zone de contact est connue. Dans ce cadre, on présente en détail trois méthodes classiques : la méthode de pénalisation (linéaire et non-linéaire), la méthode des multiplicateurs de Lagrange (avec des stratégies de «active set» différentes) et la méthode du Lagrangien augmenté. Les formes faibles pour le contact sans et avec frottement sont données pour les trois méthodes. Leurs performances sont étudiées, et les détails de l'implémentation numérique sont discutés sur un problème simple ne comportant qu'un degré de liberté.

In this chapter we give the main notions and derive the governing equations for different classes of contact problems. We start from a simple description of contact between a deformable solid and a rigid plane. Based on this simple problem, Hertz-Signorini-Moreau (non-penetration–non-adhesion) conditions are derived and the frictional contact problem is formulated. An interpretation of contact conditions is given using partial boundary conditions of Dirichlet and Neumann type. Next, Signorini's problem is formulated – contact between a deformable solid and an arbitrary rigid smooth surface with and without friction. This section is followed by a general multi-body contact problem is formulated using variational inequality. Due to the complexity of this formulation for large deformation/large sliding cases, for numerical purpose the contact problem is often formulated as a variational equality for a known contact zone. To derive the variational equality, either penalty or Lagrange multipliers methods are used, which are discussed also in this chapter. A widely used augmented Lagrangian method is also considered. A simple nonlinear contact problem is resolved by all this methods, demonstrating advantages and revealing drawbacks of each of them.

### 4.1 Unilateral contact with a rigid plane

We wish to describe the motion of a deformable body coming in contact with a rigid plane (Fig. 4.1). The points of the body are described by vector  $\underline{X}$  in the reference configuration  $\Omega^0$  and by vector  $\underline{x}(\underline{X}, t)$  in the actual configuration  $\Omega$  at time  $t$ . The motion is described relatively to a fixed spatial frame defined by orthonormal basis vectors  $\{\underline{e}_x, \underline{e}_y, \underline{e}_z\}$ . Let us suppose that the motion of the body is restricted to the upper half-space  $z \geq 0$  from the rigid plane  $z = 0$ . Then the following condition on displacements is imposed

$$g_n = \underline{x} \cdot \underline{\nu} \geq 0, \quad (4.1)$$

where  $g_n$  is a normal gap (see Chapter 2 and the next section for a more detailed definition of the gap),  $\underline{\nu} = \underline{e}_z$  is a unit vector normal to the plane and pointing to an acceptable area of motion  $z > 0$ , i.e. the points of the body at all time instants cannot penetrate under this plane  $z < 0$ . If the displacement of a point is defined as

$$\underline{u} = \underline{x} - \underline{X},$$

then expressing  $\underline{x}$  from the latter equality gives (4.1) in the form

$$g_n = \underline{u} \cdot \underline{\nu} + \underline{X} \cdot \underline{\nu} \geq 0, \quad (4.2)$$

where  $\underline{X} \cdot \underline{\nu} = g_n^0$  is a constant initial gap determined for all material points. If the body retain its integrity as well as if its deformation are consistent, the non-penetration condition (4.1), (4.2) are applied only to surface points  $\partial\Omega$ , precisely to the so-called *potential contact zone* denoted  $\Gamma_c$  in the actual configuration.  $\Gamma_c$  can be splitted into two nonintersecting sets: active  $\bar{\Gamma}_c$  (points are in contact)



and inactive  $\Gamma_c \setminus \bar{\Gamma}_c$  (points are not in contact) contact zones. The active contact zone in the actual configuration is defined by the following equality

$$\underline{x} \in \bar{\Gamma}_c \text{ if and only if } \underline{x} \cdot \underline{\nu} = 0 \quad (4.3)$$

consequently in the reference configuration the contact zone can be easily determined as

$$\underline{X} \in \Gamma_c^{0'} \text{ if and only if } \underline{X} \cdot \underline{\nu} = -\underline{u} \cdot \underline{\nu}. \quad (4.4)$$

As follows from the definition of the active contact zone,  $\bar{\Gamma}_c$  is an unknown a priori part of the potential contact zone:  $\bar{\Gamma}_c \subset \Gamma_c \subset \partial\Omega$ . To prevent the body

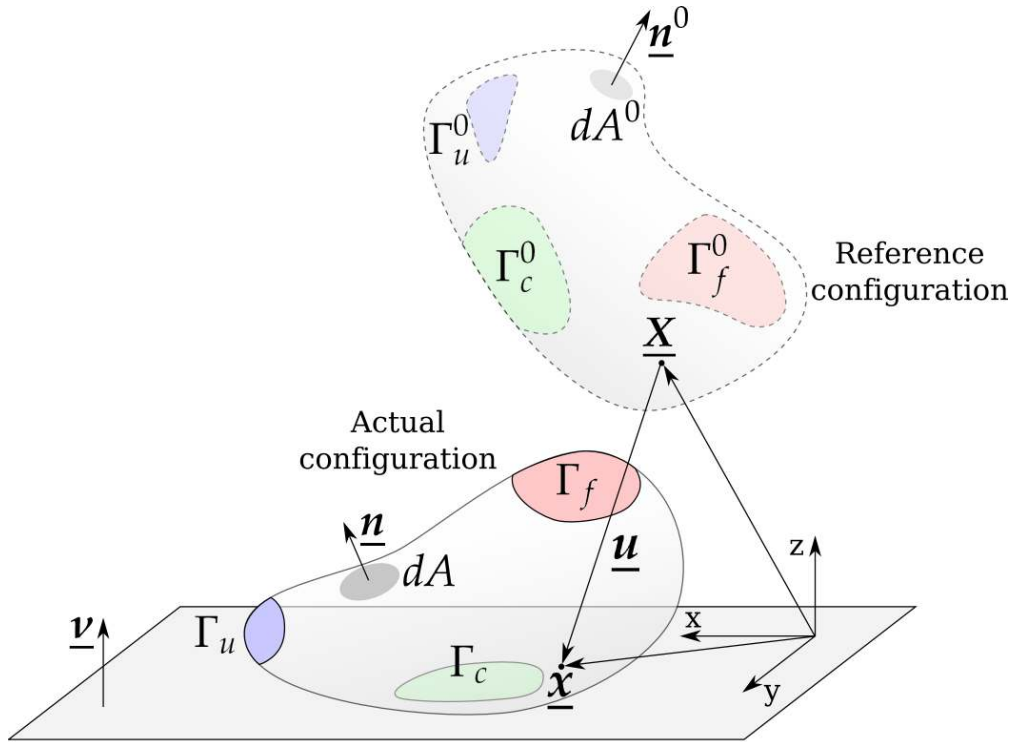


Figure 4.1: Reference and actual configuration of a deformable body in contact with a rigid plane.

from penetrating the plane, a contact pressure arises in the contact zone. If we confine ourself to the description of non-adhesive contact, then the contact pressure should be non-positive (zero in inactive and negative in active contact zones).

In the actual configuration, the stress state is described by the Cauchy stress tensor  $\underline{\underline{\sigma}}$  and in the reference configuration by the first Piola-Kirchhoff stress tensor  $\underline{\underline{P}}$ , which are connected in the following way

$$\underline{\underline{P}} = J \underline{\underline{\sigma}} \cdot \underline{\underline{F}}^{-T}, \quad (4.5)$$

#### 4.1 Unilateral contact with a rigid plane

---

where  $\underline{\underline{F}}$  is the nonsymmetric deformation gradient tensor

$$\underline{\underline{F}} = \frac{\partial \underline{x}}{\partial \underline{X}},$$

$J$  is the determinant of the deformation gradient (Jacobian), that should be positive in order to prevent self-penetration of points of the body

$$J = \det \underline{\underline{F}} > 0$$

As  $\det \underline{\underline{F}} > 0$ , its finite inverse exists

$$\underline{\underline{F}}^{-1} = \frac{\partial \underline{X}}{\partial \underline{x}}, \quad \underline{\underline{F}}^{-1} \cdot \underline{\underline{F}} = \underline{\underline{I}}.$$

The Cauchy stress vector at the surface of the body is

$$\underline{\sigma} = \underline{n} \cdot \underline{\underline{\sigma}},$$

where  $\underline{n}$  is the outward normal at the surface of the solid in the actual configuration. In the active contact zone this normal is opposite to the normal of the rigid plane  $\underline{n} = -\underline{v}$ , so the contact pressure  $\sigma_n$  in the actual configuration can be written as

$$\sigma_n = \underline{\sigma} \cdot \underline{n} = \underline{v} \cdot \underline{\underline{\sigma}} \cdot \underline{v}.$$

In the reference configuration, the normal to the surface is  $\underline{n}^0$ , so the contact pressure can be written in terms of the first Piola-Kirchhoff stress

$$\sigma_n^0 = \underline{P} \cdot \underline{n}^0 = \underline{n}^0 \cdot \underline{P} \cdot \underline{n}^0.$$

Applying the relations between the normal vectors,  $\underline{n} = \underline{\underline{F}}^{-T} \cdot \underline{n}^0$  or  $\underline{n}^0 = \underline{\underline{F}}^T \cdot \underline{n}$  and using Eq. (4.5) allows to rewrite the previous equalities as

$$\sigma_n^0 = J \underline{n} \cdot \underline{\underline{F}} \cdot \underline{\underline{\sigma}} \cdot \underline{n} = J \underline{v} \cdot \underline{\underline{F}} \cdot \underline{\underline{\sigma}} \cdot \underline{v}.$$

The tangential component of the stress vector is

$$\underline{\sigma}_t = (\underline{\underline{I}} - \underline{n} \otimes \underline{n}) \cdot \underline{\sigma} = \underline{\sigma} - \sigma_n \underline{n}. \quad (4.6)$$

It should be zero for frictionless contact  $\underline{\sigma}_t = 0$ .

Finally, for non-adhesive frictionless contact, we require that

$$\boxed{\sigma_n \leq 0 \text{ at } \Gamma_c} \quad \text{or} \quad \sigma_n^0 \leq 0 \text{ at } \Gamma_c^0. \quad (4.7)$$

In general, we suppose that if a point is not in contact, then  $\sigma_n = 0$ , meanwhile  $\sigma_n \leq 0$  in contact. This leads to the non-penetration–non-adhesion condition

$$\boxed{\sigma_n g_n = 0} \quad (4.8)$$

However, we have to keep in mind that this is a simplification. When contact is considered in a gas/liquid environment, this simplification is acceptable only if



the pressure of the environment is negligible in comparison to arising contact pressures. Another limitation is that this equality is limited to nonadhesive contact. The set of conditions (4.1), (4.7) and (4.8) are called Hertz-Signorini-Moreau or Karush-Kuhn-Tucker conditions, complemented by a non-friction condition  $\underline{\sigma}_t = 0$ :

$$g_n \geq 0, \quad \sigma_n \leq 0, \quad \sigma_n g_n = 0, \quad \underline{\sigma}_t = 0 \quad (4.9)$$

These contact conditions together with the relevant boundary conditions complement the static local balance of momentum and angular momentum ( $\underline{\underline{\sigma}} = \underline{\underline{\sigma}}^T$ )

$$\begin{cases} \nabla \cdot \underline{\underline{\sigma}} + \underline{f}_v = 0 & \text{in } \Omega, \\ \underline{\underline{\sigma}} \cdot \underline{n} = \underline{\sigma}_0 & \text{at } \Gamma_f, \\ \underline{u} = \underline{u}_0 & \text{at } \Gamma_u, \\ g_n \geq 0, \quad \sigma_n \leq 0, \quad \sigma_n g_n = 0, \quad \underline{\sigma}_t = 0 & \text{at } \Gamma_c, \end{cases} \quad (4.10)$$

where  $\underline{f}_v$  is a vector of volume forces,  $\underline{\sigma}_0$  is a prescribed traction (Neumann boundary conditions) and  $\underline{u}_0$  is a prescribed displacement (Dirichlet boundary conditions). That is the starting point for the following investigation. First, we propose to discuss in details the Hertz-Signorini-Moreau conditions imposed on the contact zone.

#### 4.1.1 Interpretation of contact conditions



##### **Idea 4.1 Replacement of contact conditions by Dirichlet-like boundary conditions.**

*Often the contact conditions are interpreted as a special type of Neumann boundary conditions (penalty, Lagrange multiplier methods) as a function of displacement, the problem being to find which contact pressure has to be applied in order to fulfill non-penetration contact condition. However, the problem can be inverse: in stead of prescribing the pressure at the contact zone, one can impose directly the displacement according to the contact constraint, i.e. apply Dirichlet-like boundary conditions depending on stress-state.*

The unilateral contact conditions of non-penetration–non-adhesion can be splitted into two parts: for active and inactive contact zones

$$g_n \geq 0, \quad \sigma_n \leq 0, \quad \sigma_n g_n = 0 \quad \text{at } \Gamma_c \Leftrightarrow \begin{cases} g_n = 0, \quad \sigma_n < 0 & \text{at } \bar{\Gamma}_c \quad (a) \\ g_n > 0, \quad \sigma_n = 0 & \text{at } \Gamma_c \setminus \bar{\Gamma}_c \quad (b) \end{cases} \quad (4.11)$$

The condition (4.11,a) can be interpreted in the following way. According to (4.2) and to the definition of the active contact zone (4.4), the first term

#### 4.1 Unilateral contact with a rigid plane

corresponds to

$$g_n = 0 \Leftrightarrow \underline{\nu} \cdot \underline{u} = -g_n^0 \Leftrightarrow \boxed{u_z = -g_n^0}, \quad (4.12)$$

the boxed term gathers the partial Dirichlet boundary condition. The condition (4.11,b) is equivalent to the free boundary outside the active contact zone. Finally, the Hertz-Signorini-Moreau conditions for the case of contact between a deformable body and a rigid plane writes as

$$\boxed{u_z = -g_n^0 \quad \text{at } \bar{\Gamma}_c, \text{ where } \underline{x} \in \bar{\Gamma}_c \text{ if } \sigma_n < 0} \quad (4.13)$$

The contact conditions have been replaced by the partial Dirichlet boundary conditions on the unknown active contact zone  $\bar{\Gamma}_c$ , which is determined by negative contact pressure. The nonlinearity of the problem consists in determining the active contact zone.

Note that only one component  $u_z$  of the displacement vector  $\underline{u}$  is prescribed, and that the other components  $u_x, u_y$  are not specified. This is easily interpreted in the Finite Element Method, where the displacement vector in 3D is splitted in 3 degrees of freedom, each of them being prescribed independently. Note that Cartesian, cylindrical or spherical basis in 3D allows to replace easily the contact with a rigid plane, cylinder or sphere respectively by simple conditions analogous to (4.13). Moreover, the limits of the possible contact zone can be restricted in the following way

$$\boxed{u_z = -g_n^0 \quad \text{at } \bar{\Gamma}_c \cap f(\underline{x}) \geq 0, \text{ where } \underline{x} \in \bar{\Gamma}_c \text{ if } \sigma_n < 0}, \quad (4.14)$$

where  $f(\underline{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}$  is an arbitrary function,  $f(\underline{x}) = 0$  determines the edge of the plane. All forementioned extensions provide us with a rather simple and multipurpose approach to the sub-class of contact problems. Intuitively, it is easier to prescribe a given displacement than an unknown contact pressure distribution on unknown active contact zones. In the first case, the active contact zone is simply determined by the sign of the contact pressure, in the second case by a zero value of the normal gap.

Another straightforward extension of this interpretation provides a solution for simple adhesion. For this purpose, an active contact zone should be determined according, for example, to

$$\bar{\Gamma}_c : \sigma_n < a_n \max_{t \in [t_0, t]} \langle -\sigma_n \rangle,$$

where  $\max_{t \in [t_0, t]} \langle -\sigma_n \rangle$  is the maximal reached contact pressure at the current material point;  $t_0$  is the time moment when this point came in contact;  $a_n \geq 0$  is the normal adhesion coefficient and  $\langle \bullet \rangle$  denotes the Macaulay brackets  $\langle x \rangle = \max(0, x)$ . If a point loses the contact, then its history variable should be set to zero ( $\max_t \langle -\sigma_n \rangle = 0$ ). The following expression describes the adhesive contact with a rigid plane

$$\boxed{u_z = -g_n^0 \quad \text{at } \bar{\Gamma}_c \cap f(\underline{x}) \geq 0, \text{ where } \underline{x} \in \bar{\Gamma}_c : \sigma_n < a_n \max_{t \in [t_0, t]} \langle -\sigma_n \rangle}, \quad (4.15)$$

Up to now we discussed the cases when the normal to the rigid plane (cylindrical or spherical surface) is parallel to one of the basis vectors in the chosen reference frame. The situation changes drastically for the Finite Element Method if the latter is not the case. Generalization of the presented approach for arbitrary rigid surface would lead to classic Signorini's problem and will be discussed in the next section. Before that, the frictional conditions will be formulated for the case of the contact with a rigid plane and will be also interpreted similarly to frictionless conditions.

#### 4.1.2 Friction

We return to the description of the motion of a deformable body in contact with a rigid plane. Friction on the contact interface can be introduced in the general form

$$\underline{\sigma}_t = \underline{\sigma}_t(\sigma_n, \dot{\sigma}_n, \underline{\dot{g}}_t, t, \dots), \quad (4.16)$$

where  $\underline{\sigma}_t$  is the tangential stress vector at the interface, according to (4.6)

$$\underline{\sigma}_t = (\underline{I} - \underline{n} \otimes \underline{n}) \cdot \underline{\sigma};$$

$\dot{\sigma}_n$  is the contact pressure rate,  $\underline{\dot{g}}_t$  is the relative sliding velocity and  $t$  is the time. For the time being, we restrict ourself to the classical non-associated Coulomb's friction law. Further some non-classical friction laws will be briefly discussed.

In the simplest case, the friction law states that the tangential resistance of the contact interface depends on the contact pressure and that the direction of the tangential stress vector is given by the sliding direction

$$\underline{\sigma}_t = \underline{\sigma}_t(\sigma_n, \underline{s}), \quad (4.17)$$

where  $\underline{s}$  is a unit vector in the contact plane, which determines the direction of sliding:

$$\underline{s} = \begin{cases} \frac{\underline{\dot{g}}_t}{\|\underline{\dot{g}}_t\|}, & \text{if } \|\underline{\dot{g}}_t\| > 0; \\ 0, & \text{if } \|\underline{\dot{g}}_t\| = 0. \end{cases} \quad (4.18)$$

So in frictional case when a body comes in contact with a rigid plane, a stress vector arises which contains not only the contact pressure but also the shear stress vector due to friction

$$\underline{\sigma} = \sigma_n \underline{n} + \underline{\sigma}_t(\underline{\sigma}_t, \sigma_n, \underline{s}).$$

At the same time, the relative motion of the point along the rigid plane is confined by the frictional shear force in the way that the point remains at the initial contact location if the shear stress vector is smaller than the critical frictional stress  $\underline{\sigma}_t^c$ , which in Coulomb's friction law is proportional to the contact pressure, so the point does not move if

$$\|\underline{\sigma}_t\| \leq \|\underline{\sigma}_t^c\| = \mu |\sigma_n|, \quad (4.19)$$

#### 4.1 Unilateral contact with a rigid plane

where the coefficient of proportionality  $\mu$  is the coefficient of friction. According to the third Newton's law in equilibrium state, the reaction should be equal by value and opposite by direction to the action, i.e. the stress vector integrated over a small surface of the body  $\underline{\sigma} dA$  should be equal by modulus and opposite to the force vector at the rigid plane  $\underline{F}^e$ . The resistance of the contact interface in the tangential direction is limited, for instance in Coulomb's friction law the limit is  $\mu|\sigma_n|$ . Up to this limit a motionless<sup>1</sup> equilibrium state is preserved  $\underline{\sigma}_t dA + \underline{F}^e = 0$ ,  $\|\underline{\sigma}_t\| < \mu|\sigma_n|$ , the maximal permitted tangential stress in equilibrium straight motion is limited by  $\underline{\sigma}_t = \mu|\sigma_n|\underline{s}$  providing the following equality

$$\underline{\sigma}_t - \mu|\sigma_n|\underline{s} = 0.$$

This equality is valid only in case of slip, but can be generalized to stick case if we multiply it by the norm of the sliding velocity  $\|\dot{\underline{g}}_t\|$

$$\boxed{\|\dot{\underline{g}}_t\|\underline{\sigma}_t - \mu|\sigma_n|\dot{\underline{g}}_t = 0}. \quad (4.20)$$

Conditions (4.19) and (4.20) can be complemented by an additional condition distinguishing stick and slip states

$$\boxed{\|\underline{s}\| \|\underline{\sigma}_t - \mu|\sigma_n|\underline{s}\| = 0}.$$

Union of all forementioned conditions forms the set of conditions for Coulomb's non-associated friction

$$\boxed{\|\underline{\sigma}_t\| \leq \mu|\sigma_n|, \quad \|\dot{\underline{g}}_t\|\underline{\sigma}_t - \mu|\sigma_n|\dot{\underline{g}}_t = 0, \quad \|\underline{s}\| \|\underline{\sigma}_t - \mu|\sigma_n|\underline{s}\| = 0}. \quad (4.21)$$

These conditions connecting  $\sigma_n$ ,  $\underline{\sigma}_t$  and  $\underline{s}$  can be represented graphically (Fig. 4.2) in a general form. In 2D case the sliding velocity and frictional stress can be represented as

$$\dot{\underline{g}}_t = \dot{g}_t \underline{t}, \quad \underline{\sigma}_t = \sigma_t \underline{t},$$

where  $\underline{t}$  is a unit tangential vector orthogonal to the normal vector  $\underline{v}$ , i.e.  $\underline{v} \cdot \underline{t} = 0$ ,  $\|\underline{t}\| = 1$ ; in this case the graphical representation of the frictional conditions can be extended as shown in Fig. 4.3. The graphical representation of frictional conditions in 3D case can be found in Fig. 4.4. The following notions have been introduced

$$\dot{\underline{g}}_t = \|\dot{\underline{g}}_t\|\underline{s} = \dot{g}_{t1}\underline{t}_1 + \dot{g}_{t2}\underline{t}_2, \quad \underline{\sigma}_t = \|\underline{\sigma}_t\|\underline{s} = \sigma_{t1}\underline{t}_1 + \sigma_{t2}\underline{t}_2,$$

where  $\underline{t}_1, \underline{t}_2$  are orthonormal vectors in the contact plane. The cone in axes  $\{\sigma_{t1}, \sigma_{t2}, \sigma_n\}$  in Fig. 4.4 is called the Coulomb's cone, an open set determined by

$$C(\sigma_n) = \sqrt{\sigma_{t1}^2 + \sigma_{t2}^2} < \mu|\sigma_n|.$$

<sup>1</sup>by motionless here we mean that there is no relative motion between sticking points in the interface, in space their motions are not limited.

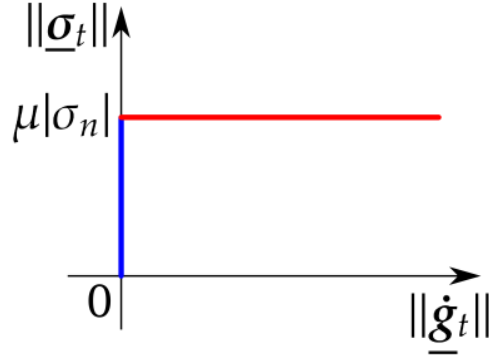


Figure 4.2: Graphical representation of Coulomb's frictional conditions (4.21): the point corresponding to a given state can be situated either on the blue line – stick state ( $\|\dot{\underline{g}}_t\| = 0$ ,  $\|\underline{\sigma}_t\| < \mu|\sigma_n|$ ) or on the red line – slip state ( $\|\dot{\underline{g}}_t\| \geq 0$ ,  $\|\underline{\sigma}_t\| = \mu|\sigma_n|$ ).

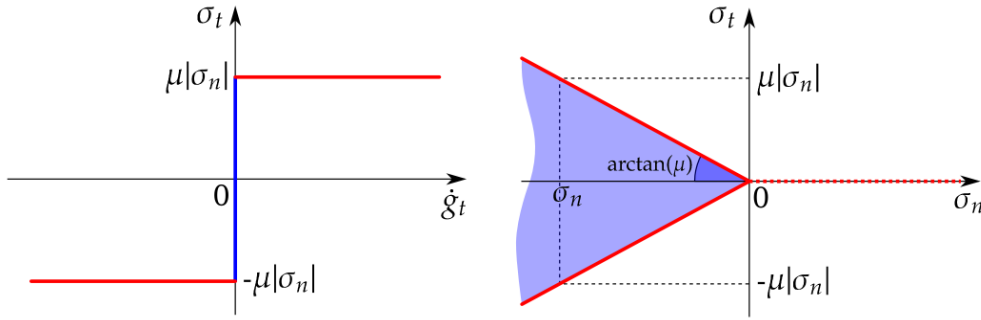


Figure 4.3: Graphical representation of Coulomb's frictional conditions for 2D contact problem: **left** – relation between the tangential velocity and the tangential stress, blue line ( $\dot{g}_t = 0$ ,  $|\sigma_t| < \mu|\sigma_n|$ ) represents stick state, red line ( $|\sigma_t| = \mu|\sigma_n|$ ) – slip state; **right** – relation between the contact pressure and the contact tangential stress: blue color ( $|\sigma_t| < \mu|\sigma_n|$ ) denotes stick state, its closure in red ( $|\sigma_t| = \mu|\sigma_n|$ ) represents slip state.

Any stress-state fulfilling frictional conditions corresponds to a unique point either in the interior of the cone

$$\underline{\sigma}_t \in C(\sigma_n)$$

or on its closure

$$\underline{\sigma}_t \in \partial C(\sigma_n).$$

The zone of positive contact pressure is excluded by the non-adhesion condition (see Hertz-Signorini-Moreau conditions, (4.9)), however, for the sake of consistency the cone can be extended to the area of positive contact pressure as represented in Fig. 4.3 and Fig. 4.4 by a dashed red line. Moreover, such

#### 4.1 Unilateral contact with a rigid plane

an extension is consistent with the description of adhesive contact. So the extended frictional conditions become

$$\|\underline{\sigma}_t\| \leq \mu \langle -\sigma_n \rangle, \quad \|\underline{\dot{g}}_t\| \underline{\sigma}_t - \mu \langle -\sigma_n \rangle \underline{\dot{g}}_t = 0, \quad \|\underline{s}\| \left\| \|\underline{\dot{g}}_t\| \underline{\sigma}_t - \mu \langle -\sigma_n \rangle \underline{\dot{g}}_t \right\| = 0, \quad (4.22)$$

where  $\langle \bullet \rangle$  denotes the Macaulay brackets.

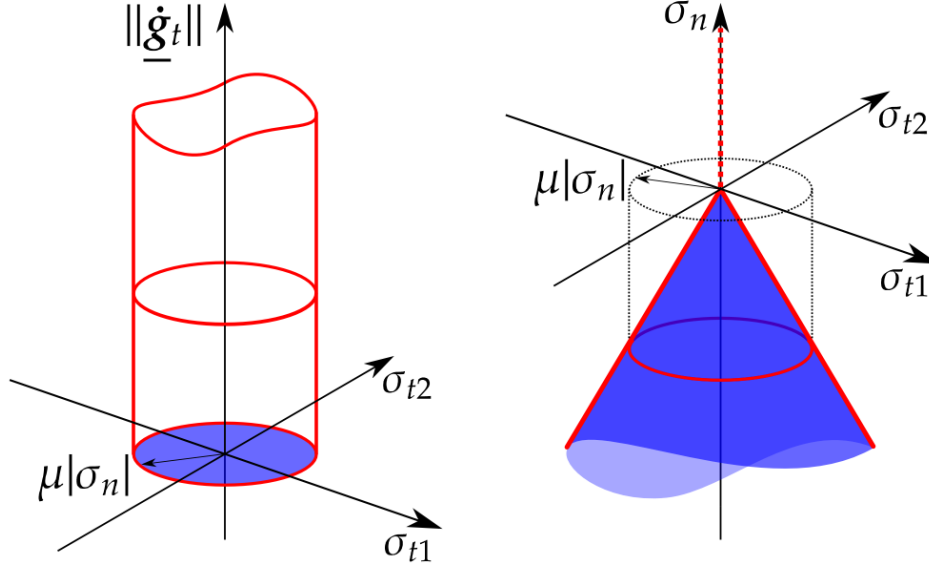


Figure 4.4: Graphical representation of Coulomb's frictional conditions for 3D contact problem: **left** – relation between the norm of the tangential velocity and the components of the tangential stress vector: the admissible points are situated either in the blue circle ( $\|\underline{\dot{g}}_t\| = 0$ ,  $\|\underline{\sigma}_t\| < \mu|\sigma_n|$ ) – stick state or on the surface of the semi-infinite cylinder ( $\|\underline{\dot{g}}_t\| \geq 0$ ,  $\|\underline{\sigma}_t\| = \mu|\sigma_n|$ ) marked with red color – slip state; **right** – relation between the contact pressure and the components of the contact tangential stress vector: in blue - interior of the Coulomb's cone ( $\|\underline{\sigma}_t\| < \mu|\sigma_n|$ ) – stick state, the surface of the cone in red ( $\|\underline{\sigma}_t\| = \mu|\sigma_n|$ ) represents the slip state.

Let us consider the case when a point slides on the plane under a constant contact pressure  $\sigma_n$  (Fig. 4.5): we plot its path in three spaces: stress, velocity and displacement. Along the  $[0 - 1)$  path, the frictional stress remains inside the circle

$$\underline{\sigma}_t \in C(\sigma_n)$$

that is the section of Coulomb's cone for a given contact pressure, so no relative sliding occurs which corresponds to **stick**. At point 1, the frictional stress reaches its limit – the boundary of the circle  $\underline{\sigma}_t \in \partial C(\sigma_n)$  and stays on it until point 3 is reached. Consequently, the relative sliding velocity follows the path  $[1 - 2 - 3]$  in the velocity space in such a way that the vector of frictional stress

Table 4.3: Analogies between friction and plasticity

Friction	Plasticity
Stick state	elastic deformation
Slip state	plastic flow
Coulomb's cone $\partial C(\sigma_n)$	yield surface
Maximal frictional stress $\ \underline{\sigma}_t\  = \mu \sigma_n $	yield strength

and the tangential velocity are collinear at every moment (see, for example, the point 2), this state corresponds to **slip**. The relative displacement  $\Delta \underline{g}_t$  is simply the integral of the velocity vector over the sliding time, as seen in Fig. 4.5 the point moves from the stick point 1 to another stick point 3 by the curved trajectory 1 – 2 – 3. At point 3, the relative velocity returns to zero, as in stress space the point dives again inside the circle. Any motion in the interior of Coulomb's cone does not result in relative tangential displacements and any relative sliding implies that the point is situated at the surface of Coulomb's cone in stress space. In general, contrary to this example, the normal pressure changes during sliding and consequently the limits of the stick zone in stress space change.

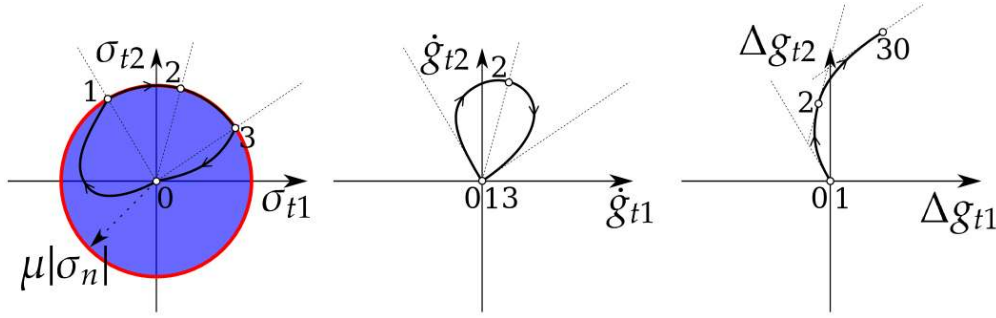


Figure 4.5: Graphical representation of frictional motion in stress, velocity and displacement spaces: path 0 – 1 corresponds to stick, path 1 – 2 – 3 – to slip, path 3 – 0 again to stick.

The derived formulation of friction is quite similar to the formulation of plastic flow [Curnier 84] (see Tab. 4.3). Note that the flow rule is non-associated since there is no irreversible slip in the normal direction and the principle of maximum dissipation is inapplicable except in the case of a prescribed constant contact pressure  $\sigma_n = \text{const}$ . So the model is somehow similar to the Drucker-Prager yield criterion for pressure dependent plasticity [de Saxce 98]. A more consistent analogy between associated/non-associated plasticity and friction is given in [Michalowski 78].

#### 4.1.3 Interpretation of frictional conditions



#### Idea 4.2 Replacement of frictional contact conditions by Dirichlet-Neumann boundary conditions.

*In case of unilateral contact with a rigid plane, Hertz-Signorini-Moreau conditions can be replaced by prescribing partial Dirichlet boundary conditions on the unknown a priori active contact zone  $\bar{\Gamma}_c$ . To take the frictional resistance into account, depending on the stress state, either Dirichlet-like conditions have to be complemented to ordinary Dirichlet conditions (stick state) or partial Neumann boundary conditions have to be applied in the contact plane (frictional slip state).*

Let us split the active contact zone  $\bar{\Gamma}_c$  into a slip  $\Gamma_c^*$  and a stick  $\Gamma_c^\bullet$  zone (in reference configuration  $\Gamma_c^{*0}$  and  $\Gamma_c^{\bullet0}$  respectively), such that  $\Gamma_c^\bullet \cup \Gamma_c^* = \bar{\Gamma}_c$  and  $\Gamma_c^\bullet \cap \Gamma_c^* = \emptyset$ . Then the frictional conditions can be rewritten as

$$\begin{cases} g_n = 0, \underline{\dot{g}}_t = 0, \sigma_n < 0, \|\underline{\sigma}_t\| < \mu|\sigma_n|, & \text{in the stick zone } \Gamma_c^\bullet \quad (\text{a}) \\ g_n = 0, \sigma_n < 0, \underline{\sigma}_t - \mu|\sigma_n|\underline{s} = 0, & \text{in the slip zone } \Gamma_c^* \quad (\text{b}) \\ g_n > 0, \sigma_n = 0, \underline{\sigma}_t = 0, & \text{in the inactive contact zone } \Gamma_c \setminus \bar{\Gamma}_c \quad (\text{c}) \end{cases} \quad (4.23)$$

Obviously, the first condition (4.23,a) for a motionless rigid plane can be replaced by an ordinary Dirichlet boundary condition

$$\underline{u} = -g_n^0 \underline{\nu} + (\underline{I} - \underline{\nu} \otimes \underline{\nu}) \cdot (\underline{x}(t^\bullet) - \underline{X}), \quad \text{if } \underline{X} \in \Gamma_c^{\bullet0}, \quad (4.24)$$

the time  $t^\bullet$  denotes the moment when the material point switches to stick state either from slip state or from non-contact state; the state of the point is determined by the stress, if it is in the Coulomb's cone (stick):

$$\Gamma_c^\bullet : \underline{\sigma}_t \in C(\sigma_n), \sigma_n < 0.$$

The boundary condition (4.24) is not transparent and it is preferable to prescribe directly the position of the point in the actual configuration, which corresponds to prescribing the boundary conditions in the actual configuration at time  $t > t^\bullet$ , then Eq. (4.24) can be rewritten simply as

$$\underline{x}(t) = \underline{x}(t^\bullet), \quad \text{if } \underline{x}(t) \in \Gamma_c^\bullet. \quad (4.25)$$

The interpretation is simple: if a point switches to a stick state it glues to the position at which this switch has occurred and cannot move before another switch to any other state happens.

The boundary conditions in the slip zone  $\Gamma_c^*$  (4.23,b) correspond to the partial Dirichlet boundary conditions for the  $z$  component of the displacement (in general, component orthogonal to the rigid surface) and to the partial Neumann boundary conditions in the contact plane  $OXY$  (in general, tangential plane), in the actual configuration, written as

$$x_z = 0, \quad \underline{\sigma}_t = \mu|\sigma_n| \frac{\underline{\sigma}_t}{\|\underline{\sigma}_t\|}, \quad \text{if } \underline{x}(t) \in \Gamma_c^* \quad (4.26)$$



where the slip region is determined according to the stress state on the surface of Coulomb's cone  $\underline{\sigma}_t \in \partial C(\sigma_n)$ , i.e.  $\|\underline{\sigma}_t\| = \mu|\sigma_n|$ .

Without any loss of generality points are allowed to come in contact only through the stick state and escape from the contact by any state, this can be summarized as follows

$$\{\Gamma_c \setminus \bar{\Gamma}_c\} \rightarrow \Gamma_c^\bullet \rightarrow \begin{cases} \{\Gamma_c \setminus \bar{\Gamma}_c\} \\ \Gamma_c^* \end{cases} \rightarrow \begin{cases} \{\Gamma_c \setminus \bar{\Gamma}_c\} \\ \Gamma_c^\bullet \end{cases}$$

This assumption allows us to correctly prescribe the Neumann boundary condition for the slip state. Suppose that a material point  $\underline{X}$  touches the plane at  $\underline{x}(t)$  and changes its state from non-contact to stick. According to Eq. (4.24), the point glues to the plane and reaction stresses appear: a contact pressure  $\sigma_n < 0$  and a tangential stress  $\underline{\sigma}_t$ . Two cases are possible (Fig. 4.6): the tangential reaction is either in the Coulomb's cone  $\underline{\sigma}_t \in C(\sigma_n) \cup \partial C(\sigma_n)$  or outside of the cone. For the latter (Fig. 4.6,c), the point switches to the slip state and a Neumann boundary condition should be applied (Fig. 4.7) that is a shear stress vector in the direction of the reaction stress vector and with the norm equal to the maximal allowed frictional stress

$$\underline{\sigma}_t^e = \mu|\sigma_n| \frac{\underline{\sigma}_t}{\|\underline{\sigma}_t\|}. \quad (4.27)$$

Tangential displacement should be set free  $\forall u_x, u_y$ ; nonzero slip will occur naturally.

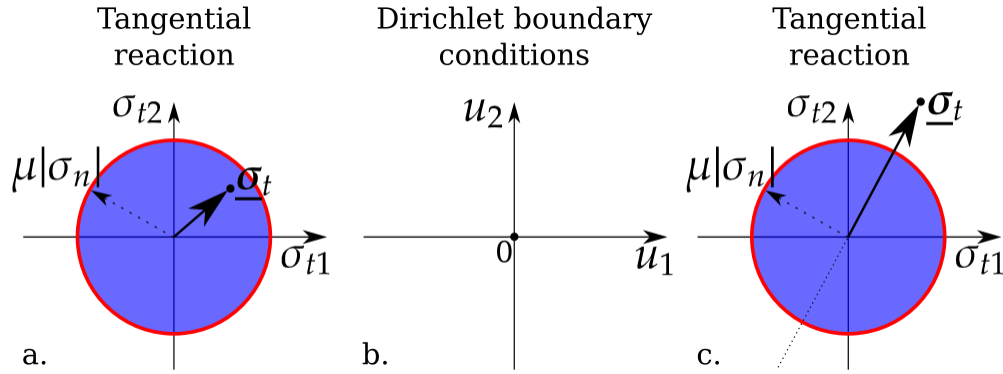


Figure 4.6: Dirichlet boundary conditions for the stick state (b): a – tangential reaction is in Coulomb's cone  $\underline{\sigma}_t \in C(\sigma_n) \cup \partial C(\sigma_n)$ ; c – tangential reaction is out of the Coulomb's cone  $\underline{\sigma}_t \notin C(\sigma_n) \cup \partial C(\sigma_n)$ , stick state should be switched to slip (Fig. 4.7).

The inactive contact zone  $\Gamma_c \setminus \bar{\Gamma}_c$  (4.23,c) remains free of boundary conditions. The source of nonlinearity comes from the presence of unknown stick  $\Gamma_c^\bullet \in \Gamma_c$  and slip  $\Gamma_c^* \in \Gamma_c$  zones ( $\Gamma_c^\bullet \cup \Gamma_c^* = \bar{\Gamma}_c$ ). As we assumed that points always pass through the stick state, this nonlinear problem can be splitted into two iteratively repeated problems:

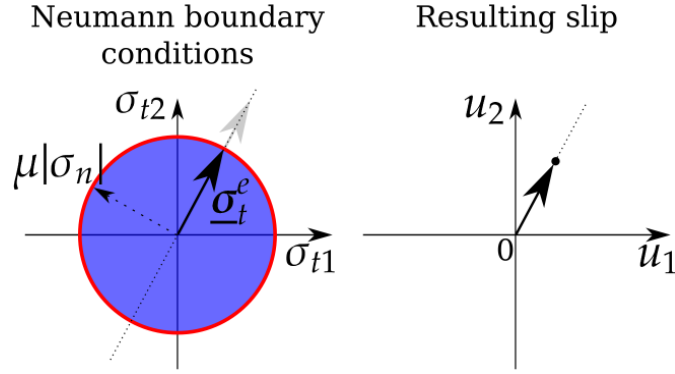


Figure 4.7: Neumann boundary conditions for the slip state: **left** – applied external force density in the direction of the tangential reaction appeared in the stick state (Fig.4.6,c) with norm equal to the radius of Coulomb's cone  $\mu|\sigma_n|$ ; **right** – resulting tangential slip.

- determination of the active contact zone  $\bar{\Gamma}_c$ ;
- determination of the slip zone  $\Gamma_c^* = \bar{\Gamma}_c \setminus \Gamma_c^\bullet$ .

Finally we get the following set of boundary conditions which replace frictional conditions formulated in (4.21):

$$\begin{cases} \underline{x} = \underline{x}(t^\bullet), & \text{if } \underline{x} \in \Gamma_c^\bullet \\ x_z = 0, \quad \underline{\sigma}_t^e = \mu|\sigma_n| \frac{\underline{\sigma}_t}{\|\underline{\sigma}_t\|}, & \text{if } \underline{x} \in \Gamma_c^* \end{cases} \quad (4.28)$$

This formulation is simple and its implementation is straightforward in a finite element code, if a dynamical update of the boundary conditions is allowed. As previously discussed, this method is applicable for a rigid surface determined by a constant value of one of the coordinates in the current reference frame

$$\zeta_i = \text{const},$$

for example, it can be a sphere in spherical coordinates  $r = \text{const}$ , a cylinder in polar coordinates  $\rho = \text{const}$ , or a plane in Cartesian coordinates  $x = \text{const}$ ,  $y = \text{const}$  or  $z = \text{const}$ . Obviously, any region of such surfaces can be chosen. The extension of such an approach for an arbitrary rigid surface will be discussed later.

#### 4.1.4 Non-classical friction and adhesion laws

In this section we discuss arbitrary non-associated friction laws written as

$$\tilde{C}(\sigma_n, \underline{\sigma}_t, \dot{\underline{\sigma}}_t, \dots).$$

All the expressions given below are valid for arbitrary contact between deformable solids. However, the case of contact with a motionless rigid plane

suggests a consistent framework for straightforward interpretation of these laws within the Dirichlet-Neumann boundary conditions and allows to avoid some details nonrelevant for their presentation (relativity of motion, nonplanar sliding paths, deformation of both contacting surfaces, etc.).

The first extension we consider is a non-associated anisotropic friction, the slip surface for a given contact pressure can be represented by a super ellipse as proposed in [Mròz 94]

$$C^{\mu_a}(\sigma_n, \underline{\sigma}_t) : \left\{ \underline{\sigma}_t \in C^{\mu_a} \text{ if } \left( \frac{\sigma_{t1}}{\mu_1} \right)^p + \left( \frac{\sigma_{t2}}{\mu_2} \right)^p < |\sigma_n|^p \right\}$$

A more general form of the anisotropic friction condition can be written as

$$C^{\mu_a}(\sigma_n, \underline{\sigma}_t) : \left\{ \underline{\sigma}_t \in C^{\mu_a} \text{ if } \|\underline{\sigma}_t\| < \mu(\phi)|\sigma_n| \right\},$$

where  $0 \leq \phi < 2\pi$  is an angle in polar coordinates on the  $OXY$  plane starting for example from the  $OX$  axis. Such definitions of friction are possible only if the anisotropy arises from the rigid plane, otherwise it is not possible to define the anisotropic slip surface according to the fixed reference frame since the rotation of the deformable body will change the axes of the super ellipse. In general case, the anisotropy function becomes  $\mu(\phi - \psi)$ , where  $\phi, \psi \in [0, 2\pi)$  are angles of orientations of the rigid plane and of the element of deformable surface with respect to the global reference frame. For references see [Michalowski 78], [Hjiaj 04b], [Hjiaj 04a] and others.

Following the same process as in (4.15), an isotropic adhesion can be introduced

$$C^{A_i} \left( \sigma_n, \max_{t \in [t_0, t]} \langle -\sigma_n \rangle \right) : \left\{ \underline{\sigma}_t \in C^{A_i} \text{ if } \|\underline{\sigma}_t\| < \mu \langle -\sigma_n \rangle + a_t \max_{t \in [t_0, t]} \langle -\sigma_n \rangle \right\}$$

where  $\max_{t \in [t_0, t]} \langle -\sigma_n \rangle$  is the absolute value of the maximal contact pressure at the current material point,  $t_0$  is the moment when this point comes in contact;  $a_t \geq 0$  is the tangential adhesion coefficient. If adhesion is assumed to be anisotropic (in tangential plane) (see, for example, [Chen 07]) then the expression for the slip surface takes the following form

$$C^{A_a} \left( \sigma_n, \underline{\sigma}_t, \max_{t \in [t_0, t]} \langle -\sigma_n \rangle \right) : \left\{ \underline{\sigma}_t \in C^{A_a} \text{ if } \|\underline{\sigma}_t\| < \mu \langle -\sigma_n \rangle + a_t(\phi) \max_{t \in [t_0, t]} \langle -\sigma_n \rangle \right\},$$

where  $a_t(\phi) \geq 0$  is the adhesion coefficient depending on the angle  $\phi$  (see Fig. 4.8,b) (as above, in general case, the angle  $\phi$  should be replaced by difference  $\phi - \psi$ ). Finally, the anisotropic friction can be coupled with an anisotropic adhesion, however this formulation is very complex and is given only for the sake of generality:

$$C^{\mu A_a} \left( \sigma_n, \underline{\sigma}_t, \max_{t \in [t_0, t]} \langle -\sigma_n \rangle \right) : \left\{ \underline{\sigma}_t \in C^{\mu A_a} \text{ if } \|\underline{\sigma}_t\| < \mu(\phi) \langle -\sigma_n \rangle + a_t(\phi) \max_{t \in [t_0, t]} \langle -\sigma_n \rangle \right\},$$

where the friction coefficient is assumed to depend on the angle  $\phi$  (in general,  $\phi - \psi$ ).

#### 4.1 Unilateral contact with a rigid plane

In Figure 4.8 several realizations of the discussed non-classical friction laws are illustrated. In Fig. 4.8,a slip surface and its section are presented for friction with isotropic adhesion for a given  $\max_{t \in [t_0, t]} \langle -\sigma_n \rangle = \text{const}$ . Other examples are given only for a single section of the slip surface: anisotropic adhesion and anisotropic friction presented in Figures 4.8,b,c respectively, the slip surface is given by the super ellipse with  $p = 2$ . A combination of both anisotropic friction and adhesion is presented in Fig. 4.8,d; an example of highly anisotropic friction is given in Fig. 4.8,e. Note that

- for adhesive friction, even for certain tensile contact pressure, the contact interface resists to tangential sliding;
- transition from stick to slip occurs at the outer boundary of the slip surface, so for a given  $\max_{t \in [t_0, t]} \langle -\sigma_n \rangle = \text{const}$  there is no difference between the blue area (due to friction) and the green one (due to adhesion).

Another important extension of the frictional rule comes from metal forming applications. The tangential and normal resistances of the contact interface are proportional to the real contact area [Rabinowicz 65], however, this is true only for moderate contact pressures: if the real area approaches the apparent contact area, the tangential resistance also approaches its limit, whereas the normal resistance keeps on increasing up to the limits determined by the compressive properties of the material. It results in deviation of the linear relation between the normal contact pressure and tangential resistance. Metal forming or metal processing with rigid tools imply very high pressures in the contact interface, that is why in such analyzes the tangential resistance of the interface has to be limited by a certain value independent of the normal pressure. A very simple model introducing a limit for the tangential resistance is Coulomb-Orowan friction law (Fig. 4.9,a),

$$\|\underline{\sigma}_t\| = \min \{ \mu |\sigma_n|, \sigma_c \}. \quad (4.29)$$

However, this purely phenomenological law does not reflect the fact that according to experiments [Bay 87] the tangential resistance saturates gradually. Another significant drawback relates to the non-smoothness of this model, which affects the convergence of numerical computations. A more sophisticated micro-mechanical based model has been proposed by Bay et al. in 1987 [Bay 87], where authors related the tangential resistance and the real contact area fraction  $A_r/A_a$

$$\|\underline{\sigma}_t\| = f \frac{A_r}{A_a} \sigma_c, \quad (4.30)$$

where  $0 \leq f \leq 1$  is a “friction factor”,  $A_r$  and  $A_a$  are the real and apparent contact areas respectively and  $\sigma_c$  is the constant shear yield strength (Fig. 4.9,b). For moderate pressures the real contact area is proportional to the contact pressure with the constant of proportionality  $\alpha$   $A_r = \alpha |\sigma_n|$ , it results in a linear relation between slip surface and contact pressure as in Coulomb’s law for small pressures

$$\|\underline{\sigma}_t\| = \left[ \frac{f \alpha \sigma_c}{A_a} \right] |\sigma_n|.$$

For higher pressures, when almost all the gaps in the contact interface are closed at all scales, the relation between the real contact area and the contact pressure becomes nonlinear and the fraction approaches gradually its limit

$$A_r/A_a \rightarrow 1$$

So this law gives the following relation

$$\|\underline{\sigma}_t\| = \frac{f\sigma_c}{A_a} A_r(|\sigma_n|) \rightarrow f\sigma_c.$$

This micro-mechanical model gives a qualitative explanation based on the real contact area growth, however, many other factors which also influence the tangential resistance of the contact interface have not been taken into account (interlocking asperities, their elasto-plastic deformation and fracture, intense plastic deformation of near surface metal layers, adhesion of newly created uncontaminated material surfaces and wear particles [Suh 81], [Suh 86]). Other models can be found in [Stupkiewicz 01] and [Wriggers 06].

An important drawback of the discussed friction models is that they do not take into account the influence of the history of relative slip, which has been clearly demonstrated by the well-known tests due to Courtney-Pratt and Eisner (1957) [Courtney-Pratt 57]. Following these experiments, more sophisticated models based on friction-plasticity analogy have been proposed, see e.g. [Anand 93] and many others. Another important point is that real contact area is time dependent due to viscous properties of solids it implies time dependence of the frictional resistance and as a consequence, its dependence on the velocity of sliding. Moreover, the temperature dependence of the sliding may be crucial for some applications. Since the frictional sliding is always related to the energy dissipation due to conversion of kinetic and potential energy in heat, the local temperature of the sliding point may be strongly affected by the locally dissipated energy or in other words by the total slip and its velocity. That is why sometimes it is important to take into account a phenomenological dependence of the friction on these quantities or explicitly on the temperature. For a general discussion on friction, adhesion and physics of contact the reader is referred to [Bowden 50], [Rabinowicz 65], [Popov 10] and many others. Discussion of different contact and friction models and their implementations in a finite element framework can be found in [Wriggers 06].

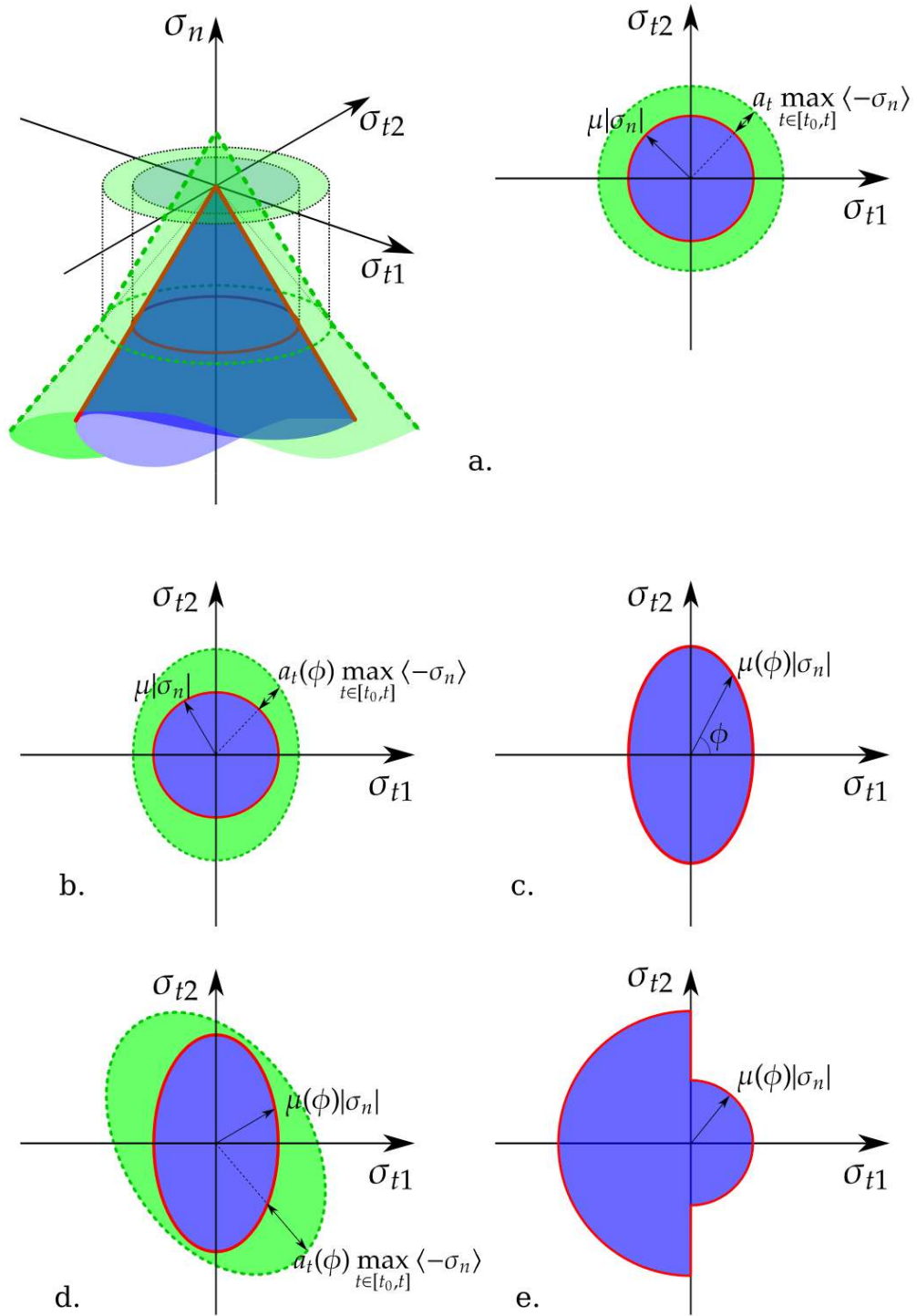


Figure 4.8: Examples of non-classical friction laws: **a** – adhesive friction, **b** – anisotropic adhesion in friction, **c** – anisotropic friction slip surface of elliptic shape, **d** – anisotropic friction combined with anisotropic adhesion, **e** – highly anisotropic friction slip surface.

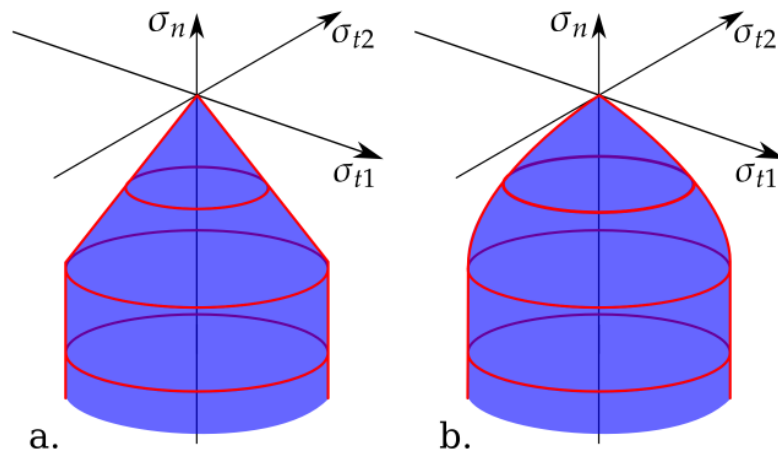


Figure 4.9: Graphical representation of friction laws for high contact pressures  
**a** – Coulomb-Orowan model, **b** – from Bay et al. [Bay 87].

## 4.2 Unilateral contact with an arbitrary rigid surface

In the previous section the treatment of frictionless and frictional problems of a deformable body with a rigid plane was discussed. A simple technique replacing the classical contact conditions by partial Dirichlet-Neumann's boundary conditions on unknown a priori parts of body's surface has been elaborated in order to demonstrate the analogy between contact and ordinary boundary conditions. Now let us generalize the formulation of the contact problem for the case of an arbitrary rigid surface: frictionless (Signorini's problem) and frictional formulation are considered. A non-selfintersecting, motionless and smooth rigid surface  $S$  is described by a vector

$$\underline{r}(\zeta) \in S,$$

where  $\zeta$  is a v-scalar of surface coordinates (see Appendix A.6). Each point of the surface has two basis vectors forming v-vectors in covariant  $\frac{\partial \underline{r}}{\partial \zeta}$  and in contravariant coordinates  $\overline{\frac{\partial \underline{r}}{\partial \zeta}}$

$$\frac{\partial \underline{r}}{\partial \zeta} = \underline{\underline{A}}_{\zeta} \overline{\frac{\partial \underline{r}}{\partial \zeta}}, \quad \underline{\underline{A}}_{\zeta} = \frac{\partial \underline{r}}{\partial \zeta} \cdot \frac{\partial \underline{r}}{\partial \zeta}^T, \quad \underline{\underline{A}}_{\zeta} \bar{\underline{\underline{A}}}_{\zeta} = \underline{\underline{I}},$$

where  $\underline{\underline{A}}_{\zeta}$  is a first fundamental surface matrix (t-scalar in s-structure notations),  $\bar{\underline{\underline{A}}}_{\zeta} = \underline{\underline{A}}_{\zeta}^{-1}$  its inverse - first fundamental contravariant surface t-scalar. For details see Chapter 2. The surface coordinates are chosen in the way that the external unit normal to the surface is determined by

$$\underline{\underline{v}} = \frac{\frac{\partial \underline{r}}{\partial \zeta_1} \times \frac{\partial \underline{r}}{\partial \zeta_2}}{\left\| \frac{\partial \underline{r}}{\partial \zeta_1} \times \frac{\partial \underline{r}}{\partial \zeta_2} \right\|}.$$

Initially, the deformable body is situated entirely on one side of the rigid surface and is not allowed to penetrate on the other side.

The surface of the body  $\partial\Omega$  can be described in exactly the same manner. Each point of the surface in the actual configuration is characterized by vector

$$\underline{\underline{\rho}}(t, \xi) \in \partial\Omega,$$

where  $\xi$  is a v-scalar of surface coordinates. The corresponding pair of covariant basis vectors  $\frac{\partial \underline{\rho}}{\partial \xi}$  are enumerated to get the outward unit normal as

$$\underline{\underline{n}} = \frac{\frac{\partial \underline{\rho}}{\partial \xi_1} \times \frac{\partial \underline{\rho}}{\partial \xi_2}}{\left\| \frac{\partial \underline{\rho}}{\partial \xi_1} \times \frac{\partial \underline{\rho}}{\partial \xi_2} \right\|}.$$

The first fundamental covariant surface matrix (t-scalar) is  $\underline{\underline{A}}_{\xi} = \frac{\partial \underline{\rho}}{\partial \xi} \cdot \frac{\partial \underline{\rho}}{\partial \xi}^T$  and its contravariant form is given by  $\bar{\underline{\underline{A}}}_{\xi} = \underline{\underline{A}}_{\xi}^{-1}$ .



#### 4.2.1 Non-penetration condition

The contact between a point  $\underline{\rho} \in \Gamma_c \subset \partial\Omega$  and a the rigid surface in the point  $\underline{r}$  implies two equalities if both surfaces are smooth in the vicinity of the contact point

$$\begin{cases} \underline{\rho}(t) - \underline{r} = 0 & (1) \\ \underline{n}(t) + \underline{\nu} = 0 & (2) \end{cases} \quad (4.31)$$

If at least for one of the surfaces the normal cannot be uniquely determined, then only the first equality holds (4.31.1). For the sake of simplicity here and afterwards, we will suppose that the contacting surfaces are locally smooth. Then the nonpenetration condition can be reformulated in two equivalent forms

$$\begin{cases} (\underline{\rho}(t + \delta t) - \underline{\rho}(t)) \cdot \underline{n}(t) \leq 0 \\ (\underline{\rho}(t + \delta t) - \underline{\rho}(t)) \cdot \underline{\nu} \geq 0 \end{cases} \quad (4.32)$$

which imply that in the “near future” ( $t + \delta t$ , where  $\delta t$  is infinitely small), the point  $\underline{\rho}$  either remains on the surface or moves outward the rigid surface or parallel to it. To get a more compact form, one can divide the inequalities (4.31) by  $\delta t$ , take the limit and get the formulation in terms of velocities

$$\begin{cases} \underline{\dot{\rho}}(t) \cdot \underline{n}(t) \leq 0 \\ \underline{\dot{\rho}}(t) \cdot \underline{\nu} \geq 0 \end{cases} \quad (4.33)$$

Note all given formulations are valid only for the active contact area.

#### Gap function

The inequalities derived above are valid only for the contact state and are local in time. A more consistent description requires a scalar gap function  $g(\underline{\rho}, S)$  for each point of the deformable surface  $\underline{\rho}$ . This function is positive if the point  $\underline{\rho}$  is not in contact, is zero if the point is in contact and is negative if the point penetrates. The gap can be defined relatively to the points of the body surface  $g(\underline{\rho} \in \partial\Omega, S)$  or to the points of the rigid surface  $g(\underline{r} \in S, \partial\Omega)$ . The two descriptions are not identical (due to asymmetry of gap definition) but equivalent, however, the first one seems to be more natural. In this case we require for non-penetration that

$$g(\underline{\rho}, S) \geq 0. \quad (4.34)$$

Some realizations of the gap function have been presented in Chapter 2:

- the classical **normal gap**

$$g_n(\underline{\rho}, S) = (\underline{\rho} - \underline{r}) \cdot \underline{\nu}(\underline{r}),$$

where  $\underline{r} \in S$  is the point on the rigid surface closest to the point  $\underline{\rho}$  and  $\underline{\nu}$  - the normal at point  $\underline{r}$ . This definition implies that the rigid surface is smooth;

## 4.2 Unilateral contact with an arbitrary rigid surface

- **the closest point gap** can be defined as the distance to the closest point with the relevant sign

$$g_c(\underline{\rho}, S) = \beta \|\underline{\rho} - \underline{r}\|,$$

where  $\beta \in \{-1; 1\}$  determines the sign of the gap; this definition does not require smoothness of the surfaces and is more appropriate for the description of discretized media, however, additional efforts have to be undertaken to determine the sign of  $g$ . For example, an average normal can be assigned, if in a small vicinity  $\rho_0$  of the point  $\underline{r}$  there are smooth open sets  $C_i(\underline{r}, \rho_0)$ ,  $i = 1, N$  of points possessing a normal vector  $\underline{v}'$

$$\underline{r}' \in C_i(\underline{r}, \rho_0) \subset S, \|\underline{r}' - \underline{r}\| < \rho_0 \text{ and } \exists \underline{v}',$$

then a normal can be assigned to the point  $\underline{r}$  in the following way

$$\tilde{\underline{v}} = \frac{1}{N} \sum_i \frac{1}{A_i(\rho_0)} \int_{C_i(\underline{r}, \rho_0)} \underline{v}' dA, \text{ where } A_i(\rho_0) = \int_{C_i(\underline{r}, \rho_0)} dA.$$

Other definitions of the normal are discussed in Chapter 3, Remark 3.2 “On normal definition for contact nodes”. Now the closest point gap  $g_c$  can be rewritten like the normal gap

$$g_c(\underline{\rho}, S) = (\underline{\rho} - \underline{r}) \cdot \tilde{\underline{v}}(C_i(\underline{r}, \rho_0)).$$

On the other hand, if the contact occurs between a nonsmooth point of one surface with a smooth point of another, due to requirement of equality of normals (4.31),(2) one can easily assign the needed normal at the nonsmooth point  $\tilde{\underline{v}} = -\underline{n}$ .

- **the shadow gap**  $g_s$  (discussed in Chapter 2) is the signed distance between the point  $\underline{\rho}$  and its “shadow projection” on the rigid surface  $S$

$$g_s(\underline{\rho}, S) = (\underline{\rho} - \underline{r}) \cdot \underline{e}(\underline{\rho}),$$

where  $\underline{e}$  is a unit vector pointing to the emitter of the light from the point  $\underline{\rho}$ .

Note that in general

$$g_n \neq g_c \neq g_s \neq g_n$$

**Remark 4.1 on the asymmetry in gap definition.**

According to (4.31), Eq. (4.32) can be rewritten as

$$\begin{cases} (\underline{\rho}(t + \delta t) - \underline{r}) \cdot \underline{n}(t) \leq 0 \\ (\underline{\rho}(t + \delta t) - \underline{r}) \cdot \underline{\nu} \geq 0 \end{cases} \quad (4.35)$$

the second line is equivalent to the definition of the normal gap  $g_n$ . By analogy, one could define the gap according to the first line of (4.35)

$$g(\underline{\rho}, S) = (\underline{r} - \underline{\rho}) \cdot \underline{n},$$

However, this is not correct and leads to wrong results. Such a definition  $g_c(\underline{\rho}, S)$  implies that  $\underline{r} \in S$  in (4.35) is the closest point to  $\underline{\rho} \in \partial\Omega$ . In general, the closest point  $\underline{r} \in S$  for  $\underline{\rho}$  does not coincide with the closest point  $\underline{\rho}' \in \partial\Omega$  for the point  $\underline{r}$ . The correct symmetric definition for the gap with respect to the surface  $\partial\Omega$  has the following form

$$g(\underline{r}, \partial\Omega) = (\underline{r} - \underline{\rho}) \cdot \underline{n}.$$

Sometimes it is convenient to introduce a gap vector  $\underline{g}(\underline{\rho}, S)$ :

- **normal gap vector**

$$\underline{g}_n(\underline{\rho}, S) = \underline{\rho} - \underline{r},$$

where  $\underline{r} \in S \in C^1$  is the closest point to the point  $\underline{\rho}$ ;  $\|\underline{g}_n\| = |g_n|$ ;

- **closest point gap vector**

$$\underline{g}_c(\underline{\rho}, S) = \underline{\rho} - \underline{r},$$

where  $\underline{r} \in S \in C^0$  is the closest point to the point  $\underline{\rho}$ ;  $\|\underline{g}_c\| = |g_c|$ ;

- **shadow gap vector**

$$\underline{g}_s(\underline{\rho}, S) = \underline{\rho} - \underline{r},$$

where  $\underline{r} = \underline{\rho} - g_s \underline{e}$  is the shadow of the point  $\underline{\rho}$  corresponding to the light emitter situated in the direction  $\underline{e}$  from the point  $\underline{\rho}$ ;  $\|\underline{g}_s\| = |g_s|$ .

For an arbitrary smooth surface  $S$  and infinitely small gaps or for a locally flat surface the following equality holds

$$\underline{g}_n = \underline{\nu} \otimes \underline{\nu} \cdot \underline{g}_s \Rightarrow \underline{\nu} \cdot \underline{g}_n = \underline{\nu} \cdot \underline{e} g_s,$$

and

$$\underline{g}_n = \underline{\nu} \cdot \underline{e} g_s \quad (4.36)$$

## 4.2 Unilateral contact with an arbitrary rigid surface

Finally, the penetration is restricted by the following inequality formulated for the points of the deformable surface  $\underline{\rho} \in \partial\Omega$  respectively to the rigid surface  $S$

$$g(\underline{\rho}, S) \geq 0, \quad \underline{\rho} \in \Gamma_c \subset \partial\Omega \quad (4.37)$$

The set of points for which the gap function is equal to zero is nothing but the active contact zone  $\bar{\Gamma}_c$

$$g(\underline{\rho} \in \bar{\Gamma}_c, S) = 0.$$

### 4.2.2 Hertz-Signorini-Moreau's contact conditions

As discussed previously, due to the restriction on penetration, a reaction stress arises at contacting points  $\underline{\rho} \in \bar{\Gamma}_c$ . For frictionless contact, the only nonzero term is the normal stress

$$\begin{aligned} \sigma_n &= \underline{n} \cdot \underline{\sigma} \cdot \underline{n}, \\ \underline{\sigma}_t &= \underline{\sigma} \cdot \underline{n} - \sigma_n \underline{n} = 0. \end{aligned} \quad (4.38)$$

If contact holds and surfaces are smooth, according to (4.31,(2))  $\underline{n} = -\underline{\nu}$  and consequently

$$\begin{aligned} \sigma_n &= \underline{\nu} \cdot \underline{\sigma} \cdot \underline{\nu} \leq 0, \\ \underline{\sigma}_t &= \underline{\sigma} \cdot \underline{\nu} - \sigma_n \underline{\nu} = 0, \end{aligned} \quad (4.39)$$

we require a non-positive contact pressure (for non-adhesive contact) and zero tangential stresses for frictionless contact. The complementary condition (switch between non-contact and contact states) holds the same as in the previous section. So the Hertz-Signorini-Moreau contact conditions for frictionless contact become

$$g \geq 0, \quad \sigma_n \leq 0, \quad g\sigma_n = 0, \quad \underline{\sigma}_t = 0 \quad (4.40)$$

The boundary value problem retains the same form as in the previous section:

$$\left\{ \begin{array}{ll} \nabla \cdot \underline{\sigma} + \underline{f} = 0 & \text{in } \Omega \\ \underline{\sigma} \cdot \underline{\nu} = \underline{\sigma}_0 & \text{at } \Gamma_f \\ \underline{u} = \underline{u}_0 & \text{at } \Gamma_u \\ g \geq 0, \quad \sigma_n \leq 0, \quad \sigma_n g = 0, \quad \underline{\sigma}_t = 0 & \text{at } \Gamma_c \end{array} \right. \quad (4.41)$$

Existence and uniqueness of the solution of this problem has been proven for linear material and small deformations by Kikuchi and Oden [Kikuchi 88]. For further numerical analysis, this problem has to be reformulated in a proper way. It will be done later for the general case of two deformable bodies. Before, an attempt to extend the approach considered in the previous section will be undertaken. The ultimate aim is to replace classical contact conditions by boundary conditions imposed on the active contact zone for an arbitrary rigid surface  $S$ .

### 4.2.3 Interpretation of contact conditions

To replace the contact conditions by a special type of Dirichlet-Neumann's boundary conditions, we return first to Eq. (4.32)

$$(\underline{\rho}(t + \delta t) - \underline{\rho}(t)) \cdot \underline{\nu} \geq 0,$$

which can be rewritten if one puts  $\underline{\rho}(t + \delta t) = \underline{\rho}(t) + \delta \underline{u}(t)$

$$\delta \underline{u}(t) \cdot \underline{\nu} \geq 0. \quad (4.42)$$

The small variation of the displacement  $\delta \underline{u}$  may be splitted in an orthonormal system of coordinates into three vectors

$$\delta \underline{u} = \delta u_1 \underline{e}_1 + \delta u_2 \underline{e}_2 + \delta u_3 \underline{e}_3.$$

In the active contact zone  $\bar{\Gamma}_c$ , the inequality (4.42) converts into an equality, which can be rewritten using the previously splitted form

$$\delta u_1 \cos(\phi_1) + \delta u_2 \cos(\phi_2) + \delta u_3 \cos(\phi_3) = 0,$$

where  $\cos(\phi_i) = \underline{e}_i \cdot \underline{\nu}$ . Since the vectors  $\underline{e}_i$  are orthonormal, at least one of the three cosine is nonzero. Without any loss of generality let us assume, that  $\cos(\phi_3) \neq 0$ , then we deduce directly

$$\delta u_3 = -\delta u_1 \frac{\cos(\phi_1)}{\cos(\phi_3)} - \delta u_2 \frac{\cos(\phi_2)}{\cos(\phi_3)}, \quad (4.43)$$

According to the later equality, one of the three displacement components of the point  $\underline{\rho} \in \bar{\Gamma}_c$  depends on two other components. Another way to express this relation consists in the following: if for  $\underline{\nu} \cdot \underline{e}_3 \neq 0$  and its vicinity the rigid surface can be locally presented as

$$\underline{r} \in S : R_3 = s(R_1, R_2), \quad (4.44)$$

where  $R_i$  are the coordinates of the point  $\underline{r}$  in the orthonormal coordinate system. Then the fact that the point is in contact  $\underline{\rho} \in \bar{\Gamma}_c$  implies that

$$\rho_3 = s(\rho_1, \rho_2) \Leftrightarrow \rho_3(t_0 + \delta t) = s(\rho_1(t_0 + \delta t), \rho_2(t_0 + \delta t)).$$

If one expands the last equality in Taylor's series, keeps the first two terms and subtracts  $\rho_3(t_0)$  from the left part and  $s(\rho_1(t_0), \rho_2(t_0))$  from the right part, then for small  $\delta t$  we get the relation between displacement components

$$\delta u_3 = \frac{\partial s}{\partial \rho_1} \delta u_1 + \frac{\partial s}{\partial \rho_2} \delta u_2, \quad (4.45)$$

which is equivalent to the relation (4.43). If the considered rigid surface is flat and orthogonal to the vector  $\underline{e}_3$ , the following equality holds

$$s(R_1, R_2) = \text{const} = R_3,$$

## 4.2 Unilateral contact with an arbitrary rigid surface

then from (4.45) we get directly the boundary condition which we imposed in the previous section

$$\delta u_3 = 0.$$

A more general connection between displacements follows directly from (4.44): if the point  $\underline{\rho}$  is in contact, it lies on the surface  $S$  and its coordinates fulfill the condition

$$\rho_3 = s(\rho_1, \rho_2) \Leftrightarrow u_3(t_0 + t) = s(\rho_1(t_0) + u_1(t_0 + t), \rho_2(t_0) + u_2(t_0 + t)) - \rho_3(t_0),$$

$$\boxed{u_3 = s(\rho_1 + u_1, \rho_2 + u_2) - \rho_3, \quad \text{if } \underline{\rho} + \underline{u} \in \bar{\Gamma}_c} \quad (4.46)$$

This is the relation between the components of the displacement for a point being in contact on the surface  $S : R_3 = s(R_1, R_2)$ , where  $u_i$  are the displacement components with respect to the configuration at time  $t \geq t_0$ , where  $t_0$  is the moment when the point comes in contact and  $\rho_i$  are the spatial coordinates of this point at time  $t_0$ . The active contact zone is determined by a negative contact pressure and a zero tangential stress at point  $\underline{\rho}$

$$\begin{cases} \sigma_n \leq 0 \\ \underline{\sigma}_t = 0 \end{cases}, \quad \text{on } \bar{\Gamma}_c.$$

The second condition is automatically fulfilled if the relations between the components of the displacement are imposed in such a way that the point slides freely along the tangential plane (4.45). The condition of negative contact pressure has to be checked to determine the active contact zone.

### 4.2.4 Frictional conditions and their interpretation

In case of a contact with an arbitrary rigid surface the frictional conditions derived in the previous section hold

$$\boxed{\|\underline{\sigma}_t\| \leq \mu|\sigma_n|, \|\underline{\dot{g}}_t\|\underline{\sigma}_t - \mu|\sigma_n|\underline{\dot{g}}_t = 0, \quad \|\underline{s}\| \left\| \|\underline{\dot{g}}_t\|\underline{\sigma}_t - \mu|\sigma_n|\underline{\dot{g}}_t \right\| = 0.} \quad (4.47)$$

As previously, the sliding direction is defined as

$$\underline{s} = \begin{cases} \frac{\underline{\dot{g}}_t}{\|\underline{\dot{g}}_t\|}, & \text{if } \|\underline{\dot{g}}_t\| > 0; \\ 0, & \text{if } \|\underline{\dot{g}}_t\| = 0. \end{cases}$$

Contrary to the case of a rigid plane, the local reference frame changes when the point  $\underline{\rho}$  slides over the surface  $S$ . Stick state implies that the point  $\underline{\rho} \in \Gamma_c^\bullet$  retains its position on the surface and since the surface is motionless

$$\underline{\rho}(t^\bullet + \delta t) = \underline{\rho}(t^\bullet), \quad \underline{\rho} \in \Gamma_c^\bullet,$$

where  $t^\bullet$  is the time when the point switches to stick state. The same statement reformulated for displacements gives simply

$$\boxed{\underline{u}(t^\bullet + \delta t) = 0, \quad \underline{\rho} \in \Gamma_c^\bullet} \quad (4.48)$$

where the stick zone  $\Gamma_c^\bullet$  is determined from the stress vector  $\underline{\sigma}$ , which should be situated in the interior of Coulomb's cone

$$\underline{x} \in \Gamma_c^\bullet : \sigma_n < 0, \|\underline{\sigma}_t\| < \mu|\sigma_n|. \quad (4.49)$$

The stick condition can then be simply replaced by an ordinary Dirichlet boundary condition.

The slip condition requires a slightly deeper analysis. For that, as previously, we suppose that locally  $\underline{\nu} \cdot \underline{e}_3 \neq 0$  and that a point coming in contact switches first to the stick state, so that a reaction stress vector  $\underline{\sigma}$  arises and if

$$(\underline{I} - \underline{n} \otimes \underline{n}) \cdot \underline{\sigma} \geq \mu|\underline{\sigma} \cdot \underline{n}| \Leftrightarrow \underline{\sigma}_t \geq \mu|\sigma_n|$$

then this point switches to the slip state. It remains to suppress the corresponding part of Dirichlet's boundary conditions to let the point slip along the rigid surface and to apply an external frictional surface force density in the direction of sliding. Additional investigations are required to ensure the consistency of the problem when Dirichlet-Neumann boundary conditions are simultaneously prescribed. A simpler approach consists in defining an independent local basis for each contact point respectively to the rigid surface  $S$ . This provides a consistent set of boundary conditions as discussed in Paragraph 4.1.3.

## 4.3 Contact between deformable solids

### 4.3.1 General formulation and variational inequality

Up to now we considered a relatively simple case when a deformable body comes in contact with a rigid smooth surface. Here, the formulation will be generalized for the case when the contact interface separates two deformable bodies: it is not important if the contact occurs between parts of one body (contact within a crack, self-contact) or between several separate bodies, the number of separate contact zones does not change the problem neither. So without any loss of generality we confine ourself to the formulation of the problem when contact occurs at a single contact zone between two separate deformable bodies  $\Omega^1$  and  $\Omega^2$ .

As previously, the vector  $\underline{X} \in \Omega^0$  denotes the position of the material point in the reference configuration and the vector  $\underline{x} \in \Omega$  in the actual configuration. To simplify the equations, we introduce the following notation

- the union of two bodies denotes two open sets  $\Omega = \Omega^1 \cup \Omega^2$ ;
- the union of their closures  $\partial\Omega = \partial\Omega^1 \cup \partial\Omega^2$ ;
- the union of surfaces, where Neumann boundary conditions are applied

$$\partial\Omega \supset \Gamma_f = \Gamma_f^1 \cup \Gamma_f^2;$$

### 4.3 Contact between deformable solids

- the union of surfaces, where Dirichlet boundary conditions are applied

$$\partial\Omega \supset \Gamma_u = \Gamma_u^1 \cup \Gamma_u^2,$$

- the potential contact surfaces of two bodies  $\Gamma_c^1 \subset \partial\Omega^1$  and  $\Gamma_c^2 \subset \partial\Omega^2$ .

The static balance of momentum states that for each point  $\underline{x} \in \Omega$  at any time the volume is in equilibrium if and only if

$$\nabla \cdot \underline{\underline{\sigma}} + \underline{f}_v = 0 \quad \text{in } \Omega, \quad (4.50)$$

where  $\underline{\underline{\sigma}}$  is the Cauchy stress tensor and  $\underline{f}_v$  is the volume force density. If (4.50) is satisfied in each point of the volume  $\Omega$  then the integral of the dot product of this equation with any arbitrary vector-function  $\underline{v}$  (also called test-function or virtual function) over the volume is zero, the converse statement is also true:

$$\nabla \cdot \underline{\underline{\sigma}} + \underline{f}_v = 0 \quad \text{in } \Omega \Leftrightarrow \forall \underline{v}, \int_{\Omega} [\nabla \cdot \underline{\underline{\sigma}} + \underline{f}_v] \cdot \underline{v} d\Omega = 0. \quad (4.51)$$

If we require that  $\underline{v} \in C^1$ , then the first term in the right part of eq.(4.51) can be integrated by parts. After using the Green's formula, we get

$$\int_{\partial\Omega} \underline{n} \cdot \underline{\underline{\sigma}} \cdot \underline{v} d\Gamma + \int_{\Omega} [\underline{f}_v \cdot \underline{v} - \underline{\underline{\sigma}} : \nabla \underline{v}] d\Omega = 0, \quad (4.52)$$

where  $\underline{n}$  is an outward unit normal vector at  $\partial\Omega$ . There is no more terms containing the derivative of the stress tensor. It implies that the requirement on smoothness of the stress vector ( $\underline{\underline{\sigma}} \in C^1(\Omega)$ ) in the differential form (4.50), has now been replaced by a weaker requirement of continuity ( $\underline{\underline{\sigma}} \in C^0(\Omega)$ ). On the other hand, according to (4.52) the test function must be smooth  $\underline{v} \in C^1(\Omega)$ . Equation (4.52) is called the weak form of the equilibrium equation.

If the abstract test functions  $\underline{v} \in C^1(\Omega)$  are replaced by arbitrary test displacements (also called virtual displacements)  $\delta \underline{u} = \delta(\underline{x} - \underline{X}) = \delta \underline{x}$ , then the weak form (4.52) is nothing but the balance of virtual work

$$\int_{\partial\Omega} \underline{n} \cdot \underline{\underline{\sigma}} \cdot \delta \underline{u} d\Gamma + \int_{\Omega} [\underline{f}_v \cdot \delta \underline{u} - \underline{\underline{\sigma}} : \delta \nabla \underline{u}] d\Omega = 0, \quad (4.53)$$

The stress vector  $\underline{n} \cdot \underline{\underline{\sigma}}$  entering in the first term in (4.53) is not zero only in the active contact zones ( $\bar{\Gamma}_c^1 \in \Gamma_c^1, \bar{\Gamma}_c^2 \in \Gamma_c^2$ ), on the surface where the stress vector is prescribed ( $\Gamma_f$ ) and on the surface where the displacement is prescribed ( $\Gamma_u$ ). By definition, since displacements are prescribed at  $\Gamma_u$ ,  $\delta \underline{u} = 0$  at  $\Gamma_u$  so we get

$$\int_{\partial\Omega} \underline{n} \cdot \underline{\underline{\sigma}} \cdot \delta \underline{u} d\Gamma = \int_{\bar{\Gamma}_c^1} \underline{n} \cdot \underline{\underline{\sigma}} \cdot \delta \underline{\rho} d\bar{\Gamma}_c^1 + \int_{\bar{\Gamma}_c^2} \underline{\nu} \cdot \underline{\underline{\sigma}} \cdot \delta \underline{r} d\bar{\Gamma}_c^2 + \int_{\Gamma_f} \underline{\sigma}_0 \cdot \delta \underline{u} d\Gamma_f, \quad (4.54)$$

where  $\underline{\sigma}_0$  is a prescribed traction (Neumann boundary conditions),  $\underline{n}$  is a unit surface normal at  $\bar{\Gamma}_c^1$ ,  $\underline{\nu}$  is a unit surface normal at  $\bar{\Gamma}_c^2$ ,  $\underline{\rho} = \underline{x}$  at  $\bar{\Gamma}_c^1$  and  $\underline{r} = \underline{x}$  at  $\bar{\Gamma}_c^2$ .



In equilibrium state it follows from the 3rd Newton's law that

$$\underline{n} \cdot \underline{\sigma} d\bar{\Gamma}_c^1 = -\underline{v} \cdot \underline{\sigma} d\bar{\Gamma}_c^1$$

So the two integrals on the contact surfaces  $\bar{\Gamma}_c^1$  and  $\bar{\Gamma}_c^2$  can be replaced by one integral over any of the two surfaces, we chose the surface  $\bar{\Gamma}_c^1$ .

$$\int_{\bar{\Gamma}_c^1} \underline{n} \cdot \underline{\sigma} \cdot \delta \underline{\rho} d\bar{\Gamma}_c^1 + \int_{\bar{\Gamma}_c^2} \underline{v} \cdot \underline{\sigma} \cdot \delta \underline{r} d\bar{\Gamma}_c^2 = \int_{\bar{\Gamma}_c^1} \underline{n} \cdot \underline{\sigma} \cdot \delta(\underline{\rho} - \underline{r}) d\bar{\Gamma}_c^1, \quad (4.55)$$

where

$$\underline{r} - \underline{\rho} = \underline{g}(\underline{r}, \bar{\Gamma}_c^1)$$

is a gap vector describing the position of the point  $\underline{r}$  relatively to its projection  $\underline{\rho}$ . So if we want to determine the integrals in (4.55) both in contact and in non-contact states, the vector  $\underline{\rho}$  becomes a projection of the slave point  $\underline{r}$  on the master surface  $\bar{\Gamma}_c^1$ , i.e.  $\underline{\rho} = \underline{\rho}(t, \underline{r})$ . The expression  $\delta(\underline{r} - \underline{\rho})$  in the integral implies that the relative motion of the independent point  $\underline{r}(t + \delta t)$  is considered relatively to the other *independent point*  $\underline{\rho}(t + \delta t)$ , which *was* the projection of the point  $\underline{r}(t)$  in the non-perturbed state at time  $t$ .

In case of a shadow projection  $\underline{g}_s$  from an infinitely remote emitter (see for details the previous section and Chapter 2)  $\underline{r} = \underline{\rho} + \underline{g}_s \underline{e}$  and directly

$$\delta(\underline{r} - \underline{\rho}) = \delta \underline{g}_s \underline{e} + \frac{\partial \underline{\rho}}{\partial \xi} \delta \xi, \quad (4.56)$$

where  $\underline{e}$  is a unit vector pointing towards the emitter. In case of a normal gap,  $\underline{g}_n$ ,  $\underline{r} = \underline{\rho} + \underline{g}_n \underline{n}$  and consequently

$$\delta(\underline{r} - \underline{\rho}) = \delta \underline{g}_n \underline{n} + \underline{g}_n \delta \underline{n} + \frac{\partial \underline{\rho}}{\partial \xi} \delta \xi, \quad (4.57)$$

where  $\delta \underline{n} = \delta \underline{n} + \frac{\partial \underline{n}}{\partial \xi} \delta \xi$  is a full variation of the unit normal vector. It is worth noting that  $\delta \xi$  describes a perturbation of the local coordinate of projection of the point  $\underline{r}$  on the surface  $\bar{\Gamma}_c^1$  and not the displacement of the material point  $\underline{\rho}$ . The surface stress vector can be splitted into normal and tangential parts, the latter can be expanded into two components in the contravariant surface basis  $\frac{\partial \underline{\rho}}{\partial \xi}$

$$\underline{n} \cdot \underline{\sigma} = \underline{n} \sigma_n + \underline{\sigma}_t = \underline{n} \sigma_n + \underline{\sigma}_t^T \frac{\partial \underline{\rho}}{\partial \xi} \quad (4.58)$$

Obviously, if the local contravariant basis  $\frac{\partial \underline{\rho}}{\partial \xi}$  is not orthonormal, then

$$\|\underline{\sigma}_t\| \neq \sqrt{\sigma_{t1}^2 + \sigma_{t2}^2},$$

### 4.3 Contact between deformable solids

where  $\sigma_{ti}$ ,  $i = 1, 2$  are components of the v-scalar  $\underline{\varrho}_t$ .

In the contact state, when points  $\underline{\rho} = \underline{r}$ , the local moment balance is automatically fulfilled, as vectors  $\underline{\sigma}_\rho = \underline{\underline{\sigma}} \cdot \underline{n} = -\underline{\underline{\sigma}} \cdot \underline{\nu} = -\underline{\sigma}_r$ . When the two points are distant, for a non-zero stress vector, we require that the moment is zero, which is equivalent to the following equality

$$(\underline{r} - \underline{\rho}) \times \underline{\sigma}_r = 0 \quad \text{or} \quad (\underline{\rho} - \underline{r}) \times \underline{\sigma}_\rho = 0.$$

It requires that when two points are distant, the contact stress vector is collinear with the gap vector

$$\underline{\sigma} \parallel \underline{g}, \quad \text{if } g \neq 0 \quad (4.59)$$

Substituting (4.56) and (4.58) into the last integral in (4.55) leads to the expression for the contact contribution to the virtual work balance of the system in case of a shadow projection from an infinitely remote emitter

$$\begin{aligned} \int_{\bar{\Gamma}_c^1} \underline{n} \cdot \underline{\underline{\sigma}} \cdot \delta(\underline{\rho} - \underline{r}) d\bar{\Gamma}_c^1 &= - \int_{\bar{\Gamma}_c^1} \left( \underline{n} \sigma_n + \underline{\varrho}_t^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \cdot \left( \delta g_s \underline{e} + \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \delta \underline{\xi} \right) d\bar{\Gamma}_c^1 = \\ &= - \int_{\bar{\Gamma}_c^1} \left( \sigma_n \delta g_s \underline{e} \cdot \underline{n} + \delta g_s \underline{\varrho}_t^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{e} + \underline{\varrho}_t^T \delta \underline{\xi} \right) d\bar{\Gamma}_c^1, \end{aligned} \quad (4.60)$$

where the dot product of contravariant and covariant bases gives the unity tensor (t-scalar)  $\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} = \underline{\underline{I}}$  and is omitted in the last term. This expression will not be investigated further in the thesis and suggests further direction of development.

To derive a similar contribution to the virtual work balance in case of the normal projection  $g = g_n$ , let us first consider the dot product of (4.57) with (4.58)

$$\begin{aligned} &\left( \underline{n} \sigma_n + \underline{\varrho}_t^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \cdot \left( \delta g_n \underline{n} + g_n \delta \underline{n} + \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \delta \underline{\xi} \right) = \\ &= \sigma_n \delta g_n + \underbrace{\sigma_n \left[ \underline{n} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \right] \delta \underline{\xi}}_{=0} + \underbrace{\underline{\varrho}_t^T \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \underline{n} \right] \delta g_n}_{=0} + \underbrace{\underline{\varrho}_t^T \left[ \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}^T}{\partial \underline{\xi}} \right] \delta \underline{\xi}}_{=\underline{\underline{I}}} + \\ &+ g_n \underbrace{\left( \underline{n} \sigma_n + \underline{\varrho}_t^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \right) \cdot \delta \underline{n}}_{=g_n \underline{\sigma} \cdot \delta \underline{n}=0, \text{ since } \underline{\sigma} \parallel \underline{n}, \text{ if } g_n \neq 0, \text{ see (4.59)}} = \sigma_n \delta g_n + \underline{\varrho}_t^T \delta \underline{\xi} \end{aligned} \quad (4.61)$$

Substituting this expression into the last integral in (4.55) gives the expression for the contact contribution to the virtual work in case of the normal projection

$$g = g_n$$

$$\int_{\bar{\Gamma}_c^1} \underline{n} \cdot \underline{\sigma} \cdot \delta(\underline{\rho} - \underline{r}) d\bar{\Gamma}_c^1 = - \int_{\bar{\Gamma}_c^1} (\sigma_n \delta g_n + \underline{\varrho}_t^T \delta \underline{\xi}) d\bar{\Gamma}_c^1, \quad (4.62)$$

Finally, the balance of virtual work (4.53) including contact contributions (in case of the normal projection) and Neumann boundary conditions from (4.54) is

$$\int_{\Omega} \underline{\sigma} \cdot \delta \nabla \underline{u} d\Omega + \int_{\bar{\Gamma}_c^1} (\sigma_n \delta g_n + \underline{\varrho}_t^T \delta \underline{\xi}) d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_f} \underline{\sigma}_0 \cdot \delta \underline{u} d\Gamma + \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} d\Omega, \quad (4.63)$$

It can be shown that in *actual configuration* the second term the part related to the normal contact contribution is negative. If the condition of non-penetration holds, then point  $\underline{r}$  either stays at the surface (sticks or slides over the tangential plane  $\delta g_n = 0$ ) or moves out of the surface, which leads to inequality

$$\delta g_n = \delta \underline{g}_n \cdot \underline{n} \geq 0.$$

Since the contact pressure is non-positive  $\sigma_n \leq 0$  we have  $\delta g_n \sigma_n \leq 0$ , which becomes in integral form

$$\int_{\bar{\Gamma}_c^1} \sigma_n \delta g_n d\bar{\Gamma}_c^1 \leq 0. \quad (4.64)$$

Frictional sliding is associated with energy dissipation, i.e. the dot product between force and displacement vectors is negative (Fig. 4.10), so naturally, a point slides in the direction opposite to the imposed frictional force. More generally, the angle between the frictional force and the sliding directions is larger than 90 degrees, i.e.

$$(\underline{I} - \underline{\nu} \otimes \underline{\nu}) \cdot \underline{\sigma} \cdot \delta \underline{g}_t \leq 0.$$

It follows that the expression in the frictional part of the integral is positive since

$$\underline{\sigma}_\tau = (\underline{I} - \underline{\nu} \otimes \underline{\nu}) \cdot \underline{\sigma} = -(\underline{I} - \underline{n} \otimes \underline{n}) \cdot \underline{\sigma} = -\underline{\sigma}_t,$$

i.e.

$$\underline{\varrho}_t^T \delta \underline{\xi} \geq 0$$

The integral form of this statement is given below

$$\int_{\bar{\Gamma}_c^1} \underline{\varrho}_t^T \delta \underline{\xi} d\bar{\Gamma}_c^1 \geq 0. \quad (4.65)$$

The virtual work of frictional forces has the same sign as the virtual work of internal forces and it is not conservative so all this energy dissipates, and cannot be recovered. Since the direction of sliding and penetration and its relations

with gap vector are relative, it is important to distinguish master and slave surfaces. Either the stress on the master has to be compared to the gap with respect to the master  $g(\underline{r}, \Gamma_c^1)$  (as in considered case) or stress on the slave has to be compared to the gap with respect to the slave,  $g(\underline{\rho}, \Gamma_c^2)$  (for the graphical interpretation see Fig. 4.10): **a** – two bodies in contact  $\underline{\rho} = \underline{r}$ ; **b** – the slave point moves out of the master  $\delta g_n \geq 0$ ,  $\sigma_n \leq 0$ ,  $\sigma_n \delta g_n \leq 0$ , where  $\sigma_n \underline{n}$  is the contact pressure on the master surface; **c** – tangential relative movement, the slave  $\underline{r}$  slides over the master, relative displacement respectively to the master  $\delta \underline{\xi}$  is the same direction as the tangential stress vector on the master, so  $\underline{\sigma}_t^T \delta \underline{\xi} \geq 0$ .

According to (4.64), the balance of virtual work (4.63) can be rewritten as a *variational inequality*. The variational inequality has to be complemented by the restrictions on the possible virtual displacement arising from Dirichlet's boundary conditions and non-penetration conditions. Classically, the variational inequality for frictional problem is written following Duvaut and Lions [Duvaut 71] and Kikuchi and Oden [Kikuchi 88], where the proof of the equivalence between the variational inequalities and the classical formulation can be found. Here we provide a generalized formulation, which does not limit the choice of the frictional law

$$\boxed{\begin{aligned} \int_{\Omega} \underline{\underline{\sigma}} \cdot \delta \nabla \underline{u} \, d\Omega + \int_{\bar{\Gamma}_c^1} \underline{\underline{\sigma}}_t^T \delta \underline{\xi} \, d\bar{\Gamma}_c^1 &\geq \int_{\Gamma_f} \underline{\underline{\sigma}}_0 \cdot \delta \underline{u} \, d\Gamma + \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} \, d\Omega, \\ \mathbb{V} &= \{ \delta \underline{u} \in \mathbb{H}^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \}, \\ \mathbb{K} &= \{ \delta \underline{u} \in \mathbb{V} \mid (\underline{r} + \delta \underline{r} - \underline{\rho} - \delta \underline{\rho}) \cdot \underline{n} \geq -g_{n0} \text{ on } \Gamma_c \} \end{aligned}} \quad (4.66)$$

In case of frictionless contact, the formulation is significantly simpler and is written as

$$\boxed{\begin{aligned} \int_{\Omega} \underline{\underline{\sigma}} \cdot \delta \nabla \underline{u} \, d\Omega &\geq \int_{\Gamma_f} \underline{\underline{\sigma}}_0 \cdot \delta \underline{u} \, d\Gamma + \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} \, d\Omega, \\ \mathbb{V} &= \{ \delta \underline{u} \in \mathbb{H}^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \}, \\ \mathbb{K} &= \{ \delta \underline{u} \in \mathbb{V} \mid (\underline{r} + \delta \underline{r} - \underline{\rho} - \delta \underline{\rho}) \cdot \underline{n} \geq -g_{n0} \text{ on } \Gamma_c \} \end{aligned}} \quad (4.67)$$

where  $\mathbb{H}^1(\Omega)$  denotes Hilbert space of the first order,  $\delta \underline{r} = \delta \underline{u}$  on  $\bar{\Gamma}_c^2$ ,  $\delta \underline{\rho} = \delta \underline{u}$  on  $\bar{\Gamma}_c^1$  - contacting material points in actual configuration,  $g_{n0} = (\underline{X}_\rho + \underline{X}_r) \cdot \underline{n}$  - initial gap. These variational inequality formulations (4.66), (4.67) are valid for any material, since constitutive law does not enter in equations. However, the presence of the inequality sign differs these expressions from the classical weak form and requires new optimization methods for its numerical treatment. No details are given here, the interested reader is referred to monographs [Duvaut 76] and [Kikuchi 88], which are dedicated to this approach.

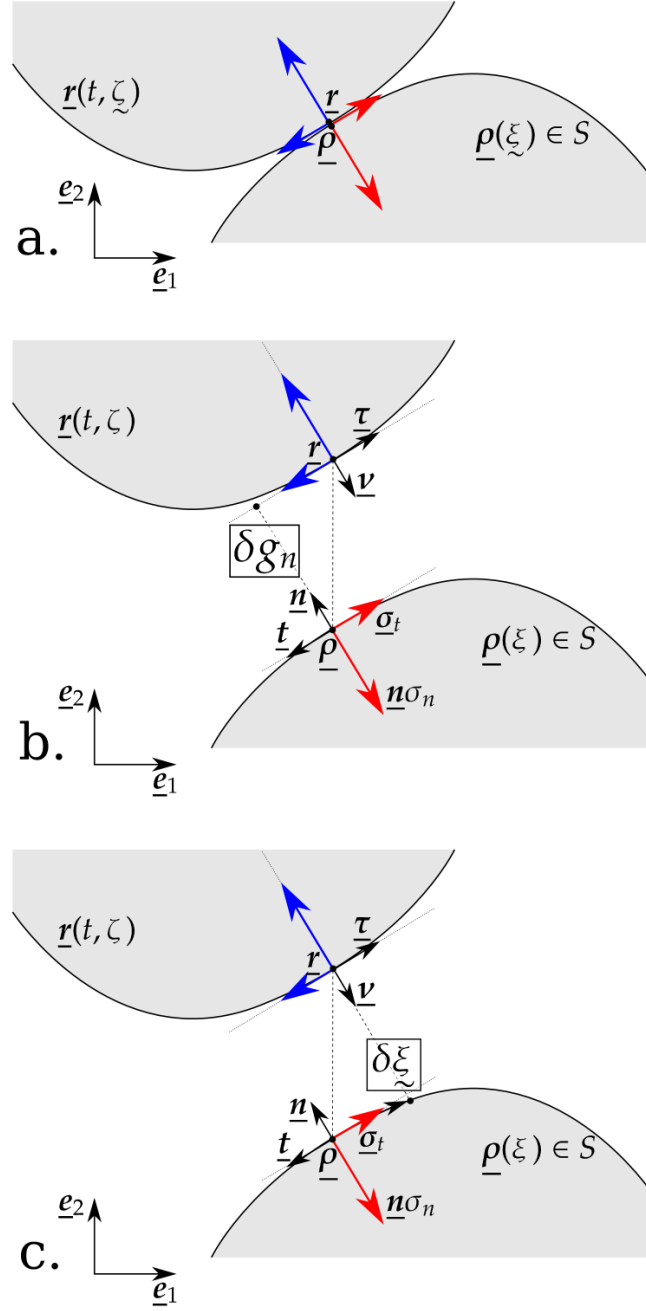


Figure 4.10: Graphical interpretation of inequalities (4.64) and (4.65) with respect to the master surface.

### 4.3.2 Remarks on Coulomb's frictional law

For a detailed analysis of inequalities arising from the formulation of frictional and frictionless cases, the reader is referred to books by Duvaut and Lions [Duvaut 76] and Kikuchi and Oden [Kikuchi 88], where

- the equivalence between the variational inequality and the classical formulation of Signorini's problem with Coulomb's friction (following Duvaut and Lions [Duvaut 71] and [Duvaut 76]) is determined;
- the existence and uniqueness of the solution is proved for small deformation frictionless contact [Kikuchi 88];
- the finite element problem with unilateral contact with and without friction for small/large deformations, incompressible and elasto-plastic materials is treated [Kikuchi 88];
- the questions of convergence of the Finite Element Method are discussed [Kikuchi 88];
- many other developments concerning contact with a rigid foundation are elaborated in the rigorous mathematical framework developed by authors [Kikuchi 88].

Following these works, let us make some remarks considering the frictional problem.

As one can see, no friction law appears in the derived variational inequality for frictional contact (4.66). If Coulomb's law is assumed

$$\|\underline{\sigma}_t\| \leq \mu \langle -\sigma_n \rangle, \quad \|\underline{\dot{g}}_t\| \|\underline{\sigma}_t - \mu \langle -\sigma_n \rangle \underline{\dot{g}}_t\| = 0, \quad \|\underline{s}\| \left\| \|\underline{\dot{g}}_t\| \underline{\sigma}_t - \mu \langle -\sigma_n \rangle \underline{\dot{g}}_t \right\| = 0,$$

where  $\underline{\dot{g}}_t = \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \underline{\dot{\xi}}$  is the relative tangential sliding velocity of the point  $\underline{r}$  over  $\underline{\rho}$  for  $g = 0$ . Then, following [Duvaut 71], the integral related to the contribution of the frictional forces can be rewritten as

$$\int_{\bar{\Gamma}_c^1} \underline{\sigma}_t^T \delta \underline{\xi} d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_c^1} \|\underline{\sigma}_t\| \|\delta \underline{g}_t\| d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_c^1} \mu \langle -\sigma_n \rangle \|\delta \underline{g}_t\| d\bar{\Gamma}_c^1 \quad (4.68)$$

and the variational inequality for the frictional problem (4.66) becomes

$$\begin{aligned} \int_{\Omega} \underline{\sigma} \cdot \delta \nabla \underline{u} d\Omega + \int_{\bar{\Gamma}_c^1} \mu \langle -\sigma_n \rangle \|\delta \underline{g}_t\| d\bar{\Gamma}_c^1 &\geq \int_{\Gamma_f} \underline{\sigma}_0 \cdot \delta \underline{u} d\Gamma + \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} d\Omega, \\ \mathbb{V} &= \left\{ \delta \underline{u} \in \mathbb{H}^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \right\}, \\ \mathbb{K} &= \left\{ \delta \underline{u} \in \mathbb{V} \mid (\underline{r} + \delta \underline{r} - \underline{\rho} - \delta \underline{\rho}) \cdot \underline{n} \geq -g_{n0} \text{ on } \Gamma_c \right\} \end{aligned} \quad (4.69)$$

However, the integral (4.68) entering in the variational inequality (4.69) has no meaning, since  $\sigma_n = \sigma_n(\underline{u})$  is a contact pressure depending on the solution  $\underline{u}$  of the problem with friction. Furthermore, the term (4.68) is nonconvex and non differentiable, consequently the questions of existence and uniqueness of the solution for the problem (4.69) remain open.

«...the absence of a complete existence theory for the general problem together with physical evidence of friction have led some investigators to question the validity of the Coulomb friction law.»

Kikuchi and Oden [Kikuchi 88]

To derive some results for the frictional problem several possible “simplifications” have been proposed: one can either assume a priori known contact pressure or tangential stress or replace Coulomb friction by a regularized law. For the case of a prescribed contact pressure, the existence and uniqueness of the solution is proven under reasonable conditions ([Duvaut 76]). So for the numerical purpose, the iterative procedure can be established using two special cases of known contact pressure and known tangential stress repeated alternately (see [Campos 82]). However, the frictional term remains nondifferentiable which presents a problem for the exact numerical treatment of frictional contact. The regularization of the contact term leads to a convex and Gâteaux differentiable integral for which the existence and uniqueness of the solution can be proved. The regularization consists in replacing the absolute value of the virtual displacement in integral (4.68) by a smooth term containing  $\varepsilon$

$$\int_{\bar{\Gamma}_c^1} \mu \langle -\sigma_n \rangle \|\delta \underline{\mathbf{g}}_t\| d\bar{\Gamma}_c^1 \longrightarrow \int_{\bar{\Gamma}_c^1} \mu \langle -\sigma_n \rangle R(\delta \underline{\mathbf{g}}_t, \varepsilon) d\bar{\Gamma}_c^1 \quad (4.70)$$

in the way that

$$R(\delta \underline{\mathbf{g}}_t, \varepsilon) \xrightarrow{\varepsilon \rightarrow 0} \|\delta \underline{\mathbf{g}}_t\|,$$

see [Kikuchi 88] for details. Such a regularization is in a good agreement with experiments on friction between metal surfaces [Courtney-Pratt 57] demonstrating that a tangential relative sliding may occur, even for a small tangential stress. This microsliding is governed by the elasto-plastic deformation of asperities in contact.

## 4.4 Variational equality and resolution methods

According to the remarks made in the previous section, the variational inequality is hard to apply for contact with finite sliding and/or rotations. That is why nowadays most of the practical studies in contact mechanics are based on the so-called *variational equalities*, which are easy to introduce in a finite element framework and does not require totally new minimization techniques. Here we derive the framework based on variational equalities constructed for known active contact zone. Due to the requirement of known contact zone, such a formulation should be coupled with an active set strategy. Active set denotes such components of the potential contact zone which are in “active” contact at the current solution step. Naturally, the inactive set contains only components of the potential contact surface which are not in contact.

Assuming the known contact zone it becomes possible to transform the nonlinear optimization problem under constraints into an unconstrained problem and to apply further classical resolution methods. Among the most popular and widely used methods in contact mechanics are those inspired from optimization theory:

- the penalty method (exterior point methods);
- the barrier method (interior point methods);
- direct elimination of constraints;
- the Lagrange multiplier method;
- the perturbed Lagrangian (valid for stick or frictionless slip) method;
- the augmented Lagrangian formulation;
- Nitsche method (weak enforcement of constraints);
- cross-constraint method (see [Wriggers 06]);

and others, which can be found in [Wriggers 06] and multiple references in it. We restrict ourself to three of them: penalty, Lagrange multipliers and augmented Lagrangian methods. All methods have their own advantages and drawbacks, which will be discussed in the following.

## 4.5 Penalty method

Let us assume that somehow the active contact zones  $\bar{\Gamma}_c^1 \in \Gamma_c^1, \bar{\Gamma}_c^2 \in \Gamma_c^2$  are known. Then the differential formulation of the contact problem can be replaced by a variational equality (4.63) complemented by restrictions on the virtual displacements  $\delta \underline{u} \in \mathbb{V}$  and contact constraints:

$$\int_{\Omega} \underline{\underline{\sigma}} \cdot \delta \nabla \underline{u} d\Omega + \int_{\bar{\Gamma}_c^1} (\sigma_n \delta g_n + \underline{\underline{\sigma}}_t^T \delta \underline{\xi}) d\bar{\Gamma}_c^1 = \int_{\Gamma_f} \underline{\underline{\sigma}}_0 \cdot \delta \underline{u} d\Gamma + \int_{\Omega} \underline{\underline{f}}_v \cdot \delta \underline{u} d\Omega, \quad (4.71)$$

$$\mathbb{V} = \{ \delta \underline{u} \in \mathbb{H}^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \}, \mathbb{K} = \{ \delta \underline{u} \in \mathbb{V} \mid g(\underline{u}, \delta \underline{u}) \geq 0 \text{ on } \bar{\Gamma}_c \}$$



We have a standard minimization problem under inequality constraints.

#### 4.5.1 Frictionless case

The motion of material points  $\underline{x} \in \Omega$  is governed by Dirichlet boundary conditions, by non-penetration conditions

$$g(\underline{\rho}, S) \geq 0$$

and by stick-slip relations which will be discussed later. First, we confine ourself to the frictionless case. To fulfill the non-penetration conditions, a normal contact pressure  $\sigma_n < 0$  arises at the contact interface. Hertz-Signorini-Moreau's conditions summarize this effect

$$g \geq 0, \quad \sigma_n \leq 0, \quad g\sigma_n = 0, \quad \underline{\sigma}_t = 0. \quad (4.72)$$

Let us construct the penalty method based on an approximate fulfillment of these conditions. For that, let us suppose that the contact pressure is a continuous function of the penetration

$$\sigma_n(g) = \epsilon_n(\langle -g \rangle) = \begin{cases} 0, & g > 0 \\ \epsilon_n(-g), & g \leq 0 \end{cases},$$

where  $\epsilon_n$  is a non-positive continuous strictly monotonically decreasing function and

$$\epsilon_n(0) = 0, \quad \epsilon_n(x) \xrightarrow{x \rightarrow \infty} -\infty$$

Then the contact condition is strictly fulfilled for non-negative gaps, however according to the relations between contact pressure and the gap function, real contact appears only for negative gaps, i.e. only if a penetration takes place

$$\begin{cases} g \geq 0, & \sigma_n = 0, & g\sigma_n = 0 \\ g < 0, & \sigma_n = \epsilon_n(-g) < 0, & g\sigma_n \neq 0 \end{cases} \quad (4.73)$$

Logically, the higher the contact stress for small penetration the better Hertz-Signorini-Moreau's conditions are fulfilled. This approximation implies that the contact surface does not restrict penetration but resists to it, the deeper the penetration the higher the resisting reaction. The physical interpretation leads to a representation of the master surface as a series of springs with zero initial length, that can elongate inside the body normally to the master surface. The reaction provided by the springs follows the law  $R = \epsilon_n(-U)$ , where  $R$  is the appearing reaction and  $U$  is the elongation of the spring. At the same time the surface of the master is supposed to be described by nondeformed springs. Since these imaginary springs are joint to the master surface and transfer the reaction forces to it, it also deforms. The energy accumulated by springs due to their deformation is

$$W_p = \int_0^{-U} \epsilon_n(U) dU = \int_0^{-\langle -g \rangle} \epsilon_n(\langle -g' \rangle) dg',$$

## 4.5 Penalty method

which in case of linear penalty method

$$\epsilon_n(\langle -g \rangle) = -\epsilon_n \langle -g \rangle, \epsilon_n \geq 0$$

becomes

$$W_p = - \int_0^{\langle -g \rangle} \epsilon_n \langle -g' \rangle dg' = \int_0^{\langle -g \rangle} \epsilon_n g' dg' = \frac{1}{2} \epsilon_n \langle -g \rangle^2,$$

which coincides with the energy accumulated by a linear spring of stiffness  $\epsilon_n$  due to an elongation or contraction  $U = \langle -g \rangle$ . Due to its simplicity and mainly due to the smooth contribution (regarding penetration) to the energy of the system, the linear penalty method is one of the most applicable methods for treatment of contact problems both in commercial and scientific finite element codes.

The assumption of dependency of the contact pressure on penetration allows to represent the frictionless contact integral as the work of the contact pressure  $\sigma_n$  on virtual penetration  $\delta g_n$ , in analogy with the work of a prescribed traction: the contact integral appears to be the virtual work due to the deformation of imaginary springs in the contact interface (Fig. 4.11). For normal projection  $g_n(\underline{r}, \Gamma_c^1)$ , the contact contribution to the balance of virtual work writes as follows

$$\delta W_c = \int_{\bar{\Gamma}_c^1} \epsilon_n(-g_n) \delta g_n d\bar{\Gamma}_c^1 = \int_{\Gamma_c^1} \epsilon_n(\langle -g_n \rangle) \delta g_n d\Gamma_c^1, \quad (4.74)$$

after the choice of the penalty function  $\epsilon_n$  (with an argument in Macaulay brackets  $\langle -g \rangle$ ) one can integrate the virtual work due to contact not only over the active contact zone but over the full contact zone  $\Gamma_c^1$ , which results in the second equality in (4.74). For shadow projection from an infinitely remote emitter  $g_s(\underline{r}, \Gamma_c^1)$  (see (4.60)), the contact integral has the form

$$\delta W_c = \int_{\bar{\Gamma}_c^1} \epsilon_n(-g_s) \delta g_s \underline{n} \cdot \underline{e} d\bar{\Gamma}_c^1 = \int_{\Gamma_c^1} \epsilon_n(\langle -g_s \rangle) \delta g_s \underline{n} \cdot \underline{e} d\Gamma_c^1, \quad (4.75)$$

where  $\underline{e}$  is the unit vector pointing towards the emitter.

The entire weak form for frictionless problems and normal projection writes as

$$\begin{aligned} \int_{\Omega} \underline{\underline{\sigma}} \cdot \delta \nabla \underline{u} d\Omega + \int_{\Gamma_c^1} \epsilon_n(\langle -g_n \rangle) \delta g_n d\bar{\Gamma}_c^1 &= \int_{\Gamma_f} \underline{\underline{\sigma}}_0 \cdot \delta \underline{u} d\Gamma + \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} d\Omega, \\ \mathbb{V} &= \{ \delta \underline{u} \in \mathbb{H}^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \} \end{aligned} \quad (4.76)$$

If the mechanical problem can be formulated as a minimization of the functional  $F(\underline{u}) : \mathbb{V} \rightarrow \mathbb{R}$  under constraints  $g(\underline{u}) \geq 0$  and if the penalty functional is  $P(\underline{u}) : \mathbb{V} \rightarrow \mathbb{R}_0^+$  then,

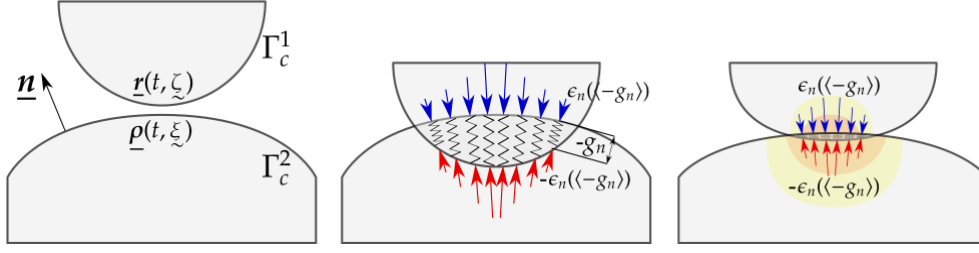


Figure 4.11: Spring interpretation of the penalty method: **left** – undeformed configuration; **middle** – configuration after penetration  $g_n(\xi)$ , which in turn results in a contact pressure  $\sigma_n(\xi) = \epsilon_n(-g_n(\xi))$  shown with red arrows on the master and in blue on the slave, contact pressure results in decreasing the penetration due to deformation and results in an **right** equilibrium state.

*«The idea behind penalty method is, roughly speaking, to append to  $F$  a "penalty functional"  $P$  which increases the magnitude accordingly to how severely the constraint is violated.»*

Kikuchi and Oden [Kikuchi 88]

#### 4.5.2 Example

Let us demonstrate the idea of the penalty method on a simple example of minimization under constraints. Let the potential energy of the system with one degree of freedom be

$$F(x) = \frac{1}{2} (2) (x + 1)^2,$$

It corresponds to a mathematical pendulum in statics (see Fig. 4.12) where  $c = (2)[N/mm]$  is the stiffness of the spring and  $x[mm]$  is the coordinate of the mass-point. We require that at equilibrium the energy is minimal or that the variation of the energy is zero:

$$\min_x \{F(x)\} \quad \text{or} \quad \delta F(x) = 0$$

It corresponds to the point  $x = -1[mm]$ .

$$\delta F(x) = 2(x + 1)\delta x = 0 \Leftrightarrow x = -1$$

Let us introduce a wall, which restricts the penetration in the zone  $x < 0$ ; now the problem is reformulated as

$$\min_{x \geq 0} \{F(x)\}.$$

Note that strictly speaking we cannot use variational formulation  $\delta F(x) = 0, x \geq 0$  since in contact state  $F(x + \delta x)$  is not determined for any  $\delta x$ . The actual position of the mass can be expressed by  $x = X + u$ , where  $X = -1$  is nothing but the

#### 4.5 Penalty method

---

equilibrium state without wall (position in the reference configuration), then the gap can be simply expressed as

$$g(x) = x \geq 0 \quad \Leftrightarrow \quad u \geq -X \quad \Leftrightarrow \quad u \geq 1.$$

then we can rewrite the problem in displacements

$$\min_{u \geq 1} \{F(u)\}$$

$$F(x) = F(X + u) = (X + u + 1)^2 = u^2 = F(u)$$

In the framework of variational equalities and the penalty method the constrained minimization problem can be rewritten as a simple minimization problem:

$$\min_{u \geq 1} \{F(u)\} \rightarrow \min_u \{F(u) + F_p(u)\},$$

where  $F_p(u)$  is the penalty term due to violation of contact constraints

$$F_p(u) = \int_0^{-\langle -g(X+u) \rangle} \epsilon(\langle -g(X+u) \rangle) dg(X+u),$$

or in case of linear penalty  $\epsilon(\langle -g(X+u) \rangle) = -\epsilon \langle -g(X+u) \rangle$

$$F_p(u) = - \int_0^{-\langle -g(X+u) \rangle} \epsilon_n \langle -g(X+u') \rangle dg(X+u')$$

where

$$g(X+u) = X+u, \quad \langle -g(X+u) \rangle = \begin{cases} -X-u, & X+u < 0 \\ 0, & X+u \geq 0 \end{cases}$$

$$F_p(u) = - \int_0^{X+u} \epsilon_n (-X-u') d(X+u') = \frac{1}{2} \epsilon_n (X+u)^2 = \frac{1}{2} \epsilon_n (u-1)^2$$

so we get the following minimization problem

$$\min_u \{F(u) + F_p(u)\} \Leftrightarrow \min_x \left\{ u^2 + \frac{1}{2} \epsilon_n (u-1)^2 \right\}$$

with the solution

$$\min_u \{F(u) + F_p(u)\} \Leftrightarrow \frac{\partial [F(u) + F_p(u)]}{\partial u} = (2 + \epsilon_n)u - \epsilon_n = 0 \Leftrightarrow u = \frac{\epsilon_n}{\epsilon_n + 2}$$

Since now the energy function  $F$  is determined for any displacements, in a more general form, the problem can be rewritten as a variational problem similar to (4.76)

$$\delta F(u) + \epsilon(\langle -g(X+u) \rangle) \delta g(X+u) = 0,$$

where

$$\delta g(X + u) = \left. \frac{\partial g(X + u)}{\partial u} \right|_{u=0} \delta u = \delta u$$

which in case of linear penalty method  $\epsilon_n(x) = -\epsilon_n x$  also provides

$$(2u - \epsilon_n(1 - u))\delta u = 0 \quad \Leftrightarrow \quad u = \frac{\epsilon_n}{\epsilon_n + 2}.$$

The functional  $F(x)$  and  $F(x) + F_p(x)$  for linear penalty are presented in Figure 4.13, which shows the energy functionals of unconstrained and constrained problems respectively. One can see that the functional is smooth and that the solution converges gradually to the solution of the problem  $x^* = 0, u^* = 1$  for increasing penalty parameter  $\epsilon_n$ .

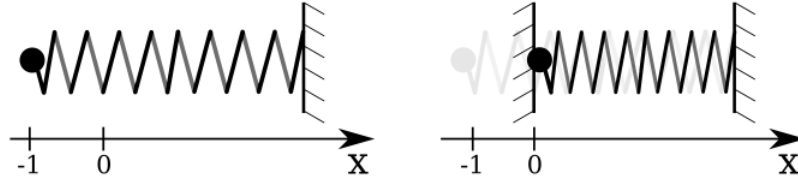


Figure 4.12: A simple example of contact problem, mathematical pendulum in statics: **left** – initial state, **right** – deformed state due to the contact with a rigid wall.

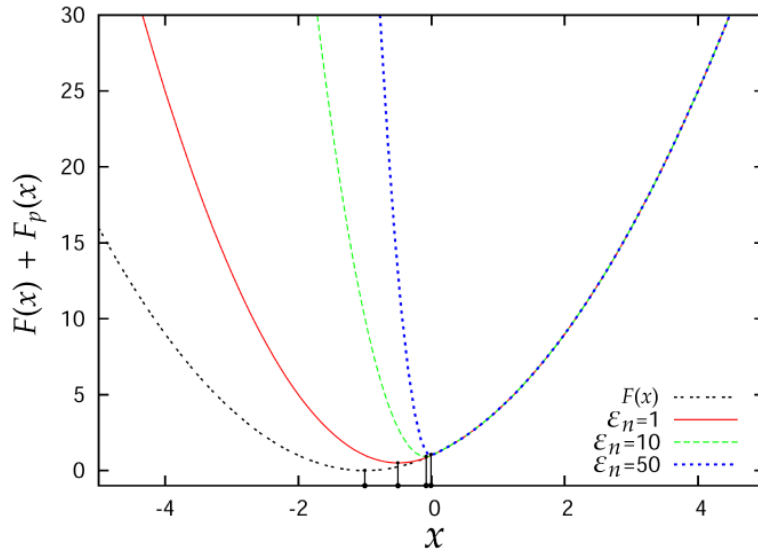


Figure 4.13: Extended energy functional  $F(x) + F_p(x)$  for different values of the penalty ( $\epsilon_n = 1, 10, 50$ ) and corresponding solutions for the problem presented in Fig. 4.12.

### 4.5.3 Nonlinear penalty functions

The linear penalty method is optimal from a numerical point of view, because it does not introduce additional nonlinearities in the problem. Moreover, its linearization is a feasible task, whereas linearization of any nonlinear penalty function leads to significant computational difficulties. However, as it can be easily shown, the linear penalty function is not optimal from the point of view of the precision. Let us compare linear penalty function with quadratic

$$\epsilon^q(x) = -\varepsilon_n(x^2 + x), \quad \int \epsilon^q(x)dx = -\varepsilon_n\left(\frac{1}{3}x^3 + \frac{1}{2}x^2\right)$$

and exponential penalty functions

$$\epsilon^e(x) = -\varepsilon_n x(2 + x)e^x, \quad \int \epsilon^e(x)dx = -\varepsilon_n x^2 e^x.$$

As one can see, the contribution of the penalty terms to the energy of the system (integrals) is smooth (value and derivative of the energy due to penalty in  $x = 0$  is zero). In Tables 4.6, 4.7, 4.8 a comparison between linear, quadratic and exponential penalty methods is presented for different values of the penalty parameter  $\varepsilon_n = 1, \varepsilon_n = 10, \varepsilon_n = 100$ . To compare these methods on a nonlinear problem, we assume that the spring energy is given by  $W = \frac{1}{2}cx^4$ , with a parameter  $c = \frac{1}{2}$ . It means that the value of penalty coefficient  $\varepsilon_n = 1$  has the same order of magnitude as the stiffness parameter. According to the tables, we see that for high penalty  $\varepsilon_n \sim 100c$  (Tab. 4.8), the linear method converges much faster than both others and the solutions of the three are quite close. For a moderate penalty  $\varepsilon_n \sim 10c$  (Tab. 4.7), the quadratic penalty function gives a faster convergence than linear and exponential ones. For a small penalty  $\varepsilon_n \sim c$  (Tab. 4.8), the exponential function gives the fastest convergence and a much better solution than the others. It is worth mentioning that in all cases quadratic and exponential penalty functions yield a better solution, and that the best precision is obtained with the exponential function in all cases.

Let us also remark, that in 2D and 3D contact problems, the variation of the normal gap function and its second variation are always nonlinear, so even for linear material and a known contact zone, the linear penalty method does not allow to achieve convergence within one iteration.

Table 4.6: Comparison of different penalty functions, penalty  $\varepsilon_n = 1$ , exact solution  $u^* = 1$ .

	Linear $\epsilon^l(x)$		Quadratic $\epsilon^q(x)$		Exponential $\epsilon^e(x)$	
Iter	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $
1	1.000	1.000	0.667	1.000	0.429	1.000
2	0.692	0.444	0.561	0.189	0.649	0.340
3	0.541	0.279	0.548	0.024	0.657	0.012
4	0.502	0.077	0.548	3.20e-4	0.657	1.81e-5
5	0.500	4.76e-3	0.548	5.67e-8	0.657	4.50e-11
6	0.500	1.71e-5	0.548	1.62e-15	0.657	1.69e-16
7	0.500	2.19e-10	0.548	2.02e-16	0.657	1.69e-16

Table 4.7: Comparison of different penalty functions, penalty  $\varepsilon_n = 10$ , exact solution  $u^* = 1$ .

	Linear $\epsilon^l(x)$		Quadratic $\epsilon^q(x)$		Exponential $\epsilon^e(x)$	
Iter	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $
1	1.0	1.0	0.667	1.0	0.429	1.0
2	0.818	0.222	0.815	0.182	0.735	0.417
3	0.798	0.026	0.816	1.7e-3	0.867	0.152
4	0.797	3.0e-4	0.816	2.5e-8	0.884	0.019
5	0.797	3.8e-8	0.816	1.4e-16	0.884	2.38e-4
6	0.797	7.0e-16	0.816	1.4e-16	0.884	3.67e-8
7	0.797	0.0	0.816	1.4e-16	0.884	8.79e-16

## 4.5 Penalty method

Table 4.8: Comparison of different penalty functions, penalty  $\varepsilon_n = 100$ , exact solution  $u^* = 1$ .

	Linear $\varepsilon^l(x)$		Quadratic $\varepsilon^q(x)$		Exponential $\varepsilon^e(x)$	
Iter	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $	Solution $u^i$	Relative error $\left  \frac{u^i - u^{i-1}}{u^i} \right $
1	1.000	1.000	0.667	1.000	0.429	1.000
2	0.964	0.037	0.918	0.274	0.746	0.425
3	0.964	1.41e-4	0.967	0.047	0.925	0.193
4	0.964	2.00e-9	0.965	1.61e-3	0.978	0.054
5	0.964	0.0	0.965	1.87e-6	0.982	4.12e-3
6	0.964	0.0	0.965	2.53e-12	0.982	2.25e-5
7	0.964	0.0	0.965	0.0	0.982	6.67e-10

### 4.5.4 Frictional case

In presence of friction the virtual work due to contact has to be complemented by the frictional term. The classical Coulomb's friction law reads as follows

$$\|\underline{\sigma}_t\| \leq \mu|\sigma_n|, \quad \underline{\sigma}_t - \mu|\sigma_n|\underline{s} = 0, \quad \|\underline{s}\| \|\underline{\sigma}_t - \mu|\sigma_n|\underline{s}\| = 0.$$

As in the case of the normal contact, this set of conditions can be fulfilled approximately using a penalty function. Let us assume that the tangential stress is zero except when a tangential sliding  $\underline{g}_t$  occurs at the interface, so the tangential stress can be considered as a function of the tangential sliding limited by Coulomb's cone surface:

$$\underline{\sigma}_t = \begin{cases} \varepsilon_t(\|\underline{g}_t\|)\underline{s}, & \varepsilon_t(\|\underline{g}_t\|) < \mu|\sigma_n| \\ \mu|\sigma_n|\underline{s}, & \varepsilon_t(\|\underline{g}_t\|) \geq \mu|\sigma_n| \end{cases}$$

The penalty function should be positive and monotonically increasing, moreover it should be zero for zero sliding

$$\varepsilon_t(x) \geq 0, \quad \varepsilon_t(0) = 0, \quad \frac{\partial \varepsilon_t(x)}{\partial x} \geq 0$$

Consequently, the contribution to the weak form due to tangential contact in the **stick state** is

$$\int_{\bar{\Gamma}_c^1} \underline{\sigma}_t \cdot \delta \underline{g}_t d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_c^1} \varepsilon_t(\|\underline{g}_t\|)\underline{s} \cdot \delta \underline{g}_t d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_c^1} \varepsilon_t \left( \left\| \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \Delta \underline{\xi} \right\| \right) \underline{s}^T \delta \underline{\xi} d\bar{\Gamma}_c^1,$$

where  $\bar{\Gamma}_c^1$  denotes the stick contact zone and

$$\underline{s} = \underline{s} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}.$$



From the derived expressions it is clear that their linearization, needed for the implicit treatment of the contact in the framework of the Finite Element Method, is a complicated task. This is why a linear penalty method is used, that gives the following expression for the stick case

$$\underline{\sigma}_t = \varepsilon_t \Delta \underline{g}_t^\bullet, \quad \|\Delta \underline{g}_t^\bullet\| < \frac{\mu |\sigma_n|}{\varepsilon_t},$$

where  $\Delta \underline{g}_t^\bullet$  is a discrepancy from the actual point and the stick point  $\underline{\rho}(\underline{\xi}^\bullet)$ . The value of this discrepancy is limited by the Coulomb's cone. The integral of the virtual work for the linear penalty can be rewritten as

$$\int_{\bar{\Gamma}_t^\bullet} \underline{\sigma}_t \cdot \delta \underline{g}_t d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_t^\bullet} \varepsilon_t \Delta \underline{g}_t^\bullet \cdot \delta \underline{g}_t d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_t^\bullet} \varepsilon_t (\Delta \underline{\xi}^\bullet)^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^T \delta \underline{\xi} d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_t^\bullet} \varepsilon_t (\Delta \underline{\xi}^\bullet)^T \delta \underline{\xi} d\bar{\Gamma}_c^1,$$

This integral can be interpreted as the work of the tangential stress vector  $\underline{\sigma}_t$  on the virtual relative sliding  $\delta \underline{\xi}$  in the contact interface. The convective coordinate  $\underline{\xi}^\bullet$  denotes the stick point, to where the slave point  $\underline{r}$  returns if the external load is removed. So  $\Delta \underline{\xi}^\bullet$  is an accumulated "slip-in-stick" over solution steps  $\Delta \underline{\xi}(t) = \underline{\xi}(t) - \underline{\xi}^\bullet$ . From a physical point of view, the "slip-in-stick" represents elastic deformations within the contact interface (elastic deformation of asperities) and should vanish when the load is removed. The direct analogy with elasto-plastic deformation is the following: the deformation inside the yield surface results in elastic deformation of the volume element, whereas pushing the yield surface leads to plastic flow. After removing the load, the accumulated plastic deformation remains unchanged and the elastic deformation vanishes. By analogy, the total slip  $\underline{g}_t$  can be splitted into a sum of "slip-in-stick"  $\underline{g}_t^\bullet$  and real slip  $\underline{g}_t^*$

$$\underline{g}_t = \underline{g}_t^\bullet + \underline{g}_t^* \quad \sim \quad \underline{\varepsilon} = \underline{\varepsilon}^e + \underline{\varepsilon}^p$$

Note that  $\underline{g}_t^\bullet$  and  $\underline{g}_t^*$  may be not collinear. To determine the slip direction, we need to reformulate this expression in velocities

$$\underline{\dot{g}}_t = \underline{\dot{g}}_t^\bullet + \underline{\dot{g}}_t^* \quad \sim \quad \underline{\dot{\varepsilon}} = \underline{\dot{\varepsilon}}^e + \underline{\dot{\varepsilon}}^p$$

Applying the incremental Euler method ( $\dot{x} = (x^{i+1} - x^i)/\Delta t$ ) to integrate this equation, we get

$$\underline{g}_t^{i+1} = \underline{g}_t^i + \Delta \underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{*i},$$

where

$$\Delta \underline{g}_t^{\bullet i} = \underline{g}_t^{\bullet i+1} - \underline{g}_t^{\bullet i}, \quad \Delta \underline{g}_t^{*i} = \underline{g}_t^{*i+1} - \underline{g}_t^{*i}.$$

So for  $\Delta t \rightarrow 0$  the frictional conditions can be reformulated as

$$\|\underline{\sigma}_t\| \leq \mu |\sigma_n|, \quad \underline{\sigma}_t - \mu |\sigma_n| \frac{\Delta \underline{g}_t^*}{\|\Delta \underline{g}_t^*\|} = 0, \quad \|\underline{g}_t^*\| \left\| \underline{\sigma}_t - \mu |\sigma_n| \frac{\Delta \underline{g}_t^*}{\|\Delta \underline{g}_t^*\|} \right\| = 0.$$

## 4.5 Penalty method

This representation is important for the numerical treatment of frictional contact problems by means of penalty method, which will be discussed in details in Section 5.2.

As proposed in [Wriggers 06], in stick state, there is no need to distinguish normal and tangential directions. The slave point has to stick to the master point at which stick occurred and, according to the penalty method, all violations of the stick condition will be penalized by a penalty stress vector  $\underline{\underline{\sigma}} \cdot \underline{n}$  in the direction of such a violation  $\underline{g} = \underline{r} - \underline{\rho}$ , i.e.

$$\underline{\underline{\sigma}} \cdot \underline{n} = \epsilon(\|\underline{g}\|) \frac{\underline{g}}{\|\underline{g}\|}, \quad \text{if } g_n \leq 0, \|\underline{\underline{\sigma}}_t\| < \mu|\sigma_n|, \quad (4.77)$$

where  $\epsilon(x) \geq 0$  is a positive penalty function of the stick constraint violation. Such a formulation can be easily integrated in the weak form by substitution (4.77) in (4.55), which leads to

$$\int_{\bar{\Gamma}_c^1} \epsilon(\|\underline{g}\|) \frac{\underline{g}}{\|\underline{g}\|} \cdot \delta(\underline{\rho} - \underline{r}) d\bar{\Gamma}_c^1,$$

and for a linear penalty function to a simpler form

$$- \int_{\bar{\Gamma}_c^1} \epsilon(\underline{r} - \underline{\rho}) \cdot \delta(\underline{r} - \underline{\rho}) d\bar{\Gamma}_c^1,$$

such an approach is formally equivalent to the standard linear penalty, if one puts  $\epsilon_n = \epsilon_t = \epsilon$ , but it yields a simpler numerical formulation as shown in [Wriggers 06].

In **slip state**, the tangential stress is determined by the contact pressure and the slip direction

$$\underline{\underline{\sigma}}_t^* = \mu|\sigma_n| \frac{\underline{\dot{g}}_t}{\|\underline{\dot{g}}_t\|} = \mu|\sigma_n| \underline{s} = \mu|\sigma_n| \underline{\underline{s}}^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}};$$

the integral due to frictional contact in slip contact zone is formulated as

$$\int_{\bar{\Gamma}_c^{1*}} \mu|\sigma_n| \underline{s} \cdot \delta \underline{\underline{g}}_t d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_c^{1*}} \mu|\sigma_n| \underline{\underline{s}}^T \delta \underline{\underline{\xi}} d\bar{\Gamma}_c^1.$$

Since  $|\sigma_n| = -\epsilon_n \langle -g_n \rangle$  we finally get the following integral due to frictional sliding, which has to be introduced in the weak form

$$\int_{\bar{\Gamma}_c^{1*}} \mu|\sigma_n| \underline{\underline{s}}^T \delta \underline{\underline{\xi}} d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_c^{1*}} -\epsilon_n \langle -g_n \rangle \mu \underline{\underline{s}}^T \delta \underline{\underline{\xi}} d\bar{\Gamma}_c^1.$$

For linear penalty  $|\sigma_n| = \epsilon_n \langle -g_n \rangle$ , the integral takes the following form

$$\int_{\bar{\Gamma}_c^{1*}} \mu|\sigma_n| \underline{\underline{s}}^T \delta \underline{\underline{\xi}} d\bar{\Gamma}_c^1 = \int_{\bar{\Gamma}_c^{1*}} \epsilon_n \langle -g_n \rangle \mu \underline{\underline{s}}^T \delta \underline{\underline{\xi}} d\bar{\Gamma}_c^1,$$

Finally, the full variational equality – balance of virtual works – for frictional contact (Coulomb's friction) and linear penalty method is

$$\begin{aligned} \int_{\Omega} \underline{\underline{\sigma}} \cdot \delta \nabla \underline{\underline{u}} d\Omega - \int_{\bar{\Gamma}_c^{1*}} \varepsilon_n \langle -g_n \rangle (\delta g_n - \mu \underline{\underline{s}}^T \delta \underline{\underline{\xi}}) d\bar{\Gamma}_c^1 - \int_{\Gamma_f} \underline{\underline{\sigma}}_0 \cdot \delta \underline{\underline{u}} d\Gamma - \int_{\Omega} \underline{\underline{f}}_v \cdot \delta \underline{\underline{u}} d\Omega + \\ + \int_{\bar{\Gamma}_c^{1\bullet}} -\varepsilon_n \langle -g_n \rangle \delta g_n + \varepsilon_t \Delta \underline{\underline{\xi}}^{\bullet T} \delta \underline{\underline{\xi}} d\bar{\Gamma}_c^1 = 0, \\ \mathbb{V} = \{ \delta \underline{\underline{u}} \in \mathbb{H}^1(\Omega) \mid \delta \underline{\underline{u}} = 0 \text{ on } \Gamma_u \}, \end{aligned} \quad (4.78)$$

where  $\bar{\Gamma}_c^{1*} \in \bar{\Gamma}_c^1$  and  $\bar{\Gamma}_c^{1\bullet} \in \bar{\Gamma}_c^1$  are respectively the slip and the stick active contact zones on the master. In case of frictionless contact, the variational equality becomes

$$\begin{aligned} \int_{\Omega} \underline{\underline{\sigma}} \cdot \delta \nabla \underline{\underline{u}} d\Omega + \int_{\bar{\Gamma}_c^1} -\varepsilon_n \langle -g_n \rangle \delta g_n d\bar{\Gamma}_c^1 - \int_{\Gamma_f} \underline{\underline{\sigma}}_0 \cdot \delta \underline{\underline{u}} d\Gamma - \int_{\Omega} \underline{\underline{f}}_v \cdot \delta \underline{\underline{u}} d\Omega = 0 \\ \mathbb{V} = \{ \delta \underline{\underline{u}} \in \mathbb{H}^1(\Omega) \mid \delta \underline{\underline{u}} = 0 \text{ on } \Gamma_u \}, \end{aligned} \quad (4.79)$$

## 4.6 Method of Lagrange multipliers

The method of Lagrange multipliers also allows to construct variational equalities for contact problems. This method, named after Joseph-Louis Lagrange, is used in optimization theory to find the extremum of a functional subjected to constraints. Briefly, one has to seek for an argument  $\underline{\underline{u}}^*$  which minimizes the scalar functional  $\Pi(\underline{\underline{u}})$  under constraints  $g(\underline{\underline{u}}) = 0$

$$\min_{g(\underline{\underline{u}})=0} \Pi(\underline{\underline{u}}).$$

This problem under certain circumstances can be replaced by the search of a stationary point (extremum, precisely saddle point) of a specifically constructed functional  $\mathcal{L}(\underline{\underline{u}}, \lambda)$  called Lagrangian, where  $\lambda$  is an additional unknown (Lagrange multiplier)

$$\min_{g(\underline{\underline{u}})=0} \Pi(\underline{\underline{u}}) \rightarrow \nabla \mathcal{L}(\underline{\underline{u}}, \lambda) = 0. \quad (4.80)$$

The Lagrangian is constructed in the following manner

$$\mathcal{L}(\underline{\underline{u}}, \lambda) = \Pi(\underline{\underline{u}}) + \lambda g(\underline{\underline{u}}), \quad \lambda \leq 0,$$

its gradient reads as

$$\nabla \mathcal{L}(\underline{\underline{u}}, \lambda) = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial \underline{\underline{u}}} \\ \frac{\partial \mathcal{L}}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} \frac{\partial \Pi(\underline{\underline{u}})}{\partial \underline{\underline{u}}} + \lambda \frac{\partial g(\underline{\underline{u}})}{\partial \underline{\underline{u}}} \\ g(\underline{\underline{u}}) \end{bmatrix} = 0 \quad (4.81)$$

The lower equation is nothing but the constraint  $g(\underline{u}) = 0$ . The upper one implies that the gradient of the functional  $\Pi$  is opposite to the gradient of the constraint with a multiplier  $\lambda$

$$\frac{\partial \Pi(\underline{u})}{\partial \underline{u}} = -\lambda \frac{\partial g(\underline{u})}{\partial \underline{u}}$$

The solution of the minimization problem is a stationary point of the Lagrangian (4.79), but in general not all stationary problems of (4.79) are solutions of the initial minimization problem. The replacement of a one argument functional  $\Pi(\underline{u})$  by a two argument functional  $\mathcal{L}(\underline{u}, \lambda)$  obviously implies that the number of unknowns is higher for the Lagrangian. From a numerical point of view, the discretized problem will contain  $N_{\mathcal{L}}$  more degrees of freedom than the initial problem, where  $N_{\mathcal{L}}$  is the number of geometrical constraints.

#### 4.6.1 Frictionless case

The method of Lagrange's multipliers can be extended for multiple and continuous constraints formulated as inequalities. Let us remind the Hertz-Signorini-Moreau conditions

$$g(\underline{u}) \geq 0, \quad \sigma_n \leq 0, \quad \sigma_n g(\underline{u}) = 0 \quad \text{on } \Gamma_c$$

On the active contact zone  $\bar{\Gamma}_c$ , we require that

$$\forall \underline{u} \in \bar{\Gamma}_c : g(\underline{u}) \geq 0.$$

The constrained minimization problem is formulated as

$$\min_{\underline{u} \in \mathbb{V}, g(\underline{u}) \geq 0} \{ \Pi(\underline{u}) \} \Leftrightarrow \exists \underline{u}^* : \forall \underline{u} \in \mathbb{V}, g(\underline{u}) \geq 0 \text{ on } \bar{\Gamma}_c : \Pi(\underline{u}^*) \leq \Pi(\underline{u}) \quad \text{in } \Omega,$$

where  $\Pi(\underline{u})$  is the energy of the mechanical system in  $\Omega$ ; this problem can be replaced by a stationary point problem for the Lagrangian

$$\nabla \mathcal{L}(\underline{u}, \lambda_n(\underline{X})) = 0 \quad \text{in } \Omega, \quad (4.82)$$

where  $\lambda_n(\underline{X}) \leq 0$  on  $\bar{\Gamma}_c^0$  and  $\lambda_n(\underline{X}) = 0$  elsewhere, then the Lagrangian is given by

$$\mathcal{L}(\underline{u}, \lambda_n) = \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^1} \lambda_n(\underline{X}) g(\underline{u}) d\bar{\Gamma}_c^1 \quad (4.83)$$

The expression  $\lambda_n(\underline{X})$ , where  $\underline{X}$  denotes a material point in the reference configuration, is rarely used and in the following we will also omit the argument. By  $\lambda_n$ , we mean a continuous set of values on the active contact zone (instead of a real number as in the discrete case). The stationary condition in derivative form (4.82) can be replaced by the variation of the Lagrangian

$$\delta \mathcal{L}(\underline{u}, \lambda_n) = \delta \Pi(\underline{u}) + \int_{\bar{\Gamma}_c} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) d\bar{\Gamma}_c = 0$$

The variation of the energy of the system  $\delta\Pi(\underline{u})$  is equivalent to the variational principle of virtual work, so the last equality can be rewritten in an extended form

$$\int_{\Omega} \underline{\sigma} \cdot \delta \nabla \underline{u} d\Omega + \int_{\bar{\Gamma}_c^1} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) d\bar{\Gamma}_c^1 - \int_{\Gamma_f} \underline{\sigma}_0 \cdot \delta \underline{u} d\Gamma - \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} d\Omega = 0$$

$$\mathbb{V} = \left\{ \delta \underline{u} \in H^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \right\} \quad \lambda_n \leq 0 \text{ on } \Gamma_c^1,$$
(4.84)

The constraint  $\lambda_n \leq 0$  has still to be fulfilled, that is why the Lagrange multiplier method does not convert a minimization problem with inequality constraints to a fully unconstrained one. For a more rigorous formulation of Lagrange multiplier method for contact problems the reader is referred to the book of Kikuchi and Oden [Kikuchi 88]. The integral due to contact for zero gap  $g(\underline{u}) = 0$  degenerates to

$$\int_{\bar{\Gamma}_c^1} \lambda_n \delta g(\underline{u}) d\bar{\Gamma}_c^1 \rightarrow \int_{\bar{\Gamma}_c} \lambda_n \delta g(\underline{u}) d\bar{\Gamma}_c^1 \sim \int_{\bar{\Gamma}_c^1} \sigma_n \delta g(\underline{u}) d\bar{\Gamma}_c^1.$$

This term is quite similar to the frictionless part of the contact integral in (4.71), where instead of  $\lambda_n$  the contact pressure  $\sigma_n$  appears, both of them should be negative. The Lagrange multiplier  $\lambda_n$  is interpreted as the contact pressure needed to fulfill the contact constraints. By analogy with a potential field  $\Pi(\underline{r})$  and the force  $\underline{F}$  it generates in point  $\underline{r}$  (gradient of the field)

$$\underline{F} = -\frac{\partial \Pi}{\partial \underline{r}}$$

the contact pressure is the gradient of the "energy"  $\mathcal{L}(\underline{u}, \lambda_n)$  with respect to the gap function

$$|\sigma_n| = |\lambda_n| = -\frac{\partial \mathcal{L}}{\partial g}$$

So we introduced a new unknown, the contact pressure  $\lambda_n$ , which will be obtained as the solution  $(\underline{u}^*, \lambda_n^*)$ , such that

$$\forall \underline{u} \in \mathbb{V}, \lambda_n \leq 0 : \mathcal{L}(\underline{u}^*, \lambda_n) \leq \mathcal{L}(\underline{u}^*, \lambda_n^*) \leq \mathcal{L}(\underline{u}, \lambda_n^*)$$

which can be shortly formulated as a so-called min-max problem or saddle point problem

$$\min_{\underline{u} \in \mathbb{V}} \max_{\lambda_n \leq 0} \mathcal{L}(\underline{u}, \lambda_n).$$

#### 4.6.2 Frictional case

In case of frictional contact, a complementary condition on tangential sliding in case of **stick** is

$$\underline{g}_t(\underline{u}) = 0, \quad \text{if } \|\underline{\sigma}_t\| \leq \mu |\sigma_n|, \quad \text{on } \bar{\Gamma}_c^{1\bullet} \quad (4.85)$$

if by analogy with frictionless contact we replace the tangential stress  $\underline{\sigma}_t$  by a Lagrange multiplier vector  $\underline{\lambda}_t$  defined on the contact surface, the stick condition can be rewritten in new terms as

$$\underline{\dot{g}}_t(\underline{u}) = 0, \quad \text{if } \|\underline{\lambda}_t\| \leq \mu|\lambda_n|, \quad \text{on } \bar{\Gamma}_c^1 \quad (4.86)$$

It is worth mentioning that the vector  $\underline{\lambda}_t$  lies in the tangential plane  $\frac{\partial \rho}{\partial \underline{\xi}}$  and  $\lambda_n \underline{n}$  is orthogonal to this plane. So if the Lagrange multipliers  $\lambda_n, \underline{\lambda}_t$  are added as degrees of freedom of the problem, contrary to the ordinary degrees of freedom determined in the global reference frame, they always correspond to the local frame and consequently are invariant to any global rotations. The Lagrangian with the stick term is

$$\mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t) = \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^1} \lambda_n g(\underline{u}) + \underline{\lambda}_t \cdot \underline{\dot{g}}_t(\underline{u}) d\bar{\Gamma}_c^1$$

and the corresponding equilibrium of virtual works  $\delta \mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t)$  gives

$$\delta \mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t) = \delta \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^1} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) + \underline{\dot{g}}_t(\underline{u}) \cdot \delta \underline{\lambda}_t + \underline{\lambda}_t \cdot \delta \underline{\dot{g}}_t(\underline{u}) d\bar{\Gamma}_c^1$$

A more straightforward formulation of this functional can be obtained if instead of  $\underline{\dot{g}}_t(\underline{u}) = 0$  it is explicitly required that the surface coordinate  $\underline{\xi} = 0$ , then

$$\delta \mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t) = \delta \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^1} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) + \underline{\xi}^T \delta \underline{\lambda}_t + \underline{\lambda}_t^T \delta \underline{\xi} d\bar{\Gamma}_c^1,$$

where  $\underline{\lambda}_t^T = \underline{\lambda}_t \cdot \frac{\partial \rho}{\partial \underline{\xi}}$  or  $\underline{\lambda}_t = \underline{\lambda}_t^T \frac{\partial \rho}{\partial \underline{\xi}}$ . Consequently the stick criterion can be rewritten in new notations

$$\sqrt{\underline{\lambda}_t \cdot \underline{\lambda}_t} < \mu|\lambda_n| \quad \Leftrightarrow \quad \sqrt{\underline{\lambda}_t^T \underline{\lambda}_t} < \mu|\lambda_n|$$

As noted in [Wriggers 06], in stick state there is no need to distinguish tangential and normal directions so that the two constraints  $g = 0$  and  $\underline{\dot{g}}_t = 0$  can be replaced by one

$$\underline{g}(\underline{u}) = \underline{r} - \underline{\rho}(\underline{X}^\bullet) = 0,$$

where  $\underline{X}^\bullet$  denotes the material point at which  $\underline{r}$  and  $\underline{\rho}$  stick together, the associated spatial Lagrange multiplier vector  $\{\underline{\lambda}\}$  represents the contact stress vector. The Lagrangian of the energy becomes

$$\mathcal{L}(\underline{u}, \underline{\lambda}) = \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^1} \underline{\lambda} \cdot \underline{g}(\underline{u}) d\bar{\Gamma}_c^1$$

The criterion of stick is

$$\left\| \underline{\lambda} \cdot (\underline{I} - \underline{n} \otimes \underline{n}) \right\| < \mu |\underline{\lambda} \cdot \underline{n}|,$$

where  $\underline{n}$  is the unit normal vector on the master surface. Finally, the variation on the Lagrangian takes the following form

$$\delta \mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t) = \delta \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^1} \underline{g}(\underline{u}) \cdot \delta \underline{\lambda} + \underline{\lambda} \cdot \delta \underline{g}(\underline{u}) d\bar{\Gamma}_c^1, \quad (4.87)$$

In case of slip, for Coulomb's friction law, the tangential stress is fully determined by the contact pressure and the sliding direction, that is why there is no more geometrical constraint on the tangential slip and so we do not need the Lagrange multiplier  $\underline{\lambda}_t$  which can be expressed through the  $\lambda_n$ :

$$\underline{\sigma}_t = \mu |\sigma_n| \underline{s} \quad \Leftrightarrow \quad \underline{\sigma}_t = \mu |\lambda_n| \frac{\underline{\xi}^T}{|\underline{\xi}|} \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = \mu |\lambda_n| \underline{s}^T \frac{\partial \underline{\rho}}{\partial \underline{\xi}}$$

and the virtual work balance takes the form

$$\delta \mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t) = \delta \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^1} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) + \mu |\lambda_n| \underline{s}^T \delta \underline{\xi} d\bar{\Gamma}_c^1,$$

Finally, for the active contact zone (union of non-intersecting stick and slip zones  $\bar{\Gamma}_c^1 = \bar{\Gamma}_c^{1*} \cup \bar{\Gamma}_c^{1\bullet}$ ) the variational formulation becomes

$$\begin{aligned} \delta \mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t) &= \delta \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^{1*}} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) + \mu |\lambda_n| \underline{s}^T \delta \underline{\xi} d\bar{\Gamma}_c^1 + \\ &\quad + \int_{\bar{\Gamma}_c^{1\bullet}} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) + \underline{\xi}^T \delta \underline{\lambda}_t + \underline{\lambda}_t^T \delta \underline{\xi} d\bar{\Gamma}_c^1, \\ \mathbb{V} &= \left\{ \delta \underline{u} \in H^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \right\} \quad \lambda_n \leq 0 \text{ on } \Gamma_c^1. \end{aligned} \quad (4.88)$$

or if one uses (4.87)

$$\begin{aligned} \delta \mathcal{L}(\underline{u}, \lambda_n, \underline{\lambda}_t) &= \delta \Pi(\underline{u}) + \int_{\bar{\Gamma}_c^{1*}} g(\underline{u}) \delta \lambda_n + \lambda_n \delta g(\underline{u}) + \mu |\lambda_n| \underline{s}^T \delta \underline{\xi} d\bar{\Gamma}_c^1 + \\ &\quad + \int_{\bar{\Gamma}_c^{1\bullet}} \underline{g}(\underline{u}) \cdot \delta \underline{\lambda} + \underline{\lambda} \cdot \delta \underline{g}(\underline{u}) d\bar{\Gamma}_c^1 \\ \mathbb{V} &= \left\{ \delta \underline{u} \in H^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \right\} \quad \lambda_n \leq 0 \text{ on } \Gamma_c^1. \end{aligned} \quad (4.89)$$

As one can see, one constraint,  $\lambda_n \leq 0$  is still remaining. To get rid of this constraint,  $\lambda_n$  is often replaced by

$$\lambda_n \rightarrow \lambda_n - \text{dist}(\lambda_n, \mathbb{R}_-) = -\langle -\lambda_n \rangle = \begin{cases} 0, & \lambda_n > 0 \\ \lambda_n, & \lambda_n \leq 0 \end{cases},$$

where  $\langle \bullet \rangle$  are the Macaulay brackets and  $\text{dist}(\lambda_n, \mathbb{R}_-^0)$  means the distance from  $\lambda_n$  to the negative half-line, i.e.

$$\text{dist}(\lambda_n, \mathbb{R}_-^0) = \begin{cases} \lambda_n, & \lambda_n \notin \mathbb{R}_-^0 \\ 0, & \lambda_n \in \mathbb{R}_-^0 \end{cases}$$

### 4.6.3 Example

Let us consider the same example as in the penalty method. The energy functional and the contact constraint are given by

$$F(x) = \frac{1}{2}(2)(x+1)^2, g(x) = x \geq 0,$$

expressed in displacement  $u = x + 1$  it writes as

$$F(u) = u^2, g(u) = u - 1 \geq 0,$$

The corresponding Lagrangian has the following form

$$\mathcal{L}(u, \lambda) = u^2 + \lambda(u - 1), \quad \lambda \leq 0$$

or using Macaulay brackets

$$\mathcal{L}(x, \lambda) = u^2 - \langle -\lambda \rangle (u - 1).$$

The contours of this Lagrangian are represented in Figure 4.14. The saddle point is easy to locate  $u = 1, \lambda = -4$ . The right plot in the figure allows to demonstrate visually the lack of smoothness on the line  $\lambda = 0$  (marked with a black dashed line), which separates contact and non-contact zones.

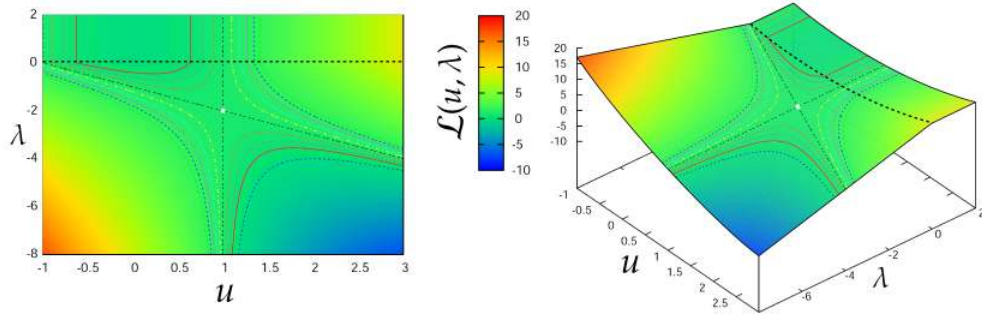


Figure 4.14: Lagrangian  $\mathcal{L}(u, \lambda)$  and several isolines, the saddle point is easy to detect  $u = 1, \lambda = -2$ ; on the black dashed line the Lagrangian lacks of smoothness by  $\lambda$ .

The variation of the Lagrangian writes

$$\delta \mathcal{L}(u, \lambda) = \begin{bmatrix} 2u + \lambda \\ u - 1 \end{bmatrix}^T \begin{bmatrix} \delta u \\ \delta \lambda \end{bmatrix} = 0, \quad \lambda \leq 0$$



As  $\delta u$  and  $\delta \lambda$  are arbitrary, for the solution we require that

$$\begin{cases} 2u + \lambda = 0 \\ u - 1 = 0 \end{cases} \Leftrightarrow \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 0$$

the solution of the last equation is  $[u \ \lambda]^T = [1 \ -2]^T$ . Since  $\lambda$  is negative,  $x = 1$  is the solution, otherwise the second equation has to be excluded from consideration and we simply get equality  $2u = 0$ .

Let us demonstrate how a numerical solution for the nonlinear problem with a Newton-Raphson's method can be obtained: as in the analysis of the nonlinear penalty method, we suppose that the spring is nonlinear with an energy functional  $F(u) = \frac{1}{4}u^4$ , where  $u = x + 1$ , the contact constraint is  $u - 1 \geq 0$ . The Lagrangian takes the form

$$\mathcal{L}(u, \lambda) = \frac{1}{4}u^4 + \lambda(u - 1), \quad \lambda \leq 0$$

or

$$\mathcal{L}(u, \lambda) = \frac{1}{4}u^4 - \langle -\lambda \rangle (u - 1)$$

its variation gives

$$\delta \mathcal{L}(u, \lambda) = \begin{bmatrix} u^3 + \lambda \\ u - 1 \end{bmatrix}^T \begin{bmatrix} \delta u \\ \delta \lambda \end{bmatrix} = 0 \quad (4.90)$$

This system of equations is nonlinear, so we choose the starting point  $[u_0, \lambda_0]$  and we resort to the help of Newton's method. So we need to linearize the equation (4.90) for a known iteration  $[u_i, \lambda_i]$  to complete the increments  $[\Delta u_i, \Delta \lambda_i]$

$$\delta \mathcal{L}(u_{i+1}, \lambda_{i+1}) = \delta \mathcal{L}(u_i, \lambda_i) + \left[ \Delta \delta \mathcal{L}(u, \lambda) \right]_{u_i, \lambda_i} \begin{bmatrix} \Delta u_i \\ \Delta \lambda_i \end{bmatrix} = 0, \quad (4.91)$$

where

$$\begin{aligned} \delta \mathcal{L}(u_i, \lambda_i) &= \begin{bmatrix} u_i^3 + \lambda_i \\ u_i - 1 \end{bmatrix} \\ \left[ \Delta \delta \mathcal{L}(u, \lambda) \right]_{u_i, \lambda_i} &= \begin{bmatrix} \frac{\partial^2 \mathcal{L}(u, \lambda)}{\partial u^2} & \frac{\partial^2 \mathcal{L}(u, \lambda)}{\partial u \partial \lambda} \\ \frac{\partial^2 \mathcal{L}(u, \lambda)}{\partial \lambda \partial u} & \frac{\partial^2 \mathcal{L}(u, \lambda)}{\partial \lambda^2} \end{bmatrix}_{u_i, \lambda_i} = \begin{bmatrix} 3u_i^2 & 1 \\ 1 & 0 \end{bmatrix}, \end{aligned}$$

Finally, from (4.91) the expression for the increments is

$$\begin{bmatrix} \Delta u_i \\ \Delta \lambda_i \end{bmatrix} = - \begin{bmatrix} 3u_i^2 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} u_i^3 + \lambda_i \\ u_i - 1 \end{bmatrix}, \quad \begin{bmatrix} u_{i+1} \\ \lambda_{i+1} \end{bmatrix} = \begin{bmatrix} u_i \\ \lambda_i \end{bmatrix} + \begin{bmatrix} \Delta u_i \\ \Delta \lambda_i \end{bmatrix} \quad (4.92)$$

For any initial conditions  $[u_0, \lambda_0]$  this procedure converges to the exact solution in 3 iterations. However, we should not forget about the restriction  $\lambda \leq 0$ , related to the geometrical constraint  $g \geq 0$ . There are several possibilities to fulfill these conditions: note that one of them is of limited use (marked with a star \*). Advantages and drawbacks will be discussed below and some examples will be given. The aim is to determine if the constraint is active or not, so we will call these approaches - local active set strategies. Below three of them are presented:

## 4.6 Method of Lagrange multipliers

---

1. If on certain iterations the geometrical constraint is inactive  $g(u_i) > 0$ , then the constraint is excluded and the reduced system (4.93) will be solved. If on the  $k$ -th iteration  $g(u_k) \leq 0$ , then we turn back to the full system (4.92) and solve it for starting from the initial guess  $[u_k, 0]$ .
- 2\*. If during iterations a positive Lagrange multiplier occurs  $\lambda_{i+1} > 0$ , then a reduced system of equation will be solved

$$\Delta u_{i+1} = -\frac{u_{i+1}^3}{3u_{i+1}^2}, \quad u_{i+2} = u_{i+1} + \Delta u_{i+1}, \quad \lambda_{i+2} = 0, \quad (4.93)$$

else the full system (4.92) is solved.

3. Both approaches can be combined: on the first iteration, we solve the full system (4.92). On the following iterations, if  $\lambda_i > 0$  or  $g(u_i) > 0$ , we solve the reduced system (4.93), else we return to the full system (4.92).

The first active set strategy, based on the check of the violation of the geometrical constraints  $g \geq 0$ , is frequently used due to its robustness. It provides the correct solution, however sometimes it can take more iterations to converge. The second strategy is based on the check of the positivity of the Lagrange multiplier, sometimes it leads to a constant switch between two systems (4.92) and (4.93), i.e. between two functionals

$$\mathcal{L}(u, \lambda) \leftrightarrow F(u)$$

which results in infinite oscillations if we fastly change the geometrical constraint so that the solution of the problem switches from constrained to unconstrained state. So this active set strategy (2\*) should be avoided. Anticipating things, we remark that when the solution of the problem remains constrained on the following step, the second strategy converges to the correct solution faster than the first one. That was the motivation for developing the third strategy – which is based on both constraints. It converges fast towards the correct solution in all considered cases. However, since on the first iteration the material point follows the contact constraint, a fast remove of contact may create a situation when the initial point for Newton's iterations is too far from equilibrium, so it may cause convergence problems. Finally, we conclude that the first strategy is robust and conditionless, but it is not always fast. The second strategy is of limited use. The third one is fast and robust, but its convergence rate strongly depends on the loading of the problem. Remark that when two initially separated points come and stay in contact, the three strategies are equivalent. However, in case of unloading they behave quite differently. Each strategy is characterized by its piecewise smooth functional, where  $u$  denotes the displacement from the reference state of the spring. Each functional is presented graphically in Fig. 4.15-4.17. A thick black dashed line separates domains of different functionals (Lagrangian in contact and original functional in non-contact domains).

It is worth mentioning that in implicit finite element codes, it is expensive to remove or add degrees of freedom during iterations, so the equation due

to the Lagrange multiplier is not eliminated, as it was done in (4.93), but the stiffness matrix and the residual vector of the problem are to be changed:

$$[K] : \begin{bmatrix} 3u_i^2 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 3u_i^2 & 0 \\ 0 & 1 \end{bmatrix}, \quad [R] : \begin{bmatrix} u_i^3 + \lambda_i \\ u_i - 1 \end{bmatrix} \rightarrow \begin{bmatrix} u_i^3 \\ 0 \end{bmatrix}$$

Two functionals  $F(u)$  and  $\mathcal{L}(u, \lambda)$  for different active set strategies occupy different domains:

## 4.6 Method of Lagrange multipliers

1. Active set strategy 1 (Fig. 4.15):

$$g(u) > 0 : F(u) \quad g(u) \leq 0 : \mathcal{L}(u, \lambda)$$

- 2\*. Active set strategy 2 (Fig. 4.16):

$$\lambda > 0 : F(u) \quad \lambda \leq 0 : \mathcal{L}(u, \lambda)$$

3. Active set strategy 3 (Fig. 4.17):

$$g(u) > 0 \text{ and } \lambda > 0 : F(u) \quad g(u) \leq 0 \text{ or } \lambda \leq 0 : \mathcal{L}(u, \lambda)$$

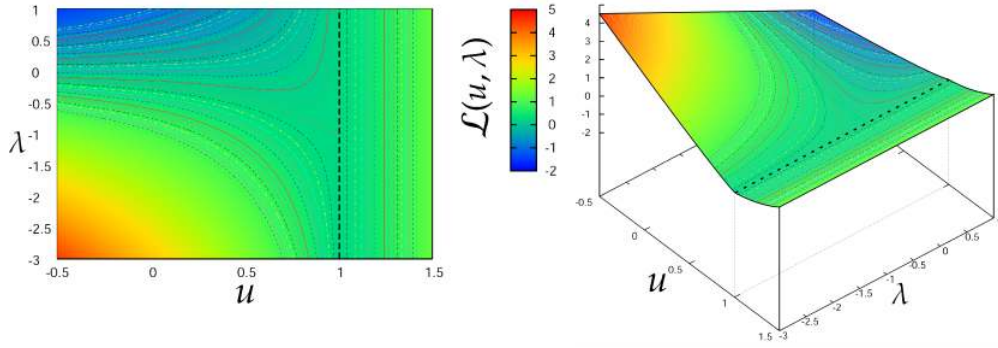


Figure 4.15: Energy functional for the active set strategy 1.

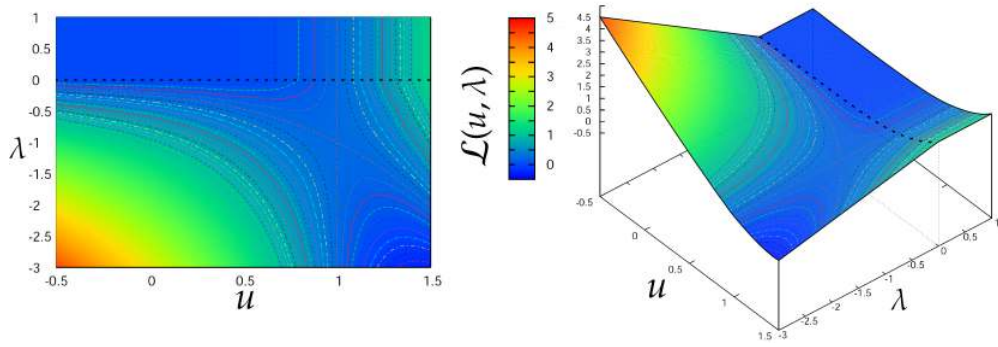


Figure 4.16: Energy functional for the active set strategy 2.

To demonstrate these three strategies let us consider the same configuration of a mass on a nonlinear spring, but for the start point we choose the previously archived solution  $[u = 1, \lambda = -1]$  – stressed spring due to the presence of the wall at  $x = 0$ . The displacement  $u$  will be related to this configuration, so  $u = x$ . We remind that the spring is free of stresses if the coordinate of the mass is  $x = -1$ . Two cases are considered:

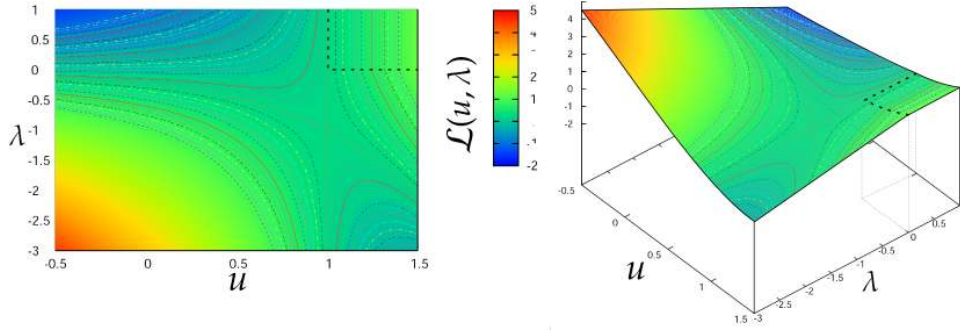


Figure 4.17: Energy functional for the active set strategy 3.

- **case I** – the rigid wall is moved from position  $x = 0$  to  $x = -0.9$ , see Table 4.9 and Fig. 4.18;
- **case II** – the rigid wall is moved from position  $x = 0$  to  $x = -1.1$ , see Table 4.10 and Fig. 4.18.

As the tangent matrix is zero on the solution point, the convergence rate of the Newton-Raphson method is only linear. The solution due to the second active set strategy in case of full unloading (case II) oscillates, and finally after iteration 20, one gets an infinite loop between  $[-1.05, 0.00]$  and  $[-1.075, 4e-4]$ . The third strategy converges in all cases faster than others, however, as it has been mentioned the number of iterations in this case strongly depends on how far the constraint has been moved. Roughly speaking, the difference between the first and third strategies depends on the concrete situation, namely, on how the solution is closer to the final state with enforced contact or to the reference state with contact.

 Table 4.9: **Case I.** Comparison of different active set strategies (ASS) in case of partial unloading, exact solution  $u^* = -0.9$ ,  $\lambda^* = -1e-3$  (converged iteration put in **bold**).

	Strategy ASS 1		Strategy ASS 2		Strategy ASS 3	
Iter, $i$	$u_i$	$\lambda_i$	$u_i$	$\lambda_i$	$u_i$	$\lambda_i$
1	-0.333	0.0	-0.900	1.70	-0.900	1.70
2	-0.556	0.0	-0.933	0.0	-0.933	0.0
3	-0.704	0.0	-0.933	-1.70	-0.933	-1.70
4	-0.802	0.0	-0.900	-7.4e-4	-0.900	-7.4e-4
5	-0.868	0.0	-0.900	-1.0e-3	-0.900	-1.0e-3
6	-0.912	0.0	<b>-0.900</b>	<b>-1.0e-3</b>	<b>-0.900</b>	<b>-1.0e-3</b>
7	-1.800	1.70	-	-	-	-
8	-0.900	-1.30	-	-	-	-
9	-0.900	-1.0e-3	-	-	-	-
10	<b>-0.900</b>	<b>-1.0e-3</b>	-	-	-	-

Table 4.10: **Case II.** Comparison of different active set strategies (ASS) in case of full unloading, exact solution  $u^* = -1.0$ ,  $\lambda^* = 0.0$ .

	Strategy ASS 1		Strategy ASS 2		Strategy ASS 3	
Iter, $i$	$u_i$	$\lambda_i$	$u_i$	$\lambda_i$	$u_i$	$\lambda_i$
1	-0.333	0.0	-1.100	2.30	-1.100	2.3
2	-0.556	0.0	-1.067	0.0	-1.067	0.0
3	-0.704	0.0	-1.067	-2.30	-1.044	0.0
4	-0.803	0.0	-1.100	7.4e-4	-1.030	0.0
5	-0.868	0.0	-1.067	0.0	-1.020	0.0
6	-0.912	0.0	-1.067	2.6e-4	-1.013	0.0
7	-0.941	0.0	-1.044	0.0	-1.009	0.0
8	-0.961	0.0	-1.078	4.8e-4	-1.008	0.0
9	-0.974	0.0	-1.051	0.0	-1.004	0.0
10	-0.983	0.0	-1.074	3.9e-4	-1.002	0.0

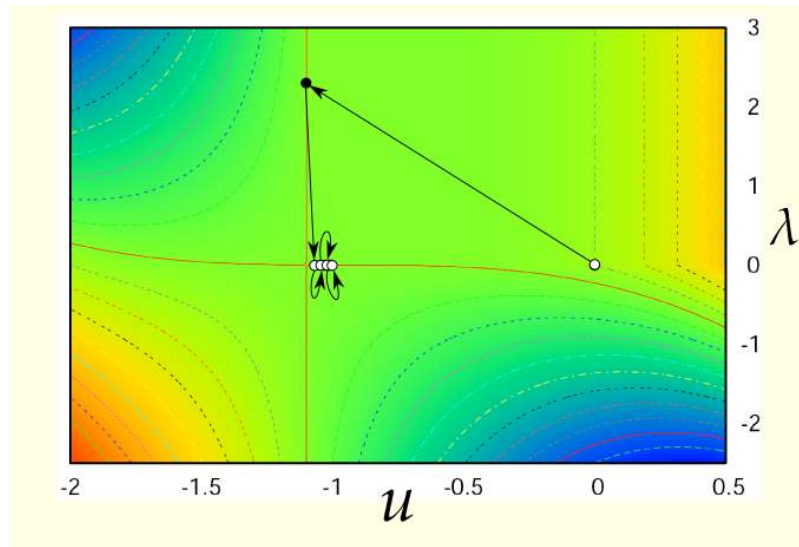


Figure 4.18: Sequence of iterations for different cases (case I – partial unloading, case II – full unloading) and different active set strategies: on the left – iteration sequence is plotted on the real functional field on the right on the Lagrangian field.

## 4.7 Augmented Lagrangian Method

### 4.7.1 Introduction

Another approach for the problem of minimization under constraints has been proposed in 1969 independently by Hestenes [Hestenes 69] and by Powell [Powell 69]. Originally, the method got the name "the multiplier method of Hestenes and Powell" or "method of multipliers". This method consists in a special combination of the penalty and Lagrange multiplier methods, it converges to the solution for finite "penalty" coefficient and provides an unconstrained minimization problem with a smooth functional, which is a great advantage from a numerical point of view.

Let us shortly outline the idea. Let  $F(x)$  be the functional to minimize under constraint  $g(x) = 0$ , in case of use of penalty or Lagrange multiplier methods the functional to minimize (or maxi-minimize) changes respectively as

$$F_p(x) = F(x) + \frac{1}{2}\varepsilon g(x)^2; \quad \mathcal{L}(x, \lambda) = F(x) + \lambda g(x)$$

For a given  $\lambda = \lambda_i$ , the Lagrangian can be considered as a function of one argument  $x$

$$\mathcal{L}(x, \lambda_i) = F(x) + \lambda_i g(x)$$

The application of the penalty method to this functional leads to

$$\mathcal{L}^a(x, \lambda_i) = \mathcal{L}(x, \lambda_i) + \frac{1}{2}\varepsilon g(x)^2 = F(x) + \lambda_i g(x) + \frac{1}{2}\varepsilon g(x)^2$$

If  $[x_*, \lambda_*]^T$  is a solution of  $\min_x \max_\lambda \mathcal{L}(x, \lambda)$ , then it can be shown that it is also the solution of the  $\min_x \max_\lambda \mathcal{L}^a(x, \lambda)$ . The original formulation was strictly iterative and supposed an independent update of  $\lambda_i$  at each iteration. In contact mechanics this approach is also known as Uzawa's algorithm [Arrow 58a] (named after *Hirofumi Uzawa*, Japanese economist). Let  $[x_i, \lambda_i]^T$  be the result of the current iteration, then the linearization procedure of the Newton's method provides us with the following expression for the solution increment

$$\begin{aligned} \mathcal{L}^a(x_i + \Delta x_i, \lambda_i) &\approx \mathcal{L}^a(x_i, \lambda_i) + \left. \frac{\partial \mathcal{L}^a(x, \lambda_i)}{\partial x} \right|_{x_i} \Delta x_i = 0 \\ \left[ F(x_i) + \lambda_i g(x_i) + \frac{1}{2}\varepsilon g(x_i)^2 \right] &+ \left[ \frac{\partial F(x)}{\partial x} + \boxed{[\lambda_i + \varepsilon g(x)]} \frac{\partial g(x)}{\partial x} \right] \bigg|_{x_i} \Delta x_i = 0 \end{aligned} \quad (4.94)$$

So according to the last equality we evaluate the increment  $\Delta x_i$ , further we need to update  $\lambda$ . Hestenes draws our attention to the boxed term in (4.94), which suggests the updating procedure

$$\lambda_{i+1} = \lambda_i + \varepsilon_i g(x_i), \quad \Delta \lambda_i = \varepsilon_i g(x_i),$$

where  $0 < \varepsilon_i \leq \varepsilon$ , which implies that if  $g(x_i) < 0$  then the Lagrange multiplier  $\lambda$  should be decreased: as  $\lambda$  can be considered as a force, so it will push  $x_i$  closer to the solution and vice versa: if  $g(x_i) > 0$ , the Lagrange multiplier

#### 4.7 Augmented Lagrangian Method

should be increased to pull  $x_i$  to the solution. If  $g(x_i)$  gradually tends to zero,  $\lambda_i$  converges to the solution  $\lambda_*$ . Also, the authors propose to choose  $\varepsilon_i = \gamma\varepsilon$ , where  $\gamma$  is positive and smaller than 1, or to choose  $\varepsilon_i$  in order to approach the constraint from one side (a kind of monotonic convergence)

$$g(x_i)g(x_{i+1}) > 0$$

In all cases, the rate of convergence of this method, where primal  $x$  and dual  $\lambda$  variables are updated independently, is linear [Powell 69]. Further this method has been generalized in 1973 by Rockafellar [Rockafellar 73b], [Rockafellar 73a] for inequality constraints, the author formulated a new unconstrained functional for the minimization of  $F(x)$  subjected to the constraint  $g(x) \geq 0$ <sup>2</sup>:

$$\mathcal{L}^a(x, \lambda_i) = F(x) - \frac{1}{2\varepsilon} \left( \lambda_i^2 - \langle -(\lambda_i + \varepsilon g(x)) \rangle^2 \right) \quad (4.95)$$

In an expanded form, it rewrites as

$$\mathcal{L}^a(x, \lambda_i) = \begin{cases} F(x) + \lambda_i g(x) + \frac{1}{2}\varepsilon g(x)^2, & \lambda_i + \varepsilon g(x) \leq 0 \\ F(x) - \frac{1}{2\varepsilon} \lambda_i^2, & \lambda_i + \varepsilon g(x) > 0 \end{cases} \quad (4.96)$$

Note that the Lagrange multiplier  $\lambda$  is not restricted to be nonpositive as in case of optimization under inequality constraints by the Lagrange multiplier method.

The variation of (4.95) with respect to  $x$  results in the following condition

$$\delta \mathcal{L}^a(x, \lambda_i) = \delta F(x) - \langle -(\lambda_i + \varepsilon g(x_i)) \rangle \frac{\partial g(x)}{\partial x} \delta x = 0 \quad (4.97)$$

from where, Rockafellar deduced

$$\lambda_{i+1} = -\langle -(\lambda_i + \varepsilon g(x_i)) \rangle \leq 0$$

or

$$\lambda_{i+1} = \lambda_i + \varepsilon \frac{\partial \mathcal{L}^a(x, \lambda)}{\partial \lambda} = \lambda_i - \varepsilon \frac{1}{\varepsilon} (\lambda_i + \langle -(\lambda_i + \varepsilon g(x)) \rangle)$$

in an extended form

$$\lambda_{i+1} = \begin{cases} \lambda_i + \varepsilon g(x), & \lambda_i + \varepsilon g(x) \leq 0 \\ 0, & \lambda_i + \varepsilon g(x) > 0 \end{cases}$$

Another possible interpretation of the update procedure can be derived if one reasonably assumes that the Lagrangian multiplier should be a "force" due to a change of the potential  $\mathcal{L}^a(x, \lambda)$  related to the change of the constraint  $g(x)$ , so we get directly the expression for the update procedure

$$\lambda_{i+1} = \frac{\mathcal{L}^a(x, \lambda)}{g(x)} = -\frac{1}{\varepsilon} \langle -(\lambda_i + \varepsilon g(x)) \rangle \varepsilon = \begin{cases} \lambda_i + \varepsilon g(x), & \lambda_i + \varepsilon g(x) \leq 0 \\ 0, & \lambda_i + \varepsilon g(x) > 0 \end{cases}$$

<sup>2</sup>in the following paragraph we follow Rockafellar [Rockafellar 73b], except for the sign of the constraint, to make the formulation adapted for contact problems within the previously introduced notations.



However, the linear instead of quadratic convergence rate is a quite high price to pay for the unconstrained minimization even if a non-infinite penalty coefficient leads to the exact solution and even if no additional unknowns<sup>3</sup> are introduced in the global convergence cycle. The advantage of the method is that the functional  $\mathcal{L}^a(x, \cdot)$  is smooth so that a standard Newton's technique is applicable. Moreover there is no need to control the non-positivity of the Lagrange multipliers.

In 1970 Fletcher [Fletcher 70] has developed a technique where both variables are adjusted simultaneously. It follows quite naturally from previous developments. Instead of the functional  $\mathcal{L}^a(x, \lambda_i)$  where  $\lambda_i$  is fixed, the author proposed to consider a full functional where both primal  $x$  and dual variables  $\lambda$  are equivalent,  $\mathcal{L}^a(x, \lambda)$ . Then we obtain the min-max problem, where solution  $[x_*, \lambda_*]$  minimizes the functional by  $x$  and maximizes by  $\lambda$ .

$$\min_x \max_{\lambda} \mathcal{L}^a(x, \lambda)$$

In case of lack of a strictly convex functional  $F(x)$  we reformulate the problem as a variational problem

$$\delta \mathcal{L}^a(x, \lambda) = \frac{\partial \mathcal{L}^a}{\partial x} \delta x + \frac{\partial \mathcal{L}^a}{\partial \lambda} \delta \lambda = 0, \quad (4.98)$$

The associated numerical iterative scheme can then be easily deduced, but another problem is generated: the new functional  $\mathcal{L}^a(x, \lambda)$  is not sufficiently smooth and so a generalization of the iterative scheme, often Newton's method, is required, which would ensure the convergence. A generalized Newton method (GNM) has been proposed by Alart and Curnier for nonsmooth potential present in contact mechanics (first for the penalty method [Curnier 88], then for the augmented Lagrangian method [Alart 88], [Alart 91]. Some generalizations can be found in [Alart 97]).

An application of the augmented Lagrangian method to frictionless contact problems can be found in Glowinski and Le Tallec [Glowinski 89] and also in a report of Wriggers, Simo and Taylor [Middleton 85]. The first application of the augmented Lagrangian method with Uzawa algorithm to frictional problems has been reported by Simo and Laursen in [Simo 92]. The first successful attempt to apply the coupled augmented Lagrangian method to large displacement frictional contact problems has been undertaken by Alart [Alart 88], Alart and Curnier [Alart 91]. The augmented Lagrangian approach has been elaborated by developing the perturbation approach to convex minimization as proposed in [Rockafellar 70] and first applied by Fortin [Fortin 76] to visco-plastic flow problems (similar to frictional contact problems). Further developments can be found in [Heegaard 93], [Pietrzak 97], [Pietrzak 99]. For the following developments of the augmented Lagrangian method for nonlinear constraints the reader is referred, for example, to Powell's survey paper [Powell 78] and for an extended theory of optimization, to the related books [Luenberger 03], [Bertsekas 03], [Bonnans 06].

<sup>3</sup>additional unknowns are nested in local update procedure.

To conclude the introductory part of this section I would like to cite Dimitri Bertsekas' statement on the augmented Lagrangian method

"The original proposal of an Augmented Lagrangian method by Hestenes (1969) and Powell (1969) may be viewed as a significant milestone in the recent history of the constrained optimization area. Augmented Lagrangian methods are not only practically important in their own right, but have also served as the starting point for a chain of research developments centering around the use of penalty functions, Lagrange multiplier iterations, and Newton's method for solving the system of necessary optimality conditions."

Dimitri P. Bertsekas [Bertsekas 81]

### 4.7.2 Application to contact problems

#### Contact conditions in subdifferential formalism

A very detailed description of the application of the augmented Lagrangian method has been given in the PhD thesis of Pietrzak [Pietrzak 97] and in the related article [Pietrzak 99]. All developments are based on the formalism applied in contributions of Alart and Curnier [Alart 91] and Heegaard and Curnier [Heegaard 93], which in turn followed Moreau's works on convex analysis, the references can be found in the cited articles. Following the cited authors, we introduce first an indicator function: indicator function of the positive half line  $x \geq 0$

$$\psi^+(x) : \forall x \in \mathbb{R} \rightarrow \psi^+(x) \in \mathbb{R}_0^+ \cup \infty$$

$$\psi^+(x) = \begin{cases} \infty, & x < 0 \\ 0, & x \geq 0 \end{cases}$$

Following [Heegaard 93], a simple interpretation of the indicator function can be given using a piecewise linear function  $\tilde{\psi}^+(x, r)$ ,  $r \geq 0$

$$\tilde{\psi}^+(x, r) = \begin{cases} -rx, & x < 0 \\ 0, & x \geq 0 \end{cases}$$

Then the indicator function is a limit of the function  $\tilde{\psi}^+(x, r)$  if  $r$  tends to infinity

$$\tilde{\psi}^+(x, r) \xrightarrow{r \rightarrow \infty} \psi^+(x)$$

The functions  $\psi^+(x)$  and  $\tilde{\psi}^+(x, r)$  are nondifferentiable in  $x = 0$ , however the last one has left and right sided derivatives. Now we can give a short formulation of the *subdifferential* of a function. For a simple locally convex function  $f(x) : \mathbb{R} \rightarrow \mathbb{R}$  the subdifferential in the point  $x^*$  is a set of all real numbers contained between left- and right-sided derivatives. This set is denoted as  $\partial f(x^*)$ . If the function is concave in a vicinity of the point  $x^*$  then the subdifferential is

confined between right- and left-sided derivatives. The subdifferential of the function  $\psi^+(0)$  is such that

$$\partial\psi^+(0) = (-\infty, 0] \quad \sim \quad \partial\psi^+(0) = \frac{d\tilde{\psi}^+(x, r)}{dx} \Big|_{x=0-}, \quad \forall 0 \leq r < \infty$$

Any element of the subdifferential is called the *subgradient*

$$s \in \partial\psi^+(0)$$

If at a given point there is only one subgradient, then the function is at least once differentiable at this point.

$$\partial\psi^+(x) = \begin{cases} \emptyset, & x < 0 \\ (-\infty, 0], & x = 0 \\ 0, & x > 0 \end{cases}$$

A slightly extended discussion of subdifferentials and subgradients in the framework of nonsmooth optimization will be given in the following chapter, see Section 5.1.4, where some useful references are also given.

The Hertz-Signorini-Moreau condition for normal contact can be rewritten as a so-called *subdifferential inclusion*

$$\sigma_n \in \partial\psi^+(g_n)$$

It means that for a positive argument  $g_n > 0$ , the contact pressure can take only a zero value, and for  $g_n = 0$ , it can take any negative value. Indeed, contact pressure is nonpositive, for negative normal gap the subdifferential is undefined, so the normal gap is nonnegative and the complementary condition is also fulfilled

$$\sigma_n \in (-\infty, 0], \quad g_n \in [0, \infty), \quad g_n \sigma_n = 0$$

Introducing a conjugate indicator function – indicator of the negative half-line  $x \leq 0$

$$\psi^-(x) : \forall x \in \mathbb{R} \rightarrow \psi^-(x) \in \mathbb{R}_0^- \cup \{-\infty\}$$

$$\psi^-(x) = \begin{cases} 0, & x \leq 0 \\ -\infty, & x > 0 \end{cases}$$

allows to reformulate the normal contact conditions as another subdifferential inclusion

$$g_n \in \partial\psi^-(\sigma_n)$$

Now let us introduce a disk  $C(R)$  of radius  $R$  centered at the origin

$$C_R = \left\{ \underline{x} \mid \|\underline{x}\| \leq R \right\}$$

The corresponding scalar indicator function of the vector argument is given by

$$\psi_R(\underline{x}) : \mathbb{T}_1^2 \rightarrow \mathbb{R}_0^+ \cup \infty$$

$$\psi_R(\underline{x}) = \begin{cases} 0, & \underline{x} \in C_R \\ \infty, & \underline{x} \notin C_R \end{cases}$$

We can extend the interpretation given in [Heegaard 93] for a two dimensional case

$$\tilde{\psi}_R(\underline{x}, r) = \begin{cases} 0, & \underline{x} \in C_R \\ r(\|\underline{x}\| - R), & \underline{x} \notin C_R \end{cases}, r \geq 0$$

The subdifferential of the disk indicator function can be interpreted as

$$\partial\psi_R(\underline{x}') = \left. \frac{d\tilde{\psi}_R(\underline{x}, r)}{d\underline{x}} \right|_{\underline{x}=\underline{x}'+}, \forall r \geq 0 = \begin{cases} r \frac{\underline{x}}{\|\underline{x}\|}, & \underline{x} \notin C_R \setminus \partial C_R \\ 0, & \underline{x} \in C_R \setminus \partial C_R \end{cases}, \forall r \geq 0,$$

where  $\partial C_R = \{ \underline{x} \mid \|\underline{x}\| = R \}$  is a closure of the disk  $C_R$ . The conditions arising from the Coulomb's friction law can be reformulated as a subdifferential inclusion

$$\underline{\dot{g}}_t \in \partial\psi_{\mu|\sigma_n|}(\underline{\sigma}_t) \quad (4.99)$$

This formulation is equivalent to the classic Coulomb's friction law. Indeed, if the tangential contact stress is inside the Coulomb's disk  $\|\underline{\sigma}_t\| \in C_{\mu|\sigma_n|} \setminus \partial C_{\mu|\sigma_n|}$ , there is no tangential sliding  $\underline{\dot{g}}_t = 0$ . The closure of the Coulomb's disk  $\partial C_{\mu|\sigma_n|}$  is called the *slip surface*. When the tangential contact stress reaches the slip surface of the disk, sliding occurs in the direction of the applied tangential stress and the velocity of sliding takes any nonnegative value  $\underline{\dot{g}}_t = r \frac{\underline{\sigma}_t}{\|\underline{\sigma}_t\|}$ ,  $\forall r \geq 0$ .

The conjugate function or Legendre-Fenchel conjugate of the disk indicator function  $\psi_{\mu|\sigma_n|}$  is constructed as a norm function

$$\psi_{\mu|\sigma_n|}^*(\underline{\dot{g}}_t) = \mu|\sigma_n|\|\underline{\dot{g}}_t\|$$

which designates a convex cone. Then the subdifferential of the conjugate function is given as

$$\partial\psi_{\mu|\sigma_n|}^*(\underline{\dot{g}}_t) = \begin{cases} \mu|\sigma_n| \frac{\underline{\dot{g}}_t}{\|\underline{\dot{g}}_t\|}, & \underline{\dot{g}}_t \neq 0; \\ \underline{\sigma}_t, & \underline{\dot{g}}_t = 0, \end{cases}$$

where  $\|\underline{\sigma}_t\| \leq \mu|\sigma_n|$ . Finally using subdifferential notations we can express the tangential contact stress vector as a subdifferential inclusion

$$\underline{\sigma}_t \in \partial\psi_{\mu|\sigma_n|}^*(\underline{\dot{g}}_t)$$

which corresponds to the classically formulated frictional conditions.

Finally for a frictional contact problem in the interface we have two possible sets of subdifferential inclusions, for kinematic arguments

$$\sigma_n \in \partial\psi^+(g_n), \quad \underline{\sigma}_t \in \partial\psi_{\mu|\sigma_n|}^*(\underline{\dot{g}}_t), \quad (4.100)$$

and for static arguments

$$g_n \in \partial\psi^-(\sigma_n), \quad \underline{\dot{g}}_t \in \partial\psi_{\mu|\sigma_n|}^*(\underline{\sigma}_t). \quad (4.101)$$

Note that the convex disc  $C_{\mu|\sigma_n|}$  is a function of the unknown contact pressure, consequently Alart and Curnier [Alart 91] proposed to call the indicator function  $\psi_{\mu|\sigma_n|}$  and its conjugate  $\psi_{\mu|\sigma_n|}^*$  *quasi-potentials* to “stress the dependence of the convex set on the pressure”. Note also, that if the gap is open  $g_n > 0 \Rightarrow \sigma_n = 0$  then the tangential contact stress is also zero  $\underline{\sigma}_t = 0$  and the subdifferential  $\partial\psi_{\mu|\sigma_n|}(\underline{\sigma}_t)$  degenerates in  $\partial\psi_0(0) \ni \forall \underline{x}$ , which means that tangential sliding is not restricted. In the incremental quasi-static analysis the tangential relative sliding velocity is replaced by increments

$$\underline{\dot{g}}_t \rightarrow \underline{g}_t^{i+1} - \underline{g}_t^i = \Delta \underline{g}_t^i = \underline{\dot{g}}_t(t^{i+1} - t^i) \Rightarrow \underline{\dot{g}}_t = \Delta \underline{g}_t / (t^{i+1} - t^i)$$

As previously, the increment of the tangential relative sliding will be denoted simply by  $\underline{g}_t$  and time will be omitted without any loss of generality

$$\underline{\dot{g}}_t \rightarrow \Delta \underline{g}_t / (t^{i+1} - t^i) \rightarrow \underline{g}_t$$

In the introduced framework of subdifferential inclusions the variational problem can be formally formulated as follows

$$\min \{ \Pi^s(\underline{u}) + \Pi^c(\underline{u}) \}, \quad (4.102)$$

where  $\Pi^s(\underline{u})$  is a smooth potential energy of the system of contacting elastic bodies or its incremental homologue in plasticity and

$$\Pi^c(\underline{u}) = \int_{\bar{\Gamma}_t^1} [ \psi^+(g_n) + \psi_{\mu|\sigma_n|}^*(\underline{\dot{g}}_t) ] d\bar{\Gamma}_c^1 \quad (4.103)$$

is a non-differentiable energy of the frictional contact interaction. However such a formulation is non-applicable in the numerical treatment of contact problems because of the non-differentiability. That is the principal motivation for the augmented Lagrangian method, however the smoothing effect is not the only advantage of the method. The augmented Lagrangian method converts the constrained minimization problem into a fully unconstrained problem, contrary to the Lagrange multiplier method which requires fulfillment of the constraint related to the Lagrange multiplier  $\lambda \leq 0$ .

The regularization of the quasi-potentials entering in the contact functional (4.103) by the augmented Lagrangian method is presented in detail in [Pietrzak 97] (Chapter 5). The reader is referred to this volume for the detailed transformation of the quasi-potentials into smooth potentials – augmented Lagrangians  $l_n$  and  $l_t$  related to normal and frictional contact respectively:

$$\psi^+(g_n) \rightarrow l_n(g_n, \lambda_n, \{\varepsilon_n\}), \quad \psi_{\mu|\sigma_n|}^*(\underline{\dot{g}}_t) \rightarrow l_t(\underline{\dot{g}}_t, \underline{\lambda}_t, \sigma_n + \varepsilon_n g_n, \{\varepsilon_t\}),$$

#### 4.7 Augmented Lagrangian Method

where  $\lambda_n$  and  $\underline{\lambda}_t$  are Lagrange multipliers representing the contact pressure and the tangential contact stress vector respectively,  $\varepsilon_n$  and  $\varepsilon_t$  are regularization or penalty coefficients for normal and tangential contact respectively and  $\sigma_n$  is the contact pressure at solution  $\underline{u}^*$ . Penalty coefficients are supposed to be constant. They are not considered as arguments of the augmented Lagrangians, so they are written in braces and further will be omitted. At solution it also holds

$$\sigma_n = \sigma_n + \varepsilon_n g_n(\underline{u}^*).$$

The augmented Lagrangian functional is constructed from (4.102) and the regularized form of (4.103):

$$\mathcal{L}^a(\underline{u}, \lambda_n, \underline{\lambda}_t, \sigma_n + \varepsilon_n g_n, \{\varepsilon_n, \varepsilon_t\}) = \Pi^s(\underline{u}) + \int_{\Gamma_c^1} l_n(g_n, \lambda_n, \{\varepsilon_n\}) + l_t(\underline{g}_t, \underline{\lambda}_t, \sigma_n + \varepsilon_n g_n, \{\varepsilon_t\}) d\Gamma_c^1,$$

The closed forms for  $l_n$  and  $l_t$  are given in the following paragraph.

##### Formulation of the virtual work principle

Here we give a formulation of the augmented Lagrangian functionals  $l_n$  and  $l_t$  which can be directly inserted as integrand in the weak form. The regularized integrand due to the nonpenetration-nonadhesion condition for the geometrical constraint  $g_n \geq 0$ , Lagrange multiplier  $\lambda_n$  (representing contact pressure) and penalty parameter  $\varepsilon_n$  is written as

$$l_n(g_n, \lambda_n) = \begin{cases} \lambda_n g_n + \frac{\varepsilon_n}{2} g_n^2, & \lambda_n + \varepsilon_n g_n \leq 0, \text{ contact} \\ -\frac{1}{2\varepsilon_n} \lambda_n^2, & \lambda_n + \varepsilon_n g_n > 0, \text{ non-contact} \end{cases}$$

using Macaulay brackets a more compact form is given as

$$l_n(g_n, \lambda_n) = -\frac{1}{2\varepsilon_n} (\lambda_n^2 - \langle -(\lambda_n + \varepsilon_n g_n) \rangle^2)$$

If following [Pietrzak 97] the augmented Lagrange multiplier is denoted by a hat

$$\hat{\lambda}_n = \lambda_n + \varepsilon_n g_n$$

then the formulation can be shorten as

$$\boxed{l_n(g_n, \lambda_n) = -\frac{1}{2\varepsilon_n} (\lambda_n^2 - \langle -\hat{\lambda}_n \rangle^2).} \quad (4.104)$$

The expanded form with hat notations takes the following form

$$l_n(g_n, \lambda_n) = \begin{cases} g_n \hat{\lambda}_n - \frac{\varepsilon_n}{2} g_n^2, & \hat{\lambda}_n \leq 0, \text{ contact} \\ -\frac{1}{2\varepsilon_n} \lambda_n^2, & \hat{\lambda}_n > 0, \text{ non-contact} \end{cases} \quad (4.105)$$

The regularized integrand due to Coulomb's friction law for the incremental tangential relative displacement  $\underline{g}_t$ , the corresponding Lagrange multiplier  $\underline{\lambda}_t$

(representing tangential contact stress vector) and the penalty parameter  $\varepsilon_t$  is written as

$$l_t(\underline{g}_t, \underline{\lambda}_t, \hat{\sigma}_n) = \begin{cases} \underline{\lambda}_t \cdot \underline{g}_t + \frac{\varepsilon_t}{2} \underline{g}_t \cdot \underline{g}_t, & \|\underline{\lambda}_t + \varepsilon_t \underline{g}_t\| \leq -\mu \hat{\sigma}_n \quad \text{stick} \\ -\frac{1}{2\varepsilon_t} \left( \underline{\lambda}_t \cdot \underline{\lambda}_t + 2\mu \hat{\sigma}_n \|\underline{\lambda}_t + \varepsilon_t \underline{g}_t\| + \mu^2 \hat{\sigma}_n^2 \right), & \|\underline{\lambda}_t + \varepsilon_t \underline{g}_t\| > -\mu \hat{\sigma}_n, \quad \text{slip} \\ -\frac{1}{2\varepsilon_t} \underline{\lambda}_t \cdot \underline{\lambda}_t, & \end{cases} \quad \begin{matrix} \hat{\sigma}_n \leq 0 \\ \hat{\sigma}_n > 0 \end{matrix},$$

where  $\hat{\sigma}_n$  is a regularized contact pressure at solution

$$\hat{\sigma}_n = \sigma_n + \varepsilon_n g_n$$

Note that the tangential regularized functional  $l_t$  is extended to the non-contact domain  $\hat{\sigma}_n > 0$ . Making use of Macaulay brackets provides a shorter form of the regularized functional  $l_t$

$$l_t(\underline{g}_t, \underline{\lambda}_t, \hat{\sigma}_n) = -\frac{1}{2\varepsilon_t} \left( \underline{\lambda}_t \cdot \underline{\lambda}_t - \|\underline{\lambda}_t + \varepsilon_t \underline{g}_t\|^2 + \langle -(\|\underline{\lambda}_t + \varepsilon_t \underline{g}_t\| - \mu \langle -\hat{\sigma}_n \rangle) \rangle^2 \right)$$

And finally, introducing the hat notations for the augmented Lagrange multiplier

$$\hat{\underline{\lambda}}_t = \underline{\lambda}_t + \varepsilon_t \underline{g}_t$$

results in an even shorter form

$$l_t(\underline{g}_t, \underline{\lambda}_t, \hat{\sigma}_n) = -\frac{1}{2\varepsilon_t} \left( \underline{\lambda}_t \cdot \underline{\lambda}_t - \|\hat{\underline{\lambda}}_t\|^2 + \langle \|\hat{\underline{\lambda}}_t\| - \mu \langle -\hat{\sigma}_n \rangle \rangle^2 \right) \quad (4.106)$$

The expanded form is given by

$$l_t(\underline{g}_t, \underline{\lambda}_t, \hat{\sigma}_n) = \begin{cases} \underline{\hat{\lambda}}_t \cdot \underline{g}_t - \frac{\varepsilon_t}{2} \underline{g}_t \cdot \underline{g}_t, & \|\hat{\underline{\lambda}}_t\| \leq -\mu \hat{\sigma}_n \quad \text{stick} \\ -\frac{1}{2\varepsilon_t} \left( \underline{\lambda}_t \cdot \underline{\lambda}_t + 2\mu \hat{\sigma}_n \|\hat{\underline{\lambda}}_t\| + \mu^2 \hat{\sigma}_n^2 \right), & \|\hat{\underline{\lambda}}_t\| > -\mu \hat{\sigma}_n, \quad \text{slip} \\ -\frac{1}{2\varepsilon_t} \underline{\lambda}_t \cdot \underline{\lambda}_t, & \end{cases} \quad \begin{matrix} \hat{\sigma}_n \leq 0 \\ \hat{\sigma}_n > 0 \end{matrix}, \quad (4.107)$$

The integration of expressions (4.104) and (4.106) over the master surface leads to the following contribution of the contact conditions to the energy of the system

$$\begin{aligned} W^c &= \int_{\Gamma_c^1} l_n(g_n, \lambda_n) + l_t(\underline{g}_t, \underline{\lambda}_t, \hat{\sigma}_n) d\bar{\Gamma}_c^1 = \\ &= \int_{\Gamma_c^1} -\frac{1}{2\varepsilon_n} \left( \lambda_n^2 - \langle -\hat{\lambda}_n \rangle^2 \right) - \frac{1}{2\varepsilon_t} \left( \underline{\lambda}_t \cdot \underline{\lambda}_t - \|\hat{\underline{\lambda}}_t\|^2 + \langle \|\hat{\underline{\lambda}}_t\| - \mu \langle -\hat{\sigma}_n \rangle \rangle^2 \right) d\bar{\Gamma}_c^1 \end{aligned} \quad (4.108)$$

For frictionless contact, the augmented functional  $l_t$  should be omitted. Variation of the integral  $W^c$  leads to

$$\delta W^c = \int_{\Gamma_c^1} \delta l_n(g_n, \lambda_n) + \delta l_t(\underline{g}_t, \underline{\lambda}_t, \hat{\sigma}_n) d\bar{\Gamma}_c^1 = \int_{\Gamma_c^1} \frac{\partial l_n}{\partial g_n} \delta g_n + \frac{\partial l_n}{\partial \lambda_n} \delta \lambda_n + \frac{\partial l_t}{\partial \underline{g}_t} \cdot \delta \underline{g}_t + \frac{\partial l_t}{\partial \underline{\lambda}_t} \cdot \delta \underline{\lambda}_t d\bar{\Gamma}_c^1 \quad (4.109)$$

Note that the contact pressure  $\hat{\sigma}_n$  is not subjected to the variation even if it contains the geometrical quantity  $\hat{\sigma}_n = \sigma_n + \varepsilon_n g_n$ . This is due to the fact that the contact pressure is assumed to be the known contact pressure at solution, so its variation is zero. Since there are three possible contact statuses (stick, slip and non-contact), it is convenient to split the contact zone into three non-intersecting zones

$$\Gamma_c^1 = \Gamma_c^{1*} \cup \Gamma_c^{1\bullet} \cup \Gamma_c^1 \setminus \bar{\Gamma}_c^1,$$

where  $\Gamma_c^{1*}$  is a slip zone,  $\Gamma_c^{1\bullet}$  is a stick zone,  $\bar{\Gamma}_c^1$  is an active contact zone and consequently  $\Gamma_c^1 \setminus \bar{\Gamma}_c^1$  is a non-contact zone.

To simplify the further derivations, four simple derivatives are given here

$$\frac{\partial \hat{\lambda}_n}{\partial g_n} = \varepsilon_n, \quad \frac{\partial \hat{\lambda}_n}{\partial \lambda_n} = 1, \quad \frac{\partial \hat{\lambda}_t}{\partial \underline{g}_t} = \varepsilon_t \underline{I}, \quad \frac{\partial \hat{\lambda}_t}{\partial \underline{\lambda}_t} = \underline{I}$$

Here and further it is more convenient to write all the derivatives in a classical way using (4.105) and (4.107), instead of short expressions with Macaulay brackets (4.104) and (4.106).

$$\frac{\partial l_n(g_n, \lambda_n)}{\partial g_n} = \begin{cases} \hat{\lambda}_n, & \hat{\lambda}_n \leq 0, \text{ contact} \\ 0, & \hat{\lambda}_n > 0, \text{ non-contact} \end{cases} \quad (4.110)$$

$$\frac{\partial l_n(g_n, \lambda_n)}{\partial \lambda_n} = \begin{cases} g_n, & \hat{\lambda}_n \leq 0, \text{ contact} \\ -\frac{1}{\varepsilon_n} \lambda_n, & \hat{\lambda}_n > 0, \text{ non-contact} \end{cases} \quad (4.111)$$

$$\frac{\partial l_t(\underline{g}_t, \underline{\lambda}_t)}{\partial \underline{g}_t} = \begin{cases} \underline{\lambda}_t, & \|\underline{\lambda}_t\| \leq -\mu \hat{\sigma}_n \text{ stick} \\ -\mu \hat{\sigma}_n \frac{\underline{\lambda}_t}{\|\underline{\lambda}_t\|}, & \|\underline{\lambda}_t\| > -\mu \hat{\sigma}_n, \text{ slip} \\ 0, & \hat{\sigma}_n > 0 \end{cases}, \quad \hat{\sigma}_n \leq 0 \quad (4.112)$$

$$\frac{\partial l_t(\underline{g}_t, \underline{\lambda}_t)}{\partial \underline{\lambda}_t} = \begin{cases} \underline{g}_t, & \|\underline{\lambda}_t\| \leq -\mu \hat{\sigma}_n \text{ stick} \\ -\frac{1}{\varepsilon_t} \left( \underline{\lambda}_t + \mu \hat{\sigma}_n \frac{\underline{\lambda}_t}{\|\underline{\lambda}_t\|} \right), & \|\underline{\lambda}_t\| > -\mu \hat{\sigma}_n, \text{ slip} \\ -\frac{1}{\varepsilon_t} \underline{\lambda}_t, & \hat{\sigma}_n > 0 \end{cases}, \quad \hat{\sigma}_n \leq 0 \quad (4.113)$$

After grouping all the derivatives (4.110)-(4.113), the contact contribution to the virtual work (4.109) takes the following form in case of frictional contact

$$\delta W^c = \begin{cases} \int_{\Gamma_c^{1\bullet}} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n + \underline{\lambda}_t \cdot \delta \underline{g}_t + \underline{g}_t \cdot \delta \underline{\lambda}_t d\Gamma_c^1, & \|\underline{\lambda}_t\| \leq -\mu \hat{\sigma}_n \text{ stick} \\ \int_{\Gamma_c^{1*}} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n - \mu \hat{\sigma}_n \frac{\underline{\lambda}_t}{\|\underline{\lambda}_t\|} \cdot \delta \underline{g}_t - \frac{1}{\varepsilon_t} \left( \underline{\lambda}_t + \mu \hat{\sigma}_n \frac{\underline{\lambda}_t}{\|\underline{\lambda}_t\|} \right) \cdot \delta \underline{\lambda}_t d\Gamma_c^1, & \|\underline{\lambda}_t\| > -\mu \hat{\sigma}_n \text{ slip} \\ \int_{\Gamma_c^1 \setminus \bar{\Gamma}_c^1} -\frac{1}{\varepsilon_n} \lambda_n \delta \lambda_n - \frac{1}{\varepsilon_t} \underline{\lambda}_t \cdot \delta \underline{\lambda}_t d\Gamma_c^1, & \hat{\sigma}_n > 0 \text{ non-contact} \end{cases} \quad (4.114)$$



and a much simpler one in case of frictionless contact

$$\delta W^c = \begin{cases} \int_{\bar{\Gamma}_c^1} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n d\bar{\Gamma}_c^1, & \hat{\sigma}_n \leq 0 \quad \text{contact} \\ \int_{\Gamma_c^1 \setminus \bar{\Gamma}_c^1} -\frac{1}{\varepsilon_n} \lambda_n \delta \lambda_n d\bar{\Gamma}_c^1, & \hat{\sigma}_n > 0 \quad \text{non-contact} \end{cases} \quad (4.115)$$

After incorporating the contact integral into the equation of the virtual work balance, the following expression is obtained for the frictional contact

$$\begin{aligned} \delta \mathcal{L}^a(\underline{u}, \lambda_n, \underline{\lambda}_t, \hat{\sigma}_n) = & \int_{\Omega} \underline{\sigma} \cdot \delta \nabla \underline{u} d\Omega - \int_{\Gamma_f} \underline{\sigma}_0 \cdot \delta \underline{u} d\Gamma - \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} d\Omega + \\ & + \int_{\Gamma_c^1} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n + \underline{\lambda}_t \cdot \delta \underline{g}_t + \underline{g}_t \cdot \delta \underline{\lambda}_t d\Gamma_c^1 + \\ & + \int_{\Gamma_c^1} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n - \mu \hat{\sigma}_n \frac{\underline{\lambda}_t}{\|\underline{\lambda}_t\|} \cdot \delta \underline{g}_t - \frac{1}{\varepsilon_t} \left( \underline{\lambda}_t + \mu \hat{\sigma}_n \frac{\underline{\lambda}_t}{\|\underline{\lambda}_t\|} \right) \cdot \delta \underline{\lambda}_t d\Gamma_c^1 + \\ & + \int_{\Gamma_c^1 \setminus \bar{\Gamma}_c^1} -\frac{1}{\varepsilon_n} \lambda_n \delta \lambda_n - \frac{1}{\varepsilon_t} \underline{\lambda}_t \cdot \delta \underline{\lambda}_t d\Gamma_c^1 = 0, \\ \mathbb{V} = & \left\{ \delta \underline{u} \in \mathbb{H}^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \right\}, \end{aligned} \quad (4.116)$$

and for frictionless contact

$$\begin{aligned} \delta \mathcal{L}^a(\underline{u}, \lambda_n, \underline{\lambda}_t, \hat{\sigma}_n) = & \int_{\Omega} \underline{\sigma} \cdot \delta \nabla \underline{u} d\Omega - \int_{\Gamma_f} \underline{\sigma}_0 \cdot \delta \underline{u} d\Gamma - \int_{\Omega} \underline{f}_v \cdot \delta \underline{u} d\Omega + \\ & + \int_{\bar{\Gamma}_c^1} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n d\Gamma_c^1 + \int_{\Gamma_c^1 \setminus \bar{\Gamma}_c^1} -\frac{1}{\varepsilon_n} \lambda_n \delta \lambda_n d\Gamma_c^1 = 0, \\ \mathbb{V} = & \left\{ \delta \underline{u} \in \mathbb{H}^1(\Omega) \mid \delta \underline{u} = 0 \text{ on } \Gamma_u \right\}, \end{aligned} \quad (4.117)$$

### 4.7.3 Example

Let us illustrate how the augmented Lagrangian work in the simple case of one primal unknown. For this purpose we return to the example considered in the previous sections (Fig. 4.12): a mass (point) on a spring subjected to a constraint due to the contact with a rigid wall  $g(u) = u - 1 \geq 0$ . The spring is considered to be nonlinear with a potential  $F(u) = \frac{1}{2}cu^4$ , as previously. In the following computations, we put  $c = 1/2$ , but we will keep this constant in derivatives to make the example more meaningful from the numerical point of view. The augmented Lagrangian takes the form

$$\mathcal{L}^a(u, \lambda_n) = \begin{cases} \frac{1}{2}cu^4 + \lambda_n(u - 1) + \frac{1}{2}\varepsilon_n(u - 1)^2, & \lambda_n + \varepsilon_n(u - 1) \leq 0 \\ \frac{1}{2}cu^4 - \frac{1}{2\varepsilon_n}\lambda_n^2, & \lambda_n + \varepsilon_n(u - 1) > 0 \end{cases}$$

The graphical representation of the functional  $\mathcal{L}^a$  in the neighborhood of the solution point  $[1; -1]$  for different penalties  $\varepsilon_n = 0.5; 1; 5; 10$  can be found in Fig. 4.19; visually it is obvious that the functional is rather smooth: isolines cross smoothly the line (white thick dashed line) dividing contact and non-contact zones  $\lambda_n + \varepsilon_n(u - 1) = 0$ , this line is almost undetectable. The solution corresponds to the saddle point which retains its position independently on the penalty parameter, but higher stiffness leads to higher concentration of isolines in the direction of the primal variable  $x$ . It is also interesting to visualize gradients (Fig. 4.20) of the augmented Lagrangian  $\nabla_x \mathcal{L}^a(x, \lambda_n)$  and  $\nabla_{\lambda_n} \mathcal{L}^a(x, \lambda_n)$ , which loose the smoothness across the contact–non-contact interface  $\lambda_n + \varepsilon_n(u - 1) = 0$ .

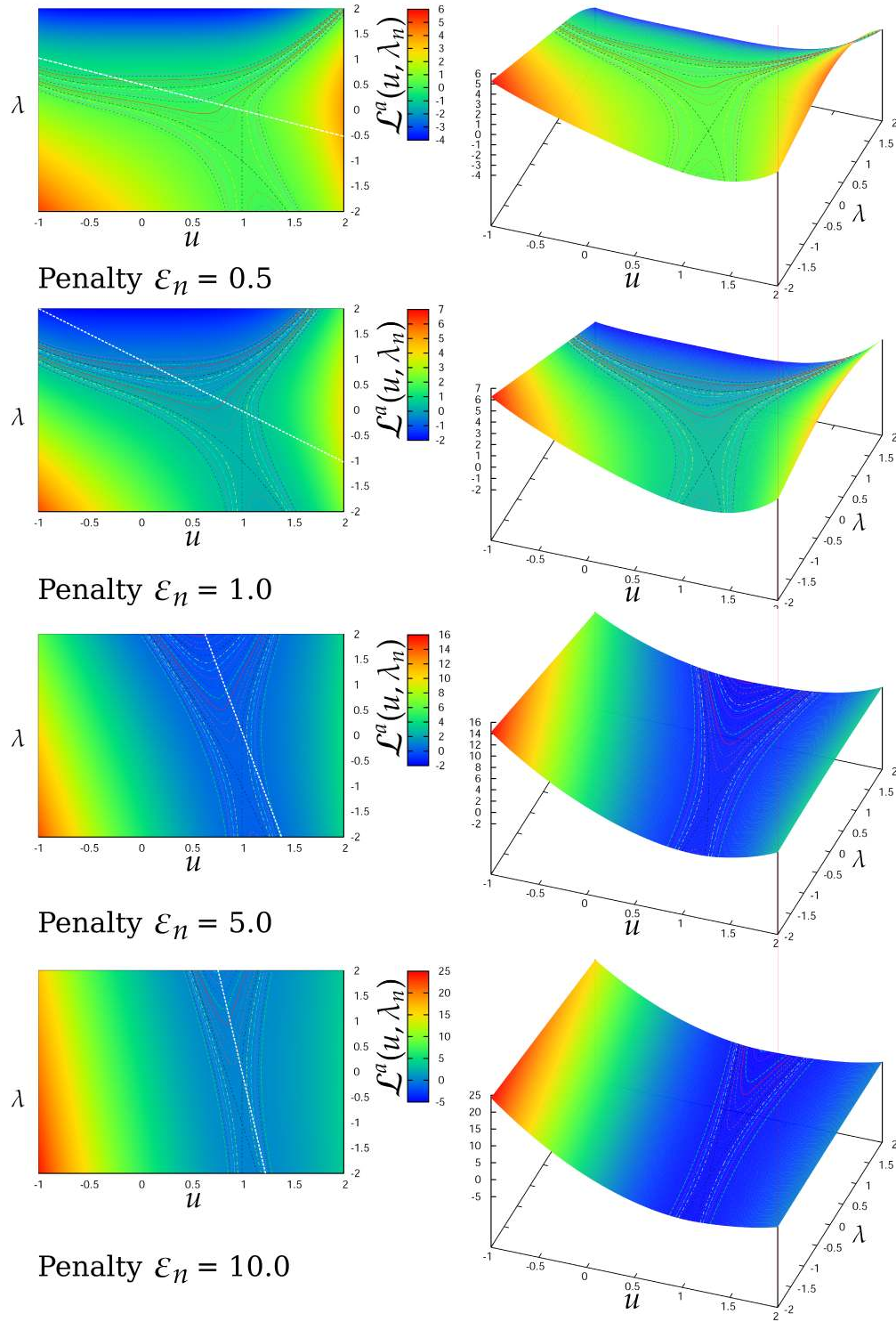


Figure 4.19: Augmented Lagrangian for a nonlinear spring compressed by a rigid wall, plotted for different penalty parameters:  $\varepsilon_n = 0.5; 1; 5; 10$ .

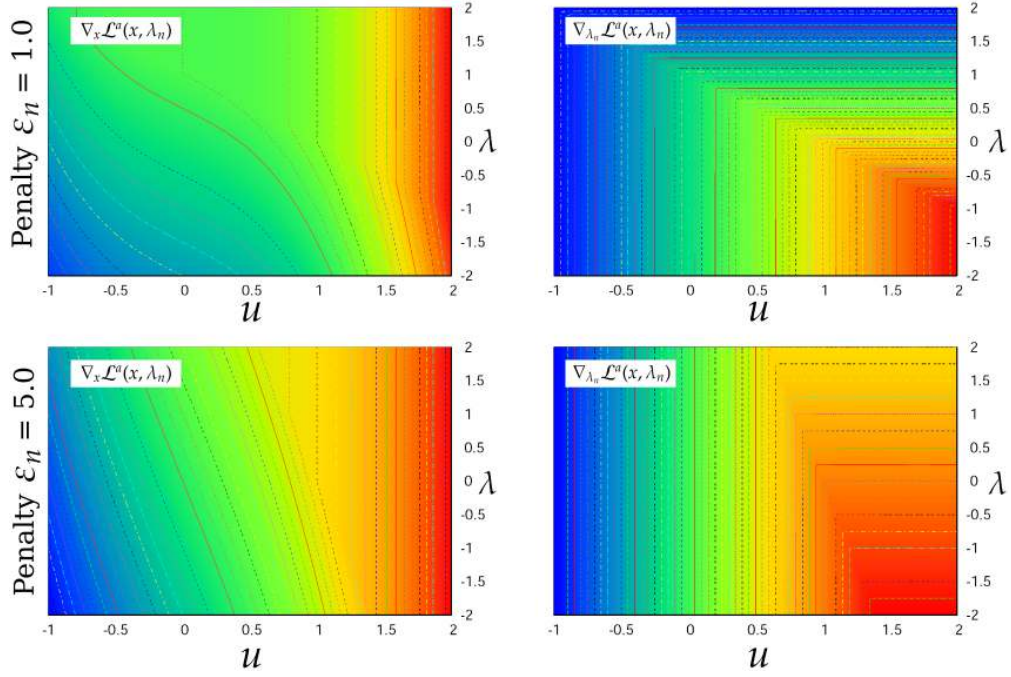


Figure 4.20: Gradients  $\nabla_x \mathcal{L}^a(x, \lambda_n)$  and  $\nabla_{\lambda_n} \mathcal{L}^a(x, \lambda_n)$  of the Augmented Lagrangian for a nonlinear spring compressed by a rigid wall are plotted for different penalty parameters:  $\varepsilon_n = 0.5; 5$ .

The variation of the augmented Lagrangian is, for the considered case

$$\delta \mathcal{L}^a(u, \lambda_n) = \begin{cases} \begin{bmatrix} 2cu^3 + \lambda_n + \varepsilon_n(u-1) \\ u-1 \end{bmatrix}^T \begin{bmatrix} \delta u \\ \delta \lambda_n \end{bmatrix} = 0, & \lambda_n + \varepsilon_n(u-1) \leq 0 \\ \begin{bmatrix} 2cu^3 \\ -\frac{1}{\varepsilon_n} \lambda_n \end{bmatrix}^T \begin{bmatrix} \delta u \\ \delta \lambda_n \end{bmatrix} = 0, & \lambda_n + \varepsilon_n(u-1) > 0 \end{cases} \quad (4.118)$$

and the second variation needed for the linearization of (4.118) is as follows

$$\Delta \delta \mathcal{L}^a(u, \lambda_n) = \begin{cases} \begin{bmatrix} \delta u \\ \delta \lambda_n \end{bmatrix}^T \begin{bmatrix} 6cu^2 + \varepsilon_n & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \delta u \\ \delta \lambda_n \end{bmatrix}, & \lambda_n + \varepsilon_n(u-1) \leq 0 \\ \begin{bmatrix} \delta u \\ \delta \lambda_n \end{bmatrix}^T \begin{bmatrix} 6cu^2 & 0 \\ 0 & -\frac{1}{\varepsilon_n} \end{bmatrix} \begin{bmatrix} \delta u \\ \delta \lambda_n \end{bmatrix}, & \lambda_n + \varepsilon_n(u-1) > 0 \end{cases} \quad (4.119)$$

One can note that the first variation of  $\mathcal{L}^a$  (4.119) (the balance of virtual work), is continuous for any values of  $u$  and  $\lambda$ , but its derivative (4.119) is not continuous. It motivates the usage of the generalized Newton's method. Note that for linear material the {11} component of the matrix in (4.119) has a form similar to  $c + \varepsilon_n$ , however in the finite element method the size of the finite element will be present. This observation leads to an appropriate choice of the penalty parameter of the order of the stiffness ( $\varepsilon_n \sim c$ ). If the

{11} component of the matrix in (4.119) has the form  $c\phi(u) + \varepsilon_n\psi(u)$ , which is the case for nonlinear material and nonlinear constraints, and if  $u$  tends to the solution  $u^*$  in the way that function  $\psi(u^*) \rightarrow 0$ , then it is reasonable to increase the penalty coefficient during iterations in order to improve the convergence of the augmented Lagrangian without danger of ill-conditioning of the matrix. Different techniques for the penalty update have been proposed (see, for example, [Bussetta 09]).

Let us demonstrate the convergence of the augmented Lagrangian method on a simple example of loading and unloading of the spring-mass system considered in the previous section. The convergence for different constant penalty coefficients and for the updated technique are compared in Table 4.11. In case of loading, the free spring deforms due to the contact with a rigid wall, the Augmented Lagrangian method converges in 3 iterations (as Lagrange multiplier method), i.e. at the second iteration the exact solution is obtained. Let us consider the case when from the obtained loaded state  $x = 0$ ,  $\lambda = -1$ , we move the wall at  $x_w = -0.9$ , so the spring remains compressed. The system of equation respectively to the deformed state becomes

$$\begin{cases} [R] = \begin{bmatrix} 2c(u^i + 1)^3 + \lambda_n^i + \varepsilon_n(u^i - x_w) \\ u^i - x_w \end{bmatrix}^T, \\ [K] = \begin{bmatrix} 6c(u^i + 1)^2 & 1 \\ 1 & 0 \end{bmatrix}, \end{cases} \quad \text{contact } \lambda_n^{i-1} + \varepsilon_n(u^{i-1} - x_w) \leq 0$$

$$\begin{cases} [R] = \begin{bmatrix} 2c(u^i + 1)^3 \\ -\frac{\lambda_n^i}{\varepsilon_n} \end{bmatrix}^T, \\ [K] = \begin{bmatrix} 6c(u^i + 1)^2 & 0 \\ 0 & -\frac{1}{\varepsilon_n} \end{bmatrix} \end{cases} \quad \text{non-contact } \lambda_n^{i-1} + \varepsilon_n(u^{i-1} - x_w) > 0$$
(4.120)

At each iteration we evaluate the solution increment in the following manner

$$\begin{bmatrix} \Delta u^i \\ \Delta \lambda_n^i \end{bmatrix} = -[K]^{-1}[R]$$

In case of non-contact, the eigen values of the tangent matrix are  $\lambda_{\max} = 6c(u^i + 1)^2$  and  $\lambda_{\min} = -\frac{1}{\varepsilon_n}$  and the condition number is

$$\text{Cond}(K) = \frac{|\lambda_{\min}|}{|\lambda_{\max}|} = 6c(u^i + 1)^2 \varepsilon_n \sim c\varepsilon_n$$

For a high penalty coefficient  $\varepsilon_n$  and high stiffness coefficient  $c$  the condition number of the tangent matrix becomes very high which is crucial for the precision of the solution and its convergence. That is the price of the fully unconstrained smooth energy functional. So even if formally the coupled augmented Lagrangian derives the precise solution, it suffers from numerical errors and its convergence is hindered. This is one of the main motivations to update the primal and dual variables separately by means of the Uzawa's algorithm. In case of simultaneous resolution, the penalty coefficients should

be chosen reasonably small at least in case of non-contact. This is especially true for the frictional problems, for which a high penalty coefficient may lead to cycling over solution. For some more details on the condition numbers and the convergence of the Newton's scheme, the reader is referred to Section 5.1.

Two different convergences can be observed, which depend on the penalty parameter  $\varepsilon_n$  (see Fig. 4.21). If  $\lambda_n^0 + \varepsilon_n(u^0 - x_w) \leq 0$ , we start from the "contact" type cycle, i.e. if  $\varepsilon_n \leq \varepsilon_n^* = -\frac{\lambda_n^0}{u^0 - x_w}$ , otherwise this is a no-contact cycle. Sometimes it is advantageous to start from a "contact" type cycle, so for that contact pressure  $\sigma_n \leftarrow \lambda_n$  should be transferred from one converged increment to the following one. In Table 4.11, for example, we make a start from  $\lambda_n^0 = -1$ , which was the solution of the spring problem when it came in contact with a rigid wall. It is also interesting to compare simultaneous (coupled) resolution of the problem for primal and dual variables with Uzawa's algorithm (nested resolution) (last four columns in Table (4.11)). The last one converges significantly more slowly than the coupled algorithm in the contact domain. Note that the initial guess of being in the contact domain (Uzawa for  $\varepsilon_n \leq \varepsilon_n^*$ ) only disturbs the solution and increases the number of necessary iterations. Linear convergence rate becomes clear when one studies the relative error (not given in the table). In case of full unloading  $x_w = -1.1$ , there are no spurious modes, and the algorithm converges properly to the minimum of the unconstrained functional both for Uzawa and coupled algorithms.

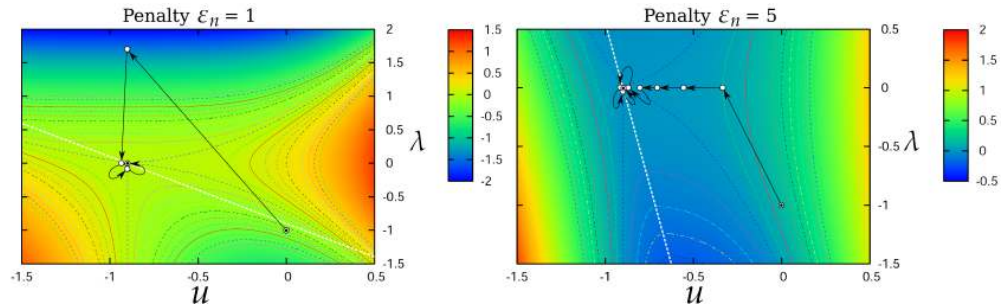


Figure 4.21: Two different types of convergence of the augmented Lagrangian method in case of partial unloading: the convergence path depends on the initial position of the point – inside contact domain or inside non-contact domain, which in turn depends upon the penalty coefficient.

Table 4.11: Augmented Lagrangian convergence in case of partial unloading.

Iter, $i$	Coupled resolution				Nested resolution (Uzawa)			
	Penalty $\varepsilon_n \leq \varepsilon_n^*$		Penalty $\varepsilon_n > \varepsilon_n^*$		Penalty $\varepsilon_n \leq \varepsilon_n^*$		Penalty $\varepsilon_n > \varepsilon_n^*$	
	$u_i$	$\lambda_i$	$u_i$	$\lambda_i$	$u_i$	$\lambda_i$	$u_i$	$\lambda_i$
0	0.0	-1.0	0.0	-1.0	0.0	-1.0	0.0	-1.0
1	-0.90	1.7	-0.333	0.0	-0.225	-0.325	-0.333	0.0
2	-0.93	0.0	-0.556	0.0	-0.483	0.0	-0.556	0.0
3	-0.90	-0.087	-0.704	0.0	-0.656	0.0	-0.704	0.0
4	-0.90	-1e-3	-0.802	0.0	-0.770	0.0	-0.802	0.0
5	-0.90	-1e-3	-0.868	0.0	-0.847	0.0	-0.868	0.0
6	-0.90	-1e-3	-0.913	0.0	-0.898	0.0	-0.912	0.0
7	-0.90	-1e-3	-0.90	-0.03	-0.932	0.0	-0.900	-9.54e-4
8	-0.90	-1e-3	-0.90	-1e-3	-0.901	-7.48e-4	-0.900	-9.99e-4
9	-0.90	-1e-3	-0.90	-1e-3	-0.900	-9.92e-4	-0.900	-9.99e-4
10	-0.90	-1e-3	-0.90	-1e-3	-0.900	-9.99e-4	-0.900	-1e-3
11	-0.90	-1e-3	-0.90	-1e-3	-0.900	-9.99e-4	-0.900	-1e-3
12	-0.90	-1e-3	-0.90	-1e-3	-0.900	-1e-3	-0.900	-1e-3





## Chapter 5

# Numerical procedures

### **Résumé de Chapitre 5 «Procédures numériques»**

*Le but de ce chapitre est de présenter toutes les méthodes numériques nécessaires pour le traitement des problèmes de contact par la méthode des éléments finis avec résolution implicite.*

*Tout d'abord la méthode de Newton et sa version généralisée (nécessaire pour le traitement des problèmes de contact) sont présentées. Quelques notions de base de la théorie des sous-différentiels sont également décrites, avec une interprétation géométrique originale. La dernière permet de raffiner l'interprétation du point le plus proche et de la projection normale pour le cas des surfaces lisses par morceaux. Les conditions de convergence de la méthode de Newton généralisée pour le problème de contact avec frottement sont présentées et interprétées par un problème purement géométrique.*

*L'algorithme de «return mapping» est celui qui est adapté pour traiter les problèmes de contact avec frottement dans le cas où on utilise la méthode de pénalisation et la méthode du Lagrangien augmenté avec l'algorithme d'Uzawa. Les détails de cet algorithme sont illustrés pour le cas 2D et 3D.*

*La discussion de ces algorithmes est suivie par une présentation du formalisme de la méthode des éléments finis dans le cadre de la nouvelle algèbre de S-structure. On présente une définition rigoureuse des éléments de contact et les règles de leur formation pour différentes discrétisations de l'interface, puis on dérive la forme faible pour le problème de contact discrétisé et sa linéarisation pour la méthode de Newton. Les vecteurs résidus et les matrices tangentes sont exprimés pour la méthode de pénalisation et la méthode du Lagrangien augmenté pour le cas de contact sans et avec frottement.*

*En conclusion on discute l'implémentation de la méthode PDN dans un code des éléments finis et plusieurs détails techniques utiles pour traiter des problèmes particuliers : cas de la surface de maître rigide, éléments de contact à faces multiples, frottement hétérogène, etc.*

This chapter is devoted to the numerical aspects of the contact mechanics.

The classical **Newton's method** is presented. Some remarks on its applicability to contact problems are made, then the multidimensional version of the method is stated and the main characteristics of the resulting system of linear equations are briefly discussed. Further we give a detailed description of the **return mapping algorithm**, which is often used for integration of frictional conditions together with the penalty method. Next the Finite Element Method formalism is briefly introduced using the new notations of s-structures. After setting up this numerical framework, the closed form expressions for the implementation of the penalty, Lagrange multiplier and coupled augmented Lagrangian methods in a Finite Element code are given. Finally, some technical details are presented.

## 5.1 Newton's method

Generally for locally convex and smooth function, the solution of nonlinear problem can be approximated by a well-known Newton or also called Newton-Raphson procedure, which replaces a nonlinear problem by a series of linear problems. If the starting point is sufficiently close to the solution (here it implies, for example, small changes of boundary conditions) the Newton-Raphson method provides a quadratic rate of convergence if the conditions of convexity and smoothness are fulfilled. However, as it has been shown in the previous chapter, it is not always possible to fulfill these conditions: the virtual work functional is not smooth. Here a short discussion of the Newton-Raphson method is given and its extension to the class of nondifferentiable problems arising from contact mechanics are given.

### 5.1.1 One-dimensional Newton's method

The main idea of Newton's method can be easily captured on a one-dimensional example. Let the solution  $x$  of an equation depend on an external parameter  $f$ , then any equation can be written in the following form

$$R(x, f) = 0,$$

where  $R$  is a scalar function of two scalar arguments  $x$  and  $f$ . To fulfill this equation, a change of  $f$  should result in a change of  $x$ . A straightforward analogy with the mechanical system in statics containing one degree of freedom is, for example, the displacement  $x$  of the mass  $m$  attached to a spring with a linear stiffness  $k$ , under a gravity force  $f = mg$ . If the system is conservative, then the solution corresponds to the minimum of its energy

$$W(x, g) = \frac{1}{2}cx^2 + mgx \rightarrow \frac{\partial W}{\partial x} = cx + mg = 0 \quad \sim \quad R(x, g) = cx + mg = 0$$

Let us consider a time discretized model, i.e. the external parameter  $f$  changes incrementally  $f_0, f_1, f_2, \dots$  and we seek the values of the argument  $x_0, x_1, x_2, \dots$  which fulfill  $R(x_i, f_i) = 0$  at least with a given precision  $|R(x_i, f_i)| \leq \varepsilon$ .

We suppose that the starting point  $x_0$  is known from the initial conditions or from a previously solved increment  $x_i$

$$R(x_i, f_i) = 0.$$

For the following increment, the external parameter  $f$  becomes  $f_{i+1}$ . The aim is to find  $x_{i+1}$ . If the problem is nonlinear, an iterative scheme should be used. The first terms of the Taylor series of  $R(x, f)$  centered at the previous iteration  $x^j$  are

$$R(x^j + \Delta x^j, f_{i+1}) = 0 \Rightarrow R(x^j + \Delta x^j, f_{i+1}) = R(x^j, f_{i+1}) + \left. \frac{\partial R}{\partial x} \right|_{x^j} \Delta x^j + r_1(x^j) = 0. \quad (5.1)$$

If higher order terms  $r_1(x^j)$  are negligibly small and if a nonzero derivative  $\left. \frac{\partial R}{\partial x} \right|_{x^j} \neq 0$  exists, then the increment  $\Delta \underline{x}^j$  can be written as

$$\Delta x^j \approx - \left. \frac{\partial R}{\partial x} \right|_{x^j}^{-1} R(x^j, f_{i+1}), \quad x^{j+1} = x^j + \Delta x^j \quad (5.2)$$

Further we put  $x^{j+1} = x^j + \Delta x^j$  and check if the new result is sufficiently close to zero  $|R(x^{j+1}, f_{i+1})| \leq \varepsilon$ , if not we repeat (5.1), where Taylor series is centered at  $x^{j+1}$ .

According to expression (5.1), we should require smoothness of the function  $R(x, \cdot)$  at least in points  $x^j$ . Since these points are arbitrary we require that

$$R(x) \in C^1(x) \quad (5.3)$$

It is also important that a small change in  $f$  leads to a small change in  $R$ , in other words

$$\exists K : 0 < K < \infty : \forall f^1, f^2 : |R(x, f^2) - R(x, f^1)| \leq K \|f^2 - f^1\| \quad (5.4)$$

the smaller the parameter  $K$ , the slower function  $R$  changes with  $f$ , i.e.  $R(\cdot, f)$  is a Lipschitz continuous function  $R(\cdot, f) \in L^K$ . As one can notice, the higher derivatives of  $R$  do appear explicitly in the derived equations, however since we neglected all terms contained in  $r_1(x^j)$ , we suppose that they are negligibly small. Let us demonstrate that this residual  $r_1(x^j)$  plays an important role. Let us rewrite the residual in Lagrange form

$$r_1(x^j) = \frac{1}{2} \left. \frac{\partial^2 R}{\partial x^2} \right|_{\xi^j} (x^* - x^j)^2,$$

where  $\xi^j = x^* + \theta(x^j - x^*)$ ,  $0 < \theta < 1$  and  $x^*$  is the exact solution. The expansion of the function  $R$  is

$$R(x^*, f_{i+1}) = R(x^j, f_{i+1}) + \left. \frac{\partial R}{\partial x} \right|_{x^j} (x^* - x^j) + \frac{1}{2} \left. \frac{\partial^2 R}{\partial x^2} \right|_{\xi^j} (x^* - x^j)^2 = 0.$$

Dividing by  $\left. \frac{\partial R}{\partial x} \right|_{x^j} \neq 0$  and grouping the terms gives:

$$\left( \left. \frac{\partial R}{\partial x} \right|_{x^j}^{-1} R(x^j, f_{i+1}) - x^j \right) + x^* = - \frac{1}{2} \left. \frac{\partial R}{\partial x} \right|_{x^j}^{-1} \left. \frac{\partial^2 R}{\partial x^2} \right|_{\xi^j} (x^* - x^j)^2$$

According to (5.2), the first brackets are nothing but  $-x^{j+1}$ , so if the distance to the solution is denoted as  $x^* - x^j = \varepsilon_j$ , the equality can be rewritten

$$\varepsilon_{j+1} = -\frac{1}{2} \frac{\partial R}{\partial x} \bigg|_{x^j}^{-1} \frac{\partial^2 R}{\partial x^2} \bigg|_{\varepsilon_j} \varepsilon_j^2,$$

from where some corollaries can be deduced

- if everywhere  $\frac{\partial^2 R}{\partial x^2} = 0$  then  $\varepsilon_{j+1} = 0$ , the function  $R$  is linear and the algorithm converges in one iteration;
- if  $\exists C, L : 0 < \{C, L\} < \infty$  such that

$$\frac{1}{2} \left| \frac{\partial^2 R}{\partial x^2} \bigg|_{\varepsilon_j} \right| \leq C < \infty \quad \text{and} \quad \left| \frac{\partial R}{\partial x} \bigg|_{x^j}^{-1} \right| \leq L < \infty$$

at least locally, then

$$|\varepsilon_{j+1}| \leq CL|\varepsilon_j|^2,$$

if  $CL|\varepsilon_j| < 1$   $\varepsilon_j$  then  $x^j$  converges to the solution  $x^*$ , and if the starting point  $x^0$  is sufficiently close to the solution  $x^*$ , i.e.  $CL|\varepsilon_0| < 1$  then the convergence is quadratic;

- if one adds to the previously formulated conditions the following ones

$$\begin{aligned} \text{sign} \left[ \frac{\partial^2 R}{\partial x^2} \bigg|_{\varepsilon_j} \right] &= \text{const}, \quad \text{sign} \left[ \frac{\partial R}{\partial x} \bigg|_{x^j}^{-1} \right] = \text{const}, \\ \text{sign}[\varepsilon_0] &= \text{sign}[x^0 - x^*] = \text{sign} \left[ \frac{\partial^2 R}{\partial x^2} \bigg|_{\varepsilon_j} \right] \end{aligned} \quad (5.5)$$

i.e.  $R(x, \cdot)$  is strictly monotonic and if the convexity does not change. Moreover, if the first point  $x_0$  is to the right from solution for convex and to the left for concave function, then  $x^j$  converges quadratically and monotonically to the solution;

- it follows, for example, that if

$$\frac{\partial R}{\partial x} \bigg|_{x^j} \xrightarrow{x^j \rightarrow x^*} 0 \quad \text{or} \quad \left| \frac{\partial^2 R}{\partial x^2} \bigg|_{x^j} \right| \xrightarrow{x^j \rightarrow x^*} \infty$$

then the convergence, if it presents, is not quadratic (as in examples considered in the previous chapter).

This short reminder allows to proceed with analysis of Newton's method for special cases. Let us note that the rigorous statement of Newton's method convergence in a  $n$ -dimensional space is given by the Kantorovich theorem [Kantorovich 48], where among other Lipschitz continuity of the first derivative is required

$$\left\| \frac{\partial R(x, \cdot)}{\partial x} - \frac{\partial R(y, \cdot)}{\partial y} \right\| \leq L \|x - y\|,$$

bold symbols mean  $\mathbf{a} \in \mathbb{R}^n$ . Carrying the results of this paragraph in mind we will shortly consider the application of Newton's method for the case of nondifferentiable functions, which is essentially the case in computational treatment of contact problems. But before let us show how Newton's method works in a multidimensional space.

### 5.1.2 Multidimensional Newton's method

In the discretized case if the problem can be stated as

$$\mathbf{R}(\mathbf{u}, \mathbf{f}) = 0,$$

where  $\mathbf{R} \in \mathbb{R}^n$  is a vector function of vector arguments  $\mathbf{u} = [u^1, u^2, \dots, u^n]^T \in \mathbb{R}^n$  and  $\mathbf{f} = [f^1, f^2, \dots, f^m]^T \in \mathbb{R}^m$ , the expression for the solution increment in the framework of Newton's method is written as

$$\Delta \mathbf{u}^i = -\mathbf{K}(\mathbf{u}^i)^{-1} \mathbf{R}(\mathbf{u}^i, \mathbf{f}_{k+1}), \quad \mathbf{u}^{i+1} = \mathbf{u}^i + \Delta \mathbf{u}^i$$

where the upper index  $i$ , as before, denotes the iteration number and the lower index the solution step or increment number. The matrix  $\mathbf{K}$  in the Finite Element method is called the *tangent stiffness matrix* for elastic problems or the *tangent matrix* for nonlinear problems, in optimization theory the *Hessian* (second order derivative of the objective function to be minimized, named after Ludwig Otto Hesse (1811-1874), German mathematician) or the *Jacobian* (named after Carl Gustav Jacob Jacobi (1804-1851), German mathematician) of the vector  $\mathbf{R}$

$$\mathbf{K}(\mathbf{u}^i) = \left. \frac{\partial \mathbf{R}(\mathbf{u}, \mathbf{f})}{\partial \mathbf{u}} \right|_{\mathbf{u}^i}$$

The inverse of this matrix exists if and only if its determinant is not zero which is equivalent to the condition that the spectrum of the matrix  $\sigma(\mathbf{K})$  does not contain any zero eigenvalue

$$\det \mathbf{K}(\mathbf{u}^i) \neq 0 \Leftrightarrow \forall \lambda_i \in \sigma(\mathbf{K}(\mathbf{u}^i)) : \lambda_i \neq 0$$

If the matrix is not Lipschitz continuous, i.e.

$$\nexists K : 0 < K < \infty : \forall \lambda_i \in \sigma(\mathbf{K}), \forall \mathbf{u}_1, \mathbf{u}_2 : |\lambda_i(\mathbf{u}_1) - \lambda_i(\mathbf{u}_2)| \leq K \|\mathbf{u}_1 - \mathbf{u}_2\|$$

then the assumptions of the Kantorovich theorem are not fulfilled and the convergence of the Newton-Raphson method may experience some problems. This case will be considered in the next paragraph.

Another important characteristic of the tangent matrix is its condition number  $\text{Cond}[\mathbf{K}]$ :

$$\text{Cond}(\mathbf{K}) = \frac{\lambda_{\max}}{\lambda_{\min}},$$

where  $\lambda_{\max}$  and  $\lambda_{\min}$  are respectively the maximal and the minimal **by moduli** eigenvalues of the matrix  $[\mathbf{K}]$ . In numerical analysis, the solution  $\mathbf{u}_{*n}$  is different from the exact solution  $\mathbf{u}_*$ , first due to the finite precision required by the user,

$\varepsilon$ , for instance like this  $\|u_{i+1} - u_i\| \leq \varepsilon$ , and secondly due to the finite number of digits in computer data types. The higher the condition number of the matrix, the lower the number of correctly evaluated digits in the solution. The number of lost digits  $N$  in accuracy can be computed according to the simple formula

$$N = \log_{10} \text{Cond}(K)$$

So ill-conditioning of the matrix (high condition number) results in loss of accuracy and may also result even in divergence of the iterative schemes. For that reason, the so-called preconditioners should be used, which replace the problem  $Ku = f$  with solution  $u_*$  by a problem  $\tilde{K}[u] = \tilde{f}$  such that the solution remains the same but  $\text{Cond}(\tilde{K}) < \text{Cond}(K)$ . In this context, it becomes evident that, if the penalty method increases the condition number of the stiffness matrix proportionally to the penalty coefficient, then a high penalty coefficient results in ill-conditioning of the matrix. It is also true for the augmented Lagrangian method.

It is worth mentioning that the procedure of the stiffness matrix calculation is the major contribution to the computational time in the implicit Finite Element Method. That is the main motivation for the group of so-called Quasi-Newton's method, for which the stiffness matrix is approximated according to some rules (see, e.g. [Bonnans 06]): DFP (Davidon–Fletcher–Powell formula), BFGS (Broyden–Fletcher–Goldfarb–Shanno method), Broyden method, etc. Another approach is to inverse stiffness matrix only once at the zeroth iteration and compute all solution increments with this matrix, which can be also updated

$$\Delta u^i = -K(u^0)^{-1} R(u^i, f_{k+1}).$$

These methods allows to avoid problems which appear if  $R$  is not strictly monotonic, i.e. possible zero determinant of the stiffness matrix. The convergence rate of such methods is slower, but each iteration is faster. There are many techniques based on Newton's method and resulting in faster convergence in physical and/or numerical sense, see [Fletcher 77], [Bonnans 06], [Bertsekas 03]. Another class of methods often used in Finite Element codes is the one of Conjugate Gradient methods, however we will not consider them here.

### 5.1.3 Application to nondifferentiable functions

Let us return to conditions formulated for the one dimensional case, i.e. nonzero and Lipschitz continuous first derivative  $L(\mathbb{U}) \ni \frac{\partial R(u,.)}{\partial u} \neq 0$  and Lipschitz continuity by parameter  $R(., f) \in L(\mathbb{F})$ , where  $\mathbb{U}$  and  $\mathbb{F}$  are allowable sets for the argument  $u$  and the parameter  $f$  respectively. The condition for the parameter holds for contact problems: any small change of boundary conditions results in a smooth change in energy of the system and consequently in continuous change of its variation. However, due to geometrical restrictions on the displacement field we cannot require that function  $R(u,.)$  (variation of the energy) would be everywhere smooth. It leads to a relatively new domain in optimization theory – Nonsmooth or Nondifferentiable Optimization.

According to the derived equations and the simple examples considered in the previous chapter, we know that the function  $R(\mathbf{u}, \cdot)$  resulting from the virtual work principle is piecewise smooth and so there is a set of points, also called "kinks"  $\mathbf{u}^s \in \mathbb{V}^s \subset \mathbb{V}$ , for which there is no unique but several one-sided derivatives

$$\forall \mathbf{u}^s \in \mathbb{V}^s : \exists \bigcup_i \left\{ \frac{\partial R}{\partial \mathbf{u}} \Big|_{\mathbf{u}^s} \right\}_i.$$

However, the measure of such a set is zero in the whole space  $\mathbb{V} \ni \mathbf{u}$

$$M(\mathbb{V}^s) = 0, \quad M(\mathbb{V}) > 0$$

It means that in practice the probability that during iterations  $\mathbf{u} \in \mathbb{V}^s$  is zero, however often such points (kinks) appear to be solutions of minimization problems, it is also true in our case. It is clear that in the vicinity of such points the derivative is not any more Lipschitz continuous

$$\nexists K : 0 < K < \infty : \forall \mathbf{u}^1, \mathbf{u}^2 \forall i, j : \left\| \left\{ \frac{\partial R}{\partial \mathbf{u}} \Big|_{\mathbf{u}^1} \right\}_i - \left\{ \frac{\partial R}{\partial \mathbf{u}} \Big|_{\mathbf{u}^2} \right\}_j \right\| \leq K \|\mathbf{u}^1 - \mathbf{u}^2\|.$$

So, a complementary convergence analysis is required, since Kantorovich theorem [Kantorovich 48] (statement of multidimensional convergence of Newton's method) requires a Lipschitz continuity of this derivative.

### 5.1.4 Subdifferentials and subgradients

In nonsmooth optimization the gradient is generalized to the class of nondifferentiable functions in the following way, in the kink point a nonempty convex set is introduced as follows

$$\partial f(\mathbf{x}) = \left\{ \mathbf{s} \mid \mathbf{s} \in \mathbb{R}^n, \forall \mathbf{y} \in \mathbb{R}^n : f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{s} \cdot (\mathbf{y} - \mathbf{x}) \right\},$$

where by dot we mean the scalar product of vectors in  $\mathbb{R}^n$ . Another more straightforward definition can be given in terms of directional derivatives, where the direction is defined by a vector  $\mathbf{d} \in \mathbb{R}^n$

$$\partial f(\mathbf{x}) = \left\{ \mathbf{s} \mid \mathbf{s} \in \mathbb{R}^n, \forall \mathbf{d} \in \mathbb{R}^n : \mathbf{s} \cdot \mathbf{d} \leq \lim_{\alpha \rightarrow 0} \frac{f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})}{\alpha} \right\}$$

This set is called a convex *subdifferential* of the function  $f$  at point  $\mathbf{x}^k$ , each component of the subdifferential of  $f$  at  $\mathbf{x}$  is called a *subgradient*.

To my opinion, it would even be preferable to use the definition given below and graphically illustrated for the onedimensional case in Fig. 5.1. Let us imagine that the kink point  $\mathbf{x}^k$  is surrounded by an hypersphere  $\mathfrak{H} \in \mathbb{R}^{n+1}$

$$\mathbf{x}, f \in \mathfrak{H} : \left( f(\mathbf{x}) - f(\mathbf{x}^k) \right)^2 + (\mathbf{x} - \mathbf{x}^k) \cdot (\mathbf{x} - \mathbf{x}^k) = R^2,$$

## 5.1 Newton's method

where  $R$  is the radius of the hypersphere  $\mathfrak{H}$  (Fig. 5.1,b.). Then if  $f(x)$  is locally convex in the vicinity of the kink point, we can confine ourself to the lower half of the hypersphere  $\mathfrak{H}_-$

$$x, f \in \mathfrak{H}_- : f_{\mathfrak{H}}(x) = f(x^k) - \sqrt{R^2 - (x - x^k) \cdot (x - x^k)} < f(x^k).$$

The gradient of such a function  $f_{\mathfrak{H}}(x)$  at all points of the lower hypersphere in all directions always exists and takes all values  $(-\infty, \infty)$

$$\nabla f_{\mathfrak{H}}(x) \cdot d = \frac{(x - x^k) \cdot d}{\sqrt{R^2 - (x - x^k) \cdot (x - x^k)}} \in (-\infty, \infty) \text{ for all } x \in \mathfrak{H}.$$

Now we can tighten this lower hypersphere into a point  $R \rightarrow 0$ , then  $f_{\mathfrak{H}}(x) \rightarrow f(x^k)$  but all gradients do not vanish but appear to be "condensed" in the kink point if one replaces

$$x - x^k = \alpha R e',$$

where  $e'$  – is a vector filled with 1 and  $\alpha \in (-1; 1)$ . Then, for any  $R \geq 0$

$$\nabla f_{\mathfrak{H}}(x, \alpha) \cdot d = \frac{\alpha e' \cdot d}{\sqrt{1 - \alpha^2}} \in (-\infty, \infty) \forall \alpha \in (-1; 1).$$

Let us denote by  $e_j = [0, \dots, 0, \underbrace{1}_j, 0, \dots, 0]^T$ . This hypersphere-point

is put in the kink point and further we can define the positive and negative subdifferential components  $j$  such that if

$$\partial f_j^+ = \nabla f_{\mathfrak{H}}(x, \alpha) \cdot e_j, \forall \alpha : 0 \leq \alpha \leq \alpha_j^+$$

$$\partial f_j^- = \nabla f_{\mathfrak{H}}(x, \alpha) \cdot e_j, \forall \alpha : \alpha_j^- \leq \alpha < 0$$

where  $\frac{\partial f}{\partial x_j^+}, \frac{\partial f}{\partial x_j^-}$  define positive and negative one sided derivative by the argument  $x_j$  respectively (Fig. 5.1,b.), and

$$\alpha_j^+ : \nabla f_{\mathfrak{H}}(x, \alpha_j^+) \cdot e_j = \frac{\partial f}{\partial x_j^+}, \quad \alpha_j^- : -\nabla f_{\mathfrak{H}}(x, \alpha_j^-) \cdot e_j = \frac{\partial f}{\partial x_j^-}$$

The subdifferential is then defined as the union of positive and negative subdifferentials

$$\partial f = \partial f^- \cup \partial f^+$$

or in a more compact form (Fig. 5.1,c.)

$$\partial f = \nabla f_{\mathfrak{H}}(x, \alpha), \forall \alpha^- \leq \alpha \leq \alpha^+,$$

where through condition  $\alpha^- \leq \alpha \leq \alpha^+$  we mean that  $\forall j \alpha_j^- \leq \alpha_j \leq \alpha_j^+$ . In other words, the subdifferential is defined as a set of derivatives of the hypersphere limited by one sided derivatives of the function  $f$ :  $\frac{\partial f}{\partial x_j^-}$  and  $\frac{\partial f}{\partial x_j^+}$ , obviously



if  $\frac{\partial f}{\partial x_j^-} = \frac{\partial f}{\partial x_j^+}$ , then  $\partial f_j = \frac{\partial f}{\partial x_j}$ . Such a definition is quite long but hopefully it provides a more intuitive and tangible visualization of the subdifferential notion. Moreover to my opinion it is easy to deal with a very small and smooth sphere than with simply a singular kink point. In case of a locally concave function  $f$ , the subdifferential defined on a hypersphere is easy to generalize. For this purpose, instead of the lower hemisphere, one should consider the upper hemisphere of the hypersphere  $f_{\mathbb{S}}$ . The generalization is also straightforward in case of a concave-convex function  $f$ .

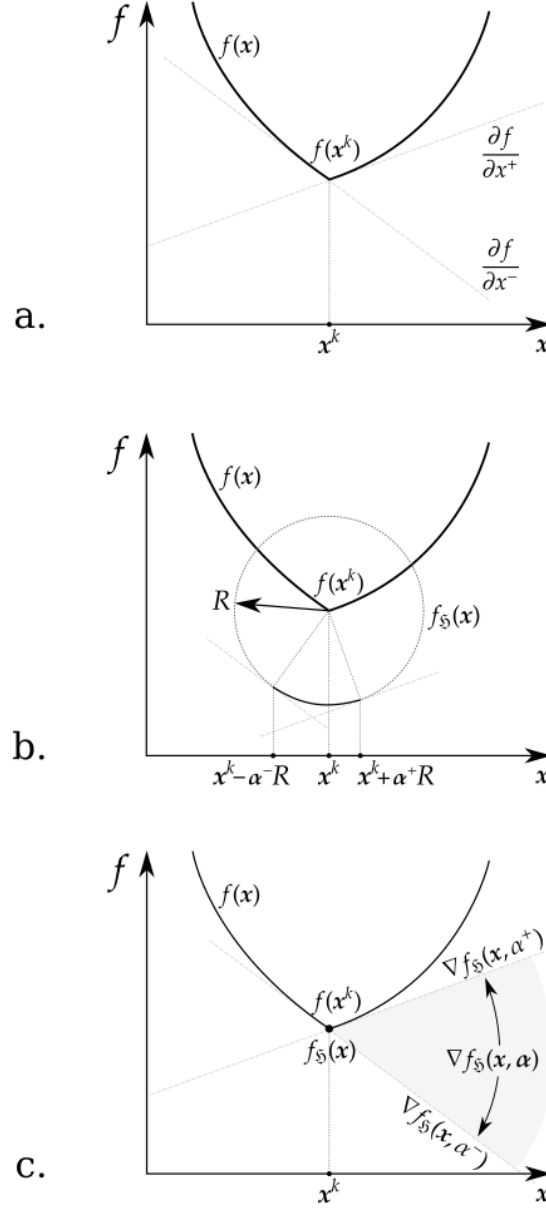


Figure 5.1: Graphical representation of a subdifferential based on hypersphere notion: **a.** – kink point  $x_k$  of the nonsmooth function  $f(x)$  and two one sided derivatives: left-sided  $\frac{\partial f}{\partial x^-}|_{x^k}$  and right-sided  $\frac{\partial f}{\partial x^+}|_{x^k}$ ; **b.** – a hypersphere  $f_S(x^k)$  of finite radius  $R$  and limits on  $\alpha$  :  $\alpha^- = \frac{\partial f}{\partial x^-}|_{x^k}$ ,  $\alpha^+ = \frac{\partial f}{\partial x^+}|_{x^k}$ ; **c.** – subdifferential  $\partial f = \nabla f_S(x, \alpha)$ ,  $\forall \alpha^- \leq \alpha \leq \alpha^+$ .

A more sophisticated description of subdifferentials applied to contact mechanics can be found in the works of Alart and Curnier [Alart 91], [Heegaard 93], [Alart 97], etc. A general subdifferential theory can be found,

for example, in [Rockafellar 70].



### Idea 5.1 Application of hyperspheres to normal projection procedure.

The idea of subdifferential and particularly of hypersphere can be directly applied to the definition of the normal projection of a slave node on the piecewise smooth master surface. For this purpose, each master node  $\underline{\rho}^i$  is replaced by a 3D sphere

$$S(\underline{\rho}^i, R) : (\underline{\rho} - \underline{\rho}^i) \cdot (\underline{\rho} - \underline{\rho}^i) - R^2 = 0$$

and each edge curve  $\underline{\rho}^j(\zeta)$  by a channel surface\* (see Fig. below)

$$C(\underline{\rho}^j(\zeta), R) : (\underline{\rho}(\zeta) - \underline{\rho}^j(\zeta)) \cdot (\underline{\rho}(\zeta) - \underline{\rho}^j(\zeta)) - R^2 = 0$$

Further we set  $R \rightarrow 0$  and  $S(\underline{\rho}^i, R) \rightarrow S(\underline{\rho}^i, 0)$ ,  $C(\underline{\rho}^j(\zeta), R) \rightarrow C(\underline{\rho}^j(\zeta), 0)$  The derivative  $\frac{\partial \underline{\rho}}{\partial \zeta}$  always exists on the master surface, but on edges and on nodes the derivative is not unique.

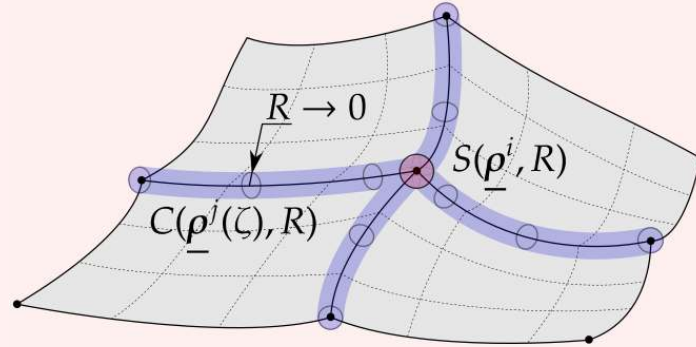


Fig. Improvement of the master surface with zero radius spheres on nodes and zero radius channel surfaces on edges in order to provide an everywhere differentiable surface.

\* – "A channel or canal surface is a surface formed as the envelope of a family of spheres whose centers lie on a space curve" [Wikipedia: [http://en.wikipedia.org/wiki/Channel\\_surface](http://en.wikipedia.org/wiki/Channel_surface)].

#### 5.1.5 Generalized Newton method

Even if the probability to meet a kink point (point where no classical derivative exists) during the iteration process is zero, the convergence of the Newton scheme is not anymore ensured by Kantorovich theorem. This is linked to the fact that the solution of the contact problem is itself a kink point. If the convergence is not monotonic the probability of convergence decreases. Moreover, for nonsmooth functional, it is sometimes not possible to fulfill the

condition of minimum  $\nabla f(x) = 0$  or even  $\|\nabla f(x)\| \leq \varepsilon$ . A simple example is a nonlinearity arising in Coulomb's friction law  $f(x) = |x|$ . The derivative of such a function is a kind of Heaviside function which takes the values  $-1$  and  $1$  and so  $\nabla f(x) < \varepsilon < 1$ . Some other possible problems: different results due to roundoff errors on different computers, approximated gradients, for example by perturbation method, may not belong to subdifferential derivative.

Since late 80s Alart [Alart 88], Alart and Curnier [Curnier 88], [Alart 91], Heegaard and Curnier [Heegaard 93] investigated an application of Newton's method to nonsmooth problems arising in contact mechanics. Authors worked out a Generalized Newton Method (GNM) and investigated its convergence properties for coupled augmented Lagrangian method [Alart 97]. Among obtained results there are

- a good convergence observed for frictionless contact in case of small and large slip;
- conditions on convergence of the GNM for frictionless contact, however as asserted by the authors the conditions appear to be too strict and are not fulfilled in real problems;
- the maximal number of iterations for the convergence is estimated by  $2^{n-1}$ , where  $n$  is number of dofs subjected to contact-type conditions, moreover at least one local status (e.g., non-contact – contact) changes without ever switch to the previous status.
- the absence of simple infinite cycling of the GNM around solution for frictionless problem, however this remains the main cause of divergence in case of frictional contact;
- a very important result is the upper estimation for penalty coefficient in augmented Lagrangian method applied to frictional problems, which avoids infinite cycling

$$0 < \varepsilon_t < 2\lambda_{\min}(K),$$

where  $\varepsilon_t$  is the penalty factor related to tangential slip and  $K$  is a stiffness matrix corresponding to the problem without contact; here the introduction of two penalty parameters, for normal  $\varepsilon_n$  and tangential  $\varepsilon_t$  contact, becomes justified.

Following the cited authors let us derive the main principles. First of all the nonlinear vector equation  $R(x, \lambda) = 0$  arising from application of the augmented Lagrangian method to contact problems is splitted into a differentiable structural part  $R^d(x)$  (due to virtual work of contact-free system) and a nondifferentiable contact part  $R^n(x, \lambda)$  (virtual work due to contact) parts

$$R(x, \lambda) = R^d(x) + R^n(x, \lambda) = 0$$

Further, the Generalized Newton method is stated as

$$\begin{bmatrix} \Delta x^{i+1} \\ \Delta \lambda^{i+1} \end{bmatrix} = \begin{bmatrix} \frac{\partial R^d(x)}{\partial x} + \nabla_x R^n(x, \lambda) \\ \nabla_\lambda R^n(x, \lambda) \end{bmatrix}^{-1} \begin{bmatrix} R^d(x^i) + R^n(x^i, \lambda^i) \end{bmatrix},$$

where  $i, i + 1$  denotes iteration numbers. The subgradients  $\nabla_x R^n$  and  $\nabla_\lambda R^n$  are components of the generalized Jacobians (here also, Hessians)  $\partial_\lambda R^n$  and  $\partial_\lambda R^n$  for primal and dual variables respectively:

$$\nabla_x R^n(x, \lambda) \in \partial_\lambda R^n(x, \lambda), \quad \nabla_\lambda R^n(x, \lambda) \in \partial_\lambda R^n(x, \lambda).$$

In other words, these notions are a generalization of the subdifferentials to the class of vector functions  $f(x)$  of vector argument  $x$ : each  $i$ -th line of the generalized Jacobian is the subdifferential of the scalar component  $f_i(x)$

$$\partial f(x) = [\partial f_1(x) \quad \dots \quad \partial f_i(x) \quad \dots \quad \partial f_n(x)]^T$$

The condition providing the convergence of the augmented Lagrangian method for frictional problem has been formulated as "sufficient stability conjecture" in [Alart 91] and investigated in more details [Alart 97]

$$0 < \varepsilon_t < 2\lambda_{\min}(\nabla_x R^n(x, \lambda))$$

A more sophisticated study of this condition can be found in theorem 6 in [Alart 97]. Let us give a short demonstration of the analysis performed by the author. Let us consider a scalar nonconvex and nondifferentiable function  $f(x)$  (Fig. 5.2). If this function is such that close to the solution its maximal derivative is  $\sup |x - x^*| < \varepsilon f' = a$ , to the right from the solution  $x > x^*$  there are points where the smallest derivative  $f'(x) = b < a$  and to the left  $\inf_{x < x^*} (f'(x)) = c < a$ , then there is a risk that Newton's method does not converge, i.e. start to cycle over two branches (Fig. 5.2,a.). Moreover, if the left branch is convex and the right branch is concave, each iteration will move away from the solution. A similar function arises from frictional contact problems even the simplest one with one degree of freedom and prescribed contact pressure. However, Newton's method does not always diverge (Fig. 5.2,b.). The question is then to check what is the condition on  $a, b, c$  which ensures the convergence of the method. P. Alart proved that Newton or Generalized Newton Methods converge if the maximal derivative of the function  $f(x)$  is not bigger than twice the minimal derivative of the function  $f(x)$  at least in a vicinity of possible iterations of the Newton's method ( $\mathbb{N}(x)$ ):

$$\sup f'(x) < 2 \inf f'(x), x \in \mathbb{N}(x).$$

In case of a nondifferentiable function, the ordinary derivatives  $f'$  are replaced by subdifferentials  $\partial f'$

$$\sup \partial f(x) < 2 \inf \partial f(x), x \in \mathbb{N}(x).$$

Further, the author generalizes this result to the  $n$ -dimensional case by comparing with convergence conditions of Uzawa's algorithm.

The onedimensional problem: which function of considered type  $a > b, a > c$  will converge can be formulated as a geometrical problem (Fig. 5.3): Newton's method converge if for three straight lines

$$L_1 : y = bx + d, \quad L_2 : y = cx + e, \quad L_3 : y = ax + g, \quad a > b, a > c$$

## 5.1 Newton's method

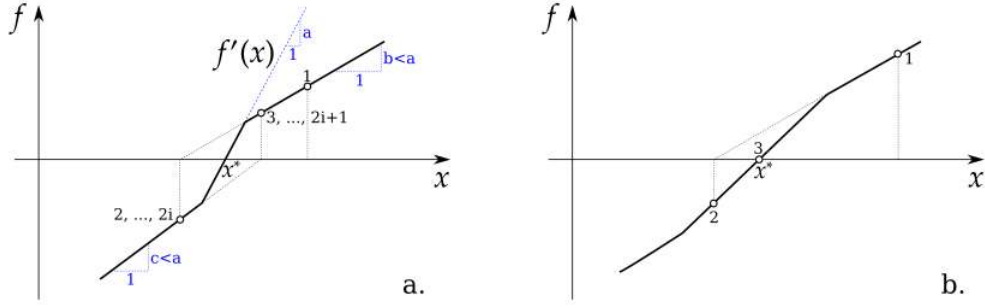


Figure 5.2: Nonconvex nondifferentiable function  $f(x)$  for which Newton's method or GNM may not converge: **a.** – methods do not converge if the starting point is located on the left or the right branch with a smaller slope, that the slope close to the solution point  $x^*$ , **b.** – methods always converge .

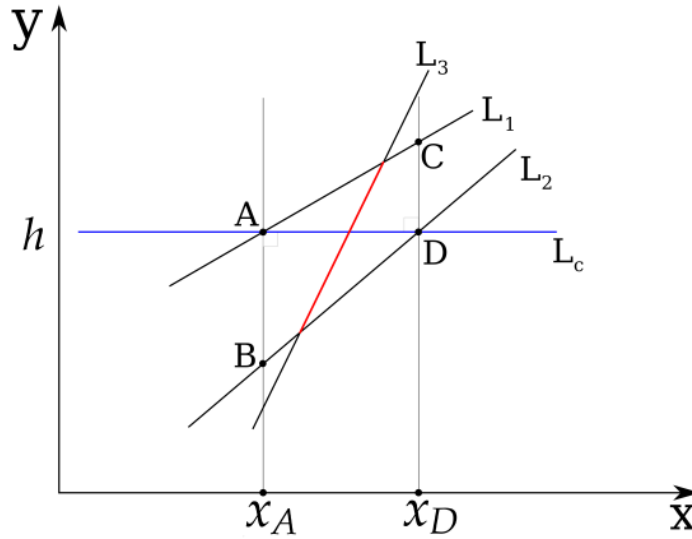


Figure 5.3: Geometrical problem arising from convergence conditions of Newton's method .

there is no horizontal line  $L_c : y = h$ , such that part of  $L_3$  enclosed between  $L_1$  and  $L_2$  lies entirely in quadrangle  $ABCD$ , where  $A = (x_A, h)$  is the intersection of  $L_1$  and  $L_c$ ,  $D = (x_D, h)$  is the intersection of  $L_2$  and  $L_c$  and  $B = L_2(x_A)$ ,  $C = L_1(x_D)$ .

## 5.2 Return mapping algorithm

The return mapping algorithm (in onedimensional case) or radial return mapping algorithm originally proposed in [Wilkins 64], is a well known scheme employed in computational elasto-plasticity [Simo 98]. Since there is a direct analogy between plasticity and friction as stated in Section 4.1.2, this method can be successfully applied to the local resolution of frictional conditions. The idea of the method in two words: elastic problem is solved, for a given strain increment, and a trial stress is evaluated. If the new stress is situated inside the yield surface, then this is the solution. Otherwise the slip rate increment is changed to return to the yield surface and the stress is updated. The application of the return mapping algorithm for frictional contact integration for the penalty method can be found in [Giannakopoulos 89], [Wriggers 90], [Simo 92]<sup>1</sup> and for the augmented Lagrangian method with Uzawa's algorithm in [Simo 92]. For the sake of completeness, the return mapping algorithm is presented below.

The contribution of the contact to virtual work principle was stated in the previous chapter as

$$I_c(\underline{u}, \delta \underline{u}) = \int_{\bar{\Gamma}_c^*} \sigma_n (\delta g_n - \mu \underline{s} \cdot \delta \underline{g}_t) d\bar{\Gamma}_c^1 + \int_{\bar{\Gamma}_c^\bullet} (\sigma_n \delta g_n + \underline{\sigma}_t \cdot \delta \underline{g}_t) d\bar{\Gamma}_c^1, \quad (5.6)$$

where  $\bar{\Gamma}_c^*$ ,  $\bar{\Gamma}_c^\bullet$  are slip and stick zones in active contact zones respectively,  $\sigma_n$  the contact pressure,  $\underline{\sigma}_t$  the tangential contact stress vector,  $\delta g_n$  the variation of the normal gap,  $\delta \underline{g}_t$  the variation of the tangential relative sliding. For a linear penalty method, contact pressure and tangential stress vector  $\sigma_n$ ,  $\underline{\sigma}_t$  are linear functions of the normal gap and of the relative sliding  $g_n$ ,  $\Delta \underline{g}_t$

$$\sigma_n = -\varepsilon_n \langle -g_n \rangle, \quad \underline{\sigma}_t = \begin{cases} \varepsilon_t \|\Delta \underline{g}_t\| \underline{s}, & \varepsilon_t \|\Delta \underline{g}_t\| \leq \mu |\sigma_n| \\ \mu |\sigma_n| \underline{s}, & \varepsilon_t \|\Delta \underline{g}_t\| > \mu |\sigma_n| \end{cases},$$

where  $\varepsilon_n, \varepsilon_t > 0$  are the penalty parameters,  $\underline{s}$  the slip direction. Substituting contact pressure and tangential stress vector into Eq. (5.6), we get

$$I_c(\underline{u}, \delta \underline{u}) = - \int_{\bar{\Gamma}_c^*} \varepsilon_n \langle -g_n \rangle (\delta g_n - \mu \underline{s} \cdot \delta \underline{g}_t) d\bar{\Gamma}_c^1 + \int_{\bar{\Gamma}_c^\bullet} (-\varepsilon_n \langle -g_n \rangle \delta g_n + \varepsilon_t \|\Delta \underline{g}_t\| \underline{s} \cdot \delta \underline{g}_t) d\bar{\Gamma}_c^1.$$

The constitutive equation for friction were given in (4.21) as

$$\|\underline{\sigma}_t\| \leq \mu |\sigma_n|, \quad \|\dot{\underline{g}}_t\| \|\underline{\sigma}_t - \mu |\sigma_n| \dot{\underline{g}}_t\| = 0, \quad \|\underline{s}\| \left\| \|\dot{\underline{g}}_t\| \underline{\sigma}_t - \mu |\sigma_n| \dot{\underline{g}}_t \right\| = 0. \quad (5.7)$$

They can be reformulated in terms of elasto-plasticity: the slip surface  $f(\sigma_n, \underline{\sigma}_t)$  is given by

$$f(\sigma_n, \underline{\sigma}_t) = \|\underline{\sigma}_t\| - \mu |\sigma_n| \leq 0$$

<sup>1</sup>Note that there is a small misprint in Eq. (3.13) in the article [Simo 92], where the penalty return mapping algorithm is stated the correct equation is given below.

## 5.2 Return mapping algorithm

---

The slip rule is

$$\underline{\dot{g}}_t = \dot{\gamma} \frac{\partial f}{\partial \underline{\sigma}_t},$$

where  $\dot{\gamma}$  is a slip rate and  $\frac{\partial f}{\partial \underline{\sigma}_t}$  determines the normal to the slip surface and also the slip direction, which for Coulomb's friction is a radial unit vector

$$\frac{\partial f}{\partial \underline{\sigma}_t} = \frac{\underline{\sigma}_t}{\|\underline{\sigma}_t\|} = \underline{s}_t$$

The numerical resolution of the contact problems is incremental, so let us suppose that on the  $i$ -th increment we know the solution  $\underline{u}^i$  and all corresponding quantities  $\sigma_n^i, \underline{\sigma}_t^i$ . On the first iteration of the next increment, we suppose that the entire active contact zone switches to stick

$$I_c(\underline{u}^{i+1}, \delta \underline{u}) = \int_{\bar{\Gamma}_t^1} (\sigma_n^{i+1} \delta g_n + \underline{\sigma}_t^{i+1} \cdot \delta \underline{g}_t) d\bar{\Gamma}_c^1,$$

where

$$\sigma_n^{i+1} = -\varepsilon_n \langle -g_n(\underline{u}^{i+1}) \rangle.$$

Further, by analogy with plasticity, we determine the trial stick tangential stress vector

$$\underline{\sigma}_t^{i+1} \text{ trial} = \underline{\sigma}_t^i + \varepsilon_t (\underline{g}_t^{i+1} - \underline{g}_t^i) = \underline{\sigma}_t^i + \varepsilon_t \Delta \underline{g}_t^i$$

A graphical interpretation of the return mapping algorithm is presented in Fig. 5.4, 5.5 for two dimensional and in Fig. 5.6 for three dimensional cases. Now we check if this stress vector does not exceed the permissible stress, i.e. is not outside of the Coulomb's slip surface for the new contact pressure  $\sigma_n^{i+1}$

$$f_{\text{trial}}^{i+1} = \|\underline{\sigma}_t^{i+1} \text{ trial}\| - \mu |\sigma_n^{i+1}|$$

If  $f_{\text{trial}}^{i+1} \leq 0$ , then the trial tangential stress is a correct one

$$\underline{\sigma}_t^{i+1} = \underline{\sigma}_t^{i+1} \text{ trial}, \quad \text{if } f_{\text{trial}}^{i+1} \leq 0$$

otherwise if  $f_{\text{trial}}^{i+1} > 0$ , the current point switches to slip state and a tangential stress is put

$$\underline{\sigma}_t^{i+1} = \underline{\sigma}_t^{i+1} \text{ trial} - \varepsilon_t \Delta \gamma^i \underline{s}_t^i, \quad \text{if } f_{\text{trial}}^{i+1} > 0$$

where  $\underline{s}_t^i$  is the slip direction determined by the trial stress and  $\Delta \gamma$  is the slip increment, determining the slip distance

$$\underline{s}_t^i = \frac{\underline{\sigma}_t^{i+1} \text{ trial}}{\|\underline{\sigma}_t^{i+1} \text{ trial}\|}$$

$$\Delta \gamma^i = \frac{f_{\text{trial}}^{i+1}}{\varepsilon_t}$$



Both results  $\forall f_{\text{trial}}^{i+1}$ , can be stated in one formula using Macaulay brackets

$$\underline{\sigma}_t^{i+1} = \underline{\sigma}_t^{i+1}{}_{\text{trial}} - \varepsilon_t \langle \Delta \gamma^i \rangle \underline{s}_t^i \quad (5.8)$$

Note that in the two dimensional case, the sliding vector  $\underline{s}^i$  and the stress unit vector  $\underline{s}_t^i$  turn simply into  $\text{sign}(\underline{g}_t)$  and  $\text{sign}(\underline{\sigma}_t)$  respectively, obviously  $\text{sign}(\underline{g}_t) = \text{sign}(\underline{\sigma}_t)$ .

As discussed in the previous chapter, the total slip can be splitted into a sum of “slip-in-stick”  $\underline{g}_t^\bullet$  and real slip  $\underline{g}_t^*$

$$\underline{g}_t = \underline{g}_t^\bullet + \underline{g}_t^*$$

This split becomes more clear in the incremental procedure. Let us briefly demonstrate by simple algebraic calculations and geometrical schemes the meaning of such a split in the frame of the return mapping algorithm. On each increment the relative sliding vector splits into an accumulated slip over the previous increments  $\underline{g}_t^{*i}$  and the total “slip-in-stick” relatively to this point  $\underline{g}_t^{\bullet i}$

$$\underline{g}_t^i = \underline{g}_t^{\bullet i} + \underline{g}_t^{*i}, \quad \underline{g}_t^{i+1} = \underline{g}_t^{\bullet i+1} + \underline{g}_t^{*i+1} \quad (5.9)$$

Further, the position vector on the increment  $i + 1$ , can be expressed as the sum of the vectors on the previous increment, of slip-in-stick  $\Delta \underline{g}_t^{\bullet i}$  and of slip increments  $\Delta \underline{g}_t^{*i}$

$$\underline{g}_t^{i+1} = \underline{g}_t^i + \Delta \underline{g}_t^{*i} + \Delta \underline{g}_t^{\bullet i}, \quad \text{or simply} \quad \Delta \underline{g}_t^i = \Delta \underline{g}_t^{*i} + \Delta \underline{g}_t^{\bullet i} \quad (5.10)$$

where

$$\Delta \underline{g}_t^{\bullet i} = \underline{g}_t^{\bullet i+1} - \underline{g}_t^{\bullet i}, \quad \Delta \underline{g}_t^{*i} = \underline{g}_t^{*i+1} - \underline{g}_t^{*i}, \quad \Delta \underline{g}_t^i = \underline{g}_t^{i+1} - \underline{g}_t^i$$

In other words, it follows from Eq. (5.10) that  $\underline{g}_t^i + \Delta \underline{g}_t^{*i}$  is the new position of stick, where the point will return if external load vanishes. However, Eq. (5.10) is not easy to interpret, because slip occurs relatively to the point of previous stick  $\underline{g}_t^{*i}$  and the direction of the total sliding increment  $\underline{\tilde{s}}^i$  does not make sense

$$\underline{\tilde{s}}^i = \frac{\Delta \underline{g}_t^i}{\|\Delta \underline{g}_t^i\|},$$

since this is the direction of the sum of “slip-in-stick” and slip vectors, but not of an occurred slip. All directions in general are different<sup>2</sup>, i.e.

$$\Delta \underline{g}_t^{*i} \nparallel \Delta \underline{g}_t^{\bullet i} \nparallel \Delta \underline{g}_t^i = \underline{\tilde{s}}^i \|\Delta \underline{g}_t^i\|$$

<sup>2</sup>All these arguments relate to three dimensional contact. In two dimensional problems, where all vectors at interface can be interpreted simply as signed real numbers  $\text{sign}(\underline{\tilde{s}}^i) = \text{sign}(\underline{s}^i) = \text{sign}(\underline{s}_t^i)$ , the present discussion is not relevant.

## 5.2 Return mapping algorithm

On the other hand if  $\underline{g}_t^i$  is splitted into “stick-in-slip” and accumulated slip parts, then Eq. (5.10) can be rewritten as

$$\underline{g}_t^{i+1} = \underline{g}_t^{*i} + \Delta \underline{g}_t^{*i} + \underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{\bullet i}, \quad (5.11)$$

where

$$\{\underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{\bullet i}\} \parallel \Delta \underline{g}_t^{*i} \parallel \{\underline{g}_t^{i+1} - \underline{g}_t^{*i}\},$$

i.e. all these vectors can be rewritten as products of norms and sliding vector  $\underline{s}^i$

$$\begin{aligned} \underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{\bullet i} &= \|\underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{\bullet i}\| \underline{s}^i, \\ \underline{g}_t^{i+1} - \underline{g}_t^{*i} &= \|\underline{g}_t^{i+1} - \underline{g}_t^{*i}\| \underline{s}^i, \\ \Delta \underline{g}_t^{*i} &= \|\Delta \underline{g}_t^{*i}\| \underline{s} = \Delta \gamma \underline{s}^i \end{aligned} \quad (5.12)$$

Eq. (5.11) can be stated in an algebraic form

$$\Delta \gamma = \|\underline{g}_t^{i+1} - \underline{g}_t^{*i}\| - \|\underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{\bullet i}\|$$

In other words slip increment is the difference between total slip from the point of the previous stick  $\|\underline{g}_t^{i+1} - \underline{g}_t^{*i}\|$  and final “slip-in-stick”  $\|\underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{\bullet i}\|$ .

Note that independently on the direction of the increment of tangential stress  $\underline{\sigma}_t^{i+1} - \underline{\sigma}_t^i$  or tangential slip  $\underline{g}_t^{i+1} - \underline{g}_t^i$ , the slip is supposed to occur in the direction of the resultant trial stress  $\underline{\sigma}_t^{i+1}$  trial. For an anisotropic friction law, the slip direction depends upon the choice of the model: associated and non-associated slip can be distinguished (see [Michalowski 78]).

Remark that in large deformation, large sliding problems a rigorous definition must be used for the relative sliding velocity expressed through the convective coordinates  $\underline{\xi}$ , as discussed in paragraph 2.2.4

$$\underline{\dot{g}}_t = \frac{\partial \underline{g}}{\partial \underline{\xi}} \circ \underline{\dot{\xi}},$$

here  $\underline{\dot{g}}_t$  denotes Lie’s derivative of the relative tangential displacement vector. For a detailed analysis of such a formulation the reader is referred to works by Laursen and Simo [Laursen 93], Laursen [Laursen 94] and recent articles by Konyukhov and Schweizerhof [Konyukhov 05], [Konyukhov 06b], [Konyukhov 07a].

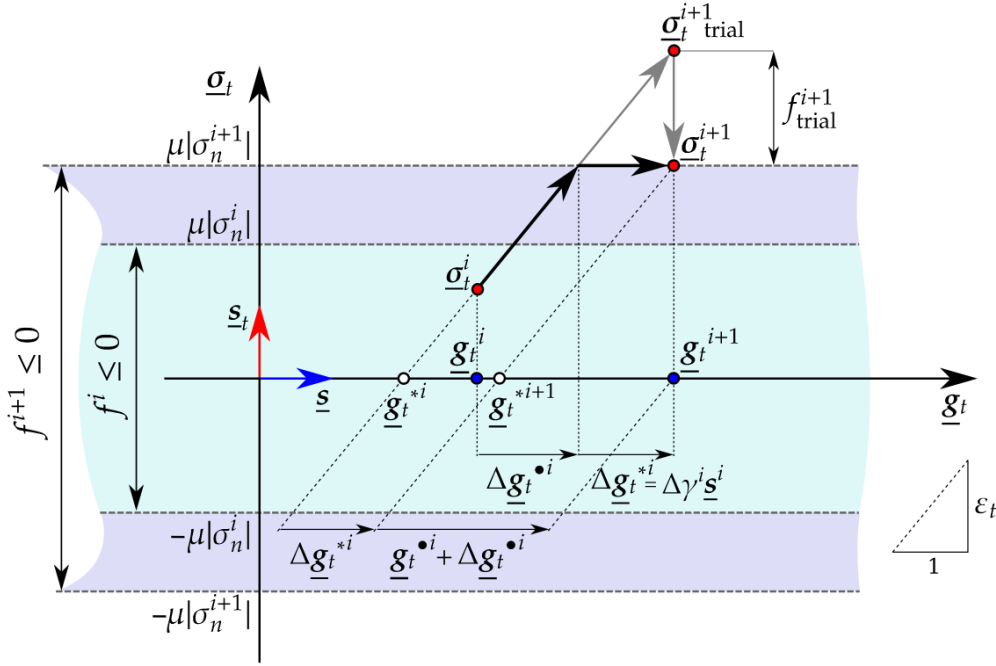


Figure 5.4: Representation of the return mapping algorithm applied to a two dimensional contact problem. A sequential set of step points is represented by red contoured circles: initial state  $\{\underline{g}_t^i, \underline{\sigma}_t^i\}$ , trial step  $\{\underline{g}_t^{i+1}, \underline{\sigma}_t^{i+1 \text{ trial}}\}$ , return mapping on the slip surface gives the solution point  $\{\underline{g}_t^{i+1}, \underline{\sigma}_t^{i+1 \text{ trial}} - \epsilon_t \Delta \gamma \underline{s}_t^i\}$ ; the initial center of stick or previously accumulated slip is  $\underline{g}_t^{*i}$ , and “slip-in-stick” on the  $i$ -th increment is given by  $\underline{g}_t^{\bullet i} = \underline{g}_t^i - \underline{g}_t^{*i}$ , the total slip during increment is  $\underline{g}_t^{i+1} - \underline{g}_t^i$ , which is the sum of “slip-in-stick” increment  $\Delta \underline{g}_t^{\bullet i}$  and real slip, accumulated during increment,  $\Delta \underline{g}_t^{*i}$ , final solution vector  $\underline{g}_t^{i+1}$  consists of accumulated slip  $\underline{g}_t^{*i+1} = \underline{g}_t^{*i} + \Delta \underline{g}_t^{*i}$  and actual “slip-in-stick”  $\underline{g}_t^{\bullet i+1} = \underline{g}_t^{\bullet i} + \Delta \underline{g}_t^{\bullet i}$ .

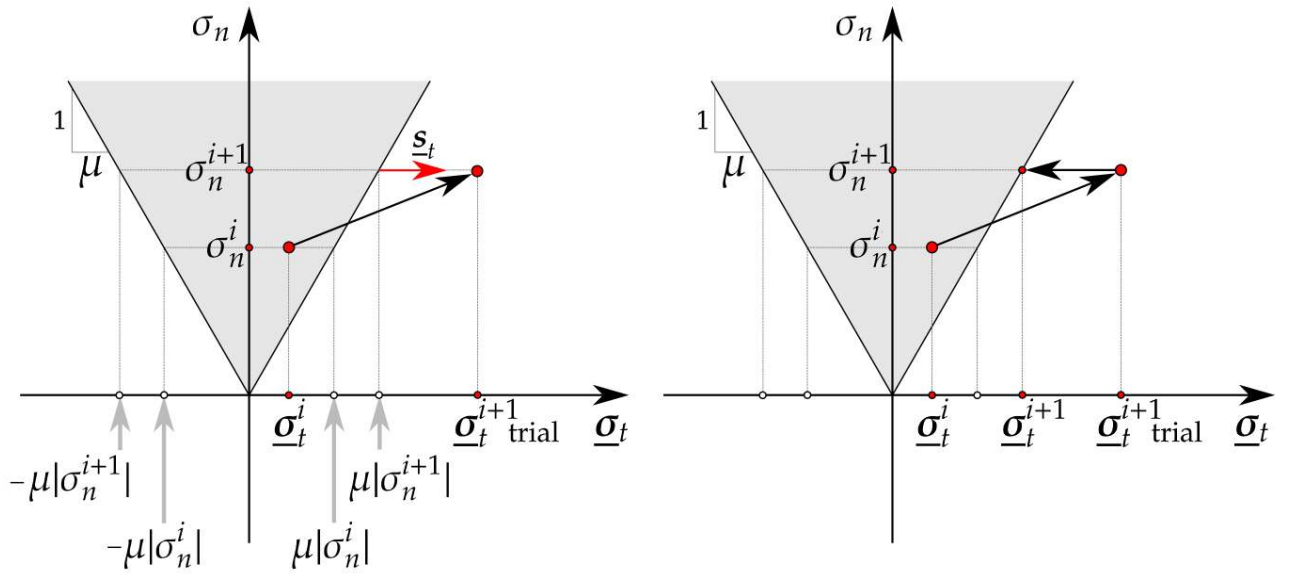


Figure 5.5: Representation of the return mapping algorithm for a two dimensional problem in stress space. A sequential set of step points is represented by red contoured circles: initial state  $\{\underline{\sigma}_t^i, \sigma_n^i\}$ , trial step  $\{\underline{\sigma}_t^{i+1} \text{ trial}, \sigma_n^{i+1}\}$  and non-associated return mapping gives the solution state  $\{\underline{\sigma}_t^{i+1}, \sigma_n^{i+1}\}$ .

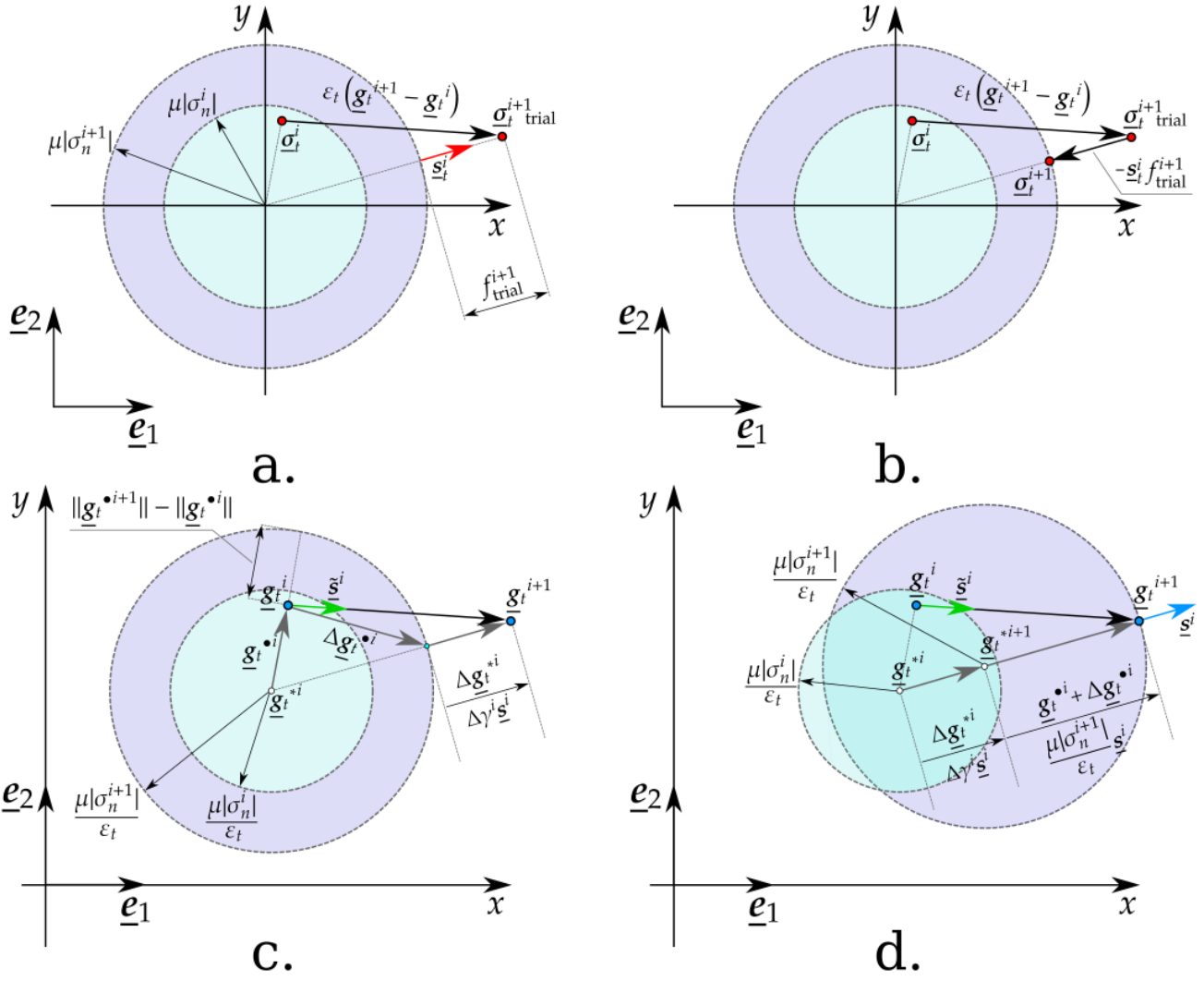


Figure 5.6: Representation of the return mapping algorithm for three dimensional contact problem,  $\{e_1, e_2\}$  orthonormal basis with coordinates  $\{x, y\}$ , figures **a.,b.** represent stress state in contact interface, figures **c.,d.** represent displacements in contact interface, sequential set of step points is represented by red contoured circles for stress (**a.,b.**) and in blue for displacement (**c.,d.**): initial state  $\{\underline{g}_t^i, \underline{\sigma}_t^i\}$ , trial step  $\{\underline{g}_t^{i+1}, \underline{\sigma}_t^{i+1, \text{trial}}\}$ , return mapping or radial return mapping on the slip surface (**b.**) gives the solution point for stress and moving of slip circle in direction  $\underline{s}^i$  by value of real slip  $\Delta\gamma^i$  gives the solution state for displacements (**d.**)  $\{\underline{g}_t^{i+1}, \underline{\sigma}_t^{i+1, \text{trial}} - \varepsilon_t \Delta\gamma^i \underline{s}_t^i\}$ ; increment of position vector  $\Delta\underline{g}_t^i = \|\underline{g}_t^{i+1} - \underline{g}_t^i\| \underline{\tilde{s}}^i$  can be presented as the sum of differently oriented vectors  $\Delta\underline{g}_t^{*i}$  and  $\Delta\underline{g}_t^{\bullet i}$  corresponding to an increment of real slip and an increment of “slip-in-stick” respectively; however as discussed above, it is more natural to consider a displacement increment starting from stick point  $\underline{g}_t^{*i}$  on the  $i$ -th increment, then  $\underline{g}_t^{*i} + \Delta\underline{g}_t^i = \Delta\underline{g}_t^{*i} + \underline{g}_t^{*i} + \Delta\underline{g}_t^{\bullet i}$ , where  $\Delta\underline{g}_t^{*i} = \Delta\gamma^i \underline{s}^i$  – real slip increment and actual “slip-in-stick” is presented by  $\underline{g}_t^{*i} + \Delta\underline{g}_t^{\bullet i} = \frac{\mu|\sigma_n^{i+1}|}{\varepsilon_t} \underline{s}^i$ , for Coulomb’s friction  $\underline{s}^i = \underline{s}_t^i \neq \underline{\tilde{s}}^i$ .

## 5.3 Finite Element Method

In this section the main notions of the Finite Element Method will be given: mesh, elements, nodes, basis functions. Non-structural contact elements will be introduced and the evaluation of contact integrals will be discussed.

### 5.3.1 Introduction

The *Finite element method* is a powerful and widely used approach for the numerical resolution of boundary value problems

$$\mathfrak{F}(f(\underline{u})) = 0 \quad + \quad \text{boundary and initial conditions}$$

where  $\mathfrak{F}(f)$  is a combination of differential operators applied to a tensor function  $f(\underline{u})$  of vector argument  $\underline{u}$  (in structural mechanical problem). The method is applicable for arbitrary geometries as well as for linear and nonlinear constitutive equations. The idea is to replace a continuous problem (i.e. infinitely dimensional) by a finite dimensional problem. The method belongs to the class of Bubnov-Galerkin methods, i.e. the solution of the problem  $\underline{u}^*$  is approximated by a decomposition over a finite number of basis function  $\phi_i(\underline{X}_i)$ ,  $i = 1, N$

$$\underline{u}^*(t, \underline{X}) \approx \underline{u}^h(t, \underline{X}) = \sum_{i=1}^N \underline{u}_i^h(t) \phi_i(\underline{X}), \quad (5.13)$$

where  $t$  is the time,  $\underline{X}$  denotes a material point vector in reference configuration,  $\underline{u}_i^h(t)$  is a coefficient at the  $i$ -th basis function. For structural problems, this coefficient is a vector. The basis functions can be chosen in such a way that  $\underline{u}_i^h(t) \phi_i(\underline{X}_i) = \underline{u}_i^h(t)$  is the solution value at the  $i$ -th *node* – material point  $\underline{X}_i$ . It implies that the  $i$ -th basis function is one in the associated node  $i$ , and zero in all other nodes  $j \neq i$ .

$$\phi_i(\underline{X}_j) = \delta_i^j.$$

The considered volume  $\Omega$  – body – at which differential equation  $\mathfrak{F}(f(\underline{u}))$  is prescribed is split into *finite elements*  $\Omega_i^h$ ,  $i = 1, N^e$  spanned onto *nodes*  $\underline{X}_i^j$ , which all together form a *finite element mesh* – a discretized representation of the body  $\Omega^h = \bigcup_{i=1, N^e} \Omega_i^h \sim \Omega$  (Fig. 5.7). If the body  $\Omega$  is continuous and non-self-intersecting, then for **almost** all points of the body  $\underline{X}$  there is a discrete form

$$\underline{X} = \sum_{j=1}^M \underline{X}_j \psi_j(\underline{X}), \quad (5.14)$$

where  $\psi_j(\underline{X})$  is another set of basis functions describing the geometry and  $N^n$  number of such functions and number of nodes. Note that for a finite number of basis functions some of the points of the continuous body  $\Omega$  have no homologue in the discretized geometry and vice versa, i.e. in general

$$\Omega \setminus \Omega^h \neq \emptyset, \quad \Omega^h \setminus \Omega \neq \emptyset,$$

This difference appears only close to the boundary, which is essential for accurately imposing boundary conditions and especially for contact treatment: in general, the real surface of the body is different from the surface of the finite element mesh

$$\partial\Omega \neq \partial\Omega^h = \bigcup_{i=1, N^s} \partial\Omega_i^h$$

Another observation: in general, the geometrical basis functions (shape functions) are compact, i.e. are zero outside the element (see Fig. 5.8, 5.9), for example, for element  $i$  spanned on  $N_i^n$  nodes, the shape functions  $\phi_i^j(\underline{\mathbf{X}})$ ,  $j = 1, N_i^n$  are defined as

$$\phi_i^j(\underline{\mathbf{X}}) = 0, \underline{\mathbf{x}} \notin \Omega_i^h \quad \text{and} \quad \phi_i^j(\underline{\mathbf{X}}_k) = \delta_k^j, \underline{\mathbf{X}}_k \in \Omega_i^h$$

Shape functions are normally infinitely differentiable  $\phi_i \in C^\infty$ , but since they are compact on the element, on the interface between elements any smooth function approximated by shape functions is in general not differentiable, i.e.

$$\nabla\phi_i^j(\underline{\mathbf{X}}_j) \neq \nabla\phi_k^j(\underline{\mathbf{X}}_j),$$

where  $\underline{\mathbf{X}}_j$  is a common node of elements  $i$  and  $k$ .

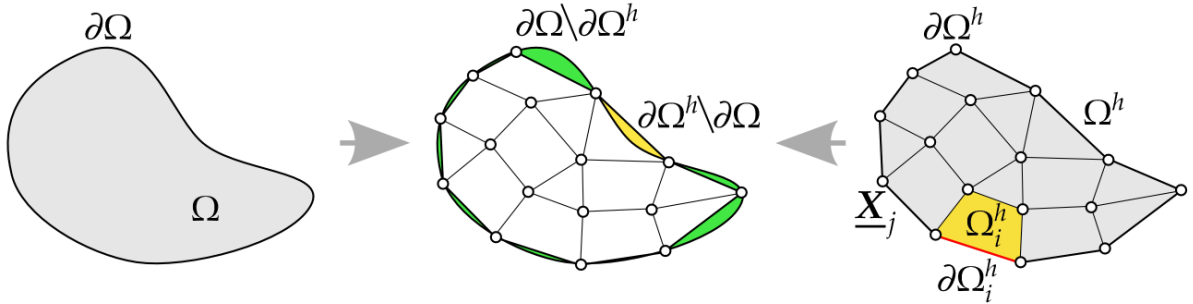


Figure 5.7: Continuous body  $\Omega$  and its discretized representation – finite element mesh  $\Omega^h$  consisting of nodes  $\underline{\mathbf{X}}_j$  and elements  $\Omega_i^h$ . Surface of the body  $\partial\Omega$ , approximated by surface  $\partial\Omega^h = \bigcup_i \partial\Omega_i^h$ .

Often the set of geometrical basis functions  $\psi_j(\underline{\mathbf{X}})$  is chosen the same as the set of basis functions approximating the solution  $\phi_i(\underline{\mathbf{X}})$  (isoparametric approximation of the problem). Further, all basis functions will be called shape functions and denote with  $\phi$

$$\underline{\mathbf{X}} = \sum_{i=1}^N \underline{\mathbf{X}}_i \phi_i(\underline{\mathbf{X}}),$$

An isoparametric choice of the basis functions is rather natural for structural problems, if we seek for the field of displacement vectors, then the actual

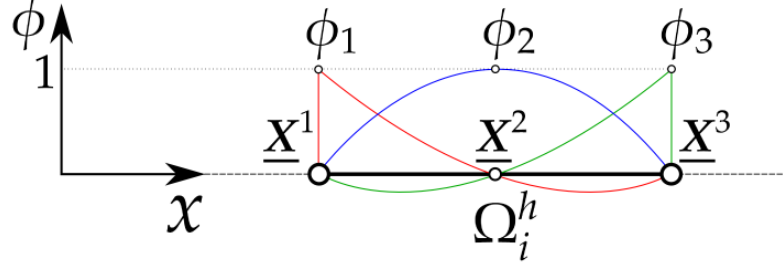


Figure 5.8: Example of compact shape functions  $\phi_i(\underline{X})$  for the one-dimensional 3-node quadratic element  $\Omega_i^h$ .

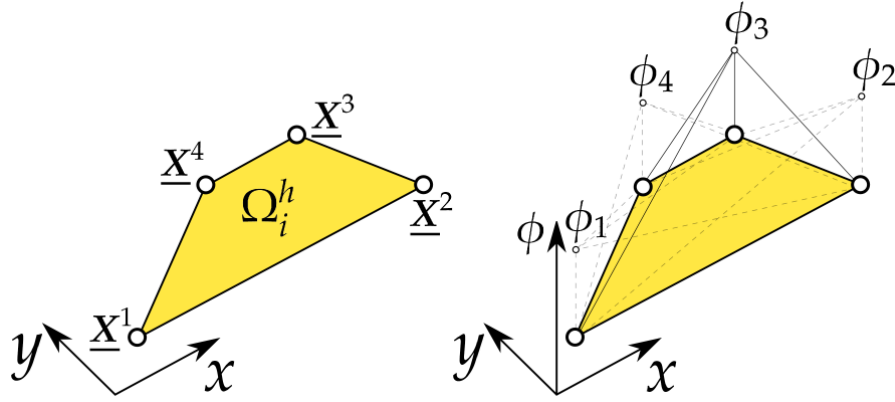


Figure 5.9: Example of compact shape functions  $\phi_i(\underline{X})$  for two dimensional 4-node linear element  $\Omega_i^h$ .

configuration of **almost** each point can be presented as

$$\underline{x} = \underline{X} + \underline{u} = \sum_{i=1}^N (\underline{X}_i + \underline{u}_i) \phi_i(\underline{X})$$

The index  $h$  often stands to demonstrate that the current decomposition by basis functions is finite dimensional, and  $h$  represents a maximal size of *finite elements* into which geometry is split, then the solution of the discretized problem tends to the precise solution if the maximal size of the finite elements  $h$  tends to zero, or equally if the number of basis functions  $N$  tends to infinity

$$\underline{u}^h \xrightarrow{h \rightarrow 0} \underline{u}^*,$$

Since further we will deal only with discretized quantities, the  $h$  index will be omitted.

It is often convenient to determine the shape functions independently of the reference coordinates  $\underline{X} \in \Omega$ . For this purpose, for each type of element, a standard simple reference configuration – the parent space  $\Omega$  – is defined



(Fig. 5.10) and then the approximation of the geometry/solution within the  $i$ -th finite element can be reformulated as

$$\forall \underline{X} \in \Omega_i, \quad \xi, \eta \in \widetilde{\Omega}_i, \quad \underline{u}(t, \underline{X}) = \sum_{j=1, N_i^n} \underline{u}_j^i(t) \phi_j^i(\xi, \eta),$$

where  $\underline{u}_j^i$  is the  $j$ -th node of the  $i$ -th element and  $\phi_j^i$  is the  $j$ -th shape function of the  $i$ -th element and  $\{\xi, \eta\}$  are *convective* coordinates of the point  $\underline{X}$  in the parent space  $\widetilde{\Omega}_i$ . For contact problems, it is particularly important to consider the mapping from the parent configuration to the actual configuration for the contact surface  $\Gamma_c$  – a part of the body  $\partial\Omega = \cup_i \partial\Omega_i$ . The closure of two-dimensional simply connected bodies is a curve and of three-dimensional bodies is a surface, so the dimension of the topology is one less compared to the dimension of the problem. In Chapter 2 we considered the mapping from a two dimensional parent space into a three dimensional vector space (3 dimensional tensor space of first order)

$$\mathbb{R}^2 \ni \{\xi, \eta\} \rightarrow \underline{\rho}(\xi, \eta) \in \mathbb{T}_1^3$$

and also the mapping from one dimensional parent space to a two dimensional vector space

$$\mathbb{R} \ni \{\zeta\} \rightarrow \underline{\rho}(\zeta) \in \mathbb{T}_1^2$$

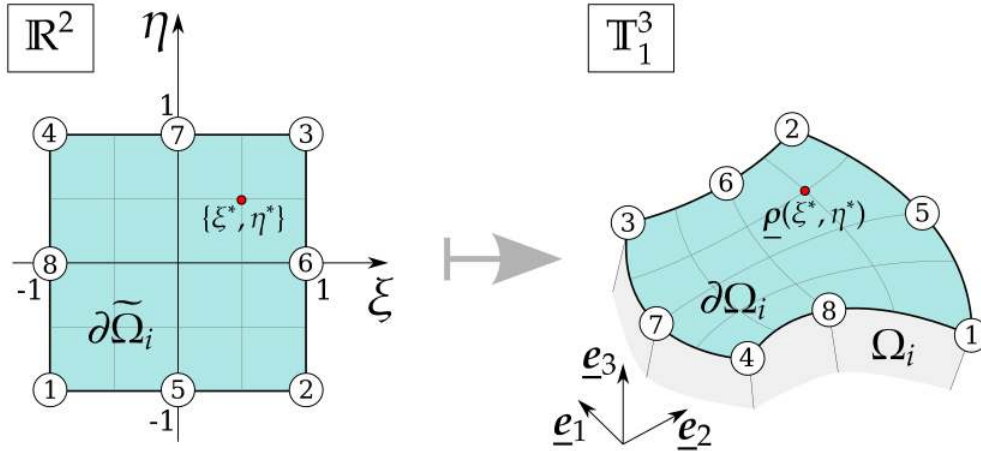


Figure 5.10: Surface segment of element: mapping from two dimensional reference space  $\mathbb{R}^2 \ni \{\xi, \eta\}$  to three dimensional vector space  $\underline{\rho}(\xi, \eta) \in \mathbb{T}_1^3$ .

In Chapter 2, we made use of the s-structures introduced in Appendix A.6. In the new formalism, instead of  $\mathbb{R}^2$  parent space, it is considered as a two dimensional v-scalar space  ${}^2\mathbb{S}_0^3$  over scalars of three dimensional space or simply as a two dimensional vector space  $\mathbb{T}_1^2$ . The mapping we need (Fig. 5.12) becomes then:

$${}^2\mathbb{S}_0^3 \ni \underline{\xi} \rightarrow \underline{\rho}(\underline{\xi}) \in \mathbb{T}_1^3$$

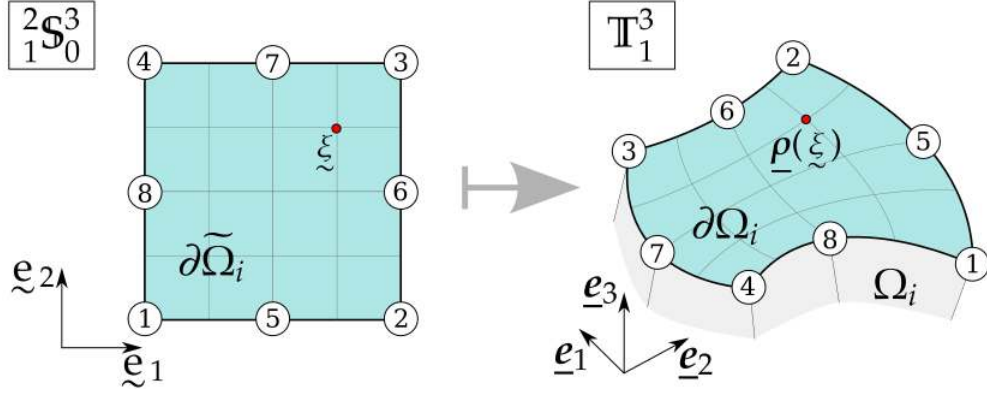


Figure 5.11: Surface segment of element: mapping from two dimensional v-scalar space  ${}^2\mathcal{S}_0^3$  to three dimensional vector space  $\mathbb{T}_1^3$ .

Moreover, the set of shape functions  $\phi_j^i(\xi)$ ,  $j = 1, N_i^n$  can be replaced by another v-scalar of dimension  $N_i^n$

$$\phi_j^i(\xi) \sim \phi^i(\xi) \in {}^N\mathcal{S}_0^3$$

and on the other hand the set of nodal vectors  $\underline{u}_j^i$ ,  $j = 1, N_i^n$  can be replaced by a v-vector

$$\underline{u}_j^i \sim \underline{u}^i \in {}^N\mathcal{S}_1^3$$

Then the finite element approximation of a vector field  $\underline{u}$  within the surface of the  $i$ -th element writes

$$\forall \underline{X} \in \partial\Omega_i, \quad \xi(\underline{X}) \in \partial\tilde{\Omega}_i, \quad \underline{u}(t, \underline{X}) = \underline{u}^i(t) \circ \phi^i(\xi)$$

Note that the dimensions of  $\xi \in {}^2\mathcal{S}_0^3$  and of  $\phi^i \in {}^N\mathcal{S}_0^3$ ,  $\underline{u}^i \in {}^N\mathcal{S}_1^3$  are different.

### 5.3.2 Contact elements

When two separated surfaces  $\Gamma_c^1$  and  $\Gamma_c^2$  come in contact from a continuum mechanics point of view, they form a continuous contact interface. Through this zone, contact stresses are transferred from one body to another according to a constitutive law due to the physical, mechanical and chemical phenomena occurring in the contact interface. From a numerical point of view, each contacting surface is represented by a number of nodes and surface or line segments – these components are not connected, so do not interact. In the previous chapter, contact surfaces and related contact stresses have been incorporated in the integral weak form – balance of virtual works; that implies that work of contact stresses has to be integrated for given virtual displacements within the contact interface. Due to the balance of forces in the contact interface, the integration can be performed over any of its sides.

The side over which the integrals will be evaluated is a master surface  $\Gamma_c^2$ . In the Finite Element Method, it is presented as a set of master segments

$$\Gamma_c^2 = \bigcup_{i=1}^M \partial\Omega_i$$

Another side of the contact interface is the slave surface, which in the node-to-segment approach is presented by nodes

$$\Gamma_c^1 = \bigcup_{i=1}^S \underline{x}_i$$

As discussed in Chapter 3: for all slave nodes, one or several master components are determined. For the sake of simplicity, let us suppose that for all contact nodes sufficiently close to the master surface there exists the closest master segment. Then the integral of the virtual work over the master surface can be split into a sum of  $M$  integrals over the master segments  $\partial\Omega_i$ ,  $i = 1, M$ . One slave node is attached at least to each segment,  $\underline{x}_j^i$ ,  $j = 1, N^i$ , where  $N^i$  is a number of slave nodes attached to the segment  $i$

$$\begin{aligned} & \int_{\Gamma_c^2} F(g_n, \underline{g}_t, \sigma_n, \underline{\sigma}_t, \delta g_n, \delta \underline{g}_t, \delta \sigma_n, \delta \underline{\sigma}_t) d\Gamma_c^2 = \\ & = \sum_{i=1}^M \int_{\partial\Omega_i} \sum_{j=1}^{N^i} F(\{g_n\}_j^i, \{\underline{g}_t\}_j^i, \{\sigma_n\}_j^i, \{\underline{\sigma}_t\}_j^i, \{\delta g_n\}_j^i, \{\delta \underline{g}_t\}_j^i, \{\delta \sigma_n\}_j^i, \{\delta \underline{\sigma}_t\}_j^i) d\partial\Omega_i, \end{aligned} \quad (5.15)$$

where  $F$  is a scalar function of the relative surface motion  $g_n, \underline{g}_t$  and of the arising contact stresses in normal and tangential directions  $\sigma_n, \underline{\sigma}_t$ . Indices  $\{.\}_j^i$  denote quantities related to the  $i$ -th master segment and the  $j$ -th master node interaction, whereas indices  $\{.\}_i$  represent the contribution of the master segment quantities. The scalar function  $F$  depends on the choice of the resolution method (see the previous chapter). Note that  $\sum_{i=1, M} N^i = S$  - number

of slave nodes. It implies that in active contact the number of slave nodes should be not smaller than the number of master segments.

The combination of the  $j$ -th slave node  $\underline{r}_s$  associated with the  $i$ -th master segment  $\underline{\rho} \in \partial\Omega_i$  presents the geometrical part of the *contact element*, which (depending on resolution method) can be complemented by a virtual node to store Lagrange multipliers – the contact stresses (Fig. 5.12, 5.13, 5.14). Since by definition a slave node can belong to only one contact element, the last can be denoted by one index  $i - \Omega_i^c$

Let us consider the contribution of one contact element to the virtual work of the system

$$\delta W_i^c = \int_{\partial\Omega_i} F(\{g_n\}_j^i, \{\underline{g}_t\}_j^i, \{\sigma_n\}_j^i, \{\underline{\sigma}_t\}_j^i, \{\delta g_n\}_j^i, \{\delta \underline{g}_t\}_j^i, \{\delta \sigma_n\}_j^i, \{\delta \underline{\sigma}_t\}_j^i) d\partial\Omega_i \quad (5.16)$$

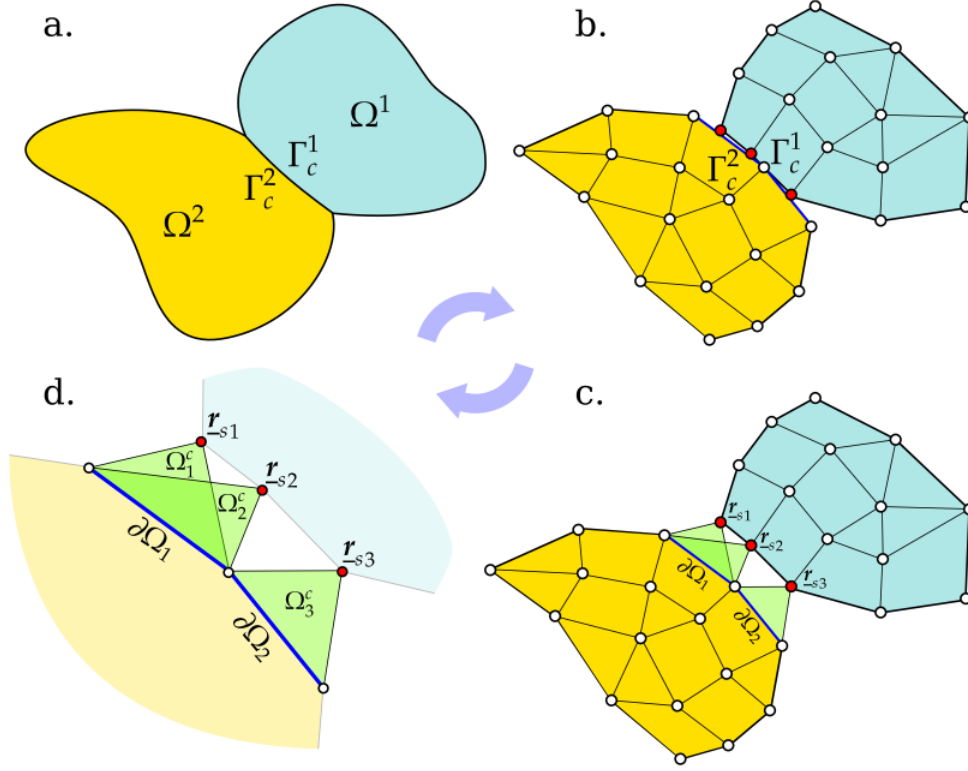


Figure 5.12: Construction of contact elements: **a** – continuous geometry two bodies  $\Omega_1$  and  $\Omega_2$  are in contact, active contact surfaces are  $\Gamma_c^1$  – slave surface and  $\Gamma_c^2$  – master surface, **b** – a discretized representation of bodies and contact interface, **c** – three contact elements spanned on three slave nodes  $\underline{r}_{s1}$ ,  $\underline{r}_{s2}$ ,  $\underline{r}_{s3}$  and two master segments  $\partial\Omega_1$ ,  $\partial\Omega_2$  **d** – zoom on three constructed contact elements  $\Omega_1^c$ ,  $\Omega_2^c$ ,  $\Omega_3^c$ .

The function  $F$  is a linear function of the variations  $\{\delta g_n\}_j^i, \{\delta \underline{g}_t\}_j^i, \{\delta \sigma_n\}_j^i, \{\delta \underline{\sigma}_t\}_j^i$ , which can be presented as follows (in the following layouts, indices will be omitted)

$$F = F_{g_n} \delta g_n + \underline{F}_{\underline{g}_t} \cdot \delta \underline{g}_t + F_{\sigma_n} \delta \sigma_n + \underline{F}_{\underline{\sigma}_t} \cdot \delta \underline{\sigma}_t = [F_{g_n}, \underline{F}_{\underline{g}_t}, F_{\sigma_n}, \underline{F}_{\underline{\sigma}_t}] \cdot \begin{bmatrix} \delta g_n \\ \delta \underline{g}_t \\ \delta \sigma_n \\ \delta \underline{\sigma}_t \end{bmatrix} \quad (5.17)$$

on the other hand

$$g_n = g_n(\underline{r}_s, \underline{\rho}), \quad \underline{g}_t = \underline{g}_t(\underline{r}_s, \underline{\rho}), \quad \underline{r}_s = \underline{x}_j^i, \quad \underline{\rho} \in \partial\Omega^i,$$

representing  $\underline{\rho}$  by nodal values  $\underline{\rho}^i$  and shape functions  $\phi_i(\underline{\xi})$ ,  $\underline{\xi} \in \partial\widetilde{\Omega}^i$

$$\begin{bmatrix} \underline{r}_s(t) \\ \underline{\rho}(t, \underline{\xi}) = \underline{\rho}^i(t) \cdot \phi_i(\underline{\xi}) \end{bmatrix} \Rightarrow \begin{bmatrix} \delta \underline{r}_s(t) \\ \delta \underline{\rho}(t, \underline{\xi}) = \delta \underline{\rho}^i(t) \cdot \phi_i(\underline{\xi}) \end{bmatrix}$$

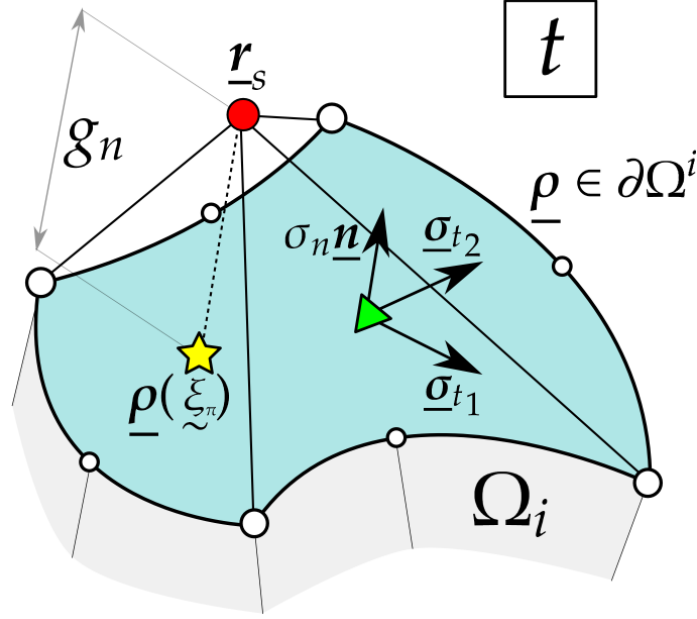


Figure 5.13: 3D contact element based on a quadratic master surface. The complementary node for Lagrange multipliers is marked with a green triangle. The projection of the slave node (big red circle) is marked by a star .

The variations of the geometrical quantities can be presented, as it was done in Section 2.4

$$\delta \underline{g}_n = [\nabla_{\underline{r}} \underline{g}_n, \nabla_{\underline{\rho}} \underline{g}_n] \cdot \begin{bmatrix} \delta \underline{r}_s \\ \delta \underline{\rho}^i \end{bmatrix}^T, \quad \delta \underline{g}_t = [\nabla_{\underline{r}} \underline{g}_t, \nabla_{\underline{\rho}} \underline{g}_t] \cdot \begin{bmatrix} \delta \underline{r}_s \\ \delta \underline{\rho}^i \end{bmatrix}^T$$

In general case, contact stresses are also split over dual shape functions  $\psi_i(\underline{\xi})$ , i.e.

$$\sigma_n(t, \underline{\xi}) = \sigma_n^i(t) \psi_i(\underline{\xi}), \quad \underline{\sigma}_t(t, \underline{\xi}) = \underline{\sigma}_t^i(t) \psi_i(\underline{\xi})$$

### 5.3.3 Discretization of the contact interface

Both shape functions for geometry  $\phi_i(\underline{\xi})$  and for contact stresses  $\psi_i(\underline{\xi})$  have to be chosen such that Babuška-Brezzi conditions (BB-conditions or also called inf-sup conditions) are fulfilled. It does not present any difficulties for small deformations and small sliding when contacting surface have matching nodes (*node-to-node discretization*). For nonconforming meshes the *mortar method* (based on Lagrange multipliers) and *Nitsche method* (purely displacement based) are commonly used approaches, which provide a stable discretization. Mortar approach was first applied in the framework of the domain decomposition methods to “glue” the solution on the interface of non-matching meshes corresponding to different subdomains. For an extensive mathematical description, the reader is referred to the book of

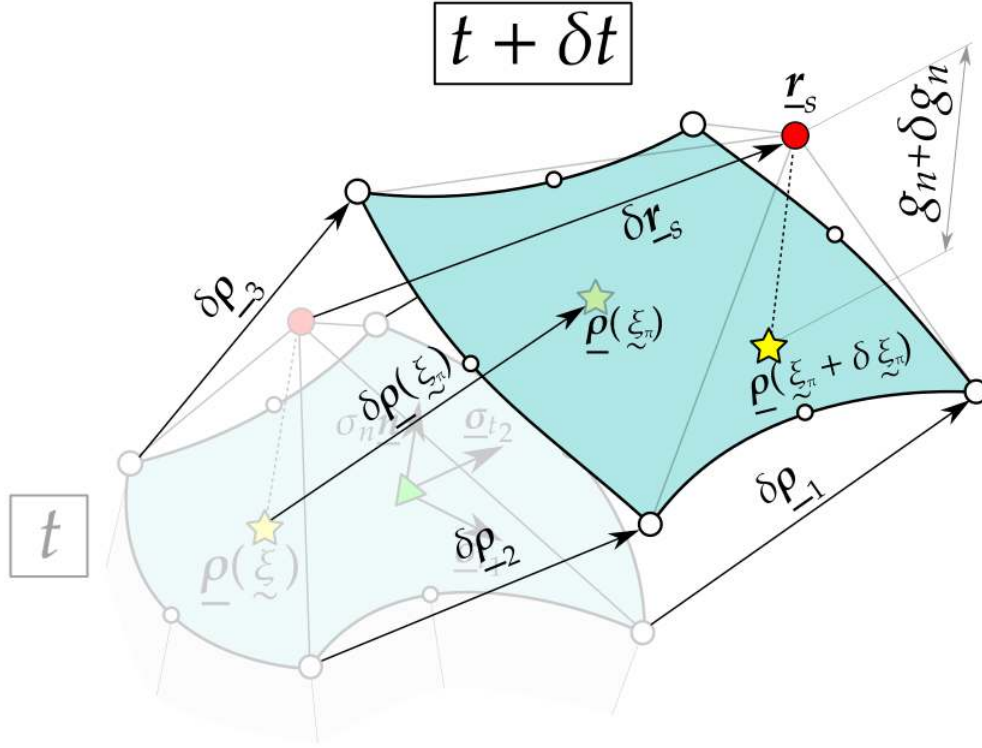


Figure 5.14: 3D contact element of Fig. 5.13 presented at two consecutive time steps  $t$  (transparent) and  $t + \delta t$ . Variational vectors are depicted as well as the variation of the main geometrical quantities: the convective surface parameter and the normal gap .

Wohlmuth [Wohlmuth 01] and references in it. For application to contact problems, see [Belgacem 98]. Further development originates to works due to McDevitt and Laursen [McDevitt 00], Puso [Puso 04], Puso and Laursen [Puso 03]. More recently, one retains contributions for two dimensional frictionless [Fischer 05] and frictional contact [Yang 05], [Fischer 06] and for three dimensional frictional contact [Yang 08b]. However, such a discretization requires the establishment of a segment-to-segment discretization as proposed by Simo, Wriggers and Taylor [Simo 85], see also [Zavarise 98], [Wriggers 06], which is a particularly complicated task for arbitrary contacting meshes in three dimensional contact problems. Another idea – *the contact domain method* – proposed recently by Oliver, Hartmann, et al. [Oliver 09], [Hartmann 09] consists in replacing the integral over the surface by an integral over the interior of the contact interface. A three dimensional version has been proposed by authors in [Oliver 10]. However, a three dimensional version requires the triangulation of the volume between two arbitrary contacting surfaces, which is not always possible.

Here we consider a simple but rather robust, widely used and multi-purpose *node-to-segment discretization* (NTS), so there is no freedom in the

choice of the discretization for the contact stresses: the contact stress within master segments is restricted to a constant value, i.e.  $\psi_i(\xi) = 1$  and  $\sigma_n(t, \xi) = \sigma_n(t)$ ,  $\underline{\sigma}_t(t, \xi) = \underline{\sigma}_t(t)$ . Note that this discretization does not fulfill the BB-conditions and so can exhibit locking, moreover, as is well known, NTS discretization fails the patch test – contact interface for nonconforming meshes cannot transfer uniform pressure see Taylor and Papadopoulos [Taylor 91]. On the other hand, in the same article, the authors demonstrated that in case of the Lagrange multiplier method and a sufficient number of slave nodes, the “two-pass” NTS approach is able to transfer correctly the uniform pressure, i.e. passes the patch test. Recently Zavarise and De Lorenzis [Zavarise 09a] proposed a modified NTS discretization passing the patch test for the penalty method.

### 5.3.4 Virtual work for discretized contact interface

According to all forementioned, the integrand of one contact element writes as

$$F = [\underline{F}_{r_s} \quad \underline{F}_{\underline{\rho}} \quad F_{\sigma_n} \quad \underline{F}_{\underline{\sigma}_t}] \cdot \begin{bmatrix} \delta \underline{r}_s \\ \delta \underline{\rho} \\ \delta \sigma_n \\ \delta \underline{\sigma}_t \end{bmatrix} \sim F = \begin{bmatrix} \underline{F}_x \\ \underline{F}_\lambda \end{bmatrix}^T \circ \begin{bmatrix} \delta \underline{x} \\ \delta \underline{\lambda} \end{bmatrix}, \quad (5.18)$$

where  $\underline{x}$  is a v-vector containing the actual vectors of all geometrical nodes of the contact element

$$\{ {}^{M+1}_1 \mathbf{S}_1^D \} \ni \underline{x} \sim [\underline{r}_s \quad \underline{\rho}_1 \quad \underline{\rho}_2 \quad \cdots \quad \underline{\rho}_M]^T.$$

$M$  is the number of master nodes,  $D$  is a dimension of the contact problem. Any point on the master surface is then determined by

$$\partial\Omega \ni \underline{\rho} = \underline{\phi}(0, \xi) \circ \underline{x}, \quad (5.19)$$

where  $\underline{\phi}(\xi)$  is a v-scalar of the contact element shape functions constructed as

$$\{ {}^{M+1}_1 \mathbf{S}_0^3 \} \ni \underline{\phi}(\zeta, \xi) \sim [\zeta \quad \phi_1(\xi) \quad \phi_2(\xi) \quad \cdots \quad \phi_M(\xi)]^T$$

Then, for example

$$g_n \underline{n} = -\underline{\phi}(-1, \xi_\pi) \circ \underline{x},$$

where  $\xi_\pi$  is a normal projection point of the slave node on the master surface. Further,  $\underline{\lambda}$  is a v-scalar containing Lagrange multipliers (contact stresses): the normal contact pressure and the contravariant coordinates of the tangential contact stress vector in the local surface basis are

$$\{ {}^D_0 \mathbf{S}_0^D \} \ni \underline{\lambda} \sim [\sigma_n \quad \underline{\sigma}_t]^T = [\lambda_n \quad \underline{\lambda}_t]^T, \quad \underline{\lambda}_t \circ \frac{\partial \underline{\rho}}{\partial \xi} = \underline{\sigma}_t$$

### 5.3 Finite Element Method

Due to the large number of s-structures of different dimensions

$$\underline{\xi}, \underline{\lambda}_t, \underline{\sigma}_t \in \left\{ \begin{smallmatrix} D-1 \\ 0 \end{smallmatrix} \mathbf{S}_0^D \right\}, \quad \underline{\lambda}, \underline{F}_\lambda \in \left\{ \begin{smallmatrix} D \\ 0 \end{smallmatrix} \mathbf{S}_0^D \right\}, \quad \underline{x}, \underline{F}_x \in \left\{ \begin{smallmatrix} M+1 \\ 1 \end{smallmatrix} \mathbf{S}_1^D \right\}$$

all computations should be carried out carefully. The remaining terms  $\underline{F}_\lambda$  and  $\underline{F}_x$  in (5.18) are v-vector and v-scalar corresponding to forces acting on the virtual geometrical displacements  $\delta \underline{x}$  and “forces” acting on the virtual Lagrange multiplier stresses (contact stresses) respectively. Finally, in new notations, from (5.16), the integral contribution of the  $i$ -th contact element to the total virtual work becomes

$$\delta W_i^c = \int_{\partial \Omega_i} \left[ \begin{smallmatrix} \underline{F}_x & \underline{F}_\lambda \end{smallmatrix} \right] \circ \left[ \begin{smallmatrix} \delta \underline{x} \\ \delta \underline{\lambda} \end{smallmatrix} \right] d\partial \Omega_i = \left[ \int_{\partial \Omega_i} \underline{F}_x d\partial \Omega_i \quad \int_{\partial \Omega_i} \underline{F}_\lambda d\partial \Omega_i \right] \circ \left[ \begin{smallmatrix} \delta \underline{x} \\ \delta \underline{\lambda} \end{smallmatrix} \right], \quad (5.20)$$

Let us demonstrate that in terms of the Finite Element Method the left term in (5.20) is a kind of residual vector and the right term is a kind of vector of dofs. For that, we consider a spatial basis  $\underline{e}$

$$\underline{e} \sim \left[ \underline{e}_1 \quad \dots \quad \underline{e}_D \right]$$

Then the dot product of  $\underline{e}$  components  $\underline{e}_k, k = 1, D$  with v-vectors from (5.20) provides the components of the scalar residual vector

$$\begin{aligned} \int_{\partial \Omega_i} \left( \underline{F}_x \cdot \underline{e}_k \right) d\partial \Omega_i &= \underline{R}_x^k \sim [R_x^i]_{k+j \times D} \in \mathbb{R}^{(M+1) \times D}, j \in [1, M+1], \\ \delta \underline{x} \cdot \underline{e}_k &\sim \delta [x^i]_{k+j \times D} \in \mathbb{R}^{(M+1) \times D}, j \in [1, M+1] \\ \int_{\partial \Omega_i} \underline{F}_\lambda d\partial \Omega_i &= \underline{R}_\lambda \sim [R_\lambda^i] \in \mathbb{R}^D, \quad \delta \underline{\lambda} \sim \delta [\lambda^i] \in \mathbb{R}^D \end{aligned} \quad (5.21)$$

and the full residual vector  $[R^i]$  of the  $i$ -th contact element and vector of its unknown  $[x^i]$  take the form

$$\mathbb{R}^{M+D+1} \ni [R^i] = \begin{bmatrix} R_x^i \\ R_\lambda^i \end{bmatrix}, \quad \mathbb{R}^{M+D+1} \ni [x^i] = \begin{bmatrix} \delta x^i \\ \delta \lambda^i \end{bmatrix}$$

This expression is given simply to have a link with classical notations of the Finite Element Method.

We remind that the form of  $\underline{F}_x$  and  $\underline{F}_\lambda$  depends on the resolution method and will be given in the following parts. Now let us assemble all contributions to the virtual work balance equation. Here we will focus on the contact elements and so the contribution of all structural finite elements will not be derived in closed form. All necessary routines can be found in books on Finite Element Method for solid mechanics: Zienkiewicz and Taylor



[Zienkiewicz 00a], [Zienkiewicz 00b], Chrisfield [Crisfield 00a], [Crisfield 00b], Bathe [Bathe 96], Belytschko et al. [Belytschko 08] and others. Let us denote by  $\delta W_j^s$  the contribution to the virtual work of the  $j$ -th structural element. To enforce equilibrium conditions we require the total virtual work on solution path to be zero

$$\sum_{j=1}^{N^e} \delta W_j^s + \sum_{i=1}^S \delta W_i^c = 0. \quad (5.22)$$

$N^e$  is the total number of structural elements and  $S$  is the number of contact elements and at the same time the number of slave nodes included in the contact elements. This equation can be presented as a set of  $\approx N^n + S$  vector equations, where  $N^n$  is a number of free nodes. The sign  $\approx$  expresses that Dirichlet boundary conditions can be imposed on any components of a displacement vector, so that the number of algebraic equations in (5.22) is equal to the number of free degrees of freedom  $N^{\text{dof}}$  plus the number of contact elements multiplied by the dimension of the problem  $SD$ . If Lagrange multiplier or coupled augmented Lagrangian methods are used, in case of use of penalty method or augmented Lagrangian with Uzawa's algorithm, the number of equations in (5.22) reduces to  $N^{\text{dof}}$

$$\sum_{j=1}^{N^e} \delta W_j^s + \sum_{i=1}^S \delta W_i^c = 0 \quad \sim \quad \begin{bmatrix} R_x^s + R_x^c & R_\lambda \end{bmatrix} \delta \begin{bmatrix} x \\ \lambda \end{bmatrix} = 0, \quad (5.23)$$

$R_x^s$  is a dense<sup>3</sup> residual vector related to structural elements,  $R^s$  a sparse residual vector related to the contact elements,  $R_i^s$  may be nonzero if the  $i$ -th dof belongs to one node of the contact elements.  $x, R_x^c, R^s, x \in \mathbb{R}^{N^{\text{dof}}}$  and  $\lambda, R_\lambda^c \in \mathbb{R}^{SD}$ . Since virtual displacements and virtual Lagrange multipliers are arbitrary, Eq. (5.23) can be satisfied only if

$$\begin{bmatrix} R_x^s + R_x^c \\ R_\lambda^c \end{bmatrix} = 0 \quad (5.24)$$

This is a set of  $N^{\text{dof}} + SD$  algebraic nonlinear equations, the nonlinearity is preserved: even if in the structural part the condition  $R_x^s = 0$  appears to be a set of linear algebraic equations, the contact problem is still nonlinear. Therefore a solution technique for nonlinear equations has to be applied, for example, Newton's method discussed in Section 5.1. Hence the linearization of this system of equations is required.

### 5.3.5 Linearization of equations

Newton's method and its generalization for nonsmooth problems have been discussed in Section 5.1. In this paragraph, this method will be applied for a system of nonlinear algebraic Eq. (5.23) derived in the previous paragraph. We assume that the solution  $\{u_k, \lambda_k\}$  for given boundary conditions  $f_k$  at time step

<sup>3</sup>here, dense vector in the sense that all components may be nonzero.

$t_k$  is given either as initial condition or as solution of the  $k$ -th increment.

$$\begin{bmatrix} R_x^s(x_k, f_k) + R_x^c(x_k, \lambda_k) \\ R_\lambda^c(x_k, \lambda_k) \end{bmatrix} \approx 0, \quad (5.25)$$

At time step  $t_{k+1}$ , a change of boundary conditions  $f_{k+1} = f_k + \Delta f_k$  makes the system loose its equilibrium state. So we establish an iterative procedure  $0, 1, 2, \dots, i$  and we wish that the sequence  $\{u^i, \lambda^i\}$  converges to the solution  $\{u_{k+1}, \lambda_{k+1}\}$  which returns the system to equilibrium for the new given boundary conditions  $f_{k+1}$ . According to Newton's method, the iteration increment is given by

$$\begin{bmatrix} \Delta x^i \\ \Delta \lambda^i \end{bmatrix} = - \begin{bmatrix} \frac{\partial R_x^s}{\partial x} + \frac{\partial R_x^c}{\partial x} & \frac{\partial R_\lambda^c}{\partial x} \\ \frac{\partial R_x^c}{\partial \lambda} & \frac{\partial R_\lambda^c}{\partial \lambda} \end{bmatrix}_{\{x^i, \lambda^i, f_{k+1}\}}^{-1} \begin{bmatrix} R_x^s(x^i, f_{k+1}) + R_x^c(x^i, \lambda^i) \\ R_\lambda^c(x^i, \lambda^i) \end{bmatrix}, \quad (5.26)$$

and the corresponding update rule is given by

$$x^{i+1} = x^i + \Delta x^i, \quad \lambda^{i+1} = \lambda^i + \Delta \lambda^i$$

The structural tangent matrix is denoted by  $\frac{\partial R_x^s}{\partial x} = K^s$  and we are particularly interested in this remaining terms which can be called *tangent contact matrix*

$$\begin{bmatrix} K^s & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \frac{\partial R_x^c}{\partial x} & \frac{\partial R_\lambda^c}{\partial x} \\ \frac{\partial R_x^c}{\partial \lambda} & \frac{\partial R_\lambda^c}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} K^s + K_{xx}^c & K_{\lambda x}^c \\ K_{x\lambda}^c & K_{\lambda\lambda}^c \end{bmatrix}$$

The full tangent contact matrix is an assembly of contact element stiffness

$$\begin{bmatrix} K_{xx}^c & K_{\lambda x}^c \\ K_{x\lambda}^c & K_{\lambda\lambda}^c \end{bmatrix} = \bigcup_{i=1}^S \begin{bmatrix} K_{xx}^{ci} & K_{\lambda x}^{ci} \\ K_{x\lambda}^{ci} & K_{\lambda\lambda}^{ci} \end{bmatrix}$$

Returning to (5.20), we can express all the terms of the elementary tangent contact matrix through derivatives of v-vector  $\underline{\underline{F}}_x$  and v-scalar  $\underline{\underline{F}}_\lambda$ , so

$$\begin{bmatrix} K_{xx}^{ci} & K_{\lambda x}^{ci} \\ K_{x\lambda}^{ci} & K_{\lambda\lambda}^{ci} \end{bmatrix} \sim \begin{bmatrix} \int_{\partial\Omega_i} \left[ \frac{\partial \underline{\underline{F}}_x}{\partial \underline{\underline{x}}} \right] d\partial\Omega_i & \int_{\partial\Omega_i} \left[ \frac{\partial \underline{\underline{F}}_\lambda}{\partial \underline{\underline{x}}} \right] d\partial\Omega_i \\ \int_{\partial\Omega_i} \left[ \frac{\partial \underline{\underline{F}}_x}{\partial \underline{\underline{\lambda}}} \right] d\partial\Omega_i & \int_{\partial\Omega_i} \left[ \frac{\partial \underline{\underline{F}}_\lambda}{\partial \underline{\underline{\lambda}}} \right] d\partial\Omega_i \end{bmatrix}$$

So, to integrate contact into the resolution, it remains only to precise the form of the functions

$$\underline{\underline{F}}_x, \underline{\underline{F}}_\lambda$$

and to evaluate their derivatives

$$\frac{\partial \underline{\underline{F}}_x}{\partial \underline{\underline{x}}}, \frac{\partial \underline{\underline{F}}_\lambda}{\partial \underline{\underline{x}}}, \frac{\partial \underline{\underline{F}}_x}{\partial \underline{\underline{\lambda}}}, \frac{\partial \underline{\underline{F}}_\lambda}{\partial \underline{\underline{\lambda}}}$$

### 5.3.6 Example

Here we give a simple example of the procedure of tangent matrix construction for a mechanical problem with contact. Let us consider two structural elements  $\Omega_1$  and  $\Omega_2$  (Fig. 5.15) which come in contact. A constant pressure is imposed on the upper segment between nodes 1 and 3. Nodes 6 and 7 are fixed so they are not included in the computations. The tangent stiffness matrix for such a configuration has the form presented in Fig. 5.15,a. The detection procedure determines that the slave node 2 may come in contact with the master segment between nodes 4 and 5. So the contact element 2 – 4 – 5 with a complementary node 8 for the Lagrange multipliers is constructed and added to the global tangent stiffness matrix (Fig. 5.15,b.)

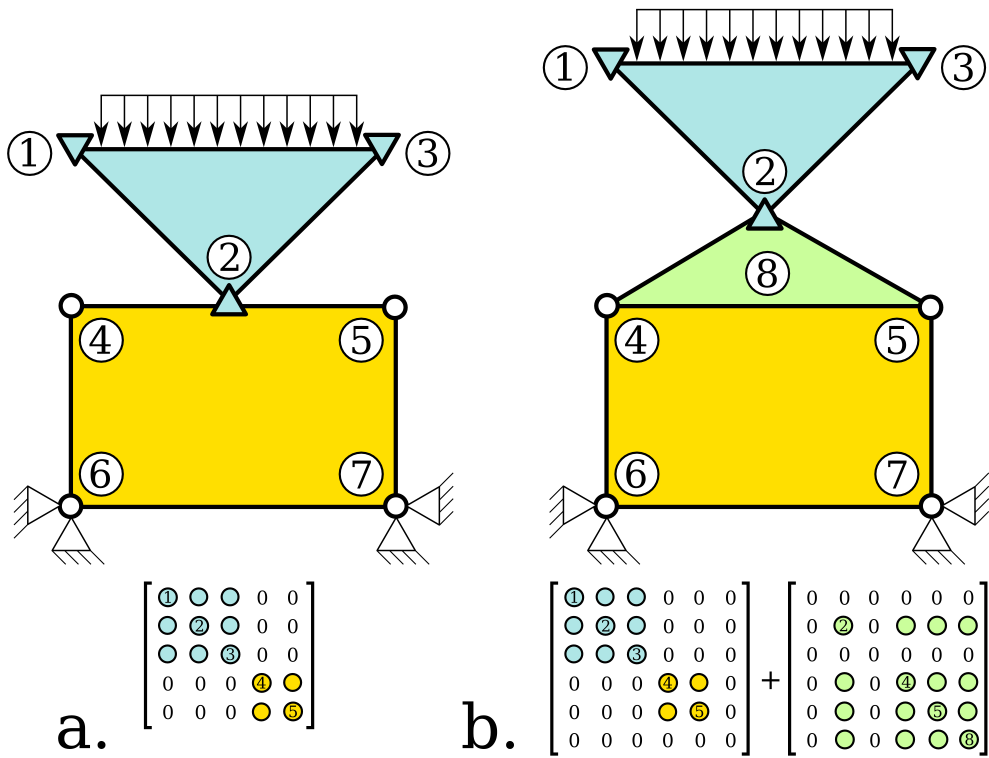


Figure 5.15: Example of tangent matrix: a – structural matrix  $K^s$  for two elements  $\Omega_1$ ,  $\Omega_2$  and given boundary conditions, b – contribution of the contact element (spanned on slave node 2 and master segment 4 – 5 with a complementary node 8 for Lagrange multipliers) to the tangent matrix .

## 5.4 Residual vectors and tangent matrices for contact elements

Residual vectors and tangent contact matrices for different resolution methods will be derived in this section. All necessary integrals have been formulated in Chapter 4 and all geometry related variations have been derived in Chapter 2. Here it remains to give closed forms for

$$\underline{\underline{R}}^c_x, \underline{\underline{R}}^c_\lambda \quad \text{and} \quad \underline{\underline{K}}^c_{xx}, \underline{\underline{K}}^c_{\lambda x}, \underline{\underline{K}}^c_{x\lambda}, \underline{\underline{K}}^c_{\lambda\lambda}$$

These vectors and matrices will be given for arbitrary discretization forms for penalty method (linear penalty) and coupled augmented Lagrangian method.

### 5.4.1 Penalty method: frictionless case

#### Residual vector

The virtual work due to frictionless contact was given in (4.79) for continuous problems. In the Finite Element framework the contribution of the  $i$ -th NTS (node-to-segment) contact element for the linear penalty method writes as

$$\delta W_i^c = \int_{\partial\Omega_i} -\varepsilon_n \langle -g_n \rangle \delta g_n d\partial\Omega_i, \quad (5.27)$$

according to (2.99) the variation of the normal gap can be presented as  $\delta g_n = [\nabla g_n]^T \cdot \delta[\underline{x}]$  or in s-structure notations

$$\delta g_n = \underline{\underline{G}}(\underline{\xi}_\pi)^\circ \delta \underline{\underline{x}},$$

where  $\underline{\xi}_\pi$  is the projection of the slave node on the master surface and  $\underline{\underline{G}}$  is a v-vector of first variation of the normal gap  $\delta g_n$  given in (2.99). So Eq. 5.27 takes the form

$$\delta W_i^c = \int_{\partial\Omega_i} -\varepsilon_n \langle -g_n \rangle \underline{\underline{G}}(\underline{\xi}_\pi)^\circ \delta \underline{\underline{x}} d\partial\Omega_i = \left[ \int_{\partial\Omega_i} \underline{\underline{F}}_x d\partial\Omega_i \right]^\circ \delta \underline{\underline{x}} = [R_x^c]^T [\delta x^i]. \quad (5.28)$$

Consequently  $\underline{\underline{F}}_x$  is given as

$$\underline{\underline{F}}_x = -\varepsilon_n \langle -g_n \rangle \underline{\underline{G}}(\underline{\xi}_\pi).$$

The integral of  $\underline{\underline{F}}_x$  over the master surface  $\partial\Omega_i$  can be simply evaluated since the integrand does not depend on the surface parameter  $\underline{\xi}$

$$\int_{\partial\Omega_i} \underline{\underline{F}}_x d\partial\Omega_i = - \int_{\partial\Omega_i} \varepsilon_n \langle -g_n \rangle d\partial\Omega_i \underline{\underline{G}}(\underline{\xi}_\pi) = P_n \underline{\underline{G}}(\underline{\xi}_\pi), \quad (5.29)$$

where the first term is nothing but a normal contact force  $P_n$  computed on the  $i$ -th iteration configuration

$$P_n = \underline{\underline{F}} \cdot \underline{\underline{n}} = \int_{\partial\Omega_i} -\varepsilon_n \langle -g_n \rangle d\partial\Omega_i.$$

Finally the residual v-vector is given by

$$\underline{\underline{R}}_x^c = P_n \underline{\underline{G}}(\xi_\pi).$$

### Tangent contact matrix

To evaluate the elemental tangent contact matrix, it is necessary to take a partial derivative of the v-vector  $\underline{\underline{F}}_x$

$$\frac{\partial \underline{\underline{F}}_x}{\partial \underline{\underline{x}}} = \begin{cases} \varepsilon_n \frac{\partial g_n}{\partial \underline{\underline{x}}} \underline{\underline{G}} + \varepsilon_n g_n \underline{\underline{H}}, & g_n < 0 \\ 0, & g_n \geq 0 \end{cases}$$

where all geometry related s-structures come from connections between the variation of the geometrical quantities and the nodal coordinate vectors

$$\delta g_n = \frac{\partial g_n}{\partial \underline{\underline{x}}} \cdot \delta \underline{\underline{x}} = \underline{\underline{G}} \cdot \delta \underline{\underline{x}} \quad \text{and} \quad \Delta \delta g_n = \Delta \underline{\underline{x}} \cdot \frac{\partial^2 g_n}{\partial \underline{\underline{x}}^2} \cdot \delta \underline{\underline{x}} = \Delta \underline{\underline{x}} \cdot \underline{\underline{H}} \cdot \delta \underline{\underline{x}}.$$

These s-structures have been derived in Section 2.4. In general form we get

$$\frac{\partial \underline{\underline{F}}_x}{\partial \underline{\underline{x}}} = \begin{cases} \varepsilon_n \underline{\underline{G}} \underline{\underline{G}} + \varepsilon_n g_n \underline{\underline{H}}, & g_n < 0 \\ 0, & g_n \geq 0 \end{cases}.$$

Integrating over the master surface gives

$$\int_{\partial\Omega_i} \frac{\partial \underline{\underline{F}}_x}{\partial \underline{\underline{x}}} d\partial\Omega_i = \varepsilon_n \underline{\underline{G}} \underline{\underline{G}} \partial\Omega_i + P_n \underline{\underline{H}}.$$

This expression can be interpreted as a tangent contact t-tensor  $\underline{\underline{K}}_{xx}^c$

$$\underline{\underline{K}}_{xx}^c = \varepsilon_n \underline{\underline{G}} \underline{\underline{G}} \partial\Omega_i + P_n \underline{\underline{H}}, \quad g_n < 0$$

or in a slightly different form

$$\underline{\underline{K}}_{xx}^c = \varepsilon_n \partial\Omega_i \left[ \underline{\underline{G}} \underline{\underline{G}} - \langle -g_n \rangle \underline{\underline{H}} \right], \quad g_n < 0$$

### 5.4.2 Penalty method: frictional case

In case of frictional contact, we have to distinguish stick and slip states, since residual vector and stiffness matrix are different in each case. The virtual work due to frictional contact was given in expanded form (4.79) for continuous problem. In the Finite Element framework contribution of the  $i$ -th NTS (node-to-segment) contact element for linear penalty method writes as

$$\delta W_i^c = \begin{cases} \int_{\partial\Omega_i} -\varepsilon_n \langle -g_n \rangle (\delta g_n - \mu \underline{s} \cdot \delta \underline{g}_t) d\partial\Omega_i, & \text{slip} \\ \int_{\partial\Omega_i} \varepsilon_n \langle -g_n \rangle \delta g_n + \varepsilon_t \underline{g}_t^\bullet \cdot \delta \underline{g}_t d\partial\Omega_i, & \text{stick} \end{cases} \quad (5.30)$$

from (2.99) and (2.102) we get

$$\delta g_n = [\nabla g_n]^T \cdot \delta [\underline{x}], \quad \delta \underline{g}_t = \frac{\partial \rho}{\partial \xi_1} \delta \xi^1 + \frac{\partial \rho}{\partial \xi_2} \delta \xi^2 = \left\{ \frac{\partial \rho}{\partial \xi_1} [\nabla \xi^1]^T + \frac{\partial \rho}{\partial \xi_2} [\nabla \xi^2]^T \right\} \cdot \delta [\underline{x}]$$

In s-structure notations these formulae can be rewritten as

$$\delta g_n = \underline{\underline{G}}(\underline{\xi}_\pi) \cdot \delta \underline{\underline{x}}, \quad \delta \underline{\underline{g}}_t = \frac{\partial \rho}{\partial \underline{\xi}} \circ \delta \underline{\xi} = \frac{\partial \rho}{\partial \underline{\xi}} \circ \underline{\underline{T}} \cdot \delta \underline{\underline{x}},$$

where  $\underline{\underline{T}}$  is a v-v-vector (two v-vectors gathered together in a v-structure). In these notations the integral in (5.31) takes the form

$$\delta W_i^c = \begin{cases} \left[ \int_{\partial\Omega_i} -\varepsilon_n \langle -g_n \rangle \left( \underline{\underline{G}}(\underline{\xi}_\pi) - \mu \underline{s} \cdot \left( \frac{\partial \rho}{\partial \underline{\xi}} \circ \underline{\underline{T}} \right) \right) \Big|_{\underline{\xi}_\pi} d\partial\Omega_i \right] \cdot \delta \underline{\underline{x}}, & \text{slip} \\ \left[ \int_{\partial\Omega_i} -\varepsilon_n \langle -g_n \rangle \underline{\underline{G}}(\underline{\xi}_\pi) + \varepsilon_t \underline{\underline{g}}_t^\bullet \cdot \left( \frac{\partial \rho}{\partial \underline{\xi}} \circ \underline{\underline{T}} \right) \Big|_{\underline{\xi}^\bullet} d\partial\Omega_i \right] \cdot \delta \underline{\underline{x}}, & \text{stick} \end{cases} \quad (5.31)$$

Note that the term  $\frac{\partial \rho}{\partial \underline{\xi}} \underline{\underline{T}}$  is evaluated at different points: in case of stick at point  $\underline{\xi}^\bullet$  and in case of slip at the current projection point  $\underline{\xi}_\pi$  (see also *Remark 9.2* in [Wriggers 06]). This difference can be explained schematically if we assume that in the contact interface there are rigid bars hinged to the master segment by a nonlinear circular spring (fig. 5.16). The slave node, penetrating under the surface at  $|g_n|$ , is assumed to be in contact with this bar, so the tangential movement of the slave produces a rotation of the bar (angle  $\phi$ ) such as

$$|\underline{\underline{g}}_t| = |g_n| \tan \phi.$$

A resulting force  $\underline{\underline{F}}_t$

$$F_t = \varepsilon_t |\underline{\underline{g}}_t| = \varepsilon_t |g_n| \tan \phi$$

acts on the slave node and on the hinge of the bar. If the external force is removed, then the slave node unloads the bar and returns to the stick position.

According to Coulomb's law, the tangential resistance is limited by the normal force multiplied by the friction coefficient

$$F_t = \mu F_n = \mu \varepsilon_n |g_n| \Leftrightarrow F_t = \varepsilon_t |g_n| \tan \phi^* = \mu \varepsilon_n |g_n| \Rightarrow \phi^* = \arctan\left(\frac{\mu \varepsilon_n}{\varepsilon_t}\right)$$

This critical angle  $\phi^*$  determines the length of the bar

$$l = \frac{g_n}{\cos \phi^*} = \frac{g_n \sqrt{\mu^2 \varepsilon_n^2 + \varepsilon_t^2}}{\varepsilon_t}.$$

When the critical angle is reached, the slave node jumps from the current bar to the next one (see Fig. 5.17), which results in a new distribution of forces in the master slave interface. Formally, this interpretation is valid only for the stick state. Remark that the fact that slave node penetrates under the master surface and slides over stick point results in appearance of a momentum.

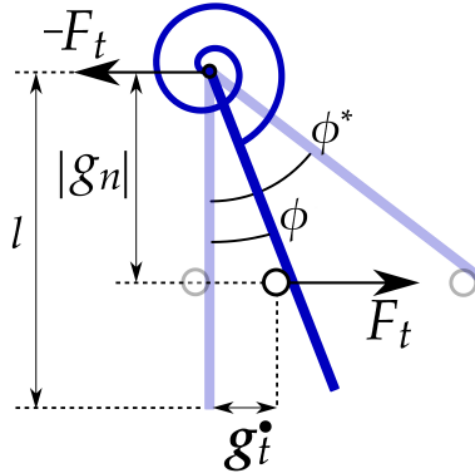


Figure 5.16: Rigid bar hinged by a spring to the master segment and its interaction with a slave node.

### Residual vector

Since in the integral (5.31) the integrands do not depend on the convective coordinate  $\xi$ , the evaluation of the residual v-vector is straightforward

$$\underline{\underline{R}}_x^c = \begin{cases} P_n \left( \underline{\underline{G}}(\xi_\pi) - \mu \underline{s} \cdot \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \circ \underline{\underline{T}} \right) \Big|_{\underline{\xi}_\pi} \right), & \text{slip} \\ P_n \underline{\underline{G}}(\xi_\pi) + \underline{F}_t \cdot \left( \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \circ \underline{\underline{T}} \right) \Big|_{\underline{\xi}_\bullet}, & \text{stick} \end{cases} \quad (5.32)$$

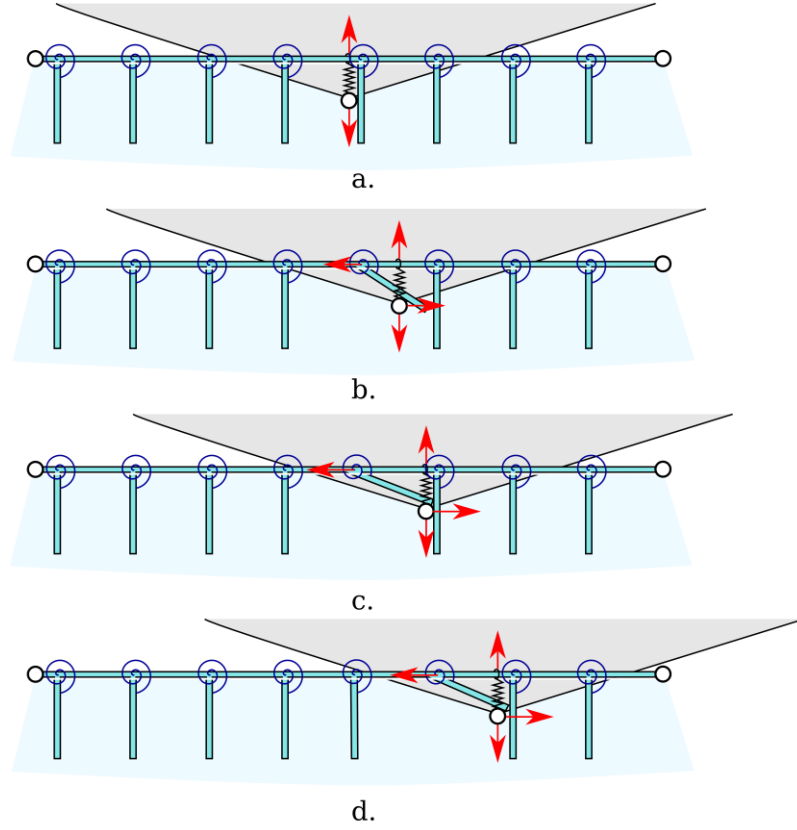


Figure 5.17: Representation of tangential interaction of the slave node with a master, as interaction with a set of rigid bars hinged to the master segment by a spring: **a-b-c** – stick state, **d** – slip.

where normal and tangential contact forces are  $P_n = \int_{\partial\Omega_i} -\varepsilon_n \langle -g_n \rangle d\partial\Omega_i$  and  $F_t = \int_{\partial\Omega_i} \varepsilon_t \underline{\mathbf{g}}_t^\bullet d\partial\Omega_i$  respectively.

Making use of the return mapping algorithm (see Section 5.2) leads to the following scheme. We make a start from the evaluation of the contact pressure for a given displacement field

$$\sigma_n^{k+1} = -\varepsilon_n \langle -g_n^{k+1} \rangle \Rightarrow P_n^{k+1} = -\varepsilon_n \langle -g_n^{k+1} \rangle \partial\Omega_i, \quad (5.33)$$

(let us remind that the lower index  $i$  denotes the  $i$ -th contact element, and the upper indices  $k, k+1$  denote solution increments) further the trial tangential contact force in stick is computed as

$$\underline{\mathbf{F}}_t^{\text{trial}} = \underline{\mathbf{F}}_t^k + \int_{\partial\Omega_i} \varepsilon_t \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \circ (\underline{\xi}^{k+1} - \underline{\xi}^k) d\partial\Omega_i = \varepsilon_t \partial\Omega_i (\underline{\xi}^{k+1} - \underline{\xi}^{\bullet k}) \circ \frac{\partial \underline{\rho}}{\partial \underline{\xi}}^{k+1}, \quad (5.34)$$

where  $\underline{\xi}^{\bullet k}$  is the convective covariant coordinate of the stick point for  $k$ -th



increment. Further, we check if this trial force is inside the Coulomb's cone

$$\|\underline{\mathbf{F}}_t^{\text{trial}}\| \leq \mu |P_n^{k+1}|.$$

If not, then the trial force has to be corrected. According to Eq. (5.8), the tangential contact force writes

$$\underline{\mathbf{F}}_t^{k+1} = \underline{\mathbf{F}}_t^{\text{trial}} - \int_{\partial\Omega_i} \varepsilon_t \langle \Delta\gamma^k \rangle \underline{\mathbf{s}}_t^k d\partial\Omega_i$$

where

$$\begin{aligned} \underline{\mathbf{s}}_t^k &= \frac{\underline{\mathbf{F}}_t^{\text{trial}}}{\|\underline{\mathbf{F}}_t^{\text{trial}}\|} \\ \langle \Delta\gamma^k \rangle &= \frac{\langle \|\underline{\sigma}_t^{\text{trial}}\| - \mu |\sigma_n^{k+1}| \rangle}{\varepsilon_t} \end{aligned} \quad (5.35)$$

The integral of the slip  $\Delta\gamma^k$  is

$$\int_{\partial\Omega_i} \langle \Delta\gamma^k \rangle d\partial\Omega_i = \frac{\langle \|\underline{\mathbf{F}}_t^{\text{trial}}\| - \mu |P_n^{k+1}| \rangle}{\varepsilon_t}$$

The stick point is updated according to the simple rule

$$\tilde{\xi}^{\bullet\{k+1\}} = \tilde{\xi}^{\bullet k} + \Delta\gamma^k \underline{\mathbf{s}}_t^k,$$

where  $\underline{\mathbf{s}}_t^k$  is the v-scalar of covariant coordinates of the sliding vector

$$\underline{\mathbf{s}}_t^k = \underline{\mathbf{s}}_t^k \cdot \frac{\partial \underline{\rho}}{\partial \xi}$$

Finally we get the following return mapping scheme

$$\begin{cases} \underline{\mathbf{F}}_t^{k+1} = \underline{\mathbf{F}}_t^{\text{trial}}, & \tilde{\xi}^{\bullet\{k+1\}} = \tilde{\xi}^{\bullet k}, & \|\underline{\mathbf{F}}_t^{k+1}\| \leq \mu |P_n^{k+1}|, & \text{stick} \\ \underline{\mathbf{F}}_t^{k+1} = \mu |P_n^{k+1}| \underline{\mathbf{s}}_t^k, & \tilde{\xi}^{\bullet\{k+1\}} = \tilde{\xi}^{\bullet k} + \Delta\gamma^k \underline{\mathbf{s}}_t^k, & \|\underline{\mathbf{F}}_t^{k+1}\| > \mu |P_n^{k+1}|, & \text{slip} \end{cases} \quad (5.36)$$

Expressing the tangential force for stick in the contravariant basis gives the following expressions (it seems to be more advantageous for the following linearization to retain vector form for tangential slip  $\underline{\mathbf{s}}_t^k$ )

$$\begin{cases} \underline{\mathbf{F}}_t^{k+1} = \left( \underline{\mathbf{F}}_t^{\text{trial}} \cdot \frac{\partial \underline{\rho}^{k+1}}{\partial \xi} \right) \circ \frac{\partial \underline{\rho}^{k+1}}{\partial \xi} = \tilde{\mathbf{F}}_t^{\text{trial}} \circ \frac{\partial \underline{\rho}^{k+1}}{\partial \xi}, & \text{stick} \\ \underline{\mathbf{F}}_t^{k+1} = \mu |P_n^{k+1}| \underline{\mathbf{s}}_t^k, & \text{slip} \end{cases} \quad (5.37)$$

According to the definition of the updated normal contact force (5.33) and the trial tangential contact force (5.34), we get

$$\begin{cases} \underline{\mathbf{F}}_t^{k+1} = \varepsilon_t \partial\Omega_i (\tilde{\xi}^{k+1} - \tilde{\xi}^{\bullet k}) \circ \frac{\partial \underline{\rho}^{k+1}}{\partial \xi}, & \text{stick} \\ \underline{\mathbf{F}}_t^{k+1} = \mu \varepsilon_n \partial\Omega_i \langle -g_n^{k+1} \rangle \underline{\mathbf{s}}_t^k, & \text{slip} \end{cases} \quad (5.38)$$

## 5.4 Residual vectors and tangent matrices

For the following linearization, we derive here the variation of the sliding unit vector  $\underline{s}_t^k$ ; from its definition (5.35) we get

$$\underline{s}_t^k = \frac{\underline{F}_t^{\text{trial}}}{\|\underline{F}_t^{\text{trial}}\|} = \frac{\underline{F}_t^{\text{trial}}}{\sqrt{\underline{F}_t^{\text{trial}} \cdot \underline{F}_t^{\text{trial}}}} \Rightarrow \Delta \underline{s}_t^k = \left( \underline{I} - \underline{s}_t^k \otimes \underline{s}_t^k \right) \cdot \frac{\Delta \underline{F}_t^{\text{trial}}}{\|\underline{F}_t^{\text{trial}}\|} \quad (5.39)$$

If the unity tensor is presented as

$$\underline{I} = \underline{n} \otimes \underline{n} + \underline{\tilde{s}}^k \otimes \underline{\tilde{s}}^k + \underline{s}_t^k \otimes \underline{s}_t^k,$$

where  $\underline{\tilde{s}}^k$  is an in-plane unit vector orthogonal to the sliding direction, since we have to take a dot product of the slip vector variation with in-plane basis vectors (see (5.32)), we can neglect the normal components  $\underline{n} \otimes \underline{n}$ , then the expression (5.39) can be rewritten as

$$\Delta \underline{s}_t^k = \underline{\tilde{s}}^k \otimes \underline{\tilde{s}}^k \cdot \frac{\Delta \underline{F}_t^{\text{trial}}}{\|\underline{F}_t^{\text{trial}}\|}, \quad (5.40)$$

where the variation of the tangential trial force follows from (5.34)

$$\Delta \underline{F}_t^{\text{trial}} = \varepsilon_t \partial \Omega_i \Delta \bar{\xi}^{k+1} \circ \frac{\partial \rho}{\partial \bar{\xi}}^{k+1} + \varepsilon_t \partial \Omega_i (\bar{\xi}^{k+1} - \bar{\xi}^{\bullet k}) \circ \Delta \frac{\partial \rho}{\partial \bar{\xi}}^{k+1}, \quad (5.41)$$

If we represent the unit vector  $\underline{\tilde{s}}^k$  in the covariant basis

$$\underline{\tilde{s}}^k = \left( \underline{\tilde{s}}^k \cdot \frac{\partial \rho}{\partial \bar{\xi}}^{k+1} \right) \circ \frac{\partial \rho}{\partial \bar{\xi}}^{k+1} = \underline{\tilde{s}}^k \circ \frac{\partial \rho}{\partial \bar{\xi}}^{k+1},$$

the variation of the sliding vector takes a simpler form

$$\Delta \underline{s}_t^k = \frac{\varepsilon_t \partial \Omega_i}{\|\underline{F}_t^{\text{trial}}\|} \underline{\tilde{s}}^k \circ \left[ \Delta \bar{\xi}^{k+1} + \frac{\partial \rho}{\partial \bar{\xi}}^{k+1} \boxtimes \Delta \frac{\partial \rho}{\partial \bar{\xi}}^{k+1} \circ (\bar{\xi}^{k+1} - \bar{\xi}^{\bullet k}) \right] \underline{\tilde{s}}^k \circ \frac{\partial \rho}{\partial \bar{\xi}}^{k+1}. \quad (5.42)$$

Since the variation of the contravariant basis is a complicated task, we express it by means of the variation of the covariant basis vectors

$$\frac{\partial \rho}{\partial \bar{\xi}} \boxtimes \frac{\partial \rho}{\partial \bar{\xi}} = \underline{I} \Rightarrow \Delta \frac{\partial \rho}{\partial \bar{\xi}} \boxtimes \frac{\partial \rho}{\partial \bar{\xi}} + \frac{\partial \rho}{\partial \bar{\xi}} \boxtimes \Delta \frac{\partial \rho}{\partial \bar{\xi}} = 0 \Leftrightarrow \frac{\partial \rho}{\partial \bar{\xi}} \boxtimes \Delta \frac{\partial \rho}{\partial \bar{\xi}} = -\Delta \frac{\partial \rho}{\partial \bar{\xi}} \boxtimes \frac{\partial \rho}{\partial \bar{\xi}}$$

Omitting increment indices in the expression (5.42), we finally get

$$\Delta \underline{s}_t = \frac{\varepsilon_t \partial \Omega_i}{\|\underline{F}_t^{\text{trial}}\|} \underline{\tilde{s}} \circ \left[ \Delta \bar{\xi} - \Delta \frac{\partial \rho}{\partial \bar{\xi}} \boxtimes \frac{\partial \rho}{\partial \bar{\xi}} \circ (\bar{\xi} - \bar{\xi}^{\bullet}) \right] \underline{\tilde{s}} \circ \frac{\partial \rho}{\partial \bar{\xi}}, \quad (5.43)$$

### Tangent contact matrix

To evaluate the elementary tangent contact matrix, it is necessary to take a partial derivative of the residual v-vector  $\underline{\underline{R}}_x^c$  from (5.32), however it is more convenient to start from the variation of the virtual work  $\Delta\delta W_i^c$

$$\Delta\delta W_i^c = \Delta P_n \delta g_n + P_n \Delta\delta g_n + \Delta \underline{\underline{F}}_t \cdot \delta \underline{\underline{g}}_t + \underline{\underline{F}}_t \cdot \Delta\delta \underline{\underline{g}}_t \quad (5.44)$$

Since the normal contact force  $P_n$  is a function of the normal gap  $g_n$ , its derivative is evaluated as follows

$$P_n = \int_{\partial\Omega^i} -\varepsilon_n \langle -g_n \rangle d\partial\Omega^i \Rightarrow \Delta P_n = \begin{cases} \int_{\partial\Omega^i} \varepsilon_n \underline{\underline{G}} \cdot \delta \underline{\underline{x}} d\partial\Omega^i, & g_n < 0; \\ 0, & g_n \geq 0 \end{cases}$$

and by replacing the integral

$$\Delta P_n = \delta \underline{\underline{x}}^\circ \left[ \varepsilon_n \partial\Omega^i \underline{\underline{G}} \right], \quad g_n < 0$$

Substituting of the last expressions into (5.44) and considering the first two terms gives

$$\Delta P_n \delta g_n + P_n \Delta\delta g_n = \delta \underline{\underline{x}}^\circ \left[ \varepsilon_n \partial\Omega^i \left( \underline{\underline{G}} \otimes \underline{\underline{G}} - \langle -g_n \rangle \underline{\underline{H}} \right) \right]^\circ \delta \underline{\underline{x}} \quad (5.45)$$

Instead of the variation of the tangential force, a dot product with the tangential sliding should be considered, slip and stick have to be also distinguished.

$$\Delta(\underline{\underline{F}}_t \cdot \delta \underline{\underline{g}}_t),$$

where

$$\delta \underline{\underline{g}}_t = \frac{\partial \rho}{\partial \underline{\underline{\xi}}} \circ \delta \underline{\underline{\xi}}$$

according to (5.38) (increment indices are omitted) we get

$$\begin{cases} \underline{\underline{F}}_t \cdot \delta \underline{\underline{g}}_t = \varepsilon_t \partial\Omega_i (\underline{\underline{\xi}} - \underline{\underline{\xi}}^\bullet) \circ \delta \underline{\underline{\xi}}, & \text{stick} \\ \underline{\underline{F}}_t \cdot \delta \underline{\underline{g}}_t = \mu \varepsilon_n \partial\Omega_i \langle -g_n \rangle \underline{\underline{s}}_t \cdot \frac{\partial \rho}{\partial \underline{\underline{\xi}}} \circ \delta \underline{\underline{\xi}}, & \text{slip} \end{cases} \quad (5.46)$$

The variation of these expressions results in the following

$$\Delta(\underline{\underline{F}}_t \cdot \delta \underline{\underline{g}}_t) = \varepsilon_t \partial\Omega_i \Delta \underline{\underline{\xi}} \circ \delta \underline{\underline{\xi}} + \varepsilon_t \partial\Omega_i (\underline{\underline{\xi}} - \underline{\underline{\xi}}^\bullet) \circ \Delta \delta \underline{\underline{\xi}}, \quad \text{stick} \quad (5.47)$$

$$\begin{aligned} \Delta(\underline{\underline{F}}_t \cdot \delta \underline{\underline{g}}_t) &= -\mu \varepsilon_n \partial\Omega_i \Delta g_n \underline{\underline{s}}_t \cdot \frac{\partial \rho}{\partial \underline{\underline{\xi}}} \circ \delta \underline{\underline{\xi}} + \mu \varepsilon_n \partial\Omega_i \langle -g_n \rangle \Delta \underline{\underline{s}}_t \cdot \frac{\partial \rho}{\partial \underline{\underline{\xi}}} \circ \delta \underline{\underline{\xi}} + \\ &+ \mu \varepsilon_n \partial\Omega_i \langle -g_n \rangle \underline{\underline{s}}_t \cdot \Delta \frac{\partial \rho}{\partial \underline{\underline{\xi}}} \circ \delta \underline{\underline{\xi}} + \mu \varepsilon_n \partial\Omega_i \langle -g_n \rangle \underline{\underline{s}}_t \cdot \frac{\partial \rho}{\partial \underline{\underline{\xi}}} \circ \Delta \delta \underline{\underline{\xi}}, \quad \text{slip}, \quad g_n < 0 \end{aligned} \quad (5.48)$$

## 5.4 Residual vectors and tangent matrices

The expression for stick appears directly in a convenient form, whereas the slip expression requires additional computations. Carrying (5.34), the first right hand term in (5.48) becomes

$$-\mu \varepsilon_n \partial \Omega_i \Delta g_n \underline{s}_t \cdot \frac{\partial \rho}{\partial \underline{\xi}} \circ \delta \underline{\xi} = -\frac{\mu \varepsilon_n \varepsilon_t \partial \Omega_i^2}{\|\underline{F}_t^{\text{trial}}\|} \Delta g_n (\underline{\xi} - \underline{\xi}^\bullet) \circ \delta \underline{\xi} \quad (5.49)$$

Using (5.43) allows to expand the second term in (5.48)

$$\mu \varepsilon_n \partial \Omega_i \langle -g_n \rangle \Delta \underline{s}_t \cdot \frac{\partial \rho}{\partial \underline{\xi}} \circ \delta \underline{\xi} = \langle -g_n \rangle \frac{\mu \varepsilon_n \varepsilon_t \partial \Omega_i^2}{\|\underline{F}_t^{\text{trial}}\|} \underline{\xi} \circ \left[ \Delta \underline{\xi} - \Delta \frac{\partial \rho}{\partial \underline{\xi}} \cdot \frac{\overline{\partial \rho}}{\partial \underline{\xi}} \circ (\underline{\xi} - \underline{\xi}^\bullet) \right] \underline{\xi} \circ \underline{\Delta} \circ \delta \underline{\xi} \quad (5.50)$$

Expansion of the third term in (5.48) follows directly from substituting (5.34) for the unit sliding vector

$$\mu \varepsilon_n \partial \Omega_i \langle -g_n \rangle \underline{s}_t \cdot \Delta \frac{\partial \rho}{\partial \underline{\xi}} \circ \delta \underline{\xi} = \langle -g_n \rangle \frac{\mu \varepsilon_n \varepsilon_t \partial \Omega_i^2}{\|\underline{F}_t^{\text{trial}}\|} (\underline{\xi} - \underline{\xi}^\bullet) \circ \frac{\overline{\partial \rho}}{\partial \underline{\xi}} \cdot \Delta \frac{\partial \rho}{\partial \underline{\xi}} \circ \delta \underline{\xi} \quad (5.51)$$

The same procedure for the last term in (5.48) yields

$$\mu \varepsilon_n \partial \Omega_i \langle -g_n \rangle \underline{s}_t \cdot \frac{\partial \rho}{\partial \underline{\xi}} \circ \Delta \delta \underline{\xi} = \langle -g_n \rangle \frac{\mu \varepsilon_n \varepsilon_t \partial \Omega_i^2}{\|\underline{F}_t^{\text{trial}}\|} (\underline{\xi} - \underline{\xi}^\bullet) \circ \Delta \delta \underline{\xi} \quad (5.52)$$

The variation of the basis vectors can be expressed directly through the variation of the nodal coordinate vectors if one makes use of (5.19)

$$\underline{\rho} = \underline{\phi}(0, \underline{\xi}) \circ \underline{x} \quad \Rightarrow \quad \Delta \frac{\partial \rho}{\partial \underline{\xi}} = \frac{\partial \underline{\phi}(0, \underline{\xi})}{\partial \underline{\xi}} \circ \Delta \underline{x} \quad (5.53)$$

Together with (5.53), we use the following expressions

$$\delta g_n = \underline{\underline{G}} \cdot \delta \underline{x}, \quad \delta \underline{\xi} = \underline{\underline{T}} \cdot \delta \underline{x}, \quad \Delta \underline{\xi} = \underline{\underline{T}} \cdot \Delta \underline{x}, \quad \Delta \delta \underline{\xi} = \Delta \underline{x} \cdot \underline{\underline{S}} \cdot \delta \underline{x}, \quad (5.54)$$

where  $\underline{\underline{S}}$  is a v-t-tensor (precisely two t-tensors gathered in v-vector) connecting nodal variation vectors with the second variation of the convective coordinate v-scalar

$$\Delta \delta \underline{\xi} = \Delta \underline{x} \cdot \underline{\underline{S}} \cdot \delta \underline{x}$$

all expressions for  $\underline{\underline{G}}$ ,  $\underline{\underline{T}}$ ,  $\underline{\underline{H}}$  and  $\underline{\underline{S}}$  can be found in Section 2.4 in Eq. (2.99), (2.102), (2.105) and (2.111) respectively. Careful grouping of terms in (5.49)–(5.52) and substituting of (5.53) and (5.54) leads to the following expression for the variation of the tangential part for stick and slip

Stick

$$\Delta(\underline{F}_t \cdot \delta \underline{g}_t) = \Delta \underline{x} \circ \left[ \varepsilon_t \partial \Omega_i \left( \underline{T}_{\otimes}^{\circ} \underline{T} + (\underline{\xi} - \underline{\xi}^{\bullet}) \circ \underline{S} \right) \right] \circ \delta \underline{x} \quad (5.55)$$

Slip

$$\begin{aligned} \Delta(\underline{F}_t \cdot \delta \underline{g}_t) = \Delta \underline{x} \circ \left[ \frac{\mu \varepsilon_n \varepsilon_t \partial \Omega_i^2}{\|\underline{F}_t^{\text{trial}}\|} \left\{ -\underline{G}_{\otimes} \underline{T} \circ (\underline{\xi} - \underline{\xi}^{\bullet}) + \right. \right. \\ \left. \left. + \langle -g_n \rangle \left( \underline{\tilde{S}} \circ \underline{T}_{\otimes} \underline{T} \circ \underline{A} \circ \underline{\tilde{S}} - \underline{\tilde{S}} \circ \frac{\partial \phi}{\partial \underline{\xi}} \boxtimes \frac{\overline{\partial \rho}}{\partial \underline{\xi}} \circ (\underline{\xi} - \underline{\xi}^{\bullet}) \underline{\tilde{S}} \circ \underline{A} \circ \underline{T} + \right. \right. \right. \\ \left. \left. \left. + (\underline{\xi} - \underline{\xi}^{\bullet}) \circ \frac{\overline{\partial \rho}}{\partial \underline{\xi}} \boxtimes \frac{\partial \phi}{\partial \underline{\xi}} \circ \underline{T} + (\underline{\xi} - \underline{\xi}^{\bullet}) \circ \underline{S} \right) \right\} \right] \circ \delta \underline{x} \quad (5.56) \end{aligned}$$

Grouping of (5.45) with (5.55) or (5.56) and getting rid of nodal variations gives the tangential contact t-tensor for stick and slip states respectively

$$\underline{K}_{xx}^{c \text{ stick}} = \varepsilon_n \partial \Omega^i \left( \underline{G}_{\otimes} \underline{G} - \langle -g_n \rangle \underline{H} \right) + \varepsilon_t \partial \Omega_i \left( \underline{T}_{\otimes}^{\circ} \underline{T} + (\underline{\xi} - \underline{\xi}^{\bullet}) \circ \underline{S} \right) \quad (5.57)$$

and

$$\begin{aligned} \underline{K}_{xx}^{c \text{ slip}} = \varepsilon_n \partial \Omega^i \left( \underline{G}_{\otimes} \underline{G} - \langle -g_n \rangle \underline{H} \right) + \frac{\mu \varepsilon_n \varepsilon_t \partial \Omega_i^2}{\|\underline{F}_t^{\text{trial}}\|} \left\{ -\underline{G}_{\otimes} \underline{T} \circ (\underline{\xi} - \underline{\xi}^{\bullet}) + \right. \\ \left. + \langle -g_n \rangle \left( \underline{\tilde{S}} \circ \underline{T}_{\otimes} \underline{T} \circ \underline{A} \circ \underline{\tilde{S}} - \underline{\tilde{S}} \circ \frac{\partial \phi}{\partial \underline{\xi}} \boxtimes \frac{\overline{\partial \rho}}{\partial \underline{\xi}} \circ (\underline{\xi} - \underline{\xi}^{\bullet}) \underline{\tilde{S}} \circ \underline{A} \circ \underline{T} + \right. \right. \\ \left. \left. + (\underline{\xi} - \underline{\xi}^{\bullet}) \circ \frac{\overline{\partial \rho}}{\partial \underline{\xi}} \boxtimes \frac{\partial \phi}{\partial \underline{\xi}} \circ \underline{T} + (\underline{\xi} - \underline{\xi}^{\bullet}) \circ \underline{S} \right) \right\} \quad (5.58) \end{aligned}$$

### 5.4.3 Augmented Lagrangian method: frictionless case

#### Residual vector

Virtual work due to frictionless contact was given in (4.117) for continuous problems. The significant difference of the coupled augmented Lagrangian method from Lagrange multiplier and penalty methods is that all constructed contact elements contribute to the virtual work of the system independently if the gap between slave node and master segment is open or closed, i.e. contact element is inactive or active. This fact ensures the smoothness of the energy potential and the continuity of the virtual work. However, as was shown on a simple example in Paragraph 4.7.3, the inactive contact elements increase significantly the condition number of the tangent matrix of the system. That is why the less useless contact elements have been formed during the detection step, the better it is for the resolution step. So a careful detection procedure is highly recommended, especially for the augmented Lagrangian method.

Remark that the difference between frictional (containing several complementary degrees of freedom for Lagrange multipliers) and frictionless (containing only one complementary degree of freedom) contact elements should be made at the stage of their creation to accelerate the solution.

Below the contribution to the virtual energy of the  $i$ -th NTS contact element is given for contact and non-contact status respectively

$$\delta W_i^c = \begin{cases} \int_{\partial\Omega_i} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n d\partial\Omega_i, & \hat{\lambda}_n \leq 0 \\ \int_{\partial\Omega_i} -\frac{1}{\varepsilon_n} \lambda_n \delta \lambda_n d\partial\Omega_i, & \hat{\lambda}_n > 0 \end{cases} \quad (5.59)$$

We remind that the hat denotes the augmented Lagrange multipliers

$$\hat{\lambda}_n = \lambda_n + \varepsilon_n g_n$$

As previously the variation of the normal gap  $\delta g_n$  is replaced by the following double dot product

$$\delta g_n = \underline{\underline{\mathbf{G}}}^\circ \cdot \underline{\underline{\delta \mathbf{x}}} \quad (5.60)$$

By grouping terms in matrices and replacing the integral by master segment area we get

$$\delta W_i^c = \begin{cases} \left[ \begin{array}{c} \hat{\lambda}_n \underline{\underline{\mathbf{G}}} \partial\Omega_i \\ g_n \partial\Omega_i \end{array} \right]^T \cdot \left[ \begin{array}{c} \underline{\underline{\delta \mathbf{x}}} \\ \delta \lambda_n \end{array} \right], & \hat{\lambda}_n \leq 0 \\ \left[ \begin{array}{c} 0 \\ -\frac{\partial\Omega_i}{\varepsilon_n} \lambda_n \end{array} \right]^T \cdot \left[ \begin{array}{c} \underline{\underline{\delta \mathbf{x}}} \\ \delta \lambda_n \end{array} \right], & \hat{\lambda}_n > 0 \end{cases}$$

The residual v-vector for the primal variables (nodal displacement vectors) and

the scalar residual component for the only dual variable takes then the forms

$$\begin{cases} \begin{bmatrix} \hat{\lambda}_n \underline{\underline{G}} \partial \Omega_i \\ g_n \partial \Omega_i \end{bmatrix}^T, & \hat{\lambda}_n \leq 0, \text{ contact} \\ \begin{bmatrix} 0 \\ -\frac{\partial \Omega_i}{\varepsilon_n} \lambda_n \end{bmatrix}^T, & \hat{\lambda}_n > 0, \text{ non-contact} \end{cases} \quad (5.61)$$

### Tangent contact matrix

To get the elementary tangent contact matrix, we take a variation of the virtual work of the  $i$ -th contact element 5.59

$$\Delta \delta W_i^c = \begin{cases} \int_{\partial \Omega_i} \Delta \hat{\lambda}_n \delta g_n + \hat{\lambda}_n \Delta \delta g_n + \Delta g_n \delta \lambda_n d\partial \Omega_i, & \hat{\lambda}_n \leq 0 \\ \int_{\partial \Omega_i} -\frac{1}{\varepsilon_n} \Delta \lambda_n \delta \lambda_n d\partial \Omega_i, & \hat{\lambda}_n > 0 \end{cases} \quad (5.62)$$

As Lagrange multiplier is an independent variable, its second variations vanishes. Expanding the variation of the augmented Lagrange multiplier

$$\Delta \hat{\lambda}_n = \Delta \lambda_n + \varepsilon_n \Delta g_n$$

and using the expression for the variation of the normal gap (5.60) we can extract the tangent matrix from the resulting expression

$$\Delta \delta W_i^c = \begin{cases} \begin{bmatrix} \Delta \underline{\underline{x}} \\ \Delta \lambda_n \end{bmatrix}^T \circ \begin{bmatrix} \varepsilon_n \partial \Omega_i \underline{\underline{G}} \otimes \underline{\underline{G}} + \hat{\lambda}_n \underline{\underline{H}} & \partial \Omega_i \underline{\underline{G}} \\ \partial \Omega_i \underline{\underline{G}} & 0 \end{bmatrix} \circ \begin{bmatrix} \delta \underline{\underline{x}} \\ \delta \lambda_n \end{bmatrix}, & \hat{\lambda}_n \leq 0 \\ \begin{bmatrix} \Delta \underline{\underline{x}} \\ \Delta \lambda_n \end{bmatrix}^T \circ \begin{bmatrix} 0 & 0 \\ 0 & -\frac{\partial \Omega_i}{\varepsilon_n} \end{bmatrix} \circ \begin{bmatrix} \delta \underline{\underline{x}} \\ \delta \lambda_n \end{bmatrix}, & \hat{\lambda}_n > 0 \end{cases}, \quad (5.63)$$

where  $\underline{\underline{H}}$  is a t-tensor connecting nodal variation vectors with the second variation of the normal gap

$$\Delta \delta g_n = \Delta \underline{\underline{x}} \circ \underline{\underline{H}} \circ \delta \underline{\underline{x}}$$

All needed expressions  $\underline{\underline{G}}$  and  $\underline{\underline{H}}$  for arbitrary discretization can be found in Section 2.4. The resulting elemental tangent contact matrices for contact and non-contact statuses are given below. These matrices are symmetric.

$$K_{\text{contact}}^c = \begin{bmatrix} \varepsilon_n \partial \Omega_i \underline{\underline{G}} \otimes \underline{\underline{G}} + \hat{\lambda}_n \underline{\underline{H}} & \partial \Omega_i \underline{\underline{G}} \\ \partial \Omega_i \underline{\underline{G}} & 0 \end{bmatrix}, \quad \hat{\lambda}_n \leq 0, \text{ contact} \quad (5.64)$$

$$K_{\text{non-contact}}^c = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{\partial \Omega_i}{\varepsilon_n} \end{bmatrix}, \quad \hat{\lambda}_n > 0, \text{ non-contact} \quad (5.65)$$

## 5.4 Residual vectors and tangent matrices

The structure of these matrices can be depicted by blocks

$$K^c = \begin{bmatrix} \begin{bmatrix} \nabla_x \nabla_x & \nabla_{\lambda_n} \nabla_x \\ \nabla_x \nabla_{\lambda_n} & \nabla_{\lambda_n} \nabla_{\lambda_n} \end{bmatrix} \\ \begin{bmatrix} \nabla_x \nabla_x & \nabla_{\lambda_n} \nabla_x \\ \nabla_x \nabla_{\lambda_n} & \nabla_{\lambda_n} \nabla_{\lambda_n} \end{bmatrix} \end{bmatrix} (l_n + l_t) \quad (5.66)$$

$\begin{matrix} [N \times N] & [N \times 1] \\ [1 \times N] & [1 \times 1] \end{matrix}$

where  $n \times m$  in each block designates the number of “strings”  $n$  and the number of “columns”  $m$ , where in s-structure representation  $N = D(M + 1)$  and in scalar representation  $N = D(M + 1)$ :  $D$  is the dimension of the problem and  $M$  is the number of nodes on the master segment,  $M + 1$  is a total number of geometrical nodes of the contact element.

### 5.4.4 Augmented Lagrangian method: frictional case

#### Residual vector

The balance of virtual work for frictional contact coming from the variation of the augmented Lagrangian functional was stated in (4.116). The contribution of the frictional contact interface extracted from (4.116) being splitted into three integrals over slip, stick and non-contact zones is stated below

$$\begin{aligned} \delta W^c = & \int_{\Gamma_c^1} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n - \mu \hat{\sigma}_n \frac{\hat{\lambda}_t}{\|\hat{\lambda}_t\|} \cdot \delta \underline{g}_t - \frac{1}{\varepsilon_t} \left( \underline{\lambda}_t + \mu \hat{\sigma}_n \frac{\hat{\lambda}_t}{\|\hat{\lambda}_t\|} \right) \cdot \delta \underline{\lambda}_t d\Gamma_c^1 + \\ & + \int_{\Gamma_c^1} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n + \underline{\lambda}_t \cdot \delta \underline{g}_t + \underline{g}_t \cdot \delta \underline{\lambda}_t d\Gamma_c^1 + \int_{\Gamma_c^1 \setminus \bar{\Gamma}_t^1} -\frac{1}{\varepsilon_n} \lambda_n \delta \lambda_n - \frac{1}{\varepsilon_t} \underline{\lambda}_t \cdot \delta \underline{\lambda}_t d\Gamma_c^1 \end{aligned} \quad (5.67)$$

A given contact element can be in one of three states: slip, stick or non-contact. Depending on the status, its contribution to the virtual work of the system is given by one of the three possible integrals stated below.

$$\delta W_i^c = \begin{cases} \int_{\partial\Omega_i} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n - \mu \hat{\sigma}_n \frac{\hat{\lambda}_t}{\|\hat{\lambda}_t\|} \cdot \delta \underline{g}_t - \frac{1}{\varepsilon_t} \left( \underline{\lambda}_t + \mu \hat{\sigma}_n \frac{\hat{\lambda}_t}{\|\hat{\lambda}_t\|} \right) \cdot \delta \underline{\lambda}_t d\partial\Omega_i, & \underbrace{\|\hat{\lambda}_t\| > -\mu \hat{\sigma}_n}_{\text{slip}} \\ \int_{\partial\Omega_i} \hat{\lambda}_n \delta g_n + g_n \delta \lambda_n + \underline{\lambda}_t \cdot \delta \underline{g}_t + \underline{g}_t \cdot \delta \underline{\lambda}_t d\partial\Omega_i, & \underbrace{\|\hat{\lambda}_t\| \leq -\mu \hat{\sigma}_n}_{\text{stick}} \\ \int_{\partial\Omega_i} -\frac{1}{\varepsilon_n} \lambda_n \delta \lambda_n - \frac{1}{\varepsilon_t} \underline{\lambda}_t \cdot \delta \underline{\lambda}_t d\partial\Omega_i, & \underbrace{\hat{\sigma}_n > 0}_{\text{non-contact}} \end{cases} \quad (5.68)$$



To extract the residual s-structures, we make use of the relations between the variation of the nodal coordinate vectors of the contact element and the corresponding variations of the geometrical quantities: see (5.60) and

$$\delta \underline{g}_t = \frac{\partial \underline{p}}{\partial \underline{\xi}} \circ \delta \underline{\xi} = \frac{\partial \underline{p}}{\partial \underline{\xi}} \circ \underline{T} \circ \delta \underline{x} \quad (5.69)$$

Further let us replace the sliding direction by  $\hat{\underline{s}}$

$$\hat{\underline{s}} = \frac{\hat{\underline{\lambda}}_t}{\|\hat{\underline{\lambda}}_t\|}, \quad (5.70)$$

this sliding can be presented in the contravariant basis as well as the augmented Lagrange multiplier vector corresponding to the tangential contact stress

$$\hat{\underline{s}} = \left( \hat{\underline{s}} \cdot \frac{\partial \underline{p}}{\partial \underline{\xi}} \right) \circ \frac{\partial \underline{p}}{\partial \underline{\xi}} = \hat{\underline{s}} \circ \frac{\partial \underline{p}}{\partial \underline{\xi}}, \quad \hat{\underline{\lambda}}_t = \left( \hat{\underline{\lambda}}_t \cdot \frac{\partial \underline{p}}{\partial \underline{\xi}} \right) \circ \frac{\partial \underline{p}}{\partial \underline{\xi}} = \hat{\underline{\lambda}}_t \circ \frac{\partial \underline{p}}{\partial \underline{\xi}} \quad (5.71)$$

Regardless possible confusions, the Lagrange multiplier  $\underline{\lambda}_t$  should be presented in the contravariant basis and its variation in the covariant one

$$\underline{\lambda}_t = \left( \underline{\lambda}_t \cdot \frac{\partial \underline{p}}{\partial \underline{\xi}} \right) \circ \frac{\partial \underline{p}}{\partial \underline{\xi}} = \underline{\lambda}_t \circ \frac{\partial \underline{p}}{\partial \underline{\xi}}, \quad \delta \underline{\lambda}_t = \underline{\lambda}_t \circ \frac{\partial \underline{p}}{\partial \underline{\xi}} \quad (5.72)$$

and finally the incremental change of the tangential relative sliding vector is also given in the contravariant basis

$$\underline{g}_t \sim \underline{g}_t(t^{i+1} - t^i) \Rightarrow \underline{g}_t = (\underline{\xi}^{i+1} - \underline{\xi}^i) \circ \frac{\partial \underline{p}}{\partial \underline{\xi}}. \quad (5.73)$$

Note that this vector should be zero in stick state. Note also that this vector appears in split of the augmented Lagrange multiplier  $\hat{\underline{\lambda}}_t$ . After substituting all these expressions into (5.68) and replacing the integral by the area of the master segment  $\partial\Omega_i$ , we get

$$\begin{aligned} \delta W_{i \text{ slip}}^c &= \begin{bmatrix} \left( \hat{\underline{\lambda}}_n \underline{G} - \mu \hat{\sigma}_n \hat{\underline{s}} \circ \underline{T} \right) \partial\Omega_i \\ g_n \partial\Omega_i \\ -\frac{\partial\Omega_i}{\varepsilon_t} (\underline{\lambda}_t + \mu \hat{\sigma}_n \hat{\underline{s}}) \end{bmatrix}^T \circ \begin{bmatrix} \delta \underline{x} \\ \delta \lambda_n \\ \delta \underline{\lambda}_t \end{bmatrix}, \quad \|\hat{\underline{\lambda}}_t\| \geq \hat{\sigma}_n, \hat{\sigma}_n \leq 0, \text{ slip} \\ \delta W_{i \text{ stick}}^c &= \begin{bmatrix} \left( \hat{\underline{\lambda}}_n \underline{G} + \hat{\underline{\lambda}}_t \circ \underline{T} \right) \partial\Omega_i \\ g_n \partial\Omega_i \\ (\underline{\xi}^{i+1} - \underline{\xi}^i) \partial\Omega_i \end{bmatrix}^T \circ \begin{bmatrix} \delta \underline{x} \\ \delta \lambda_n \\ \delta \underline{\lambda}_t \end{bmatrix}, \quad \|\hat{\underline{\lambda}}_t\| < \hat{\sigma}_n, \hat{\sigma}_n \leq 0, \text{ stick} \end{aligned}$$

$$\delta W_{i \text{ non-contact}}^c = \begin{bmatrix} 0 \\ -\frac{\partial \Omega_i \lambda_n}{\varepsilon_n} \\ -\frac{\partial \Omega_i \tilde{\lambda}_t}{\varepsilon_t} \end{bmatrix}^T \circ \begin{bmatrix} \delta \underline{x} \\ \delta \lambda_n \\ \delta \tilde{\lambda}_t \end{bmatrix}, \quad \|\hat{\lambda}_t\| < \hat{\sigma}_n, \hat{\sigma}_n \leq 0, \text{ stick}$$

The left hand matrices in these expressions represent the residual vectors for slip, stick and non-contact respectively.

### Tangent contact matrix

To derive the tangent contact matrix, we start from the variation of the virtual work contributions arising from one contact element (5.68). Slip, stick and non-contact states will be considered separately. As Lagrange multipliers are independent variables, their second variations vanish. It is worth mentioning that the augmented contact pressure  $\hat{\sigma}_n = \sigma_n + \varepsilon_n g_n$ , previously assumed known, now becomes the unknown variable and is replaced by the augmented Lagrange multiplier  $\hat{\lambda}_n = \lambda_n + \varepsilon_n g_n$ .

### Slip

$$\begin{aligned} \Delta \delta W_{i \text{ slip}}^c &= \int_{\partial \Omega_i} \Delta \hat{\lambda}_n \delta g_n + \hat{\lambda}_n \Delta \delta g_n + \Delta g_n \delta \lambda_n - \mu \Delta \hat{\lambda}_n \hat{\underline{s}} \cdot \delta \underline{g}_t - \mu \hat{\lambda}_n \frac{\Delta \hat{\lambda}_t}{\|\hat{\lambda}_t\|} \cdot (\underline{\underline{I}} - \hat{\underline{s}} \otimes \hat{\underline{s}}) \cdot \delta \underline{g}_t - \\ &- \mu \hat{\lambda}_n \hat{\underline{s}} \circ \Delta \delta \underline{\xi} - \frac{1}{\varepsilon_t} \Delta \underline{\lambda}_t \cdot \delta \underline{\lambda}_t - \frac{\mu \hat{\lambda}_n}{\varepsilon_t} \frac{\Delta \hat{\lambda}_t}{\|\hat{\lambda}_t\|} \cdot (\underline{\underline{I}} - \hat{\underline{s}} \otimes \hat{\underline{s}}) \cdot \delta \underline{\lambda}_t - \frac{1}{\varepsilon_t} \mu \Delta \hat{\lambda}_n \hat{\underline{s}} \cdot \delta \underline{\lambda}_t d\partial \Omega_i \end{aligned} \quad (5.74)$$

Let us get rid of dot products replacing them by s-dot products by means of the expressions derived in (5.69)–(5.73)

$$\begin{aligned} \hat{\underline{s}} \cdot \delta \underline{g}_t &= \hat{\underline{s}} \circ \delta \underline{\xi}, \quad \frac{\Delta \hat{\lambda}_t}{\|\hat{\lambda}_t\|} \cdot (\underline{\underline{I}} - \hat{\underline{s}} \otimes \hat{\underline{s}}) \cdot \delta \underline{g}_t = \frac{\Delta \hat{\lambda}_t}{\|\hat{\lambda}_t\|} \circ (\underline{\underline{A}} - \hat{\underline{s}} \boxtimes \hat{\underline{s}}) \circ \delta \underline{\xi} \\ \Delta \underline{\lambda}_t \cdot \delta \underline{\lambda}_t &= \Delta \tilde{\lambda}_t \circ \underline{\underline{A}} \circ \delta \tilde{\lambda}_t, \quad \frac{\Delta \hat{\lambda}_t}{\|\hat{\lambda}_t\|} \cdot (\underline{\underline{I}} - \hat{\underline{s}} \otimes \hat{\underline{s}}) \cdot \delta \underline{\lambda}_t = \frac{\Delta \hat{\lambda}_t}{\|\hat{\lambda}_t\|} \circ (\underline{\underline{A}} - \hat{\underline{s}} \boxtimes \hat{\underline{s}}) \circ \delta \tilde{\lambda}_t \\ \hat{\underline{s}} \cdot \delta \underline{\lambda}_t &= \hat{\underline{s}} \circ \delta \tilde{\lambda}_t \end{aligned}$$

Expanding of the augmented Lagrange multipliers under the variation sign

$$\delta \hat{\lambda}_n = \delta \lambda_n + \varepsilon_n \delta g_n, \quad \delta \hat{\lambda}_t = \delta \underline{\lambda}_t + \varepsilon_t \delta \underline{g}_t \quad \Leftrightarrow \quad \delta \hat{\tilde{\lambda}}_t = \delta \tilde{\lambda}_t + \varepsilon_t \delta \underline{\xi}$$

Now we state the variation of the elemental virtual work in the expanded form

$$\begin{aligned}
 \Delta \delta W_{i \text{ slip}}^c = & \int_{\partial \Omega_i} \Delta \lambda_n \delta g_n + \varepsilon_n \Delta g_n \delta g_n + \hat{\lambda}_n \Delta \delta g_n + \Delta g_n \delta \lambda_n - \mu \Delta \lambda_n \hat{\underline{\underline{\mathbf{g}}}} \circ \delta \underline{\underline{\mathbf{x}}} - \mu \varepsilon_n \Delta g_n \hat{\underline{\underline{\mathbf{g}}}} \circ \delta \underline{\underline{\mathbf{x}}} - \\
 & - \mu \hat{\lambda}_n \frac{\Delta \underline{\underline{\lambda}}_t}{\|\hat{\underline{\underline{\lambda}}}_t\|} \circ \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right) \circ \delta \underline{\underline{\mathbf{x}}} - \mu \hat{\lambda}_n \varepsilon_t \frac{\Delta \underline{\underline{\mathbf{x}}}}{\|\hat{\underline{\underline{\lambda}}}_t\|} \circ \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right) \circ \delta \underline{\underline{\mathbf{x}}} - \mu \hat{\lambda}_n \hat{\underline{\underline{\mathbf{g}}}} \circ \Delta \delta \underline{\underline{\mathbf{x}}} - \\
 & - \frac{1}{\varepsilon_t} \Delta \underline{\underline{\lambda}}_t \circ \underline{\underline{\mathbf{A}}} \circ \delta \underline{\underline{\lambda}}_t - \frac{\mu \hat{\lambda}_n}{\varepsilon_t} \frac{\Delta \underline{\underline{\lambda}}_t}{\|\hat{\underline{\underline{\lambda}}}_t\|} \circ \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right) \circ \delta \underline{\underline{\lambda}}_t - \mu \hat{\lambda}_n \frac{\Delta \underline{\underline{\mathbf{x}}}}{\|\hat{\underline{\underline{\lambda}}}_t\|} \circ \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right) \circ \delta \underline{\underline{\lambda}}_t - \\
 & - \frac{1}{\varepsilon_t} \mu \Delta \lambda_n \hat{\underline{\underline{\mathbf{g}}}} \circ \delta \underline{\underline{\lambda}}_t - \mu \Delta g_n \hat{\underline{\underline{\mathbf{g}}}} \circ \delta \underline{\underline{\lambda}}_t d\partial \Omega_i
 \end{aligned} \tag{5.75}$$

Finally the geometrical variations are expressed through the variations of the nodal coordinate vectors and the integral is replaced simply by the area of the master segment, which results in a “matrix”

$$\Delta \delta W_{i \text{ slip}}^c = \begin{bmatrix} \Delta \underline{\underline{\mathbf{x}}} \\ \Delta \lambda_n \\ \Delta \underline{\underline{\lambda}}_t \end{bmatrix}^T \circ \begin{bmatrix} \underline{\underline{\mathbf{K}}}_{xx} & \underline{\underline{\mathbf{K}}}_{x\lambda_n} & \underline{\underline{\mathbf{K}}}_{x\lambda_t} \\ \underline{\underline{\mathbf{K}}}_{\lambda_n x} & \underline{\underline{\mathbf{K}}}_{\lambda_n \lambda_n} & \underline{\underline{\mathbf{K}}}_{\lambda_n \lambda_t} \\ \underline{\underline{\mathbf{K}}}_{\lambda_t x} & \underline{\underline{\mathbf{K}}}_{\lambda_t \lambda_n} & \underline{\underline{\mathbf{K}}}_{\lambda_t \lambda_t} \end{bmatrix} \circ \begin{bmatrix} \Delta \underline{\underline{\mathbf{x}}} \\ \Delta \lambda_n \\ \Delta \underline{\underline{\lambda}}_t \end{bmatrix}. \tag{5.76}$$

The components of this “matrix” are stated below. Blue terms in braces denote unsymmetrical components of the matrix

$$\begin{aligned}
 \underline{\underline{\mathbf{K}}}_{xx} &= \partial \Omega_i \left( \varepsilon_n \underline{\underline{\mathbf{G}}} \boxtimes \underline{\underline{\mathbf{G}}} + \hat{\lambda}_n \underline{\underline{\mathbf{H}}} - \left\{ \varepsilon_n \mu \underline{\underline{\mathbf{G}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \circ \underline{\underline{\mathbf{T}}} \right\} \right) - \\
 & - \frac{\mu \hat{\lambda}_n \varepsilon_t \partial \Omega_i}{\|\hat{\underline{\underline{\lambda}}}_t\|} \underline{\underline{\mathbf{T}}} \circ \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right) \circ \underline{\underline{\mathbf{T}}} - \mu \hat{\lambda}_n \partial \Omega_i \hat{\underline{\underline{\mathbf{g}}}} \circ \underline{\underline{\mathbf{S}}} \\
 \underline{\underline{\mathbf{K}}}_{x\lambda_n} &= \partial \Omega_i \underline{\underline{\mathbf{G}}} \\
 \underline{\underline{\mathbf{K}}}_{x\lambda_t} &= - \frac{\mu \hat{\lambda}_n \partial \Omega_i}{\|\hat{\underline{\underline{\lambda}}}_t\|} \underline{\underline{\mathbf{T}}} \circ \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right) - \left\{ \mu \partial \Omega_i \underline{\underline{\mathbf{G}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right\} \\
 \underline{\underline{\mathbf{K}}}_{\lambda_n x} &= \partial \Omega_i \underline{\underline{\mathbf{G}}} - \left\{ \partial \Omega_i \mu \hat{\underline{\underline{\mathbf{g}}}} \circ \underline{\underline{\mathbf{T}}} \right\} \\
 \underline{\underline{\mathbf{K}}}_{\lambda_n \lambda_n} &= 0 \\
 \underline{\underline{\mathbf{K}}}_{\lambda_n \lambda_t} &= - \left\{ \frac{\mu \partial \Omega_i}{\varepsilon_t} \hat{\underline{\underline{\mathbf{g}}}} \right\} \\
 \underline{\underline{\mathbf{K}}}_{\lambda_t x} &= - \frac{\mu \hat{\lambda}_n \partial \Omega_i}{\|\hat{\underline{\underline{\lambda}}}_t\|} \circ \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right) \circ \underline{\underline{\mathbf{T}}} \\
 \underline{\underline{\mathbf{K}}}_{\lambda_t \lambda_n} &= 0 \\
 \underline{\underline{\mathbf{K}}}_{\lambda_t \lambda_t} &= - \frac{\partial \Omega_i}{\varepsilon_t} \underline{\underline{\mathbf{A}}} - \frac{\mu \hat{\lambda}_n \partial \Omega_i}{\varepsilon_t \|\hat{\underline{\underline{\lambda}}}_t\|} \left( \underline{\underline{\mathbf{A}}} - \hat{\underline{\underline{\mathbf{g}}}} \boxtimes \hat{\underline{\underline{\mathbf{g}}}} \right)
 \end{aligned} \tag{5.77}$$

## 5.4 Residual vectors and tangent matrices

All s-structures arising from variations of the geometrical quantities, namely  $\underline{\underline{G}}, \underline{\underline{T}}, \underline{\underline{H}}, \underline{\underline{S}}$ , can be found in Section 2.4.

**Stick** In case of stick the elemental tangent contact matrix has a more simple form. As previously, we start from the variation of the contact element contribution (5.68, *stick*) to the virtual work of the system

$$\begin{aligned} \Delta \delta W_{i \text{ stick}}^c = & \int_{\partial \Omega_i} \Delta \lambda_n \delta g_n + \varepsilon_n \Delta g_n \delta g_n + \hat{\lambda}_n \Delta \delta g_n + \Delta g_n \delta \lambda_n + \Delta \hat{\lambda}_t \circ \underline{\underline{A}} \circ \delta \underline{\underline{\xi}} + \\ & + \varepsilon_t \Delta \underline{\underline{\xi}} \circ \underline{\underline{A}} \circ \delta \underline{\underline{\xi}} + \hat{\lambda}_t \circ \Delta \delta \underline{\underline{\xi}} + \Delta \underline{\underline{\xi}} \circ \underline{\underline{A}} \circ \delta \hat{\lambda}_t d\partial \Omega_i \end{aligned} \quad (5.78)$$

The replacement of the geometrical variations by dot products of s-structures with variation of the nodal coordinate vectors and replacement of the integral by the area of the master segment leads to the following structure

$$\Delta \delta W_{i \text{ stick}}^c = \begin{bmatrix} \Delta \underline{\underline{x}} \\ \Delta \lambda_n \\ \Delta \hat{\lambda}_t \end{bmatrix}^T \circ \begin{bmatrix} \partial \Omega_i \left( \varepsilon_n \underline{\underline{G}} \otimes \underline{\underline{G}} + \varepsilon_t \underline{\underline{T}} \circ \underline{\underline{A}} \circ \underline{\underline{T}} + \partial \Omega_i \hat{\lambda}_n \underline{\underline{H}} + \hat{\lambda}_t \circ \underline{\underline{S}} \right) & \partial \Omega_i \underline{\underline{G}} & \partial \Omega_i \underline{\underline{A}} \circ \underline{\underline{T}} \\ \partial \Omega_i \underline{\underline{G}} & 0 & 0 \\ \partial \Omega_i \underline{\underline{A}} \circ \underline{\underline{T}} & 0 & 0 \end{bmatrix} \circ \begin{bmatrix} \Delta \underline{\underline{x}} \\ \Delta \lambda_n \\ \Delta \hat{\lambda}_t \end{bmatrix} \quad (5.79)$$

This matrix is symmetric; for expressions of  $\underline{\underline{G}}, \underline{\underline{T}}, \underline{\underline{H}}, \underline{\underline{S}}$ , see Section 2.4.

**Non-contact** In case of non-contact, the tangent contact “matrix” can be stated immediately

$$K_{i \text{ non-contact}}^c = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{\partial \Omega_i}{\varepsilon_n} & 0 \\ 0 & 0 & -\frac{\partial \Omega_i}{\varepsilon_t} \underline{\underline{I}} \end{bmatrix} \quad (5.80)$$

Obviously this “matrix” is symmetric.

The structure of the derived matrices for frictional case can be depicted by

blocks

$$K^c = \begin{bmatrix} \begin{bmatrix} \nabla_x \nabla_x \\ \text{\textcolor{red}{[N \times N]}} \end{bmatrix} & \begin{bmatrix} \nabla_{\lambda_n} \nabla_x \\ \text{\textcolor{red}{[N \times 1]}} \end{bmatrix} & \begin{bmatrix} \nabla_{\lambda_t} \nabla_x \\ \text{\textcolor{red}{[N \times D]}} \end{bmatrix} \\ \begin{bmatrix} \nabla_x \nabla_{\lambda_n} \\ \text{\textcolor{red}{[1 \times N]}} \end{bmatrix} & \begin{bmatrix} \nabla_{\lambda_n} \nabla_{\lambda_n} \\ \text{\textcolor{red}{[1 \times 1]}} \end{bmatrix} & \begin{bmatrix} \nabla_{\lambda_n} \nabla_{\lambda_t} \\ \text{\textcolor{red}{[1 \times D]}} \end{bmatrix} \\ \begin{bmatrix} \nabla_x \nabla_{\lambda_t} \\ \text{\textcolor{red}{[D \times N]}} \end{bmatrix} & \begin{bmatrix} \nabla_{\lambda_n} \nabla_{\lambda_t} \\ \text{\textcolor{red}{[D \times 1]}} \end{bmatrix} & \begin{bmatrix} \nabla_{\lambda_t} \nabla_{\lambda_t} \\ \text{\textcolor{red}{[D \times D]}} \end{bmatrix} \end{bmatrix} (l_n + l_t) \quad (5.81)$$

where  $n \times m$  in each block designates the number of “strings”  $n$  and the number of “columns”  $m$ , where in s-structure representation  $N = D(M + 1)$  and in scalar representation  $N = D(M + 1)$ .  $D$  is the dimension of the problem and  $M$  is a number of nodes on the master segment,  $M + 1$  is a total number of geometrical nodes of the contact element.



**Remark 5.1 on the residual vector and the tangent matrix for the multi-face contact elements.**

If one uses multi-face contact elements proposed by [Heegaard 93], [Barboteu 02] and mentioned in Section 3.3.4, then the structure of the residual vector and the tangent matrix are slightly different. Among  $N^m$  master segments of the  $i$ -th contact element there is one active segment  $\partial\Omega_i^a$  on the current iteration

$$\partial\Omega_i^a \in \bigcup_{j=1}^{N^m} \partial\Omega_i^j$$

All master nodes of the contact element can be splitted into active nodes (denoted by an upper index  $a$ ), which are attached to the active master segment and passive – all other nodes denoted by an upper index  $p$ ). Active master nodes should be complemented by the slave node. Then the  $v$ -vector of the nodal coordinate vectors of the contact element can be written as

$$\underline{\underline{x}} \sim [\underline{\underline{x}}^a \quad \underline{\underline{x}}^p]^T$$

The residual contact vector and tangent contact matrix take the following forms

$$R_i^c \sim \begin{bmatrix} R_{x^a}^c \\ 0 \\ R_{\lambda_n}^c \\ R_{\lambda_t}^c \end{bmatrix}, \quad K_i^c \sim \begin{bmatrix} K_{x^a x^a} & 0 & K_{x^a \lambda_n} & K_{x^a \lambda_t} \\ 0 & 0 & 0 & 0 \\ K_{\lambda_n x^a} & 0 & K_{\lambda_n \lambda_n} & K_{\lambda_n \lambda_t} \\ K_{\lambda_t x^a} & 0 & K_{\lambda_t \lambda_n} & K_{\lambda_t \lambda_t} \end{bmatrix}$$

In case of smoothing of the master surface (see e.g. [Pietrzak 97],[Wriggers 01], more references can be found in [Wriggers 06]), the structure of the residual vectors and tangent matrices does not change, the only difference is that the  $v$ -vector of nodal coordinate vectors will consists of all the nodes of the contact element. To make use of the smoothed master surface, it remains only to construct a new set of shape functions  $\tilde{\phi}$ , corresponding to a smoothed master surface spanned on all master nodes, further this  $v$ -scalar of shape functions should be substituted into the expressions of  $\underline{\tilde{G}}, \underline{\tilde{T}}, \underline{\tilde{H}}, \underline{\tilde{S}}$  from Section 2.4. No additional computational efforts are needed.

## 5.5 Method of partial Dirichlet-Neumann boundary conditions

Here we give a brief explanation of the technique described in Paragraphs 4.1.1, 4.1.3, 4.2.3, 4.2.4 in the framework of the Finite Element Method. The main idea is to replace the geometrical constraints due to normal and frictional contact by *partial Dirichlet-Neumann* boundary conditions. Further we will refer to this method as Partial Dirichlet-Neumann (PDN) method. This technique is very advantageous in comparison to the standard methods described in the previous sections of this chapter (penalty, Lagrange multipliers, augmented Lagrangian method) because there is no need to evaluate the residual vectors and the tangent matrices. Here we confine ourself to the contact between a deformable body and a rigid surface. In this case, the rigid surface can be described by a smooth function and there is no need for contact detection or contact elements. This method can be extended to the case of two-body contact [Wriggers 06]. Moreover, coupled with a Lagrange multiplier method, it is equivalent to a *mortar method*, but contrary to Lagrange multiplier and augmented Lagrangian methods, which increase the number of unknowns in the system, the PDN method reduces the number of unknowns. Moreover the PDN method is trivial to integrate in a parallelized Finite Element code.

### 5.5.1 Description of the numerical technique

Standard methods are based on a search for the contact stress vectors ensuring the fulfillment of geometrical inequality constraints and equilibrium of the bodies – this approach leads to an inequality variational problem which has a limited use. If inequality constraints are replaced by equality constraints, i.e. active contact zone is assumed to be known, then the standard weak form can be constructed and implemented in a Finite Element code, as previously shown in this chapter. In any case, the problem consists in determining unknown contact stresses. The advantage of the PDN method is that, instead of seeking for contact stress vectors, the geometrical inequality constraints are imposed directly as partial Dirichlet boundary conditions and consequently the contact stress arises directly as a reaction. It remains to determine the active contact zone, this task will be discussed in the following section.

Four different classes of deformable-rigid contact can be distinguished:

1. frictionless contact with a rigid plane<sup>4</sup>;
2. frictional contact with a rigid plane;
3. frictionless contact with an arbitrary rigid surface (Signorini's problem);
4. frictional contact with an arbitrary rigid surface.

<sup>4</sup>by plane here we mean any surface at which any coordinate of problem reference frame takes a constant value, i.e. plane in Cartesian coordinates, cylinder in cylindrical coordinates or sphere in spherical coordinates.

To solve the first two classes (1,2), the Finite Element code should include the possibility of a dynamical update of boundary conditions. In addition to this feature, solving of remaining classes (3,4) needs to make use of the so-called Multi-Point boundary conditions (MPC). Further we will consider only the last classes of Signorini problem since they include the rigid plane as a subcase.

Briefly the Multi-Point boundary condition implies that a chosen degree of freedom – “slave” dof  $u^s$  is replaced by a linear combination of “master” dofs  $u_i^m$ ,  $i = 1, M$

$$u^s = \alpha^i u_i^m + \beta,$$

where  $\alpha^i$  and  $\beta$  are scalar coefficients. To replace the Hertz-Signorini-Moreau conditions in case of an arbitrary rigid surface, the MPC is used in such a way that one dof of the contacting node – slave dof – is expressed through the other dofs of the same contacting node (2 in 3D, 1 in 2D):

$$u^s = \underline{\alpha} \circ \underline{u}^m + \beta,$$

where  $\underline{u}^m$  is a v-scalar of master dofs,  $\underline{\alpha}$  is a v-scalar of the scalar coefficients and  $\beta$  is also a scalar coefficient. The slave dof can be chosen arbitrarily for each contacting node, however it is required that locally the coefficients

$$\underline{\alpha} < \infty.$$

According to Paragraph 4.2.3, this condition is equivalent to the requirement that  $\underline{e}^s \cdot \underline{n} \neq 0$ , where  $\underline{e}^s$  is a unit vector, along which the coordinate of the slave dof is measured, and  $\underline{n}$  is a normal to the rigid surface at projection point.

### 5.5.2 Frictionless case

As it was discussed in details in the cited paragraphs, the Hertz-Signorini-Moreau conditions can be replaced by MPC boundary conditions, which allow sliding of the contacting node only in the tangential plane. After introducing some notations, a general algorithm will be given. Let  $\underline{x}^i$  be a coordinate vector of a contacting node on the  $i$ -th iteration and  $\underline{x}^0$  the coordinate vector of this node at the previous increment, then the incremental displacement vector is given by

$$\underline{u}^i = \underline{x}^i - \underline{x}^0$$

The increment of the degrees of freedom is given by splitting of the vector  $\underline{u}$  into the reference frame basis

$$u_j^i = \underline{u}^i \cdot \underline{e}_j,$$

where  $\underline{e}_j$  is a set of basis vectors (normally in the Finite Element code an orthonormal set of basis vectors is used, so we will not distinguish basis and dual basis). The rigid surface in general can be described by a vector  $\underline{\rho}(\underline{\xi}) = \rho_j(\underline{\xi})\underline{e}_j$  with a normal  $\underline{n}$  pointing to the permitted area of body motion. Without any loss of generality, let us suppose that locally it exists a function  $f$  such that

$$z = f(x, y),$$



where  $x, y, z$  are coordinates of the surface points in the chosen coordinate system

$$x = \underline{\rho} \cdot \underline{e}_1, \quad y = \underline{\rho} \cdot \underline{e}_2, \quad z = \underline{\rho} \cdot \underline{e}_3.$$

Then the geometrical constraint can be rewritten in the classical form

$$x_3^i \geq f(x_1^i, x_2^i) \Leftrightarrow u_3^i \geq f(x_1^0 + u_1^i, x_2^0 + u_2^i) - x_3^0.$$

The tangential plane at the given point  $\{x^*, y^*\}$  is determined by the following equality

$$P : z = \frac{\partial f}{\partial x} \Big|_{\{x^*, y^*\}} (x - x^*) + \frac{\partial f}{\partial y} \Big|_{\{x^*, y^*\}} (y - y^*) + f(x^*, y^*),$$

after introducing the following notations

$$a = \frac{\partial f}{\partial x} \Big|_{\{x^*, y^*\}}, \quad b = \frac{\partial f}{\partial y} \Big|_{\{x^*, y^*\}}, \quad c = f(x^*, y^*)$$

the equation for the tangential plane can be rewritten as

$$P : z = a(x - x^*) + b(y - y^*) + c,$$

and the MPC to be imposed is given as

$$u_3 = au_1 + bu_2 + c - x_3^0,$$

where  $x_3^0$  is a  $z$  coordinate of the node at the beginning of the increment (see Fig. 5.18). Further it is necessary to check that there is no tension forces in the created contact interface. The reaction force  $\underline{R}$  appearing in the nodes, where the MPC boundary conditions have been imposed, has to be checked

$$\underline{R} \cdot \underline{n} \geq 0,$$

precisely the normal contact force should point in the same direction as the normal to the rigid surface, otherwise the imposed MPC at adhering node should be removed (see Fig. 5.19). The full algorithm is stated in the box below.

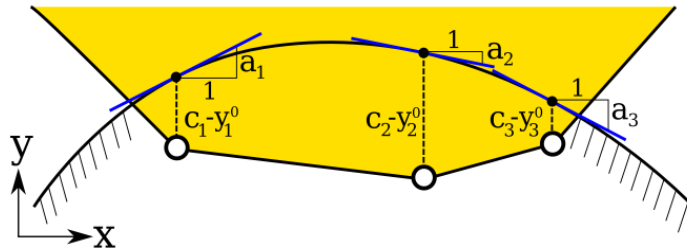


Figure 5.18: MPC boundary conditions  $y = ax + x - y^0$ .

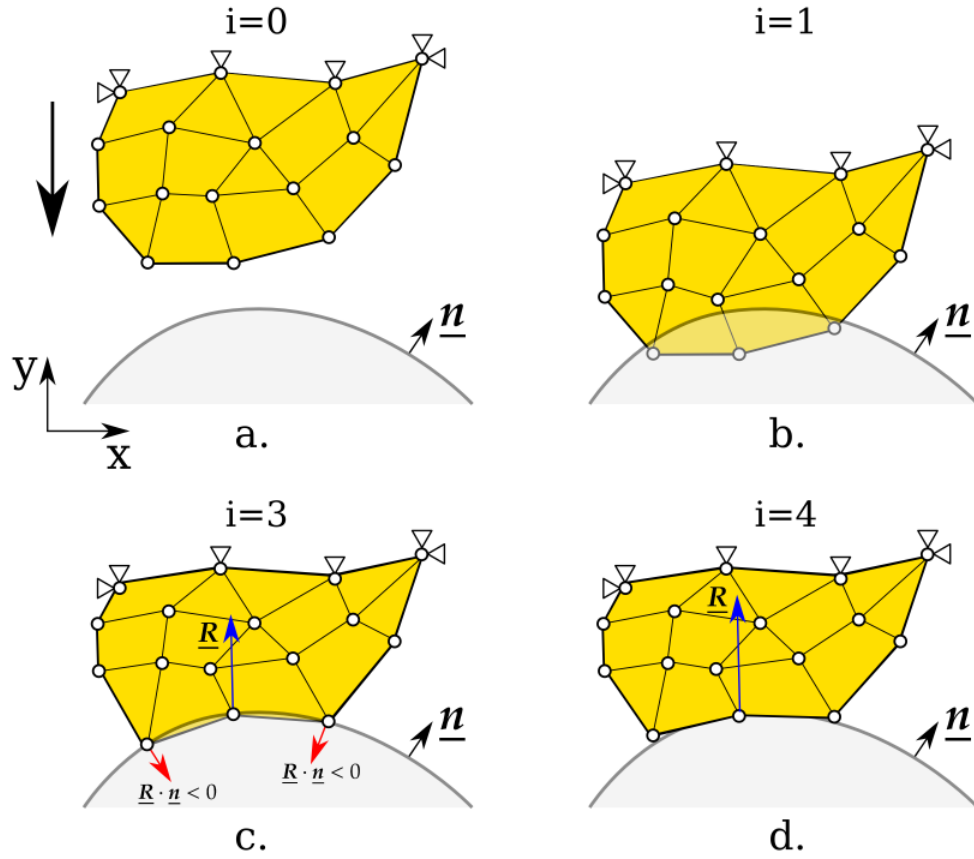


Figure 5.19: Illustration of the iterative process of MPC update procedure: **a** – initial configuration, **b** – first iteration, three MPC boundary conditions are imposed at penetrated nodes, **c** – after solution some nodes adhere to the surface, MPC at these nodes are removed, **d** – final solution.

```

i=0;
i<n;
i++;

```

**PDN algorithm for frictionless contact**

- Iterations ( $i = 0; i < i_{\max}; i++$ )
  1. Loop over possibly contacting nodes  $j$ 
    - **IF** penetration  $(x_3^i < f(x_1^i, x_2^i))$  **THEN**
      - \* impose MPC boundary condition
$$u_3^{i+1} = a^i u_1^{i+1} + b^i u_2^{i+1} + c^i - x_3^i$$
  2. Solve  $i \rightarrow i + 1$
  3. Check convergence:
    - loop over possibly contacting nodes  $j$ 
      - **IF** penetration  $(x_3^{i+1} < f(x_1^{i+1}, x_2^{i+1}))$  **THEN**
        - \* impose MPC boundary condition
      - $$u_3^{i+2} = a^{i+1} u_1^{i+2} + b^{i+1} u_2^{i+2} + c^{i+1} - x_3^{i+1}$$
      - **IF** adhesion  $\underline{R}^{i+1} \cdot \underline{n}^{i+1} < 0$  **THEN**
        - \* remove MPC
      - **IF** penetration or adhesion
        - \* **GOTO** 1
      - **ELSE** finish

**Active set search**

For simple geometries of contacting body and rigid plane: for example, sphere on plane or sphere on sphere, for reasonable time step this algorithm converges in one iteration. However, for certain geometries and/or boundary conditions the cycling between contact–non-contact states is possible. This is why we propose a simple algorithm which provides a fast determination of the active contact zone and ensures the absence of the cycling.

In case of undesirable adhesion the proposed technique suppresses the imposed MPC boundary conditions gradually. The idea consists in sorting of all nodes in the contact interface by their statuses (normal contact, adhesion and non-contact) and further analysis of the contact topology. Below the algorithm is explained in details.

Contacting surface consists of a set of segments

$$\Gamma_c = \bigcup_i \Gamma_{ci}.$$

Each segment has several nodes

$$\underline{x}_j^i \in \Gamma_{ci}$$

After the solution step (step 3 in the box above) to each node in the contact interface the corresponding status  $\alpha_j^i$  is assigned: normal contact  $\alpha_j^i < 0$ , adhesion  $\alpha_j^i > 0$  or non-contact  $\alpha_j^i = 0$ . If there are penetrations of nodes then new MPC boundary conditions are imposed. If there is an adhesion the detection algorithm is executed: it checks all contact segment  $\Gamma_{ci}$ , which contain at least one node in adhesion  $\alpha_j^i > 0$ , if all other nodes of this segment are in adhesion or in normal contact

$$\alpha_j^i \alpha_k^i > 0 \text{ or } \alpha_j^i \alpha_k^i < 0, \quad k = 1, N^i, k \neq j$$

the algorithm goes to the next contact segment, otherwise if there is at least one node with non-contact status

$$\exists k : \alpha_j^i \alpha_k^i = 0,$$

or if the adhering node is situated on the edge of the contact zone, then the MPC boundary conditions are removed on the given contact segment  $i$ . In other words it means that the algorithm unsticks nodes layer by layer starting from the edge of the adhesion zone (see Fig. 5.20).

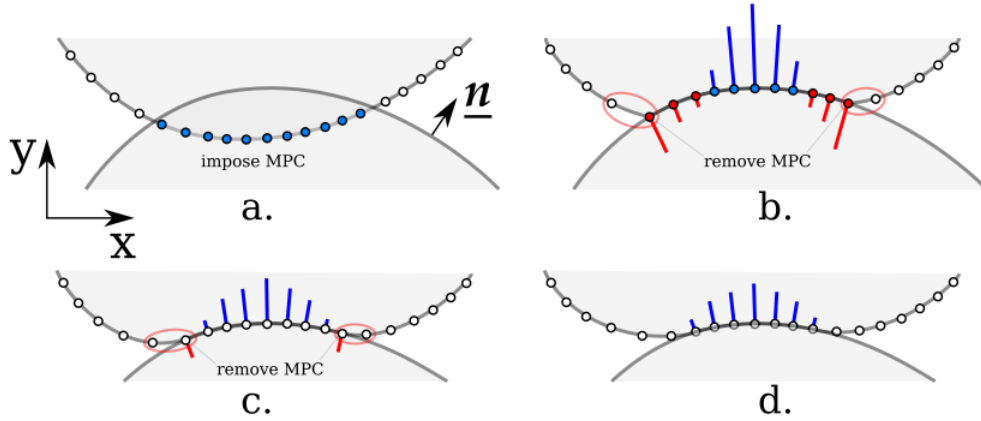


Figure 5.20: Active set strategy based on the gradual suppression of the MPC boundary conditions starting from the edges of the contact zone: **a** – MPC boundary conditions are imposed at all penetrated nodes, **b** – arising reactions in blue normal contact in red adhesion, MPC is removed at adhering nodes on the edges of the contact zone, **c** – still some nodes adhere, MPC is removed at edge nodes, **d** – converged solution.

The proposed technique is adapted for globally convex geometry of contacting interfaces, in a more complicated situation the full topology of contact, non-contact and adhesion zones should be analyzed. Since this technique depends on the number of nodes in contact interface, it is not directly applicable to large problems. An adaptation is required: instead of detaching node-by-node one should detach entire groups of nodes determined according to the topology. This is especially crucial for frictional contact, where all nodes may slip.

### 5.5.3 Frictional case

The frictional case has been discussed in details for the case of contact with a rigid plane orthogonal to one of the basis vectors in paragraph 4.1.3. The idea consists in replacing the stick conditions by full Dirichlet boundary conditions, i.e. if a node  $\underline{x}$  penetrates the rigid plane, it should be returned to the penetration point  $\underline{x}^\bullet$  and stuck to this point by imposing

$$\underline{u} = \underline{x}^\bullet - \underline{x}.$$

Further the reaction  $\underline{R}$  arising at the node should be splitted into the normal  $\underline{R}_n$  and tangential  $\underline{R}_t$  parts and the non-adhesion condition

$$\underline{R}_n \cdot \underline{n} \geq 0$$

should be checked as well as the stick-slip condition

$$\begin{cases} \underline{R}_t \in C_f(\underline{R}_n, \underline{R}_t, \dots), & \text{stick} \\ \underline{R}_t \notin C_f(\underline{R}_n, \underline{R}_t, \dots), & \text{slip} \end{cases} ,$$

where  $C_f(\underline{R}_n, \underline{R}_t, \dots)$  denotes the stick zone. In case of Coulomb's friction law, the stick-slip conditions take the following form

$$\begin{cases} \|\underline{R}_t\| \leq \mu \|\underline{R}_n\|, & \text{stick} \\ \|\underline{R}_t\| > \mu \|\underline{R}_n\|, & \text{slip} \end{cases}$$

In case of slip, the full Dirichlet boundary condition has to be replaced by an MPC boundary condition (in case of arbitrary surface) or by a partial Dirichlet boundary condition in the direction orthogonal to the rigid plane (in the case discussed in Paragraph 4.1.3). Moreover, in the tangent plane, the external force  $\underline{F}^e$  should be applied to the sliding node in the opposite direction to the arisen tangential reaction  $\underline{R}_t$ , and the magnitude of the applied load is evaluated according to

$$\underline{F}^e = -\mu \|\underline{R}_n\| \frac{\underline{R}_t}{\|\underline{R}_t\|}$$

Elaboration of this scheme to an arbitrary rigid surface has not been yet achieved.

### 5.5.4 Remarks

Many engineering contact problems can be approximated by contact between a deformable body and a rigid surface: metal forming and metal processing, tire-road contact, polymer-metal water seals, indentation tests and, in general, any contact occurring between materials with significantly different stiffnesses. The PDN method seems to be very useful for the mentioned problems: first, because of its simplicity and second due to its stability and accurate fulfillment of the geometrical constraints. It is worth mentioning that the method is applied

## 5.5 Method of partial Dirichlet-Neumann boundary conditions

---

both for linear and higher order discretization<sup>5</sup> of the Finite Element mesh in the contact interface. The practice demonstrates that the most challenging part is the determination of the active contact zone. Further investigation is still required on the subject. The PDN method can be also efficiently used in case of possible contact through the planes of symmetry, i.e. can be considered as a “symmetry contact boundary condition”, see examples in Chapter 6.

---

<sup>5</sup>the method is applicable to second order discretizations in two dimensional case; in three dimensional case additional developments are required.

## 5.6 Technical details

This section contains short descriptions of different contact related features and improvements which have been incorporated in the Finite Element code *ZéBuLoN* (Zset) [Besson 97].

### 5.6.1 Rigid master surface

Sometimes it is useful to approximate one of the contacting surfaces by a rigid surface. Frequently the master is chosen to be rigid, since that provides a better convergence of the numerical scheme. For that purpose it is convenient to organize one directional transfer of contact stresses, i.e. contact stress arises only on the slave surface due to violation of geometrical constraints but are not transferred through the contact interface to the master surface.

From a technical point of view, such an approximation leads to much simpler forms of the geometrical variations  $\underline{\underline{G}}, \underline{\underline{T}}, \underline{\underline{H}}, \underline{\underline{S}}$  and consequently to simpler forms of the residual vector and tangent matrix. As the reader may remember these s-structures connect the variations of geometrical quantities with variations of nodal coordinate vectors of the contact element

$$\delta g_n = \underline{\underline{G}}^\circ \delta \underline{\underline{x}}, \quad \delta \xi = \underline{\underline{T}}^\circ \delta \underline{\underline{x}}, \quad \Delta \delta g_n = \Delta \underline{\underline{x}}^\circ \underline{\underline{H}}^\circ \delta \underline{\underline{x}}, \quad \Delta \delta \xi = \Delta \underline{\underline{x}}^\circ \underline{\underline{S}}^\circ \delta \underline{\underline{x}}$$

To obtain a simplified discretization, one has simply to put all variations of the master nodes to zero, i.e., if  $\underline{\underline{x}}$  is a v-vector of contact element nodal coordinate vectors

$$\underline{\underline{x}} \sim \begin{bmatrix} \underline{r}_s & \underline{\rho}^1 & \dots & \underline{\rho}^M \end{bmatrix},$$

where  $\underline{r}_s$  and  $\underline{\rho}^i$  are coordinate vectors of the slave and master nodes respectively. Then the v-vector of variations has the following form

$$\underline{\underline{x}} \sim \begin{bmatrix} \delta \underline{r}_s & 0 & \dots & 0 \end{bmatrix}$$

It allows to reduce the order of geometrical s-structures

$$\underline{\underline{G}} \rightarrow \underline{G}, \quad \underline{\underline{T}} \rightarrow \underline{T}, \quad \underline{\underline{H}} \rightarrow \underline{H}, \quad \underline{\underline{S}} \rightarrow \underline{S},$$

$$\delta g_n = \underline{G}^\circ \delta \underline{r}_s, \quad \delta \xi = \underline{T}^\circ \delta \underline{r}_s, \quad \Delta \delta g_n = \Delta \underline{r}_s^\circ \underline{H}^\circ \delta \underline{r}_s, \quad \Delta \delta \xi = \Delta \underline{r}_s^\circ \underline{S}^\circ \delta \underline{r}_s$$

To obtain simplified expressions for the variations of the geometrical quantities, one can simply put the corresponding shape functions to zero:

$$\phi_i = 0$$

in (2.99) and (2.102), and

$$[\Phi'_i] = 0, \quad [\Phi''_{ij}] = 0, \quad [\Phi'''_{ijk}] = 0$$

in (2.105) and (2.111). Or directly put

$$\delta \underline{\rho} = 0, \quad \delta \frac{\partial \rho}{\partial \xi} = 0, \quad \Delta \frac{\partial \rho}{\partial \xi} = 0, \quad \delta \frac{\partial^2 \rho}{\partial \xi^2} = 0, \quad \Delta \frac{\partial^2 \rho}{\partial \xi^2} = 0$$

in continuous form of the geometrical variations of the normal gap (2.24), surface parameter (2.31), second variations of the normal gap (2.67) and surface parameter (2.78). These settings results in simple expressions given below

- **Variation of the normal gap**

$$\delta g_n = \underline{n} \cdot \delta \underline{r}_s \Rightarrow \underline{G} = \underline{n}$$

- **Variation of the surface parameter** (convective coordinate)

$$\delta \xi = \left( \underline{\underline{A}} - g_n \underline{\underline{H}} \right)^{-1} \circ \frac{\partial \rho}{\partial \xi} \cdot \delta \underline{r}_s \Rightarrow \underline{\underline{T}} = \left( \underline{\underline{A}} - g_n \underline{\underline{H}} \right)^{-1} \circ \frac{\partial \rho}{\partial \xi}$$

- **Second variation of the normal gap**

$$\Delta \delta g_n = \Delta \underline{r}_s \cdot \left[ -\underline{\underline{T}} \circ \underline{\underline{H}} \circ \underline{\underline{T}} + g_n \underline{\underline{T}} \circ \underline{\underline{H}} \circ \underline{\underline{\bar{A}}} \circ \underline{\underline{H}} \circ \underline{\underline{T}} \right] \cdot \delta \underline{r}_s$$

$$\underline{\underline{H}} = -\underline{\underline{T}} \circ \underline{\underline{H}} \circ \underline{\underline{T}} + g_n \underline{\underline{T}} \circ \underline{\underline{H}} \circ \underline{\underline{\bar{A}}} \circ \underline{\underline{H}} \circ \underline{\underline{T}}$$

- **Second variation of the surface parameter**

$$\Delta \delta \xi = \Delta \underline{r}_s \cdot \left[ \left( g_n \underline{\underline{H}} - \underline{\underline{A}} \right)^{-1} \circ \left\{ \underline{\underline{T}} \circ \left( \frac{\partial \rho}{\partial \xi} \boxtimes \frac{\partial^2 \rho}{\partial \xi^2} + g_n \underline{n} \cdot \frac{\partial^3 \rho}{\partial \xi^3} \right) \circ \underline{\underline{T}} - \right. \right.$$

$$\left. - \underline{n} \otimes \underline{\underline{T}} \circ \underline{\underline{H}} - \underline{\underline{H}} \circ \underline{\underline{T}} \otimes \underline{n} + \right.$$

$$\left. + g_n \underline{\underline{T}} \circ \left( \frac{\partial^2 \rho}{\partial \xi^2} \boxtimes \frac{\partial \rho}{\partial \xi} \circ \underline{\underline{\bar{A}}} \circ \underline{\underline{H}} + \underline{\underline{H}} \circ \underline{\underline{\bar{A}}} \circ \frac{\partial \rho}{\partial \xi} \boxtimes \frac{\partial^2 \rho}{\partial \xi^2} \right) \circ \underline{\underline{T}} \right] \cdot \delta \underline{r}_s \quad (5.82)$$

These variations are relatively simple and can be easily integrated in any Finite Element code as a special case for rigid master surfaces.

### 5.6.2 Multi-face contact elements and smoothing techniques

Multi-face contact elements have been discussed in Section 3.3.4 and some remarks concerning the residual vector and tangent matrix are made in Section 5.4.4. Here a brief discussion of the multi-face contact element will be given in the framework of the Finite Element Analysis.

A standard contact element consists of one slave node and one master segment. What happens when the slave node during the iterative process slides out of the master segment? It has no more projection on the master segment, but excluding this slave node from the solution would result in a penetration. So a possible way out is to create a new contact element, but this procedure is quite time consuming, because it requires an update of the tangential stiffness matrix and its re-factorization. There are two possibilities:



1. extend the parent domain of the master segment (see 5.21, a);
2. extend the standard contact element by adding the adjacent master segments, i.e. create a *multi-face* contact element in advance, before starting the iterative solution procedure.

The first possibility results in continuous sliding of the slave node along the extension of the current master segment before the equilibrium is reached. Such a sliding results either in penetration if the surface is locally concave or in opening of the gap in case of convex surface. These drawbacks are negligible if the sliding increment is small in comparison to the size of master segment, otherwise the solution becomes unrealistic. In this case the use of multi-face elements is highly desirable: when the slave node passes from one master segment to another, the residual vector and tangent matrix are reformulated respectively to the new active master segment and the corresponding master nodes (see Fig. 5.21, b). Note also that in frictional case the described situation results in loss of the sliding path tracking, i.e. the relative sliding increment cannot be approximated by

$$\Delta \underline{g}_t = \frac{\partial \rho}{\partial \underline{\xi}} \circ \Delta \underline{\xi}$$

and hence some special techniques are required, see e.g. [Konyukhov 07a] and [Wriggers 06]. Another drawback is the presence of possible oscillations in the neighborhood of concave angles – constant switching between two adjacent segments. Such oscillations are also possible within the first approach but in this case the oscillations take place from one increment to another and the local convergence is often achieved (see 5.22). In other words, to avoid this undesirable effect the node-to-node, node-to-edge and node-to-vertex elements should be available in the Finite Element code, however it does not always solve the problem, another approach is required.

An extension of the multi-face contact elements leads to an even more advantageous treatment of the problem: instead of distinguishing between active and inactive master segments, one can construct a  $C^1$ -smooth surface based on all master nodes included in the multi-face contact element (see Fig. 5.23). One of the first successful applications of the smoothing procedures to large sliding contact problems can be found in [Pietrzak 97],[Pietrzak 99], where authors use Bézier curves and Splines to smooth a structured surface mesh consisting of regular quadrilateral segments. However, in case of arbitrary surface mesh, the construction of a  $C^1$ -smooth surface is not evident and presents an up to date research topic in computational contact mechanics; many details, examples and references on this topic can be found in books of Wriggers [Wriggers 06] and Laursen [Laursen 02]). Hermitian, Spline and Bézier polynomials as well as Gregory patches are used for smoothing the master surface in two and three dimensional contact problems.

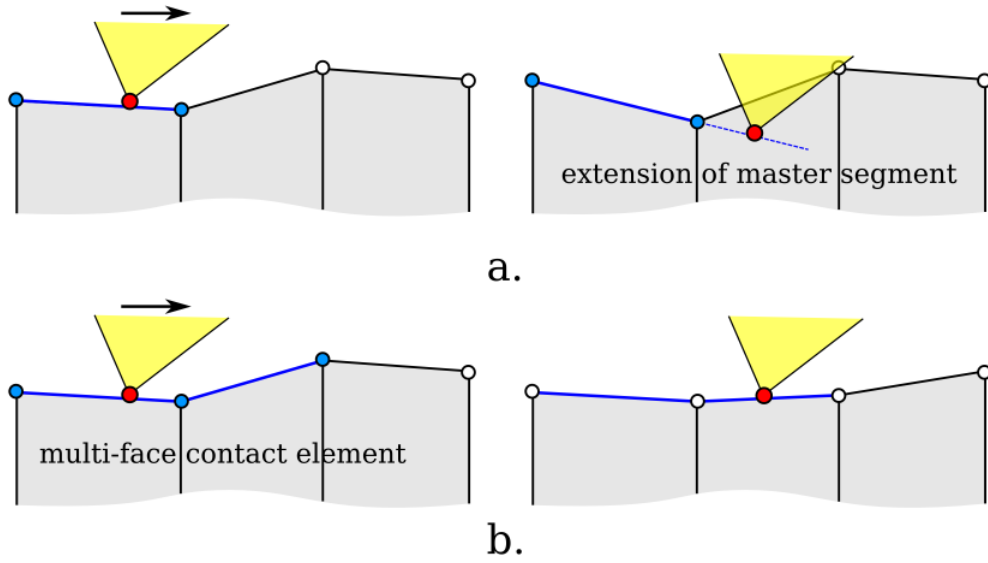


Figure 5.21: Example of two dimensional large sliding contact: extension of the master segment (a) and multi-face contact element (b).

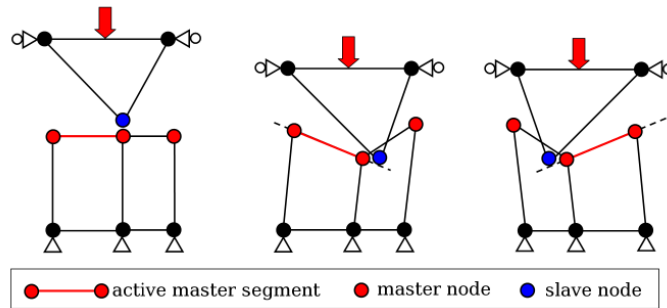


Figure 5.22: Example of frictionless oscillations in a neighborhood of the concave vertex.

### 5.6.3 Heterogeneous friction

Contact modeling of heterogeneous materials (composites, alloys with an explicit representation of the microstructure) leads to a problem with heterogeneous friction. For engineering aspects of the problem, see [Dick 06a], [Dick 06b], [Dick 08]. The computational aspects of this type of problems are shortly discussed below. Two types of friction heterogeneity can be distinguished:

- friction properties are determined by domains (e.g. composite material, polycrystalline surface, Fig. 5.24);
- friction properties change smoothly along the contact interface (e.g., due to temperature distribution).

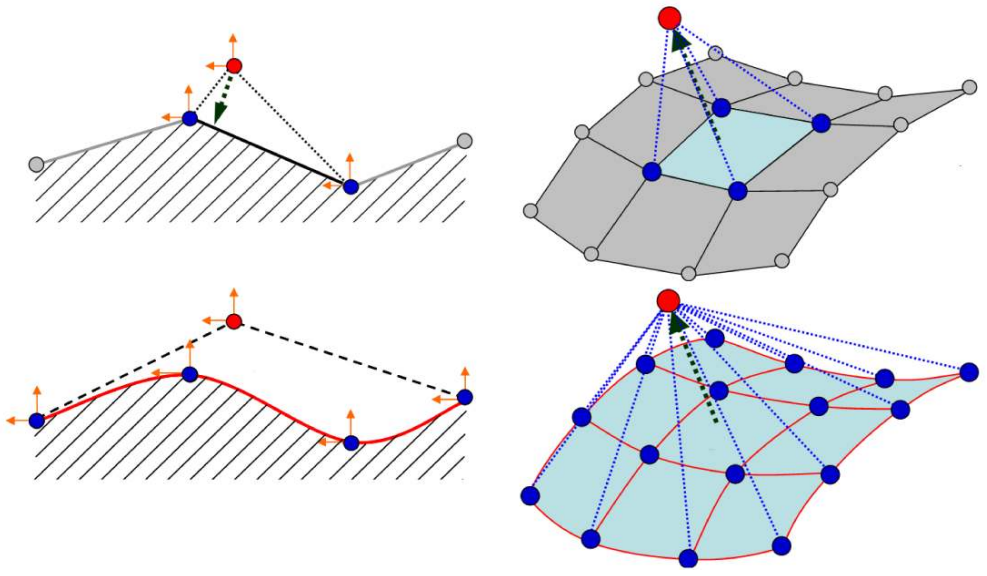


Figure 5.23: Example of master surface smoothing by a cubic Bézier curve and surface, based on the structured quadrilateral surface mesh.

A rigorous treatment of the friction change within one master segment leads to a quite complicated linearization of the virtual work, however it seems to be always possible to assume a constant friction within one contact element.

Note that the friction is an interface phenomenon, i.e. both contacting surfaces determine the friction properties. For example, in case of contact between two bi-phase composites  $f_1^1, f_2^1$  and  $f_1^2, f_2^2$ , four local coefficients of friction (in case of Coulomb's friction law) have to be defined if self-contact is excluded:

$$\mu_1 = \mu_1(f_1^1, f_1^2), \mu_2 = \mu_2(f_1^1, f_2^2), \mu_3 = \mu_3(f_2^1, f_1^2), \mu_4 = \mu_4(f_2^1, f_2^2)$$

Treatment of such problems requires a proper definition of the friction coefficient for each contact element. A standard Node-to-Segment contact element consists of one slave node and master segment, so to each slave node and master segment friction IDs  $f_i^s$  and  $f_j^m$  should be assigned and a corresponding rule for friction coefficients:

$$\mu_{ij} = \mu_{ij}(f_i^s, f_j^m).$$

Another possibility leading to a symmetric treatment of contacting surfaces consists in assignment of friction IDs to both slave and master nodes:  $f_i^s, f_{j_1}^m, f_{j_2}^m, \dots, f_{j_M}^m$  and a corresponding rule, for example, as proposed below

$$\mu_{ij} = \frac{1}{M} \sum_{k=1}^M \alpha_k \mu_{ij}(f_i^s, f_{j_k}^m), \quad \sum_{k=1}^M \alpha_k = 1.$$

Such an approach is valid both for normal and self-contact and includes the previous possibility as a subcase, for properly chosen  $\alpha_k$  coefficients.

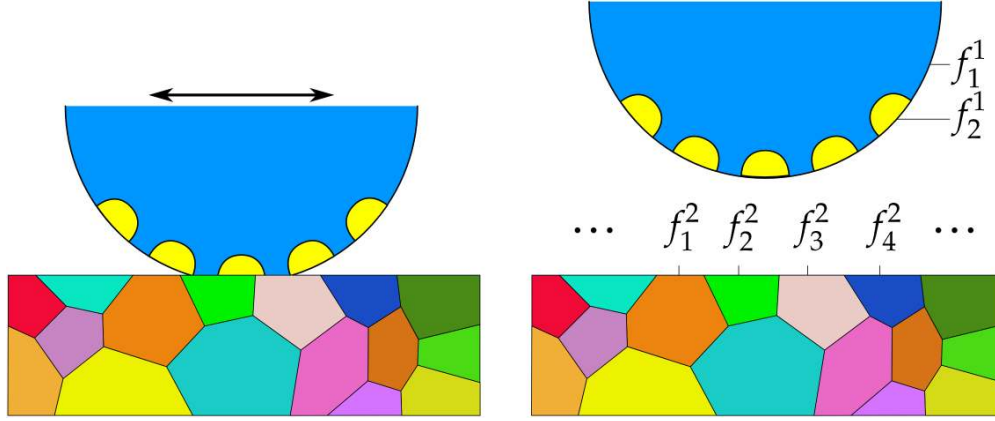


Figure 5.24: Example of heterogeneous friction between a composite and a polycrystalline structure.

#### 5.6.4 Short remarks

Below some short remarks concerning the code organization are listed:

- It is often advantageous to organize the code in a way that **independent contact zones** can be treated separately but within a common contact framework; it leads to an efficient organization of the contact detection procedure (as described in the Chapter 3) and also to a smart update of the tangent matrix.
- The components of contact elements may change from one increment to another, however sometimes it is advantageous to **keep the values of Lagrange multipliers** from the previous time step and use them as an initial guess. It is easy to save these values directly for each slave node as well as the sliding path; the initial guess is especially important in case of use of Uzawa's algorithm.
- Regardless all advantages of the augmented Lagrangian method, its convergence strongly depends on the choice and **update of the penalty parameters**  $\varepsilon_n$  and  $\varepsilon_t$  related to normal and tangential contact respectively. We did not elaborate any consistent and theoretically based updating technique, but some general rules can be deduced. From remarks given in Section 5.1 according to the investigations of Alart [Alart 97] and Alart and Curnier [Alart 91] on the convergence of the augmented Lagrangian in frictional contact it follows that if a frictional cycling occurs, then the tangential penalty parameter should be decreased

$$\text{stick-slip cycling} \Rightarrow \varepsilon_t^{i+1} = \alpha_t \varepsilon_t^i, \quad 0 < \alpha_t < 1.$$

On the other hand if a slave node loses the contact, then high penalty coefficients  $\varepsilon_n$  and  $\varepsilon_t$  may result in a poor conditioning of the tangent

matrix, so both of them should be decreased:

$$\text{non-contact} \Rightarrow \varepsilon_n^{i+1} = \beta_n \varepsilon_n^i, \quad \varepsilon_t^{i+1} = \beta_t \varepsilon_t^i, \quad 0 < \beta, \beta_t < 1$$

And finally if the contact status of the contact element remain constant (slip or stick) then the penalty parameters should be increased in order to enforce the fulfillment of geometrical constraints

$$\text{stick or slip} \Rightarrow \varepsilon_n^{i+1} = \gamma_n \varepsilon_n^i, \quad \varepsilon_t^{i+1} = \gamma_t \varepsilon_t^i, \quad \gamma_n, \gamma_t > 1$$

A more elaborated update scheme may also take into account the values of normal gap, tangential sliding and of the corresponding Lagrange multipliers. An example of phenomenological update scheme can be found in [Bussetta 09].

- A FE problem, containing, for example, two initially separated bodies, has a tangent matrix (sparse matrix) with a block structure and the minimal thickness of the bandwidth of each block is ensured by a smart numbering of nodes and dofs. If contact occurs between these bodies the **optimality of the sparse structure** may be lost, i.e. the band-width can be significantly enlarged. Hence the storage of the sparse matrix becomes non-optimal and requirements on the needed memory space can reach the amount needed for storage of the full matrix. There are at least two solutions:
  - renumbering of nodes after construction of contact elements at each change of mesh topology;
  - one preliminary renumbering of nodes in the way that all possible mesh graphs (due to presence of contact element) retain a narrow bandwidth.
- According to our experience, the **user friendliness** of the FEA code is very important, namely, the FEA tool should check automatically the **orientation** of the master surface and change it if needed, also it should choose **automatically** the optimal “maximal detection distance” parameter and the detection cell size, the two latter are possible in the framework of the bucket sort detection algorithm (see Chapter 3).
- In case of complex geometry, large deformations and/or self-contact, it is complicated or even impossible to know the master-slave discretization a priori. This is why Benson and Hallquist [Benson 90] introduced the **single contact surface**. From the programming point of view, such a surface is nothing but a master surface. It is necessary simply to adapt the the detection algorithm, as described in Chapter 3, the rest of the code remains unchanged. It is worth mentioning also that, in case of simple buckling simulations, the contact may occur only after a certain amount of time. In order to accelerate the solution, the contact detection procedure should not be executed before that time.

- **Parallel treatment** of contact problems requires a special detection technique. From the programming point of view, it is easy to construct a single contact detection procedure which takes as arguments master and slave surfaces. Further in the parallel framework it remains only to joint all the parts of the slave and master surfaces from different subdomains and to pass them to the detection procedure. This approach was called SDMR (Single Detection, Multiple Resolution). If there are several independent contact zones, then the SDMR detection procedure can be carried out independently for each zone at different processors. The same detection procedure is valid for the MDMR (Multiple Detection, Multiple Resolution) technique, in this case each processor does not receive the entire master and slave surfaces, but their close parts, which should be simply passed to the detection procedure (for details, see Chapter 3).

## Chapter 6

# Numerical examples

### **Résumé de Chapitre 6 «Exemples numériques»**

Ce chapitre réunit neuf simulations numériques obtenues à l'aide des différentes méthodes de contact implémentées dans le code de calcul par éléments finis Z-set (ZéBuLoN). Trois méthodes sont considérées : méthode de pénalisation, méthode du Lagrangien augmenté et méthode PDN. On présente six problèmes 2D et trois problèmes 3D. Deux des exemples 2D sont des problèmes de contact frottant pour lesquels on dispose d'une solution analytique : indentation d'un demi-espace par un cylindre rigide, et contact entre un espace troué et un cylindre situé à l'intérieur. On compare les solutions analytiques avec les simulations obtenues par la méthode du Lagrangien augmenté.

Le troisième exemple est un problème 2D d'indentation d'une pièce rectangulaire par un cylindre rigide, les résultats de cette simulation sont obtenus par la méthode PDN, avec des éléments linéaires et quadratiques.

Le quatrième exemple est un problème d'emboutissage qui réunit trois non-linéarités fortes : matériau élasto-visco-plastique avec écrouissage non linéaire en grande déformation et contact avec frottement. Le problème est traité par la méthode du Lagrangien augmenté.

Le cinquième problème considéré est le repassage peu profond d'une plaque par un indenteur déformable. Il s'agit donc de contact multi-corps avec frottement en présence des grands glissements et grandes déformations avec un matériau Neo-Hookéen. Nos résultats, obtenus par la méthode du Lagrangien augmenté, sont comparés à ceux de la littérature, obtenus soit avec une discrétisation «Nœuds-à-Segment» [Fischer 06] par la méthode de «sliding cone» et la méthode «mortar», soit par la méthode des domaines de contact [Hartmann 09]. La méthode du Lagrangien augmenté fonctionne mieux que la méthode de domaine de contact, et permet de retrouver les résultats de référence obtenus par [Fischer 06].

## 6.1 Two dimensional problems

*Le dernier des problèmes 2D considérés présente le comportement en post-flambage d'un tube mince élasto-plastique, dans le cadre des grandes déformations, avec auto-contact. Ce problème a été considéré dans [Laursen 92]. On utilise la même méthode que dans le travail original (méthode de pénalisation) et deux maillages différents. Les résultats démontrent la sensibilité importante de la solution au paramètre de pénalisation.*

*Le premier exemple 3D est un cas de post-flambage non-symétrique avec auto-contact d'un tube mince élasto-plastique, qui a été calculé par la méthode de pénalisation. Le deuxième exemple traite du début d'extrusion d'une plaque à travers un trou circulaire, qui a été simulé par la méthode PDN et différents modèles de matériau en petites et grandes déformations. Le dernier exemple considère le glissement avec frottement d'un cube sur un plan rigide. L'effet de détachement qui est classiquement observé pour les grands coefficients de frottement est bien capté par la méthode PDN. Cet effet est conduit à une intéressante réduction du coefficient de frottement global.*

Some simple and not so simple examples are demonstrated in this chapter. Penalty, augmented Lagrangian and PDN methods are used. Two and three dimensional frictional and frictionless contact problems are considered, in small and large deformations for linear and nonlinear material models. The more contact elements the problem contains, the more complicated its numerical treatment, so in order to demonstrate the performance of the contact implementation in the finite element code (ZéBuLoN) we use quite fine discretizations, even when a coarser mesh would provide a reliable estimation of the solution. A full input data is given for many considering problems.

## 6.1 Two dimensional problems

### 6.1.1 Indentation by a rigid flat punch

Indentation of an elastic half space by a rigid axisymmetric flat punch is considered in presence of finite friction. The analytical solution can be found, for example, in [Spence 75]. An attempt of numerical solution within the Finite Element Method and the augmented Lagrangian method was undertaken in [Pietrzak 97],[Pietrzak 99], however the numerical solution appears to be significantly different than the analytical one. This error can be explained by the coarse mesh used by the authors. Here a more precise numerical solution is given, but the discrepancy between numerical and analytical solutions still exists.

A rigid flat axisymmetric punch of radius  $a$  is indented into an elastic half space by a force  $P$  (see Fig. 6.1). The Coulomb's friction law with the coefficient of friction  $\mu$  is considered. The frictionless contact pressure distribution is described by the following function (see e.g. [Johnson 94])

$$\sigma_n(\rho) = \frac{P}{\sqrt{a^2 - \rho^2}},$$



where  $\rho$  is a distance from the axis of symmetry. Note that due to this equation, the mesh convergence cannot be obtained for this problem, because of the singularity on the edge of the flat punch  $\rho = a$ . However, a rather good numerical estimation can be given.

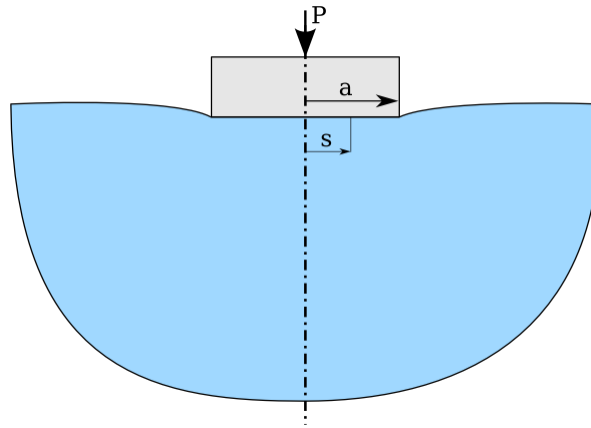


Figure 6.1: Indentation of an elastic half space by a rigid axisymmetric flat punch:  $a$  – radius on the punch,  $s$  – radius of the stick zone.

The problem is solved with an axisymmetric mesh. All relevant data are listed in the box below, all quantities are given in force units (f.u.) and length units (l.u.), which consequently will be omitted. The finite element mesh of the halfspace consists of a half circle fixed along the radial perimeter  $R = 9a$ . The indenter (master) is constructed in such a way that its nodes coincide with the nodes on the opposite contact surface of the half space (slave). Such a construction of the finite element mesh allows to avoid the inherent problem of the node to segment discretization – inability to transfer a uniform pressure through the contact interface for an arbitrary discretization. Three finite element meshes were used with 41, 201 and 401 nodes on the active slave surface, the finest mesh is represented in Fig. 6.2. The meshes contain 4183, 60533 and 95872 nodes respectively.

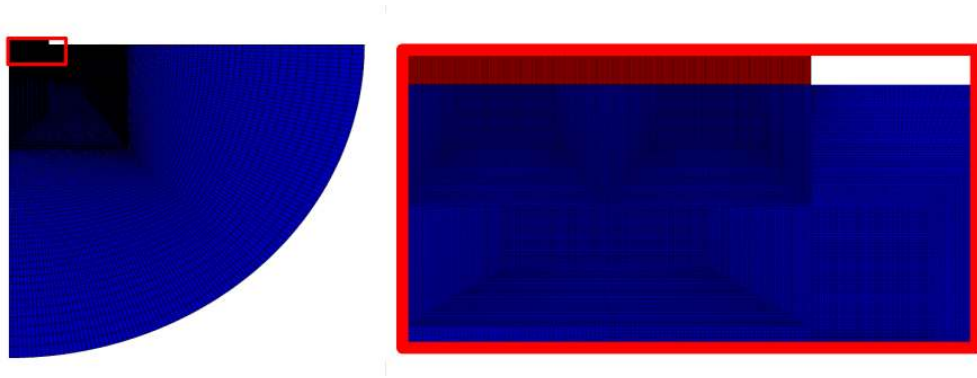


Figure 6.2: Whole finite element mesh and zoom on the contact region, 400 slave nodes in contact.

- Linearly elastic material, small deformations:

Young's modulus  $E = 1000$  [f.u./l.u.<sup>2</sup>]

Poisson's ration  $\nu = 0$ .

- Boundary conditions and geometry:

Punch radius  $a = 1$ . [l.u.]

Pressure on the indenter  $p = 1.61$  [f.u./l.u.<sup>2</sup>]

Friction coefficient  $\mu = 0.2063, 0.2986, 0.4013$

- Solution conditions:

Method: augmented Lagrangian

Penalty coefficients  $\varepsilon_n = \varepsilon_t = 100$  [f.u./l.u.<sup>2</sup>]

Increment 1

- Finite element mesh:

linear full integration quadrilateral axisymmetric elements  
(4 nodes, 4 integration points)

In his article [Spence 75], the author gives semi-analytical estimations of the friction coefficients  $\mu$  for different stick regions  $s/a$  and different Poisson's ratios  $\nu$ . His estimations are presented for  $\nu = 0$  in Tab. 6.2. Here the problem is inverse and for the given friction coefficient, the stick region is measured, the results for different finite element meshes are presented in Tab. 6.3. The convergence is quite slow and the correct values of the stick zone taken from [Spence 75] have not been reached. Contour plots of shear stress  $\sigma_{12}$ , stress  $\sigma_{22}$ , von Mises stress  $\sigma_v$  and horizontal displacement  $u_1$  distributions are represented in Fig. 6.3.

Table 6.2: Semi-analytical estimations (from [Spence 75]) for the friction coefficient ensuring stick zone of radius  $c$  for  $\nu = 0$ .

$\mu$	0.2063	0.2986	0.4013	0.4862
$s/a$	0.24	0.5	0.7	0.8

Table 6.3: Finite element estimation of the stick region radii for given friction coefficients ( $\nu = 0$ ), convergence by mesh.

Friction $\mu$	0.2063	0.2986	0.4013
FE mesh $N_s$	stick zone $s/a$		
41	0.300	0.550	0.750
201	0.260	0.534	0.735
401	0.255	0.520	0.717
exact	0.240	0.50	0.700

The stress distribution (extrapolation from Gauss points to nodes) in the contact interface is represented in Fig. 6.3

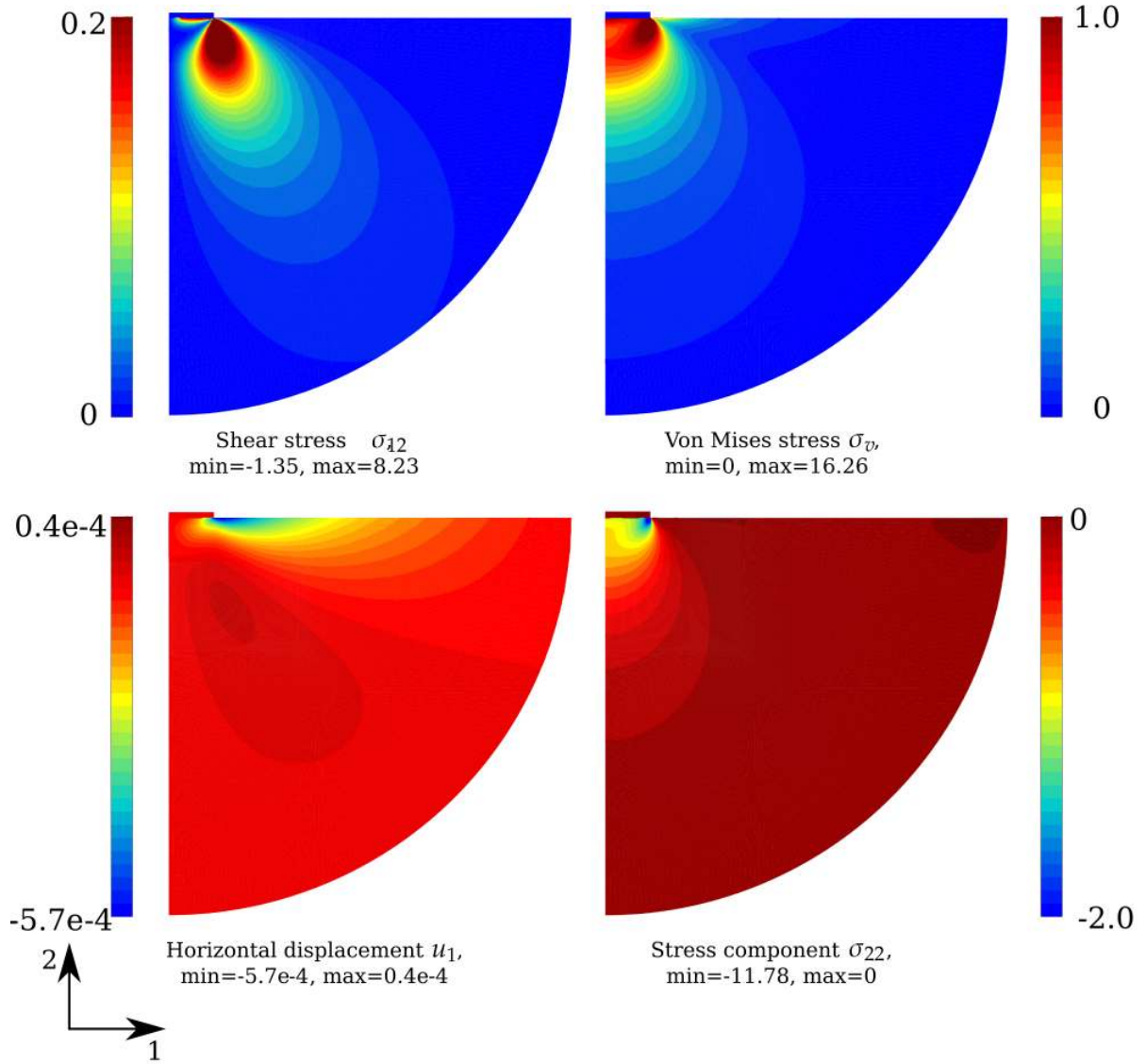


Figure 6.3: Contour plots of shear stress  $\sigma_{12}$ , von Mises stress  $\sigma_v$ , horizontal displacement  $u_1$  and stress component  $\sigma_{22}$  for  $\mu = 0.2063$ .

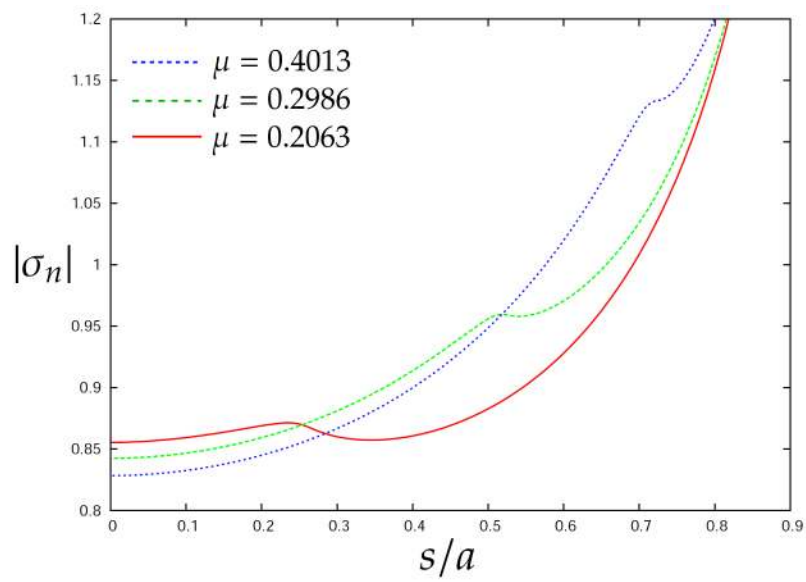


Figure 6.4: Distribution of the contact pressure for different coefficients of friction.

### 6.1.2 Elastic disk embedded in an elastic bored plane

The geometry presents a thin elastic infinite plane with a hole of almost the same radius as a thin elastic disk, embedded in the hole. A concentrated force is applied in the middle of the disk, the value of the force is chosen such that the contact occurs at one third ( $120^\circ$ ) of the interface between the hole and the disk (Fig.6.5).

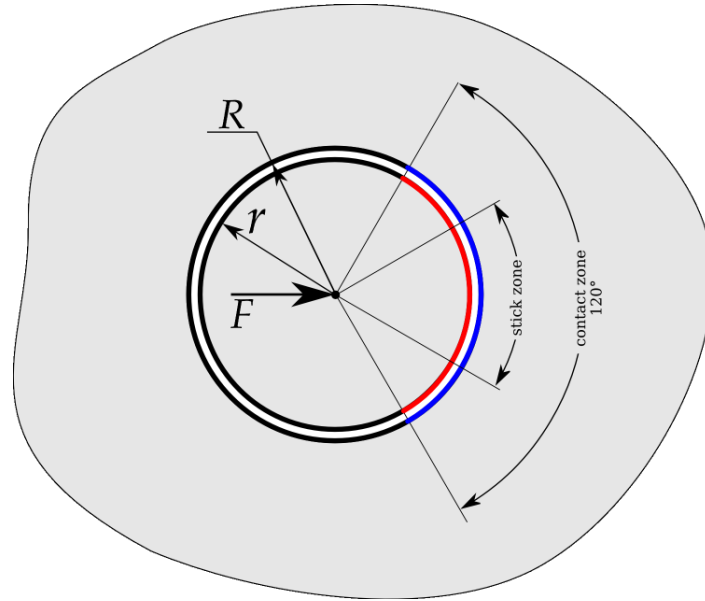


Figure 6.5: Setting of the problem: a thin elastic disk embedded in a thin elastic infinite plane with a circular hole.

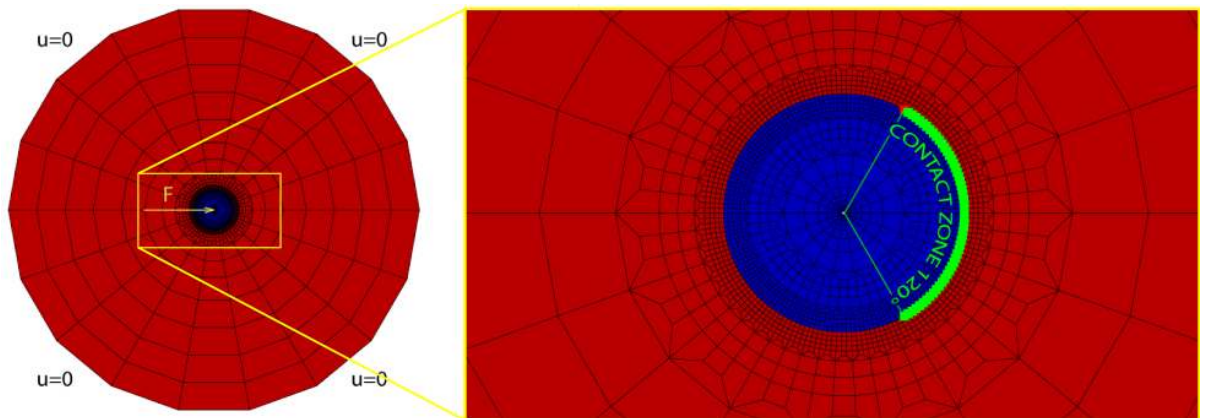


Figure 6.6: Finite element mesh (2502 nodes, 54 active slave nodes in the interface) and zoom on the contact region.

Coulomb's friction law is assumed in the interface. Again due to the

inability of the Node-to-segment discretization to transfer a uniform pressure through an arbitrary discretization, the finite element mesh is created in the way that master and slave nodes coincide in the interface.

A solution for this problem has been given by Klang [Klang 79], numerical treatments can be found in [Alart 91], [Pietrzak 97] or in [Pietrzak 99]. The finite element mesh is represented in Fig. 6.6. As one can see in the middle of the embedded disk, there is a circular hole in the center, where the concentrated force is applied to a node on the axis of symmetry and MPC boundary conditions are applied to all other nodes of this hole to retain its circular form. The concentrated load is applied in several increments. As proposed in [Alart 91] the force value changes quadratically with increments:

$$F_i = F_{\max} \frac{t_i^2}{t_{\max}^2},$$

where  $F_{\max}$  is the load and  $t_i \in [0; t_{\max}]$  is the incremental time with  $t_i = i \frac{t_{\max}}{n}$ , where  $n$  is a number of increments and  $t_{\max}$  is the total solution time. The remaining data are given in the box below.

- Linear elastic material, small deformations:

Young's modulus  $E = 2.1e7$  [f.u./l.u.<sup>2</sup>]

Poisson's ration  $\nu = 0.3$

- Boundary conditions and geometry:

plane stress

Disk radius  $r = 5.999$

Hole radius  $R = 6$

Concentrated force  $F = 9375t^2/t_0^2$

Friction coefficient  $\mu = 0.4$

- Solution conditions:

Method: augmented Lagrangian

Penalty coefficients  $\varepsilon_n = \varepsilon_t = 500$  [f.u./l.u.<sup>2</sup>]

Increment 1, 5, 10, 25, 50, 100

- Finite element mesh:

linear full integration quadrilateral axisymmetric elements  
(4 nodes, 4 integration points); 2502 nodes, 54 active slave  
nodes in the interface

The experience shows that the shear contact stress distribution in the slip zone and the limit of the stick zone ( $-40^\circ < \alpha_s < 40^\circ$ ) as well as the contact zone width  $-60^\circ < \alpha_s < 60^\circ$  are easily achieved in at least two increments. However, the shear stress distribution in the stick zone appears to be strongly dependent on the mesh density and on the number of increments. The correct shear stress distribution in the stick zone for the considered finite element mesh is achieved only when the load is applied in more than 50 increments. The distribution of the shear stress in the contact zone for different number of increments is represented in Fig. 6.7. Our experience demonstrated, that the precision of the mesh node positioning in the interface must be very accurate: for example the relative error (respectively to the size of element) in node positioning is about  $5 \cdot 10^{-3}\%$ , which seems to be negligible leads to severe oscillations of the solution in the interface, the maximal allowed precision in node positioning ( $5 \cdot 10^{-7}\%$ ) yields a smooth result.



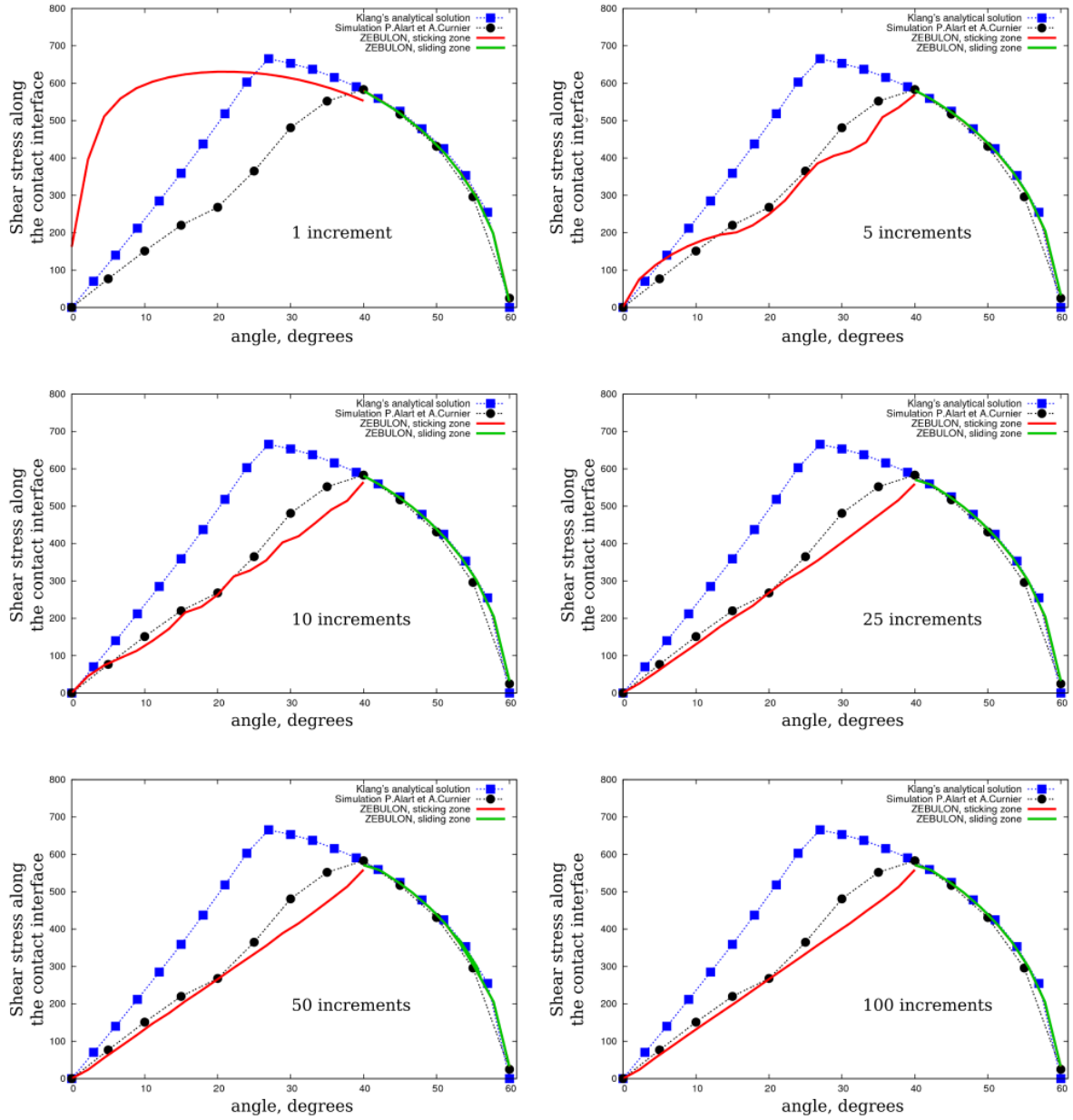


Figure 6.7: Shear stress distribution in the contact interface for different number of increments are compared with analytical results and numerical results from [Alart 91].

### 6.1.3 Indentation of an elastic rectangle by a circular indenter

This short example illustrated the PDN method performance for a non flat rigid surface in contact with a deformable body. The geometry of the problem is depicted in Fig. 6.8. All data are given in the box below.

- Linearly elastic material, small deformations:

Young's modulus  $E = 210$

Poisson's ration  $\nu = 0.3$

- Boundary conditions and geometry:

Indenter radius  $R = 1$

Rectangle  $L = 4, H = 2$

Friction coefficient  $\mu = 0$ .

Bottom of the block is fixed and moved towards the indenter

$u_x = 0, u_y = 1$

- Solution conditions:

Method: PDN

Increments 10

- Finite element mesh:

linear and quadratic full integration quadrilateral plane-strain elements (4 nodes, 4 integration points for linear and 8 nodes, 9 integration points); 200 elements

The von Mises stress distributions in Gauss points and nodal reactions in the interface due to MPC boundary conditions, arising from application of the PDN method, are represented in Fig. 6.9 for linear and quadratic meshes.

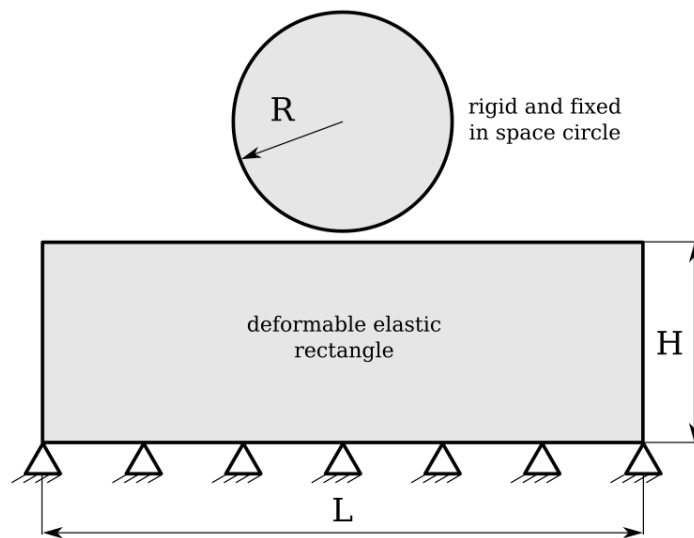


Figure 6.8: Scheme of the problem: indentation of an elastic rectangle by a rigid circle.

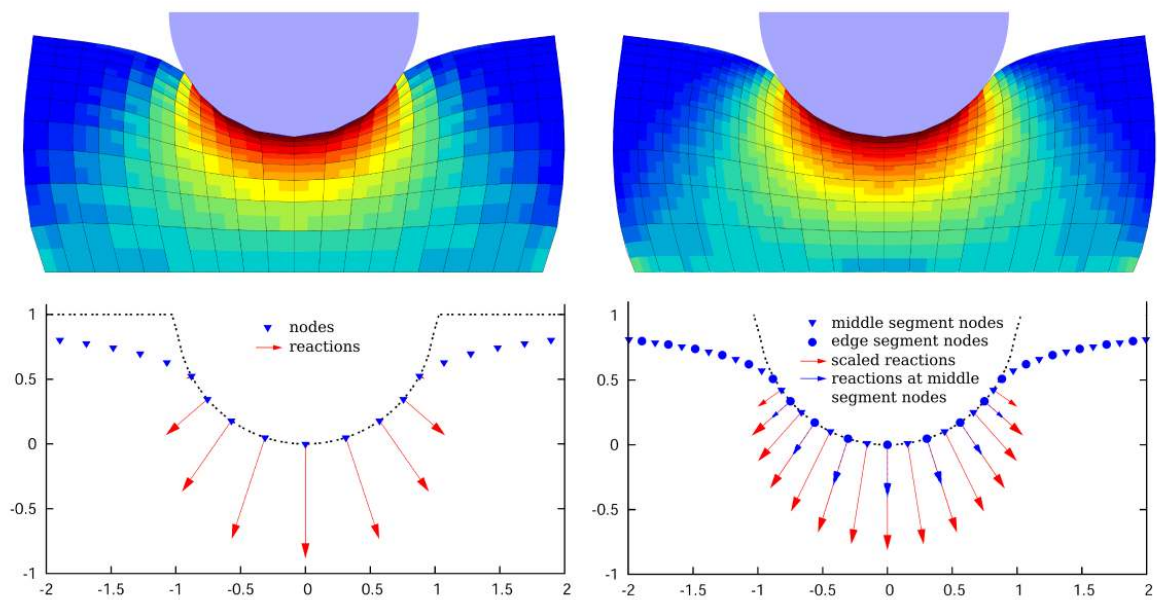


Figure 6.9: Von Mises stress distribution and nodal reactions for linear (**left**) and quadratic (**right**) meshes.

### 6.1.4 Axisymmetric deep cup drawing

Metal forming process is one of the important engineering applications of the computational contact mechanics, many examples can be found in scientific literature. Here we consider an axisymmetric deep cup drawing problem, stated in [Rousselier 09], see Fig. 6.10. This problem engenders many nonlinearities: large deformations within the updated Lagrangian framework, frictional contact and nonlinear material model (elasto-visco-plastic with exponential hardening). All parameters of the problem are stated in the box below. The problem is mixed force-displacement driven: the holding pressure  $p$  is applied linearly within the first 100 time units, and the die is gradually moved down up to  $t = 335$ , then removed. In order to stabilize the solution a soft spring (700 times softer than the material of the sheet) is attached to the edge of the sheet, as shown in Fig. 6.10.

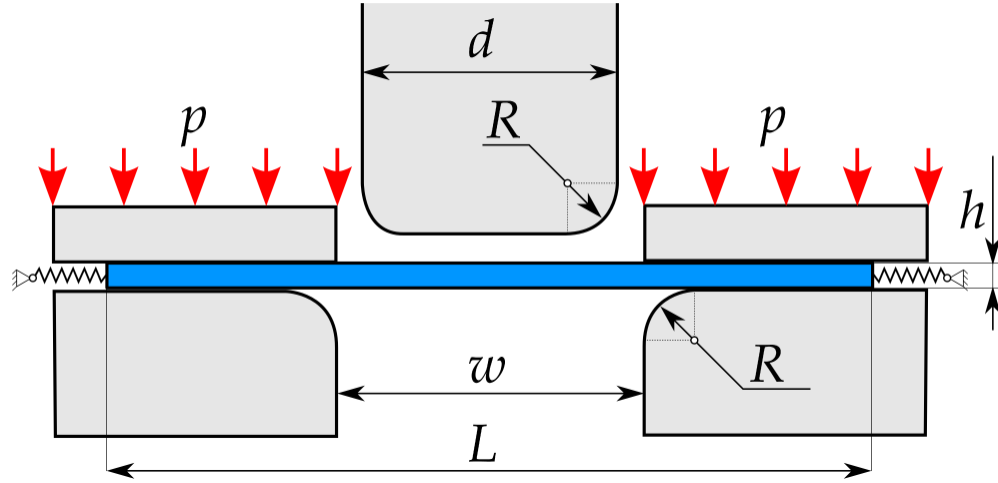


Figure 6.10: Geometrical setting of the deep cup drawing problem: rounded cylindrical die and metal sheet fixed within the tool.

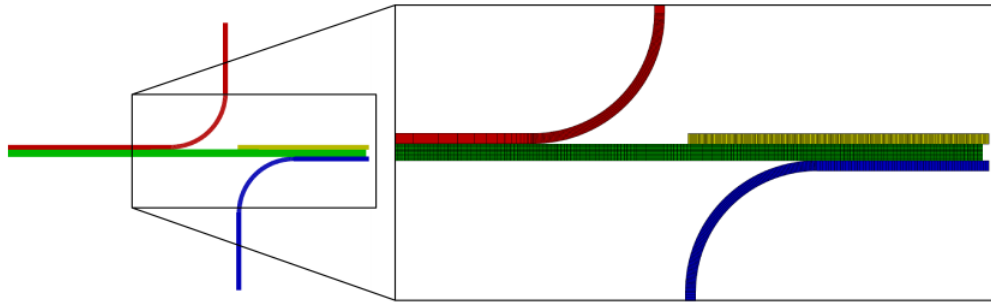


Figure 6.11: Finite element mesh for the axisymmetric deep cup drawing problem.

- Finite deformation elasto-plasticity (updated Lagrangian):

Young's modulus  $E = 69$ .

Poisson's ration  $\nu = 0.33$

Yield criterion: von Mises  $R_0 = 0.22$

Norton creep power law:

$$\dot{\lambda} = \langle f/K \rangle^n, K = 0.5, n = 7$$

Isotropic power hardening law

$$R = R_0 + K(e_0 + p)^n, K = 0.99, e_0 = 7e^{-4}, n = 7$$

- Boundary conditions and geometry:

Applied pressure  $p = 1.86e-2$

Die is gradually moved down  $u_z = -33.5t/t_l$  and removed  
 $u_z = -33.5 + 5.5 \frac{t-t_l}{t_f-t_l}$

Die diameter  $d = 97.46$

Rounding radius  $R = 12.7$

Die opening  $w = 101.48$

Sheet thickness  $h = 1.6$

Sheet diameter  $L = 158.76$

Friction coefficient  $\mu = 0.1$

- Solution conditions:

Method: augmented Lagrangian

Penalty coefficients  $\varepsilon_n = \varepsilon_t = 200$

Increments 1414

Full solution time: loading  $t_l = [0, 335]$  and unloading  $t_f = [335, 420]$

- Finite element mesh:

linear full integration quadrilateral axisymmetric elements  
 (4 nodes, 4 integration points for linear and 8 nodes, 9  
 integration points); 6063 elements (Fig. 6.11)

In average, on each time step 1000 contact elements are created, 400 of them are active. In average convergence on each time step is achieved within 10-15 iterations. The distribution of accumulated plastic strain for different time

## 6.1 Two dimensional problems

steps is represented in Fig. 6.12. The force-displacement curve is represented in Fig. 6.13. Note that even in presence of viscosity the reaction force oscillates (see Fig. 6.14). This is due to the fact that at this stage the sheet slides along the surface of the holder, which is not smooth. That is why in case of large sliding, it is more advantageous to use either smoothing of the master surface or determine the holder and the punch as analytical surfaces, or use the PDN method described in previous chapters.

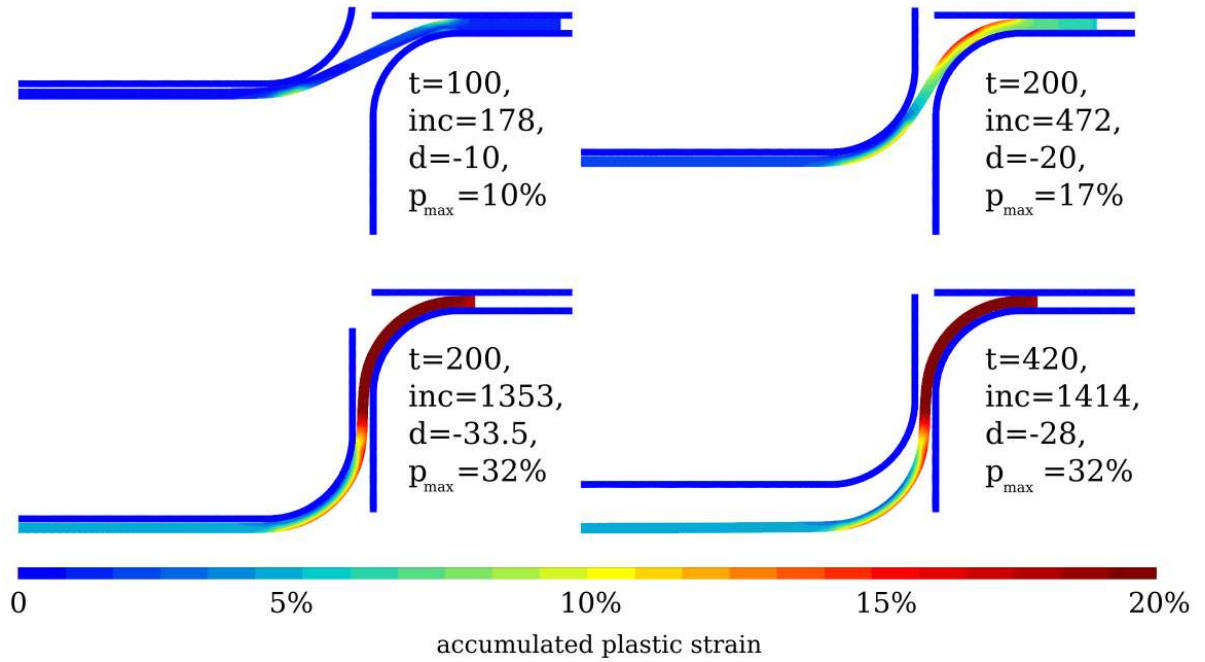


Figure 6.12: Distribution of accumulated plastic strain (in Gauss points) on different time steps.

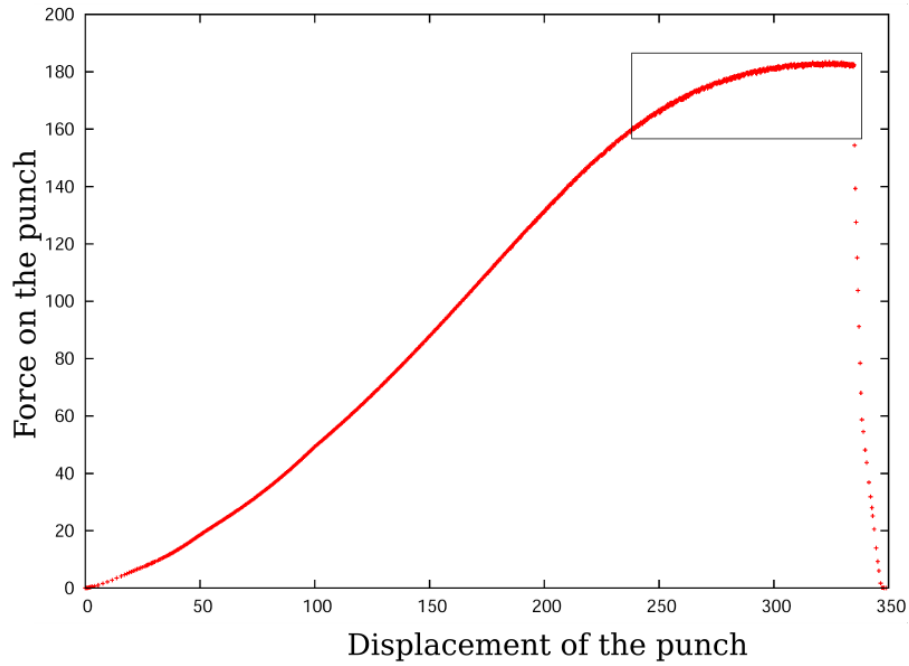


Figure 6.13: Evolution of the reaction on the punch with time, time =  $10 \times$  vertical displacement of the punch before  $t = 335$ .

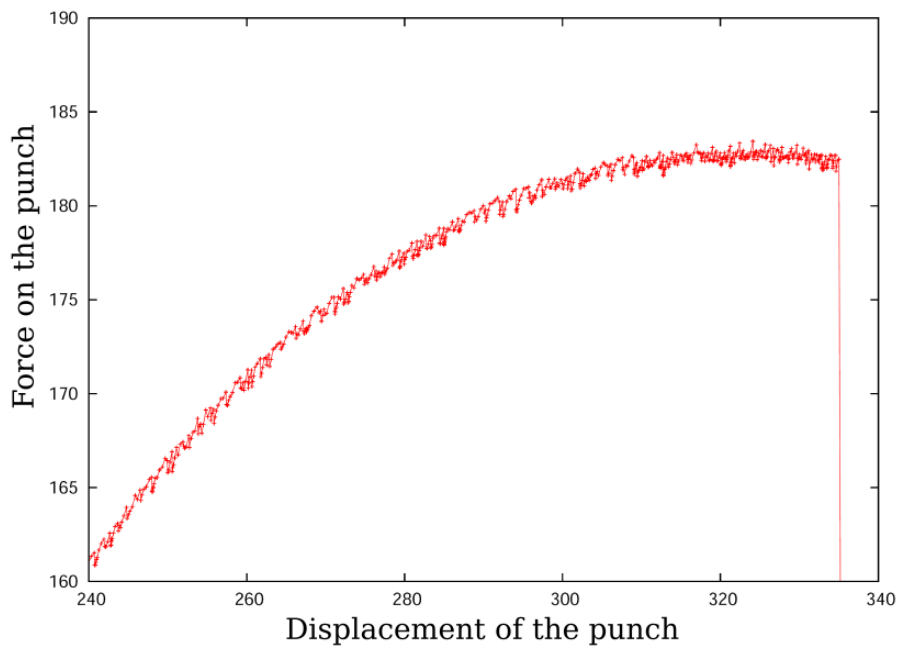


Figure 6.14: Oscillations of the reaction on the punch, zoom on the region marked in Fig. 6.13 by a rectangle.

### 6.1.5 Shallow ironing

The example shown in this paragraph consist in sliding of a deformable indenter along a deformable rectangle. The geometrical setting of the problem as well as the finite element mesh are presented in Fig. 6.15. This example has been proposed and solved in [Fischer 06] with the mortar based method using moving friction cone (see [Wriggers 06]), and resolved with contact domain method in [Hartmann 09]. Here the problem is solved using the augmented Lagrangian method and the Node-to-Segment discretization. Our results are compared with results of the cited authors in Fig. 6.16. Contrary to [Fischer 06], here and in [Hartmann 09] linear elements are used, however the total number of dofs is in average preserved.

The surface of the rectangle is set as slave and the surfaces of the indenter are considered as masters. Note that both the lower surface of the indenter and the front come in contact as depicted in Fig. 6.15. A two dimensional plane strain displacement driven problem is considered, the indenter is pushed down on  $u_y = 1$  in 10 increments then to the right  $u_x = 10$  within 500 increments, the rectangle is ten times softer than the indenter. A zoom on the contact topology in the vicinity of indenter's corner is given in Fig. 6.17 for several consecutive increments. This behavior explains the oscillations in the tangential reaction (Fig. 6.16). The distribution of shear stress for several increments is represented in Fig. 6.18. All input data is summarized in the box below.



- Neo-Hookean material, large deformations:

Indenter: Young's modulus  $E^* = 68.96$ , Poisson's ratio  $\nu^* = 0.32$

Rectangle: Young's modulus  $E = 6.896$ , Poisson's ratio  $\nu = 0.32$

- Boundary conditions and geometry:

Displacement on the indenter

$$u_y = t, \quad t = [0; 1], \quad u_y = 1, \quad t > 1$$

$$u_x = 0, \quad t = [0; 1], \quad u_x = 10(t - 1), \quad t = [1; 2]$$

is prescribed on the top surface

The rectangle is fixed in all directions on the lower edge

Geometry:  $d_1 = 0.2, d_2 = 1.2, d_3 = 10.6$

$r = 0.75, h_1 = 0.95, h_2 = 1.2, a_1 = 0.3, a_2 = 0.2$

Friction coefficient  $\mu = 0.3$

- Solution conditions:

Method: augmented Lagrangian

Penalty coefficients  $\varepsilon_n = \varepsilon_t = 0.5$

Increments: 10 for  $t \in [0; 1]$  and 500 for  $t \in [1; 2]$

- Finite element mesh:

linear full integration quadrilateral plane strain elements (4 nodes, 4 integration points); 3672 elements

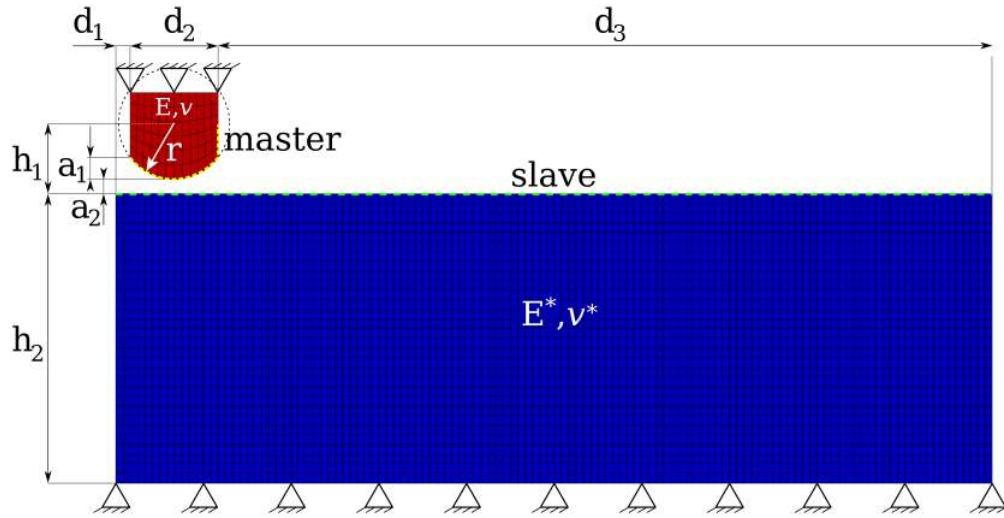


Figure 6.15: Geometrical setting of the problem and finite element mesh.

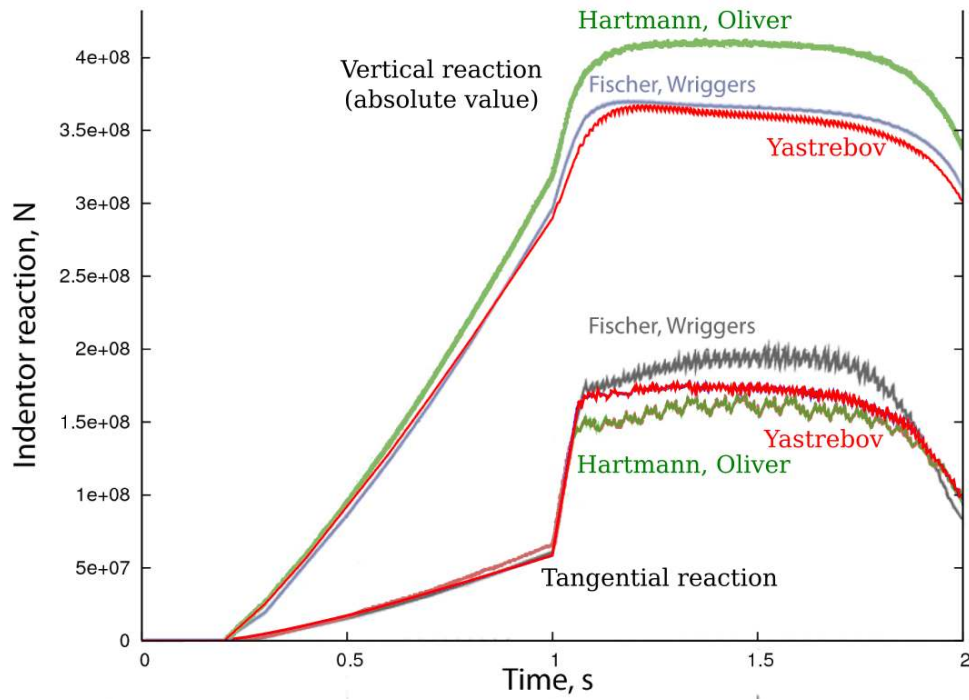


Figure 6.16: Results of the finite element simulation: comparison of vertical and horizontal reactions on the indenter, data of Fischer, Wriggers [Fischer 06], Hartmann and Oliver [Hartmann 09] and our results.

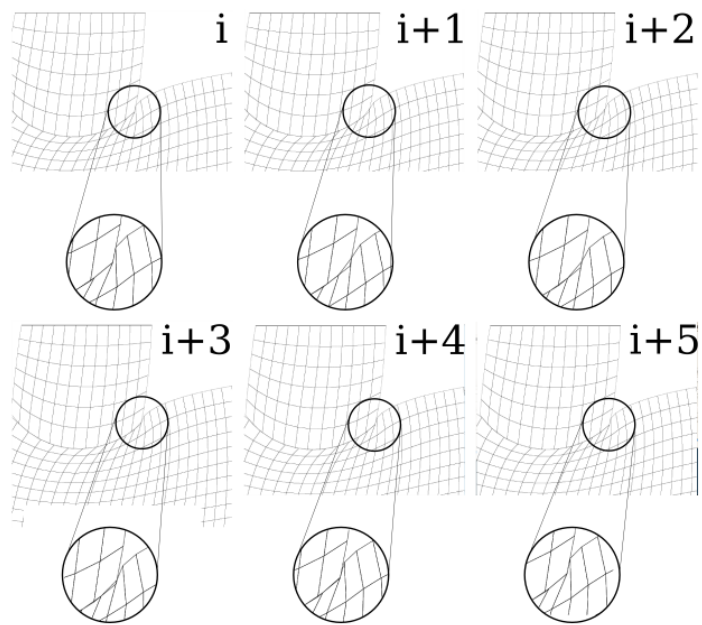


Figure 6.17: Zoom on the contact zone in the neighborhood of the indenter's angle, 6 consecutive increments.

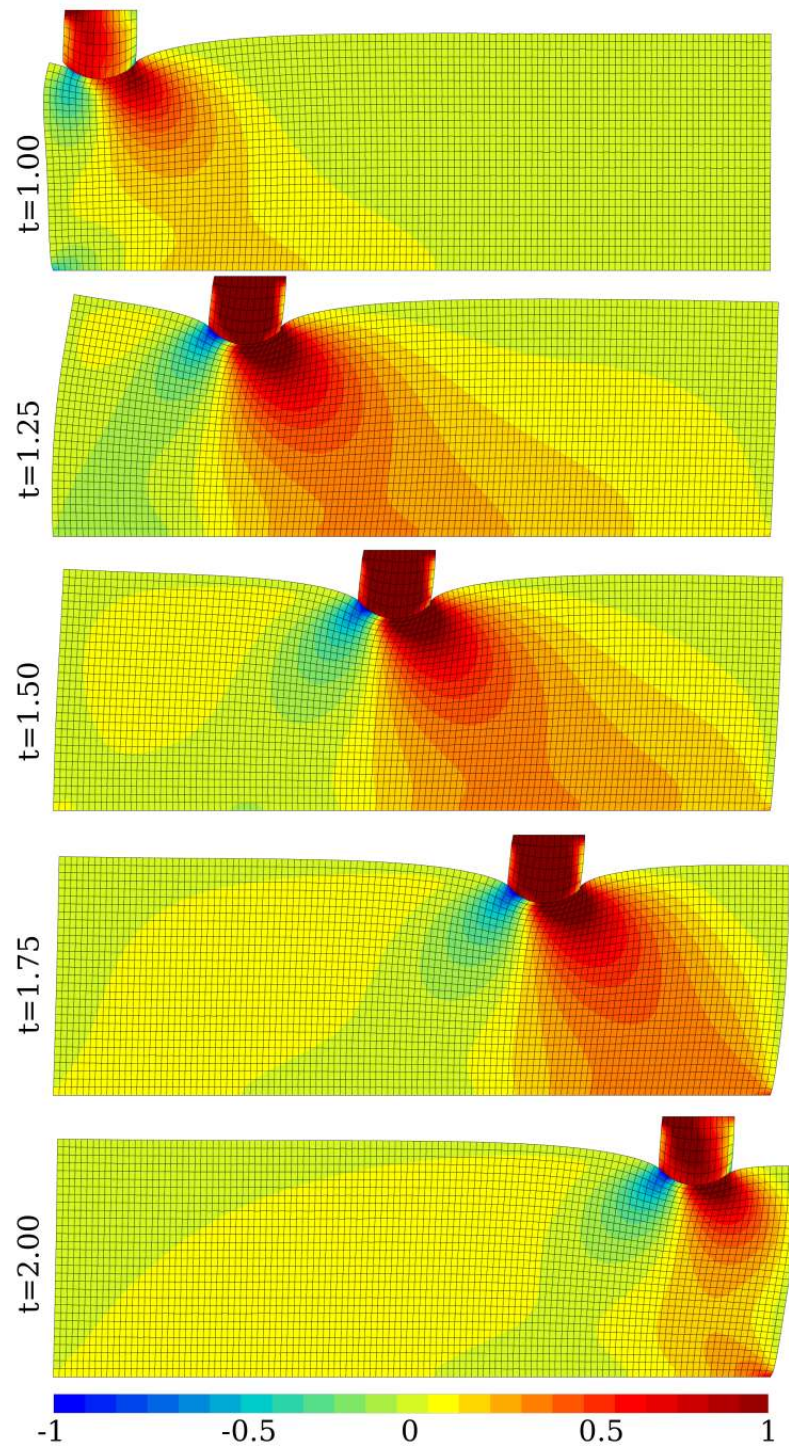


Figure 6.18: Contour plot of shear stress at time  $t = 1, 1.25, 1.5, 1.75, t = 2$ .

### 6.1.6 Axisymmetric post buckling of a thin-walled cylinder

This axisymmetric problem combines large plastic deformations, buckling, Signorini contact and self-contact phenomena. A force driven elasto-plastic cylinder is forced towards a rigid fixture, where, after some deformations, it gets stuck, and further a series of buckles occur; besides the contact between the cylinder and parts of the fixture, neighboring bends also come in contact. Since the master-slave discretization is not known a priori, the Single Surface algorithm (proposed in [Benson 90] and discussed in Chapter 3) has to be employed. The original numerical solution of this problem was given in [Laursen 92], here we use the same penalty method, but a finer finite element mesh. All details of the problem setting are given in the box below, geometry and a part of the finite element mesh are given in Fig. 6.19. Four independent contact zones can be distinguished a: self-contact within interior and exterior surfaces of the cylinder and Signorini contact between interior and exterior parts of the fixture and the corresponding surfaces of the cylinder.

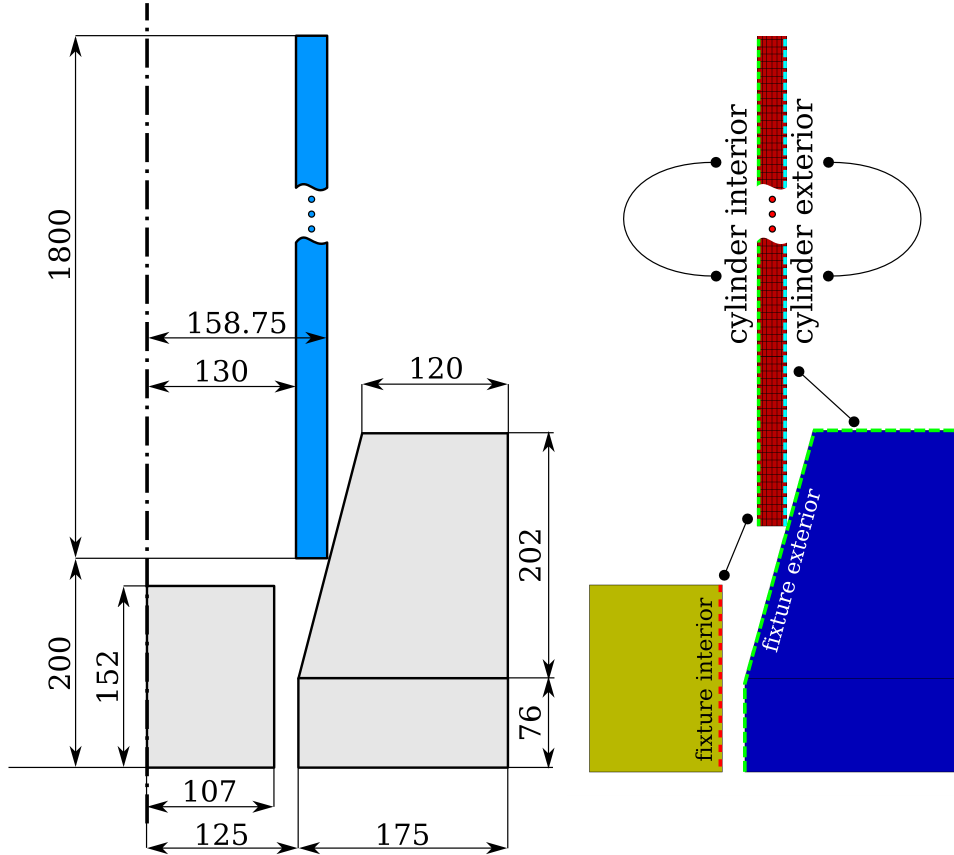


Figure 6.19: Geometrical setting of the problem and finite element mesh.

An updated Lagrangian formulation is used. The loading conditions are very severe. The accumulated plastic deformation in the interior of the bends reaches 190%. In post buckling regime, the contact topology changes

significantly during forming of each new bend. Due to the redistribution of momentum, self-contact between already formed bends vanishes on a part of the surface; Fig. 6.21 depicts the accumulated plastic strain and post buckling shapes on different time steps. It is interesting to note that the penalty parameter influences significantly the post-buckling behavior. The number of contact elements in the final stage of the post buckling deformation reaches 90 elements. PDN method, instead of penalty, is used to enforce the nonpenetration conditions on the interior surface of the fixture. The contact is assumed frictionless.

- Finite deformation elasto-plasticity (updated Lagrangian):

Young's modulus  $E = 210$ , Poisson's ratio  $\nu = 0.3$

Von Mises yield strength  $R_0 = 0.7$

Nonlinear isotropic hardening:

$$R = R_0 + Q(1 - e^{-bp}), Q = 10, b = 10$$

- Boundary conditions and geometry:

Displacement on the top of the cylinder  $u_y = -100t$ ,  $t \in [0; 13]$

Friction coefficient  $\mu = 0$ .

- Solution conditions:

Methods: PDN, penalty

Penalty coefficients  $\varepsilon_n = 5 \cdot 10^4$  for selfcontact,  $\varepsilon_n = 1 \cdot 10^5$  for contact with fixture

Time  $t \in [0; 13]$  increments  $N = 15$ ,  $t \in [0; 1.5]$ ,  $N = 240$ ,  $t \in [1.5; 10]$ ,  $N = 80$ ,  $t \in [10; 13]$

- Finite element mesh:

linear full integration quadrilateral plane strain elements (4 nodes, 4 integration points); 2117 nodes, 1803 elements

A totally different post buckling behavior (see Fig. 6.21) is obtained in the simulation with a two times higher penalty coefficient ( $\varepsilon_n = 10^5$  instead of  $\varepsilon_n = 5 \cdot 10^4$ ) and a slightly more coarse finite element mesh (765 nodes, 603 elements). No contact occurs between the cylinder and the horizontal surface of the fixture. This example emphasizes that the penalty method should be used very carefully for such error depending simulation as buckling. It is possible that assuming a nonzero friction in the self-contact interface would lead to a more stable solution.



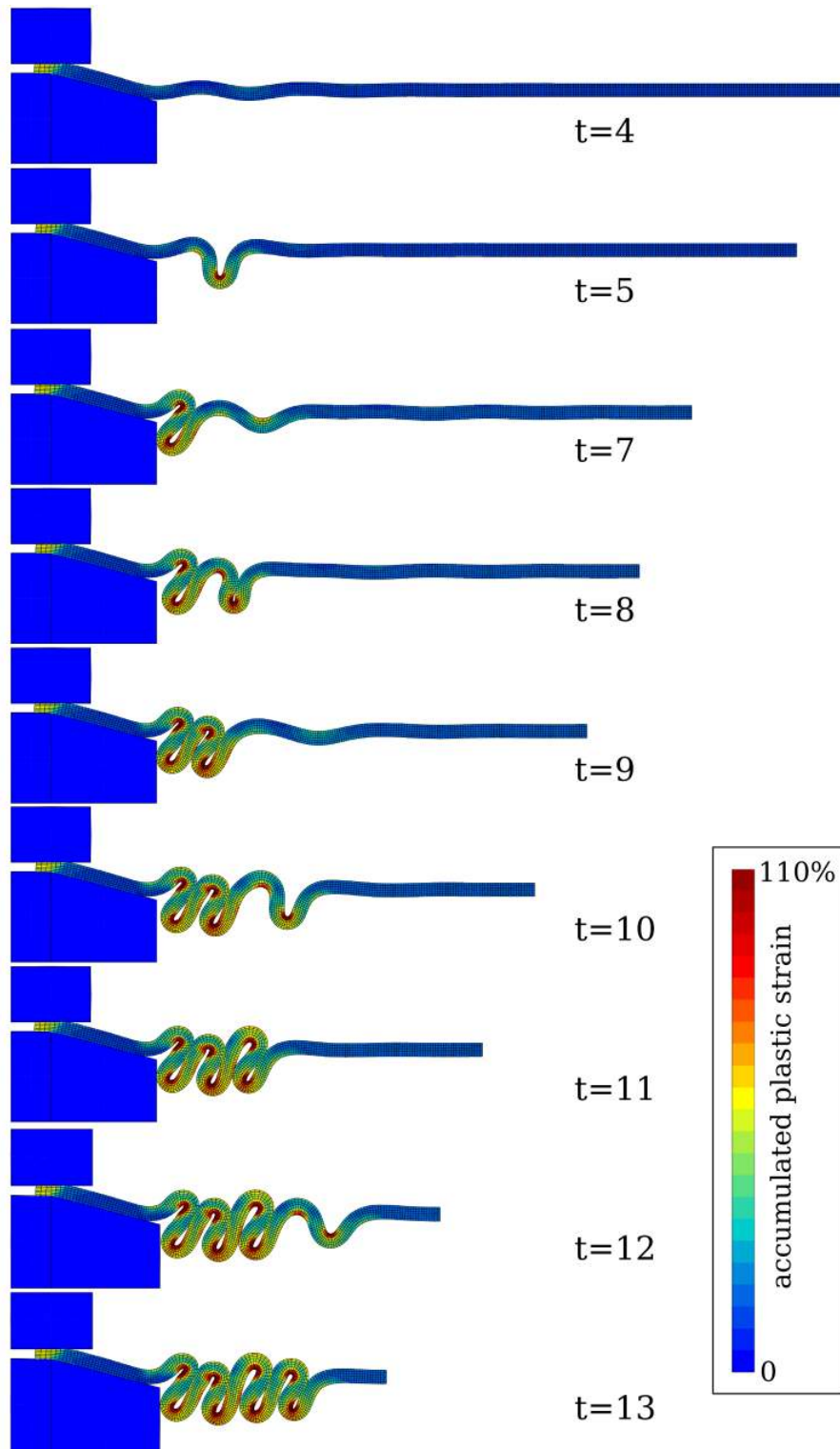


Figure 6.20: Accumulated plastic strain (Gauss points) at different time steps and corresponding geometry, scale bar is limited by 110% of accumulated plastic strain, however the maximal extrapolated value reaches 194%, maximal value at Gauss points is 181%; penalty value  $\varepsilon_n = 5 \cdot 10^4$ .

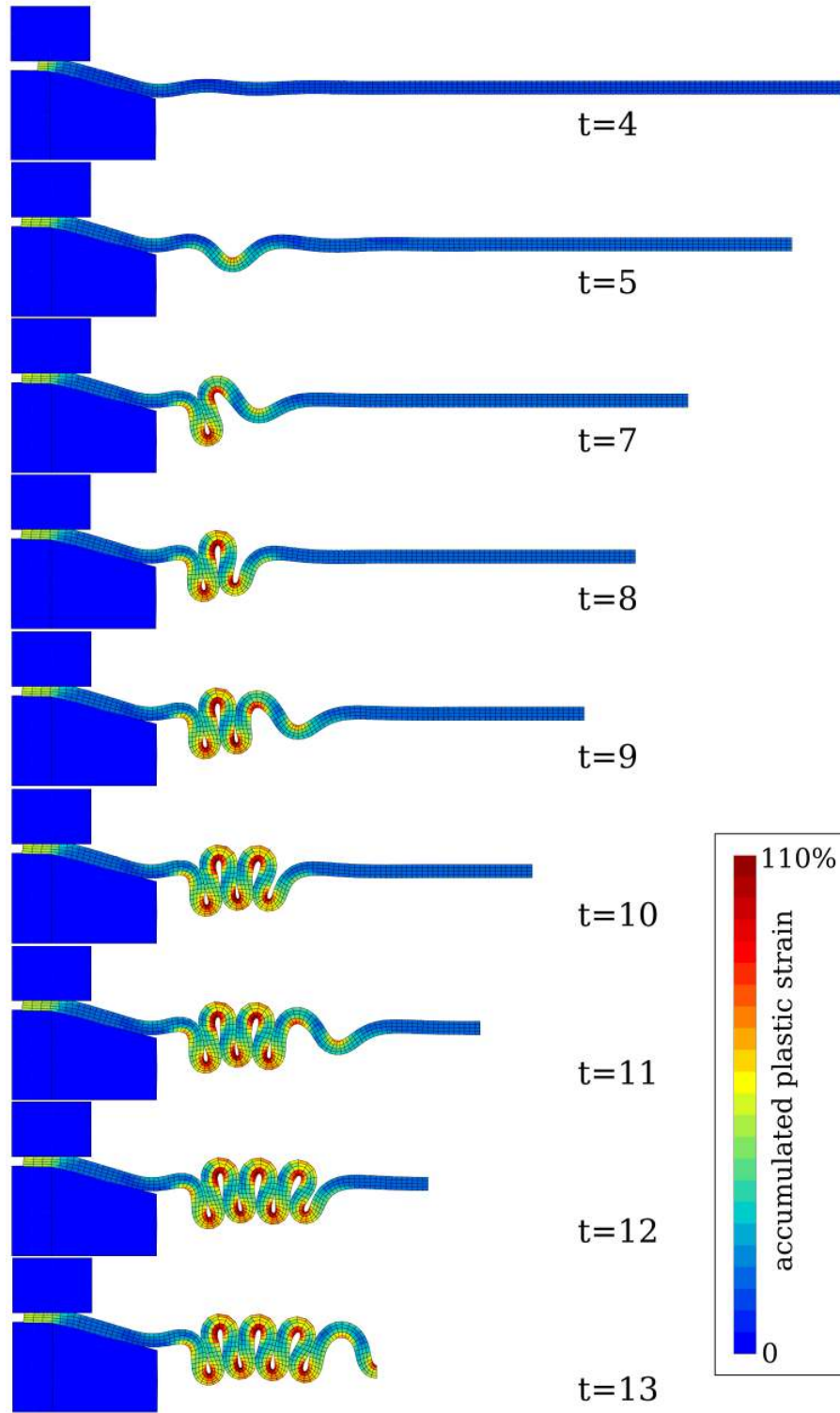


Figure 6.21: Another post buckling behavior (penalty value  $\varepsilon_n = 1 \cdot 10^5$ ): accumulated plastic strain (Gauss points) at different time steps and corresponding geometry, scale bar is limited by 110% of accumulated plastic strain, however the maximal extrapolated value reaches 156%, maximal value at Gauss points is 148%.



## 6.2 Three dimensional problems

### 6.2.1 Accordion post buckling folding of a thin-walled tube

This problem is similar to the axisymmetric post buckling behavior considered in the previous section. One quarter of an elasto-plastic thin walled tube, fixed on both edges, is compressed beyond the critical load, so that it starts to buckle first in axisymmetric mode, then in accordion mode. The problems considered in the two last sections are encountered in crashworthiness research; during post buckling folding of metal thin walled structures, large amounts of energy due to an impact are consumed. Such structures are thus used for improvement of vehicles safety.

The axisymmetry of the problem is disturbed by small errors due to numerical precision, that is why classical axisymmetric post buckling changes quickly to an accordion mode. Note that the numerical treatment of this problem is impossible without using symmetric contact boundary conditions which prevent the forming bends from penetration beyond the planes of symmetry, so the penalty method is combined with the PDN method. The geometrical setting and the finite element mesh of the problem are presented in Fig. 6.22. All necessary data are listed in the box below.

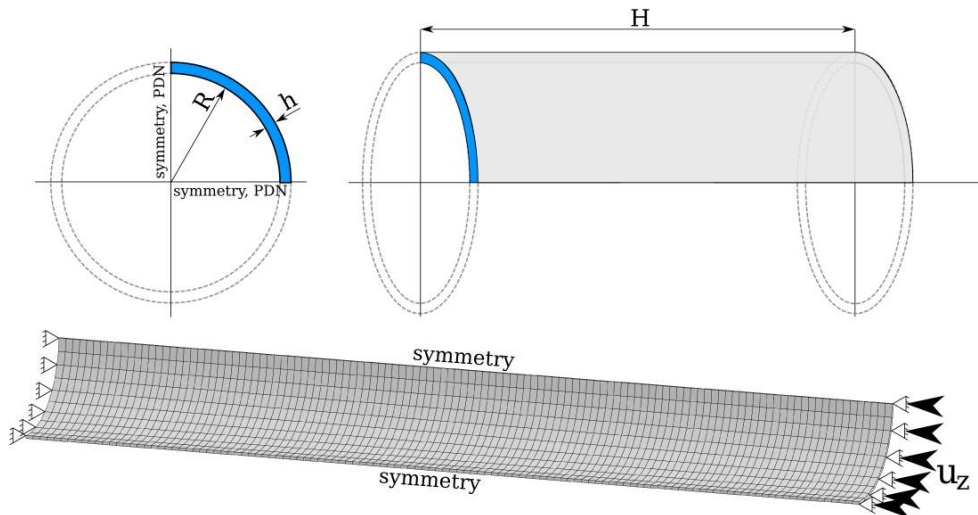


Figure 6.22: Geometric setting of the problem, boundary conditions and finite element mesh.

- Finite deformation elasto-plasticity [Weber 90] (updated Lagrangian):

Young's modulus  $E = 69$ , Poisson's ration  $\nu = 0.33$

Von Mises yield strength  $R_0 = 0.25$

Isotropic hardening constant  $Q = 1.5$

- Boundary conditions and geometry:

Inner radius  $R = 14.5$

Tube thickness  $h = 0.5$

Tube height  $H = 150$

Displacement on the top of the cylinder

$$u_y = -7t, \quad t \in [0; 10], \quad u_y = -70 - 3(t - 10), \quad t \in [10; 15]$$

$$u_y = -85 - 13(t - 15), \quad t \in [15; 16]$$

Friction coefficient  $\mu = 0$ .

- Solution conditions:

Methods: PDN, penalty (with updating coefficients)

Initial penalty coefficients  $\varepsilon_n^0 = 10^2$ , tolerable penetration  $g_n = 0.01$

Time  $t \in [0; 16]$ ; increments  $N = 100$ ,  $t \in [0; 4]$

$$N = 1500, \quad t \in [4; 15], \quad N = 300, \quad t \in [15; 16]$$

- Finite element mesh:

linear full integration brick elements (8 nodes, 8 integration points); 2662 nodes, 1200 elements

Fig. 6.23 depicts post buckling geometry and the corresponding accumulated plastic strain at several time steps, the minimal accumulated plastic strain reaches 7% and the maximal 73% at final time. The reaction-displacement curve is represented in Fig. 6.24.

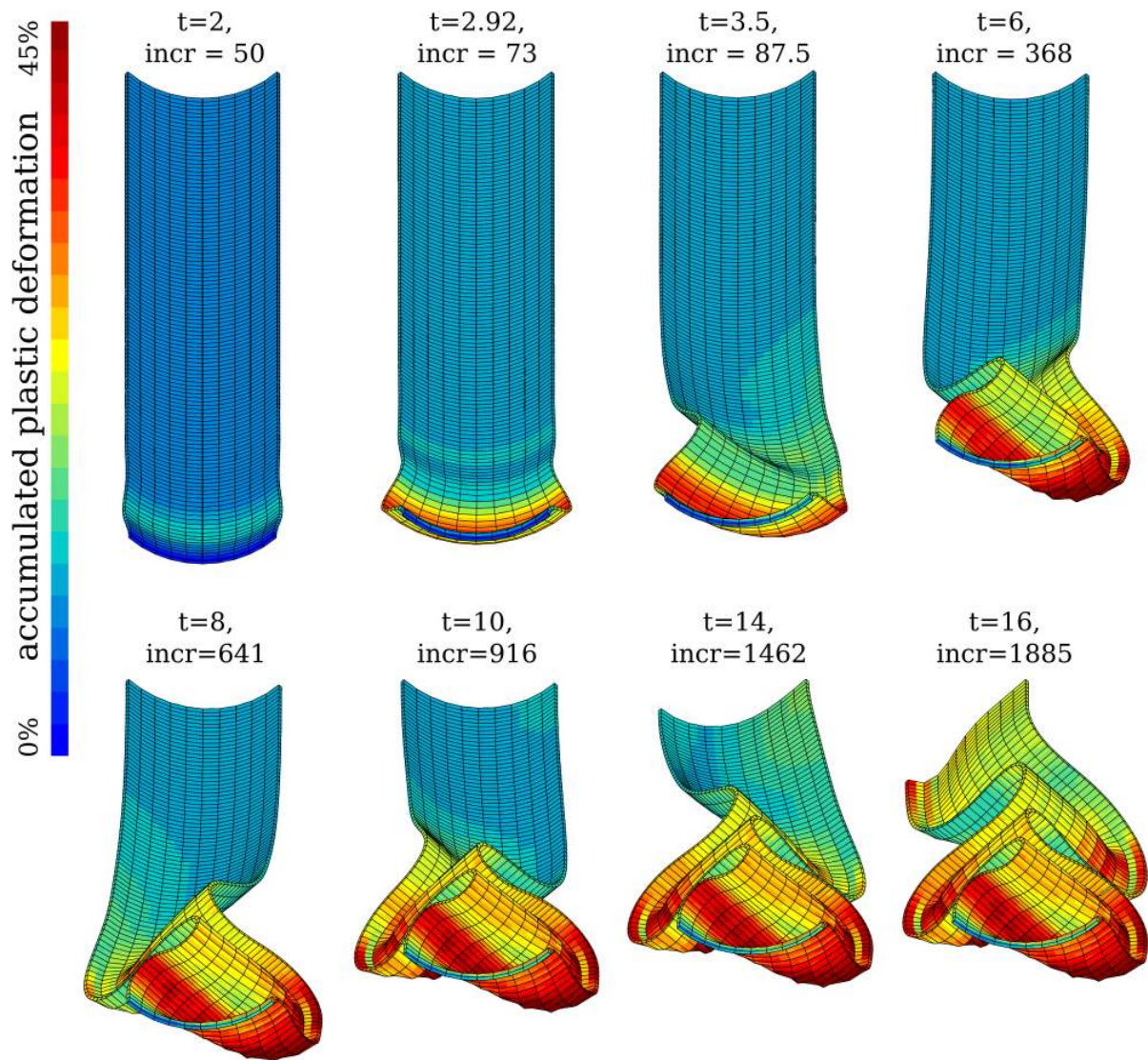


Figure 6.23: Post buckling geometry and corresponding accumulated plastic strain.

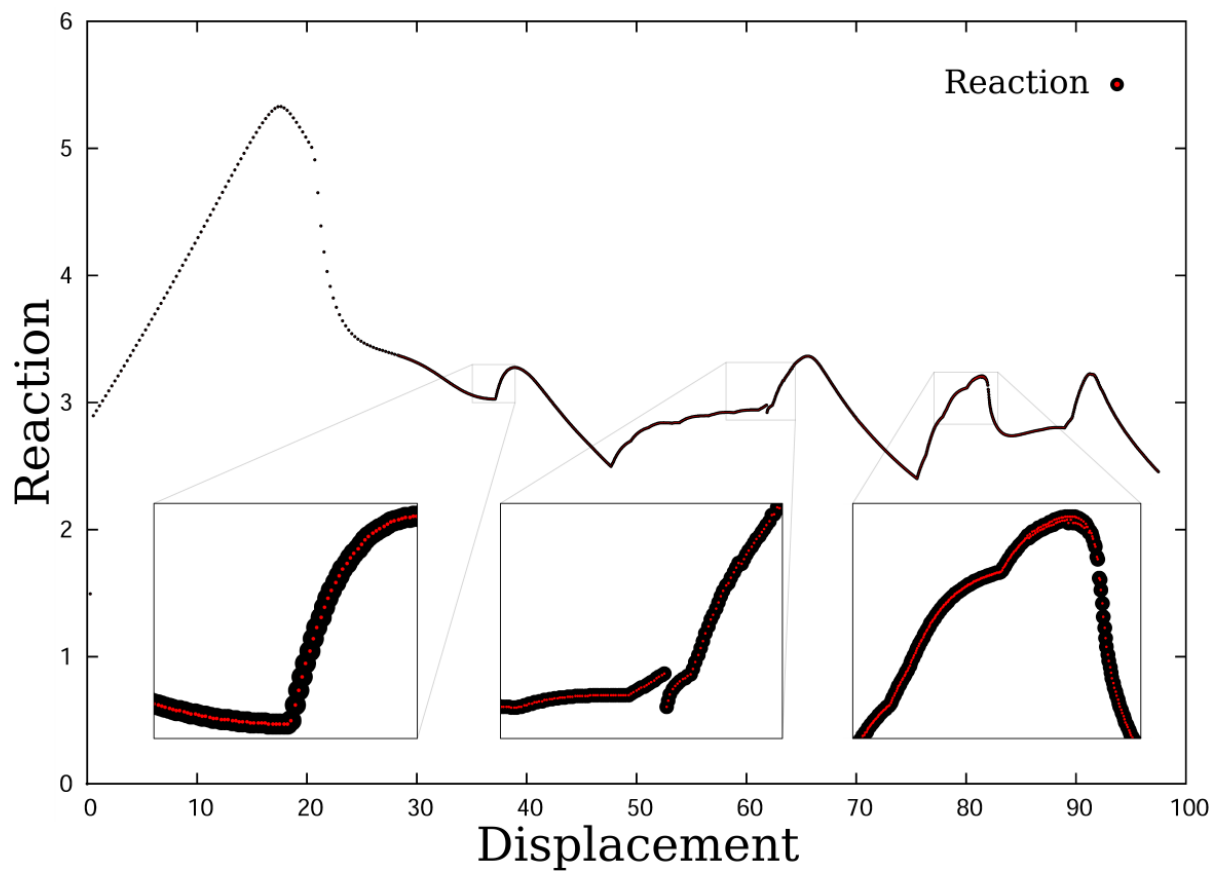


Figure 6.24: Reaction on the top of the folding cylinder.

### 6.2.2 Hydrostatic extrusion of a square plate through a circular hole

This problem deals with an artificial hydrostatic extrusion process: a thin square plate is loaded by a hydrostatic pressure from one side. On the other side, the motion of the plate is limited by a rigid foundation with a hole. Elastic and elasto-plastic material models, small and large deformations and different configurations are considered. We do not relate this simulation with any particular industrial problem, it presents simply a demonstration of the PDN method performance in case of large three dimensional contact problems between a deformable body and a flat rigid foundation with an edge, Fig. 6.25. Two finite element meshes (coarse – 900 linear element and fine – 13946 linear elements) are presented in Fig. 6.26.

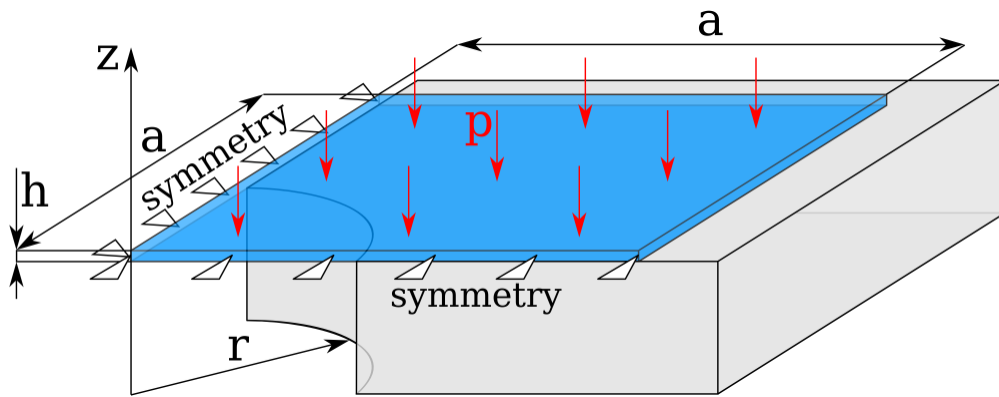


Figure 6.25: Geometrical setting of the problem.

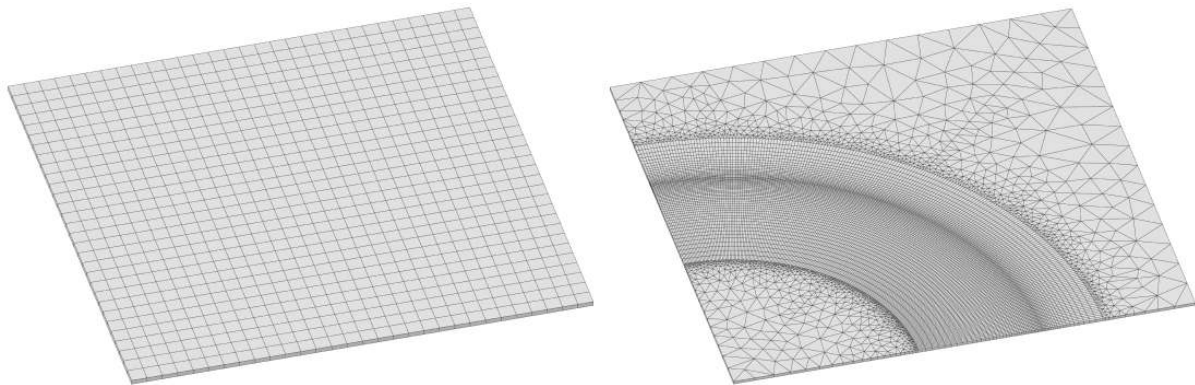


Figure 6.26: Finite element meshes: coarse and fine.

Note that the problem is not axisymmetric, but there is a symmetry of 8th order. However, to demonstrate the performance of the algorithm a quarter of the problem is considered. Regardless a high hydrostatic pressure, a detachment in the contact zone close to the hole can be observed. Due to this detachment the determination of the active contact zone takes up to 11

## 6.2 Three dimensional problems

---

iterations for the fine mesh. The problem with the coarse mesh has been solved for an elastic and elasto-plastic material models within small deformations, ratio of the plate side to the diameter of the hole is 5/2. A fine mesh is associated with a problem in large deformations and linear material, to highlight the lack of axisymmetry the ratio of the plate side to the diameter of the hole is chosen 3/2. All input data can be found in the box below:

- Finite deformation elasticity (updated Lagrangian) for fine mesh; small deformation elasticity and elasto-plasticity for coarse mesh.

Young's modulus  $E = 210$ , Poisson's ration  $\nu = 0.3$

Von Mises yield strength  $R_0 = 0.25$

Nonlinear isotropic hardening:

$$R = R_0 + Q(1 - e^{-bp}), Q = 10, b = 10$$

- Boundary conditions and geometry:

Hole radius  $r = 10$

Plate thickness  $h = 0.1$

Plate half-side  $a = 25$  (coarse mesh),  $a = 15$  (fine mesh)

Hydrostatic pressure  $p = 0.0125$  (coarse mesh),  $p = 0.0008$  (fine mesh)

Friction coefficient  $\mu = 0$ .

- Solution conditions:

Methods: PDN

Time  $t \in [0; 1]$ ; increments  $N = 1$  (coarse mesh),  $N = 12$  (fine mesh)

- Finite element mesh:

coarse mesh – linear full integration brick elements (8 nodes, 8 integration points); 1922 nodes, 900 elements

fine mesh – linear full integration brick (8 nodes, 8 integration points) and prismatic elements (6 nodes, 6 integration points); 25350 nodes, 13946 elements

The comparison of the deformed shape and the vertical displacement in the detachment zone are presented in Fig. 6.27. The maximal vertical displacement of the middle point is about  $-3$  for the case of elasto-plastic material model. It is about 30% lower for the elastic material  $-2.15$ . Detachment (positive vertical displacement) seems to be similar in both cases. A detailed zoom on



the detachment zone shows a slight difference in Fig. 6.28. The displacements in the radial plane are represented in Fig. 6.29 for coarse meshes, besides the quantitative difference, a qualitative difference can be observed in the detachment zone. Note that for the given configuration (ratio of the plate side to the diameter of the hole is 5/2) the entire lower surface of the sheet is expanding due to the applied hydrostatic pressure. Von Mises stress distribution and accumulated plastic strains are plotted on the nonscaled deformed geometry in Fig. 6.30. The maximal value of the accumulated plastic strain is situated on the edge of the contact zone and does not overpass 2.4%.

A qualitatively different result has been obtained for the fine mesh, a central region of radial expansion is followed by a contraction region, see Fig. 6.31 and in plot Fig. 6.32. Also for the fine mesh the vertical displacement both in extrusion and detachment zones appears to be axisymmetric, see Fig. 6.33. The distribution of von Mises stress for this case is given in Fig. 6.34. Slight oscillations appear in the stress distribution due to numerical errors. Since the edges of the circular hole are sharp and the detachment effect takes place, the contact zone in the vicinity of the hole edge is given by several nodes with radial coordinate  $r = 10$ , then  $r = 10 - \varepsilon$  are not in contact and  $r = 10 + \varepsilon$  are normally in the detachment region. In Fig. 6.35 only nodes which are in contact are plotted, this plot allows to visualize two things: the nonregularity of the contact on the edge of the hole and the size of the detachment zone.

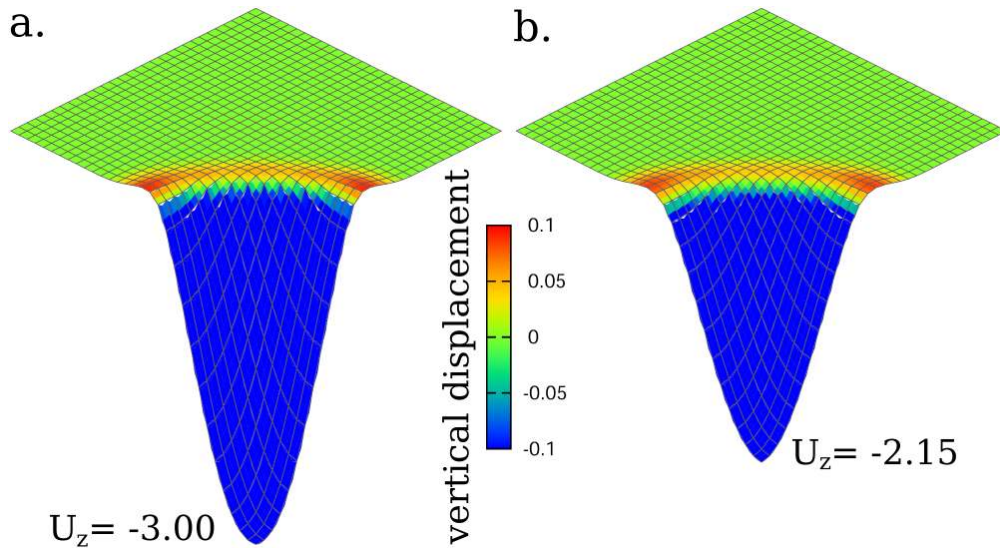


Figure 6.27: Vertical displacement on the lower surface of the deforming plate: **a** – elasto-plastic material model; **b** – elastic material model; scale of displacement 500%.

## 6.2 Three dimensional problems

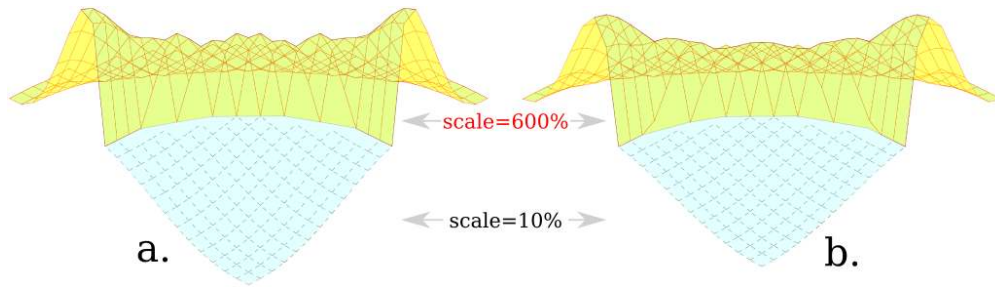


Figure 6.28: Zoom on the vertical displacement on the lower surface of the deforming plate; different scales for extrusion zone  $r < 10$  (10%) and detachment zone  $r > 10$  (600%): **a** – elasto-plastic, **b** – elastic material models.

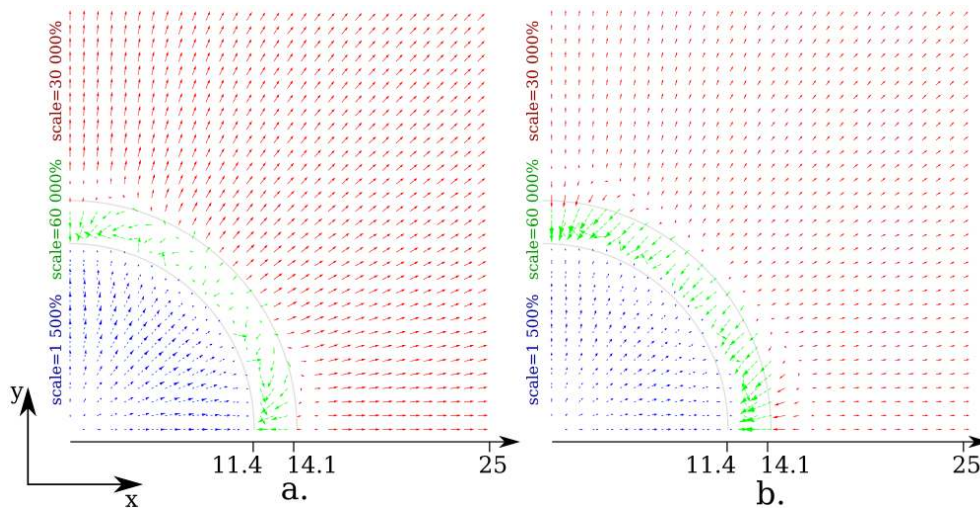


Figure 6.29: Displacement on the lower surface of the deforming plate in the horizontal plane OXY, different scales are chosen for different zones :**a** – elasto-plastic, **b** – elastic material models.

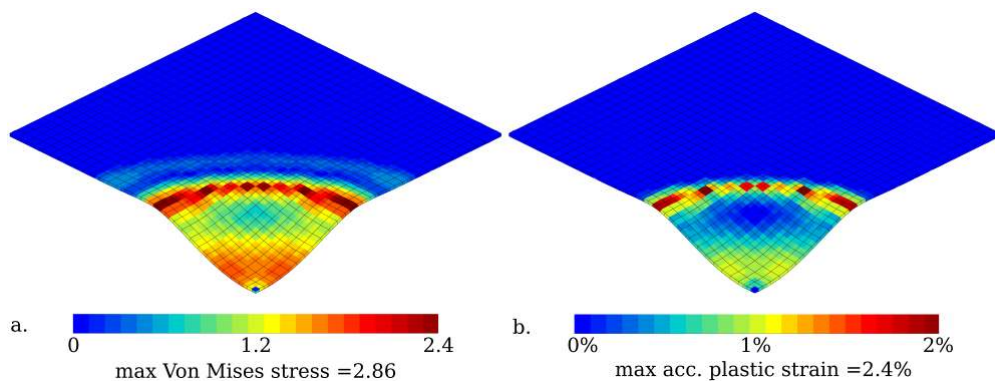


Figure 6.30: Distribution of von Mises stress at Gauss points for the elasto-plastic material (**a**), accumulated plastic strain (**b**); scale 100% .



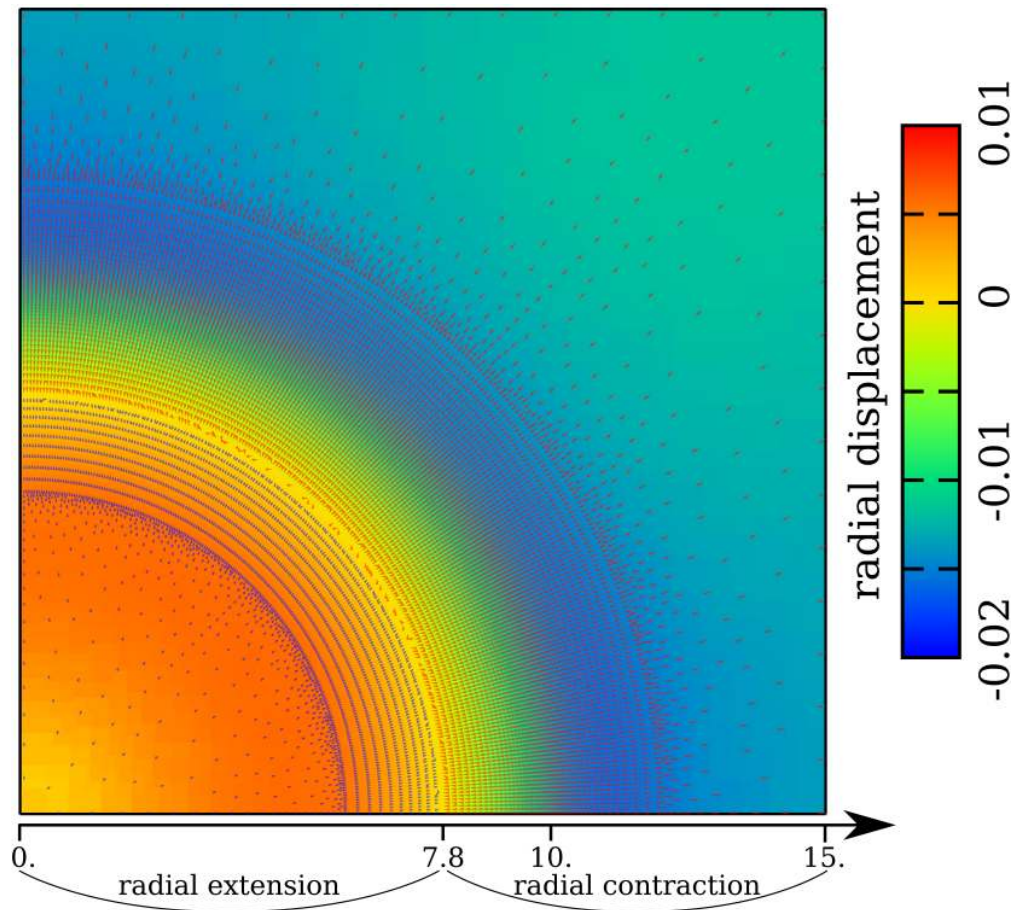


Figure 6.31: In plane displacements on the lower surface of the deforming plate in case of finite elastic deformations and fine mesh, zones of radial extension and contraction can be distinguished .

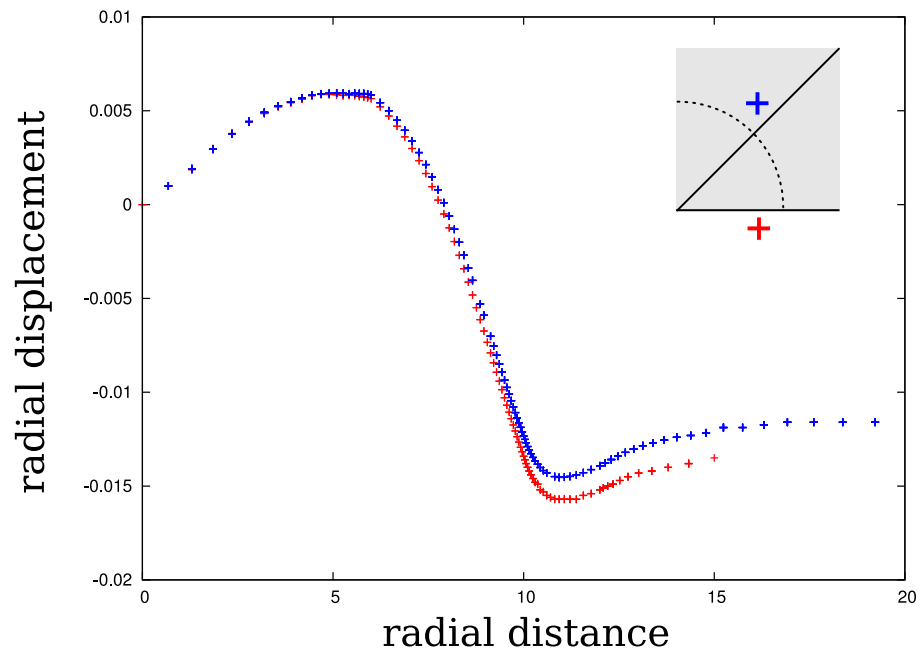


Figure 6.32: Radial displacements along two axes of symmetry.

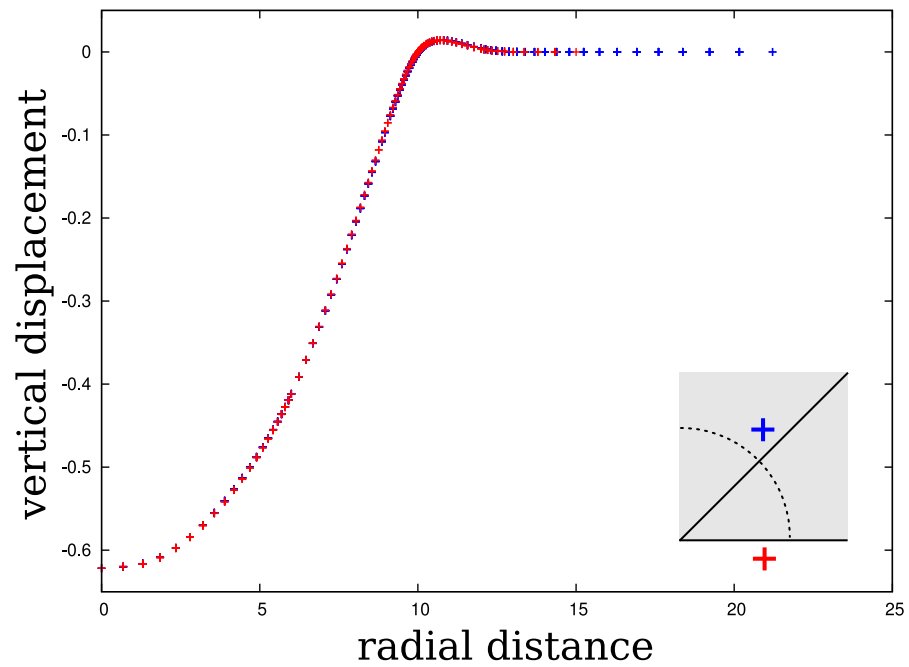


Figure 6.33: Vertical displacements along two axes of symmetry.

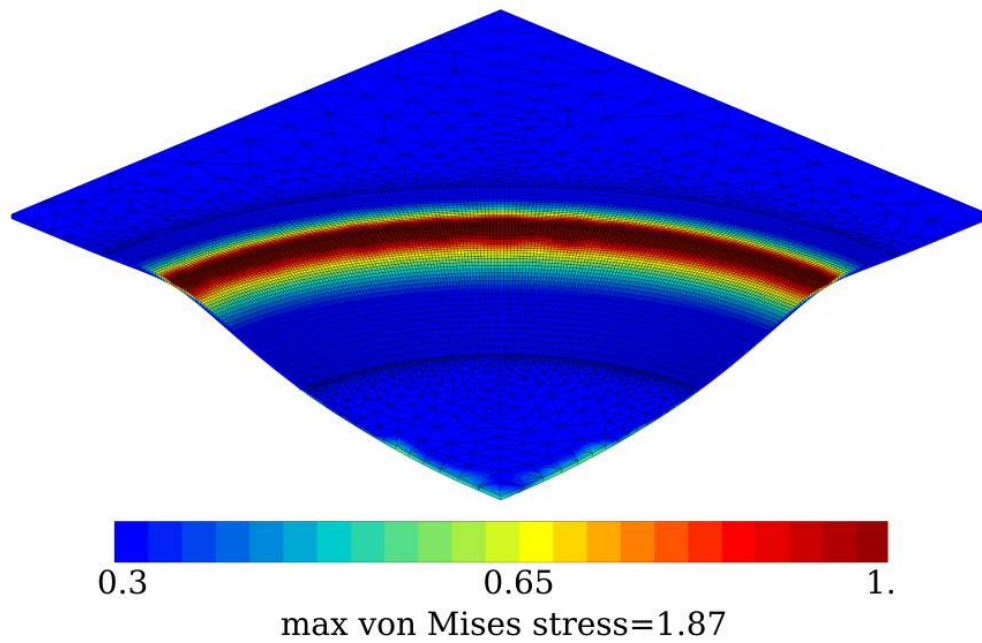


Figure 6.34: Distribution of von Mises stress (contour plot) for elastic material, large deformations and fine mesh, scale 300% .

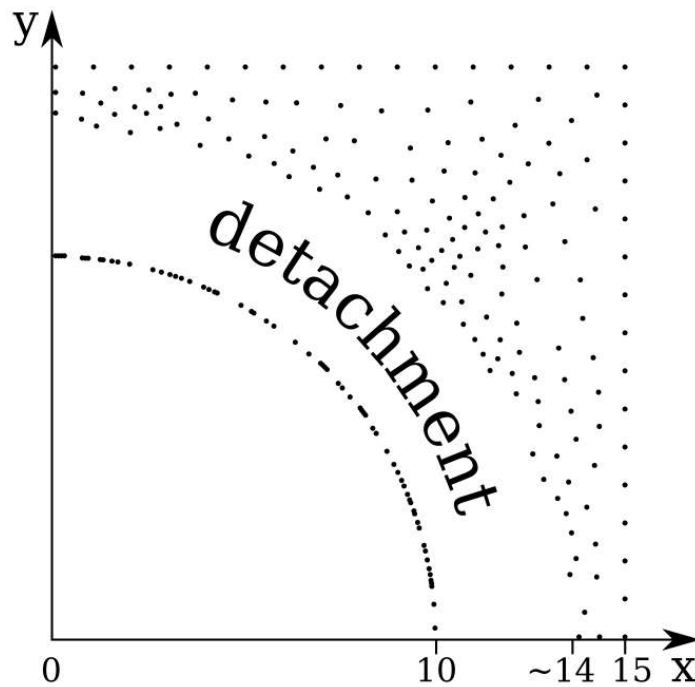


Figure 6.35: Nodes in contact (lower surface of the deforming plate in case of finite elastic deformations and fine mesh) .

### 6.2.3 Frictional sliding of a cube on a rigid plane

A frictional sliding of a deformable block on a rigid plane demonstrates the difference between local and global (measured) coefficients of friction. The geometry and the finite element mesh are presented in Fig. 6.36. A deformable cube is moved towards a rigid plane and further is moved along the plane, three coefficients of friction are considered:  $\mu = 0.2, 0.5, 0.8$ ; for the latter a detachment of a part of a cube takes place in the sliding motion. Rigorously this problem is not well posed, because of sharp angles resulting in infinite stresses; a given finite element discretization can be considered as a regularization of the problem. The mesh of the cube contains 1000 equal linear brick elements. All data are given in the box below.

- Small deformation elasticity:

Young's modulus  $E = 210$ , Poisson's ratio  $\nu = 0.3$

- Boundary conditions and geometry:

Cube side  $a = 1$

Friction coefficient  $\mu = 0.2, 0.5, 0.8$

Vertical displacement on the top of cube  $u_z = -0.05t$ ,  $t \in [0; 1]$ ,  $u_z = -0.05$ ,  $t > 1$

Horizontal displacement on the top of cube  $u_x = 0$ ,  $t \in [0; 1]$ ,  $u_x = 1/6(t - 1)$ ,  $t \in [1; 3]$

- Solution conditions:

Methods: PDN

Time  $t \in [0; 3]$ , increments:

$\mu = 0.2$ :  $t \in [0; 1.2]$ ,  $N = 12$ ,  $t \in [1.2; 3]$ ,  $N = 36$

$\mu = 0.5$ :  $t \in [0; 3]$ ,  $N = 30$

$\mu = 0.8$ :  $t \in [0; 2]$ ,  $N = 20$ ,  $t \in [2; 3]$ ,  $N = 20$

- Finite element mesh:

linear full integration quadrilateral plane strain elements (4 nodes, 4 integration points); 2117 nodes, 1803 elements

Deformed geometries and the corresponding shear stress  $\sigma_{xz}$  distributions are assembled in Fig. 6.37 for the three considered friction coefficients  $\mu = 0.2, 0.5, 0.8$  and three time points  $t = 1, 2, 3$ . The case of friction  $\mu = 0.8$  is qualitatively different from the two other, since detachment occurs in the contact zone. As a consequence, there is some stress redistribution in the contact interface and a change of the sliding velocity. Contrary to expectations,

due to this detachment, the box with a higher friction  $\mu = 0.8$  slides further than the box with a lower friction  $\mu = 0.5$ . The total vertical reaction scaled by the friction coefficient  $\mu|P|$  and the tangential reaction  $|T_x|$  are plotted in Fig. 6.38. Remark that in slip state the absolute value of the tangential reaction is not equal to the absolute value of the contact pressure multiplied by a friction coefficient  $\mu|P| \neq |T_x|$ , but

$$\mu|P| > |T_x|.$$

This is due to the fact that the shear stress in the direction orthogonal to sliding  $\sigma_{zy}$  is not zero. Nevertheless, the problem remains symmetric respectively to the XOZ plane passing through the middle of the cube, so the integral of  $\sigma_{zy}$  over the contact interface is zero. To make it evident, nodal tangential reactions in the contact zone are represented in Fig. 6.39. The difference between the global coefficient of friction  $\bar{\mu}$  and the predefined local one  $\mu$  remains low. Relative error  $(\mu - \bar{\mu})/\mu$  is 3% for  $\mu = 0.2$ , 2% for  $\mu = 0.5$  and only 1.1% for  $\mu = 0.8$ .

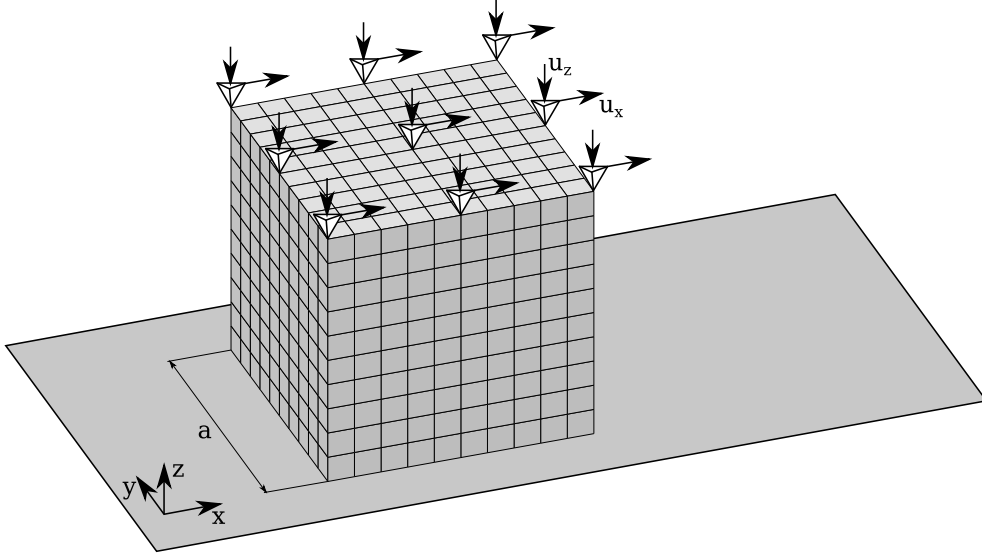


Figure 6.36: Geometrical setting of the sliding cube problem: finite element mesh and boundary conditions.

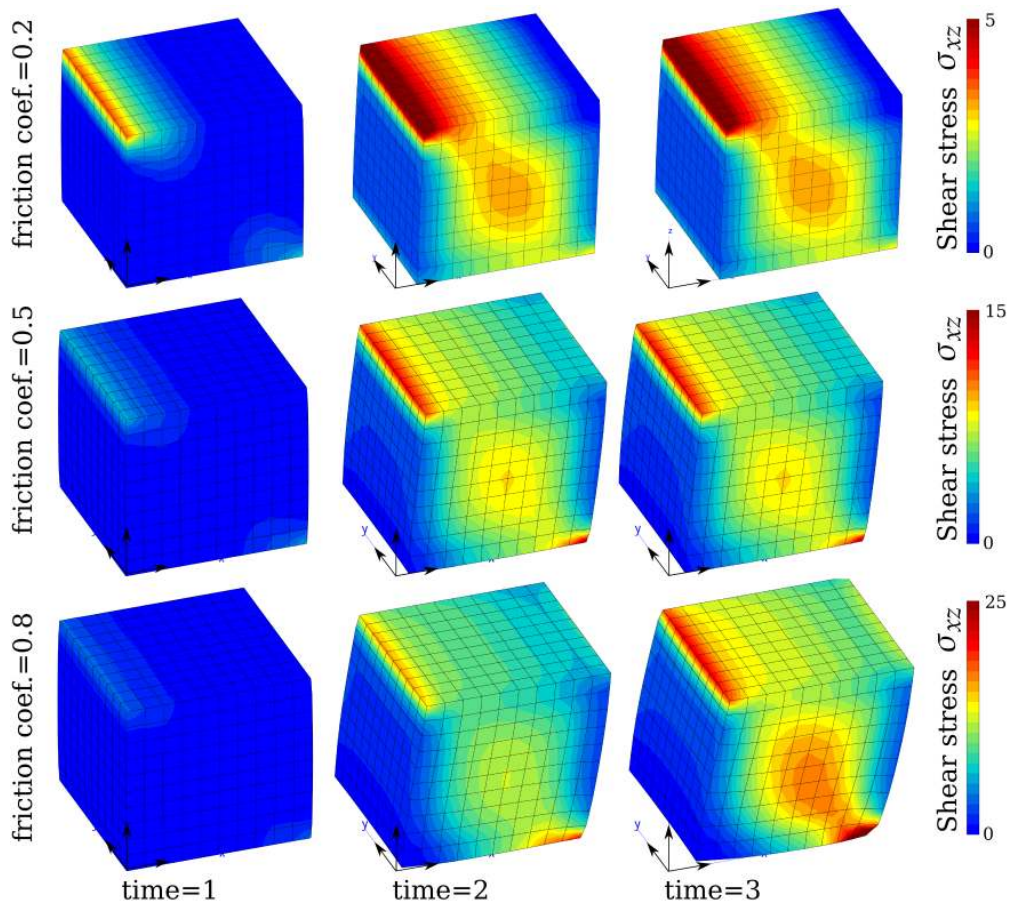


Figure 6.37: Contour plots of shear stress  $\sigma_{xz}$  for different friction coefficients and time moments.

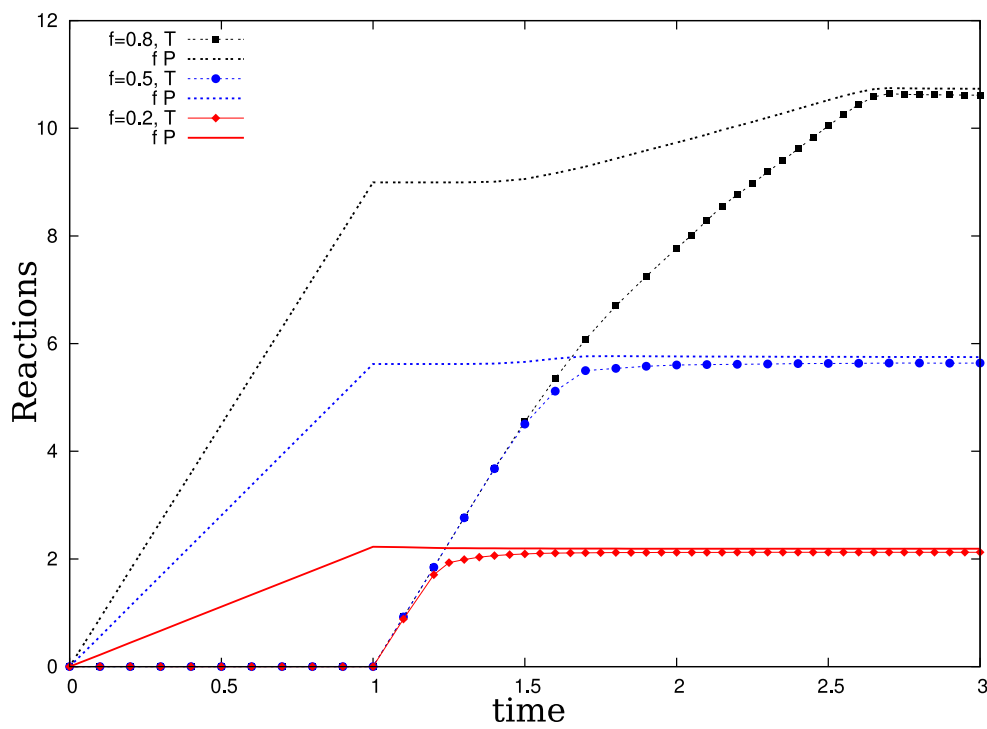


Figure 6.38: Evolution of the scaled normal reaction  $\mu|P|$  and tangential reaction  $|T_x|$ .

## 6.2 Three dimensional problems

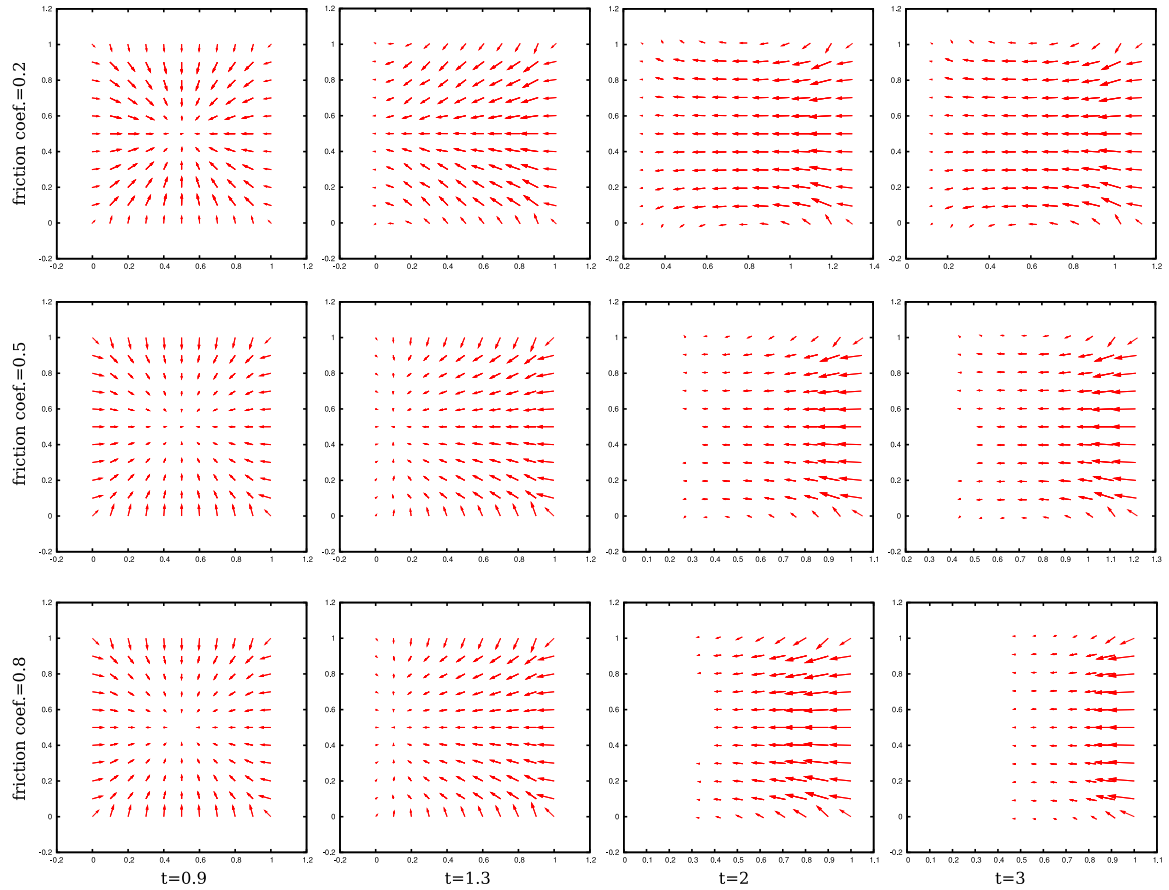


Figure 6.39: Tangential reactions on the bottom of the cube for different friction coefficients and time moments  $t = 0.9, 1.3, 2, 3$ ; different scales are used for different plots.



## Chapter 7

# Conclusions and perspectives

### **Résumé de Chapitre 7 «Conclusions»**

*Ce chapitre énumère les contributions originales de la thèse et donne des perspectives pour les travaux futurs. Les contributions originales peuvent être distinguées par les thèmes dans lesquels elles se situent : géométrie, détection et résolution.*

- **Géométrie :**

- définition rigoureuse du point le plus proche pour le cas d'une surface lisse par morceaux ;
- nouvelle procédure de projection de type «ombre portée» ;
- expressions «prêtes-à-implémenter» pour les quantités cinétiques sont dérivés pour la discrétisation arbitraire de la surface maître ;
- méthode d'enrichissement de la géométrie de contact ;
- nouvel algèbre de s-structure.

- **Détection :**

- utilisation de boîtes englobantes pour accélérer le processus ;
- parallélisation de la méthode de «bucket sort» ;
- établissement des paramètres optimaux pour le «bucket sort» ;
- nouvelle procédure de détection basée sur une définition rigoureuse du point le plus proche.

- **Résolution :**

- nouvelle approche de résolution des problèmes de contact unilatéral (PDN) ;
- considération de fonctions de pénalisation non linéaires et de différentes stratégies «active set» pour la méthode des multiplicateurs de Lagrange.

*Les méthodes développées dans cette thèse sont applicables à une très grande variété de problèmes de contact. L'une des prolongations en cours concerne le problème de contact entre surfaces rugueuses, pour lequel une discrétisation très fine est indispensable, afin de bien représenter le caractère fractal de la surface réelle. La résolution de ce type de problème demande un traitement sur machine parallèle [Yastrebov 11b]. Une autre application qui est prévue dans un avenir proche présente l'analyse sous chargement cyclique d'un composant aéronautique constitué de trois grandes pièces assemblées par 45 boulons.*

The goal of this work is to derive a consistent framework for the treatment of contact problems within the Finite Element Method using the Node-to-Segment discretization. Three main components of the computational contact have been considered: geometry, detection and resolution techniques. For the sake of completeness, the mechanical aspects of contact as well as numerous numerical algorithms and methods have been discussed. A new mathematical formalism called “s-structures” has been employed through the entire dissertation. It results in a comprehensive coordinate-free notations and provides an elegant apparatus, available for other mechanical and physical applications. Several original ideas and extensions of standard techniques have been proposed and implemented in the finite element software ZéBuLoN (Z-set). Numerical case studies, presented in the dissertation, demonstrate the performance and robustness of the employed detection and resolution schemes.

## 7.1 Original contributions

### 7.1.1 Geometry

The classical normal projection technique for searching for the closest point is widely used in computational contact; however this formulation has some inherent problems: the objective distance function used for the determination of the closest point may be not smooth and the minimum point may be not unique. The latter results in a discontinuous projection when the point slides along a surface. To provide an always existing normal projection 1) *we have proposed to seek for an infimum of the minimal distance on the union of open sets of master edges and segments as well as on the discrete set of master nodes*, see Section 2.2.1. 2) *A new technique, shadow projection, has been proposed* (see Section 2.2.3), which ensures an always existing and unique projection of any point onto the surface. Moreover this projection is continuous, i.e. a small change of the geometry results in a small change of the projection point. However, this method is not applicable for arbitrary surfaces and the imaginary “emitting light” point has to be chosen carefully.

In Sections 2.2.5 and 2.2.6 the first and second order variations of the geometrical quantities are derived for continuous geometries. Arbitrary gaps, normal and 3) *shadow projections* have been considered. The accuracy of derived expressions have been validated by several numerical tests (see Section 2.3). 4)

*Statistical analysis* of these tests allowed to demonstrate the difference between the real variations of geometrical quantities and the designed analytical expressions. Moreover the tests allowed to compare the error limits and their deviations between the rigorous expressions and their truncated versions with and without use of the assumption of infinitely small gaps.

In Section 2.4 the 5) *ready-to-implement expressions for variations of geometrical quantities have been derived*. The expressions are valid for any discretization of the master surface including smoothing techniques. It remains only to choose the appropriate shape functions.

In order to enrich the contact geometry on the submesh scale or to take into account its change due to loading, 6) *an analytical description of the enriched geometry and needed variations have been derived* in Section 2.5. An example of closed form expressions is given for the two-dimensional linear frictionless element.

All geometrical computations have been carried out using 7) *the new mathematical formalism of s-structures*. The s-structures or set-structures consist in a generalization of a set of tensor components with related internal and external operations. In Appendix A the notion of s-structures is extended and an attempt to construct invariant structures has been undertaken. Moreover 8) *an extended introduction to tensor algebra* is given within the s-structure formalism. The s-structures have been initially introduced in order to carry out computations with objects of different dimensions and to provide a coordinate-free framework.

### 7.1.2 Detection

A comprehensive framework for the NTS contact detection inspired from the bucket sort method is derived. The original approach proposed in [Benson 90] has been complemented by the 9) *bounding box concept* and 10) *extended to the parallel case* [Yastrebov 09]. Two approaches for parallel detection have been proposed in Section 3.6 and implemented in the finite element code: the Single Detection, Multiple Resolution (SDMR) and the Multiple Detection, Multiple Resolution (MDMR). The MDMR scheme has been inspired from the Linked Cell Method widely used in the Molecular Dynamic simulations with short range potentials. The parallel test of a contact between two rough surfaces has demonstrated a good scalability of the approach.

For the closest node based detection 11) *a strong connection between the master mesh, the maximal detection distance and the optimal bucket cell size has been established and validated on various numerical tests* in Section 3.3.2. Contact detection based on the closest node requires additional verifications due to blind angles and passing by nodes. The techniques allowing to overcome these difficulties have been discussed. The derived numerical scheme is valid both for known and unknown a priori master-slave discretizations [Yastrebov 11a].

Next we have demonstrated that the contact searching algorithms based on the closest node detection are not robust (see Remark 3.1), that is why in Section 3.3 we have proposed to 12) *construct a detection procedure based on the rigorous definition of the closest point*. The performance of the method has

been tested on large contact problems (up to 2 million nodes in contact), the designed algorithm in sequential case works up to 160 000 times faster than a simple all-to-all technique.

### 7.1.3 Resolution

Based on a simple case of contact between a deformable solid and a rigid plane, the classical Hertz-Signorini-Moreau contact conditions and frictional conditions are replaced by partial Dirichlet-Neumann boundary conditions on the active contact zone. Further this approach has been extended for the case of an arbitrary rigid surface. Within this representation, *13) a simple and robust approach to treat frictionless and frictional contact problems has been derived.*

After consideration of the variational inequalities approach, we focus on the implementation of three basic methods in the framework of variational equalities: penalty (PM), Lagrange multiplier (LMM) and augmented Lagrangian (ALM) methods. The variational formulation for each method has been derived; the principle of operation has been demonstrated on a simple test. *14) Nonlinear penalty functions are considered within the PM (Section 4.5.3) and different local active set strategies are compared within the LMM (Section 4.6.3).*

For the sake of completeness the basic ingredients of the Newton's scheme (Section 5.1) and the return mapping algorithm (Section 5.2) have been presented in the context of the computational contact. *15) A geometrical interpretation of subdifferential formalism has been proposed in Section 5.1.4.* This interpretation is also valid to establish an always existing normal projection. Further the finite element formalism has been presented using the s-structures (Section 5.3).. Elementary contribution of the NTS contact elements to the virtual work of the system and its linearization has been derived for the PM (Sections 5.4.1, 5.4.1 for frictionless and frictional cases) and the ALM (Sections 5.4.3, 5.4.4)). Some relevant technical details have been discussed in Section 5.6.

The accuracy, performance and robustness of the implemented techniques and methods have been validated on nine two- and three-dimensional contact problems, among them:

- two with available analytical solutions (Section 6.1.1 – indentation of a half space by a rigid cylindrical punch, Section 6.1.2 – elastic disk embedded in an elastic bored plane);
- two problems with numerical solutions obtained by other authors (Section 6.1.5 – shallow ironing, Section 6.1.6 – post-buckling folding of a thin-walled tube);
- two highly nonlinear problems with frictional contact and finite strain plasticity (Section 6.1.4 – axisymmetric deep cup drawing, Section 6.2.1 – accordion post-buckling folding of a thin-walled tube);

## 7.2 Intermediate results and perspectives

The fast and robust parallel framework established in the PhD dissertation allows to perform large parallel computations including frictional contact. Two large applications are to be considered:

1. analysis of the normal contact between rough metallic surfaces;
2. simulation of the aeronautic turbine disk consisting of three separate parts assembled by 45 bolts.

Further investigations are required to improve the stability of the parallel scheme. Different types of contact element distribution among the subdomains have to be investigated. A standard test procedure has to be established in order to analyze the performance of detection/resolution steps in parallel treatment of contact problems.

### 7.2.1 Normal contact of rough surface

The problem of the normal contact between two metallic rough surfaces arises, for example, in metal-to-metal water seals. The aim of the study, carried out in close collaboration with Julian Durand (PhD student at the Centre des Matériaux, Mines ParisTech) and Electricité de France (EDF), is to predict the leakage through such water seals for a specific material, used in nuclear power plants.

An accurate measurements of various rough surface 3D profiles have been carried out by C. Vallet [Vallet 09]. Based on these measurement ( $610\ \mu\text{m} \times 460\ \mu\text{m}$  and  $736 \times 480$  pixels), a Representative Surface Element has been chosen ( $54\ \mu\text{m} \times 63\ \mu\text{m}$ ) based on the height distribution. Further the resulting surface profile has been enriched by a bi-cubic Bézier surface. The resulting profile has been projected on a flat finite element mesh. The finite element analysis has been carried out on a mesh containing 964 000 nodes. More than 945 000 of these nodes are located in 8 layers forming a regularly meshed zone adjacent to the contacting surface (Fig. 7.1). An elasto-plastic material model with an isotropic exponential hardening has been used within the updated Lagrangian formulation for finite strain plasticity. Symmetric boundary conditions have been imposed on the lateral sides of the specimen. The maximum applied normal load generates a contact area equal to 6% of the nominal area. The PDN method for contact and the FETI (Finite Element Tearing and Interconnecting, [Farhat 94]) method have been used to perform the full computation. The finite element analysis demonstrated:

- a linear increase of the real contact area with increasing normal load;
- a strong interaction between asperities;
- the evolution of the real contact area and of the residual volume as well as their topologies to analyze the percolation;

- a statistical study of the three-dimensional rough surface has shown the that distribution of asperity relative heights is not Gaussian.

The maximal accumulated plastic deformation in the test on the Representative Surface Element reached 200%, i.e. since the metal cannot sustain such a large deformations, a damage and/or a fracture mechanism have to be introduced in the material model in order to make the simulation more realistic. The von Mises stress distribution on the surface and the vertical displacement for the ultimate load are depicted in Fig. 7.2.

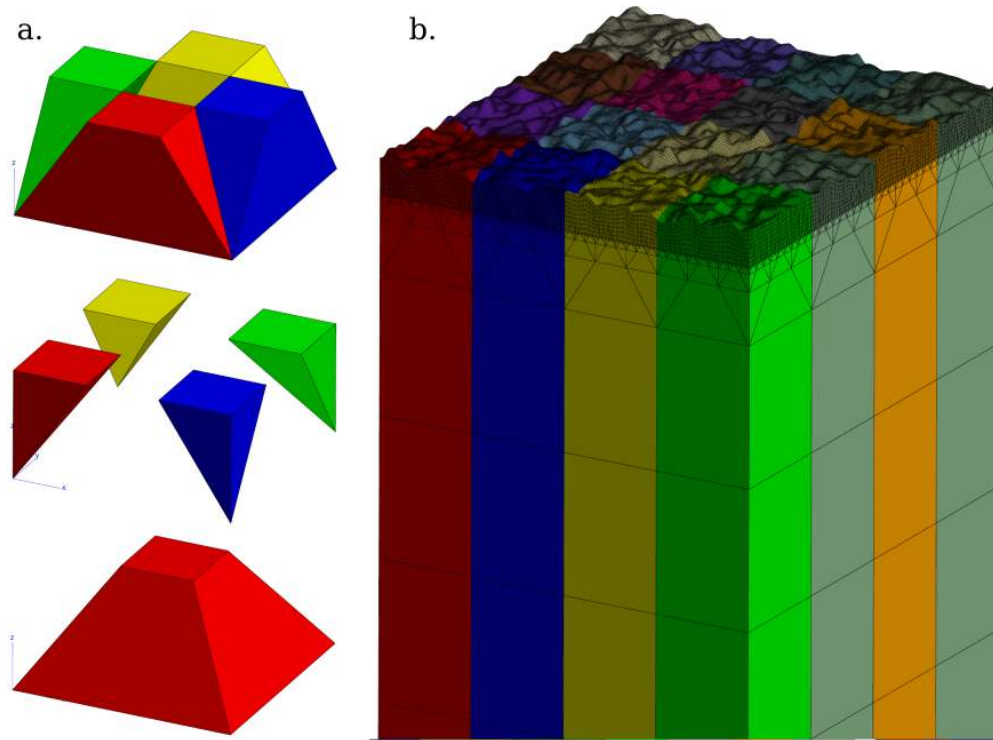


Figure 7.1: The finite element mesh used for the analysis of rough surface in contact: **a** – assembly of 9 elements to get a fast transition from coarse mesh to fine mesh; **b** – full finite element mesh containing 964 000 nodes, more than 945 000 of these nodes are located in 8 regularly meshed layers close to the surface; different color designate subdomains for the parallel computation .

At the same time a reduced model has been proposed based on the finite element simulations of a single asperity and periodic structures of similar asperities. Different asperity shapes and geometrical characteristics, coefficients of friction and meshes have been used. The reduced model is based on the accurate analysis of the rough surface (location and geometrical characteristics of each asperity) and in the reconstruction of the surface consisting of super-elements – asperities with a prescribed force-displacement behavior. The interaction between asperities has been also taken into account by an analytical function of the relative displacement, distance and geometrical



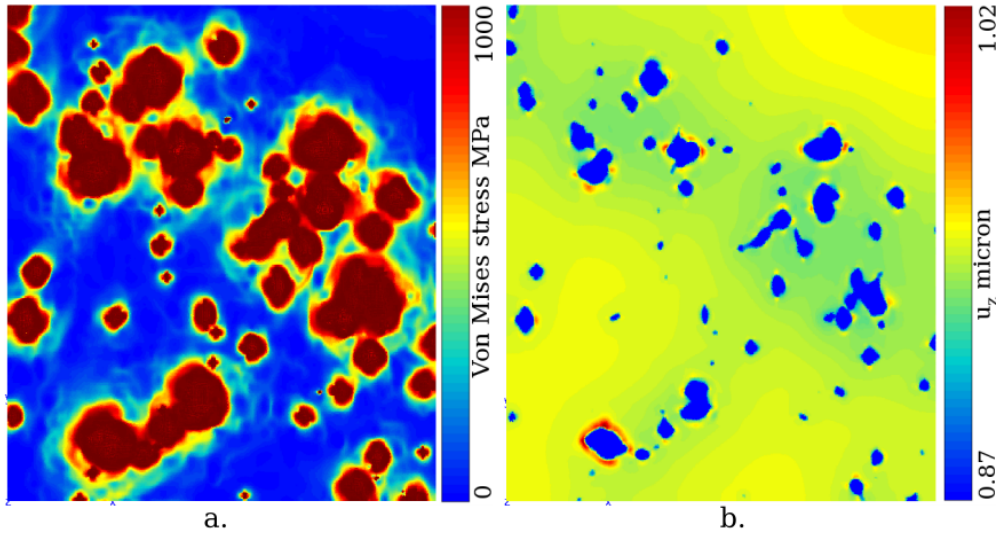


Figure 7.2: Results of the FEA of the normal contact between a metallic rough surface and a rigid plane: **a** – von Mises stress distribution, **b** – vertical displacement (blue color correspond to the real contact area) .

characteristics of each asperity. This function has been calibrated by means of a series of finite element analyses. The displacement–force, force–contact area and force–residual volume curves, predicted by the reduced model are in good agreement with large parallel finite element simulations as well as the topology of the real contact area and the residual volume. For more details, see [Yastrebov 11b].

In perspective, the parallel framework and the augmented Lagrangian method will be used to simulate a contact between two rough deformable surfaces. Such an analysis will allow to determine the application limits of the Johnson’s assumption<sup>1</sup>.

### 7.2.2 Aeronautical applications

A large aeronautical engineering problem has been considered. Three parts of a turbine disk are assembled by 45 bolts and loaded by a centrifugal load. The difficulty of the problem arises, first, from the fact that all parts of the disk allow rigid-body motions, which should be removed by contact constraints and second, due to the relatively large size of the mesh (180 000 nodes coarse mesh, 820 000 fine mesh) the number of contact elements and the number of independent contact zones. Moreover the aim of this study is to investigate the cycling loading of a turbine disk due to a) the presence of a small eccentricity, which results in an oscillation of the load and b) the start-up and shut-down cycles. That is why many time steps are needed. The need to consider the full

<sup>1</sup>Johnson’s assumption affirms that the contact between two elastic frictionless rough surfaces can be replaced by the contact between a rigid plane and a body with effective material properties and with a roughness obtained by the “sum” of the roughnesses of two undeformed surfaces.

assembly instead of a sector, stems out of the need to investigate the behavior of precracked disks.

The assembly of a disk without cracks has been analyzed by means of a representative sector (the stress states for different time steps are presented in Fig. 7.3). The next step is the parallel treatment of the entire assembly. The partition on subdomains is depicted in Fig. 7.4.

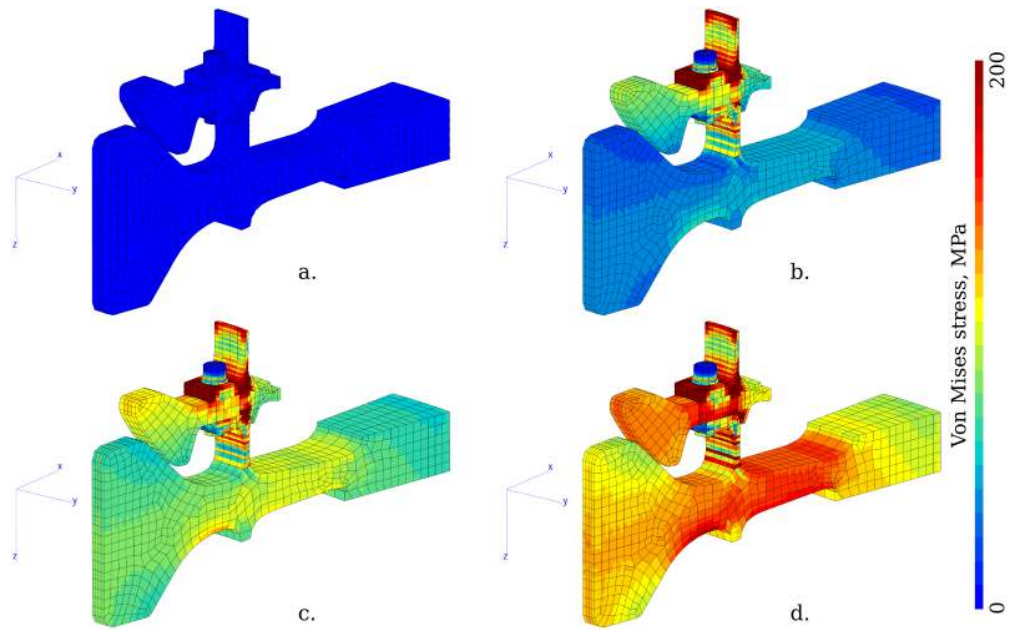


Figure 7.3: Von Mises stress distribution for different rotational frequency in a representative sector of the full mesh.



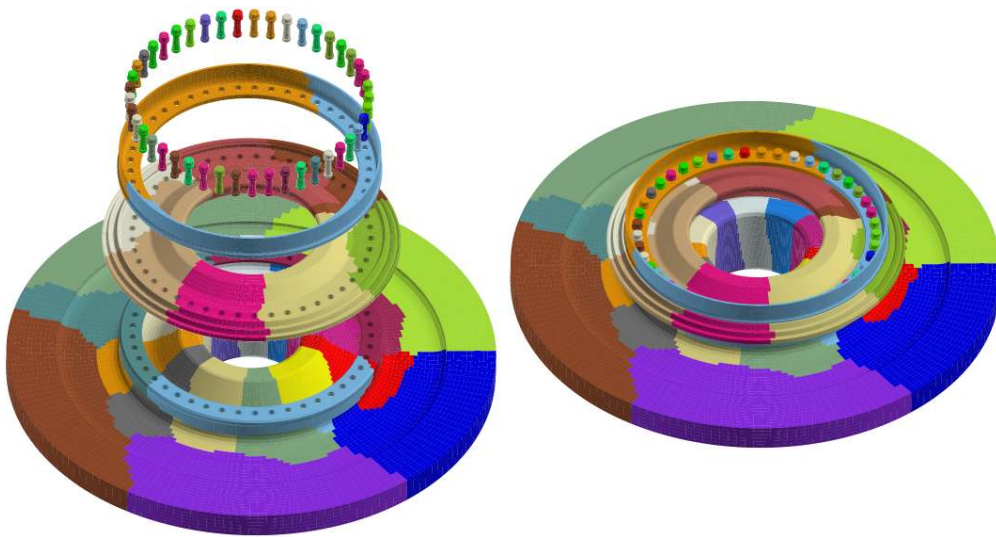


Figure 7.4: Domain decomposition of the disk assembly.



## Appendix A

# Vectors, tensors and s-structures

Here the vector-tensor formalism employed through the dissertation is briefly presented. All computations have been carried out using a rigorous vector-tensor apparatus and the associated notations – the direct “tensor language”. In my honest opinion the used formalism - the direct (component-free) tensor language – is an elegant and intuitive tool which can be easily employed in mechanics; moreover it allows to decrease significantly the probability of errors and/or misprints in comparison to index notations. In spite of all forementioned, sometimes a rigorous proof can be more easily obtained in the coordinate form of tensor with respect to a given basis. So here almost all operations will be duplicated in index notations.

Generally in literature bold symbols or explicit indices are used both for vectors and tensors; sometimes the “array” and “matrix” notations are used. The notations used here are intentionally different and may look unusual. Vectors are underlined by one line, tensors of 2nd and higher orders by two lines. It is due to the fact that we will distinguish two different vector spaces and their elements. This task is hard to accomplish using standard notations.

The first systematic exposition of the tensor language was given by Josiah Willard Gibbs [Gibbs 84], improved and extended in [Gibbs 60]. We follow the course of lectures given by Pavel A. Zhilin at the Saint-Petersburg State Polytechnical University [Zhilin 01]. The originality of our description of the tensor language consists in the introduction of new constructions – s-structures, which results in fruitful extension of the tensor language and its application in mathematics and physics.

### A.1 Fundamentals

- A vector  $\underline{a}$  of dimension  $n$  is an element of the vectorial oriented space

$$\underline{a} \in \mathbb{T}_1^n$$

and should be associated with an oriented segment in this space, but not with a set of  $n$  real numbers, which depend upon the choice of the basis. This is the main difference between the direct (component-free) tensor language and the component form.

- A special construction

$$\mathbb{T}_2^n = \mathbb{T}_1^n \otimes \mathbb{T}_1^n$$

– the tensor product of two vector spaces – will be called the second-order tensor space and its elements second-order tensors

$$\underline{\underline{a}} \in \mathbb{T}_2^n$$

The simplest element of this tensor space is a diad [Gibbs 84] – an ordered pair of vectors

$$\underline{a} \otimes \underline{b}.$$

Every diad is a second-order tensor but not every second-order tensor is a diad. Any second order tensor is a formal sum of a finite number of diads

$$\underline{\underline{a}} = \underline{a} \otimes \underline{b} + \underline{c} \otimes \underline{d} + \cdots + \underline{m} \otimes \underline{n}$$

- In the same manner, higher order tensor spaces can be introduced

$$\mathbb{T}_m^n = \underbrace{\mathbb{T}_1^n \otimes \mathbb{T}_1^n \otimes \cdots \otimes \mathbb{T}_1^n}_{m \text{ times}}$$

For example, the third-order tensor is a formal finite sum of triads

$${}^3\underline{\underline{a}} = \underline{a} \otimes \underline{b} \otimes \underline{c} + \underline{d} \otimes \underline{e} \otimes \underline{f} + \cdots + \underline{m} \otimes \underline{n} \otimes \underline{o}.$$

In the following, the notation of tensor spaces without upper index  $\mathbb{T}_m$  will define vectors and tensors in the three dimensional space.

- To move to component notations the oriented vector space  $\mathbb{T}_1^n$  has to be complemented by a basis  $B$ ; every vector can be associated with a unique set of coordinates in the basis  $\mathcal{B}$  and vice versa, i.e. spaces  $\mathbb{T}_1^n$  and  $\mathbb{R}^n$  are bijective. By basis we mean a material point in space, with associated linearly independent axes, and a clock<sup>1</sup> – *reference frame* – and established coordinate system.
- Scalars can be considered as zero-order tensors; a scalar is determined by one real number which does not depend on the choice of the coordinate system. So coordinates of a vector cannot be considered a scalar. Scalars may depend on the reference frame (kinetic energy) or not (temperature, internal energy, etc.)
- For the sake of generality here and below the expressions in the direct tensor language are followed by the corresponding expressions in component form in dark gray. Each vector is associated with  $n$  real numbers – its components in the given coordinate system in the reference frame. In this paragraph, for simplicity, the coordinate system is supposed to be orthonormal. The basis unit vectors of the system are

---

<sup>1</sup>clock is not necessary for the tensor formalism, but for physics for which this formalism is used, the clock is necessary.

$\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$  such as  $\underline{e}_i \cdot \underline{e}_j = \delta_i^j$ , where “ $\cdot$ ” denotes the dot product (defined below), and  $\delta_i^j$  is Kronecker delta

$$\delta_i^j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

For such a choice any vector can be splitted as follows

$$\underline{a} = a^1 \underline{e}_1 + a^2 \underline{e}_2 + \dots + a^n \underline{e}_n = a^i \underline{e}_i,$$

where  $a^i = \underline{a} \cdot \underline{e}_i$  are the components of the vector. So each vector for a fixed coordinate system has a one-to-one correspondence with  $n$  real numbers

$$\underline{a} \leftrightarrow a^i, \quad i = 1, n$$

The Einstein summation from 1 to  $n$  by identical upper-lower indices is used, e.g.  $a^i b_i = \sum_{i=1}^n a^i b_i$ . In details, the basis and all related questions will be considered in the following section.

- Each vector is characterized by a direction and by a length – the vector’s norm

$$\|\cdot\| : \mathbb{T}_1^n \rightarrow \mathbb{R}_0^+ : \|\underline{a}\| \geq 0.$$

Let  $\|\cdot\|$  be the Euclidean norm, so in component form

$$\|\underline{a}\| = \sqrt{\underline{a}^2} = \sqrt{\underline{a} \cdot \underline{a}}; \quad \|\underline{a}\| = \sqrt{a_i a_i}$$

- Two types of vectors are distinguished: straight vectors and spin-vectors (denoted with  $*$ , e.g.  $\underline{a}_*$ ). The first ones intuitively describe forward motions (translation), while the second ones characterize proper rotations. There is a correspondence between straight vectors and spin-vectors: this correspondence is uniquely defined if the reference frame is oriented (left-hand or right-hand oriented). A straight vector is called *polar* if it does not change its direction when the orientation of the reference frame changes. An *axial vector* is a straight vector which changes its direction to the opposite if the orientation is changed. Tensors of any order can be polar or axial. The type of tensor is determined by the sensitivity to orientation change: axial tensors are multiplied by  $-1$ , polar remain the same, if the orientation changes (independently from left-hand oriented to right-hand or vice versa). Axial tensors of zero order are called pseudoscalar.

- **Vectors summation.**

Contrary to mathematics, in physics, vectors summation is not abstract and is determined by the well known triangle or parallelogram rules. The result of the summation is a vector.

- Commutativity:  $\underline{a} + \underline{b} = \underline{b} + \underline{a}$ ;  $a_i + b_i = b_i + a_i$
- Associativity:  $(\underline{a} + \underline{b}) + \underline{c} = \underline{a} + (\underline{b} + \underline{c})$ ;  $(a_i + b_i) + c_i = b_i + (a_i + c_i)$
- Zero vector 0:  $\underline{a} + 0 = 0 + \underline{a} = \underline{a}$ ;  $a_i + 0 = 0 + a_i = a_i$

- **Product of a vector and a scalar**

The product of a vector  $\underline{a}$  and a scalar  $\alpha \in \mathbb{R}$  is a vector  $\underline{b}$

$$\underline{b} = \alpha \underline{a}; \quad b_i = \alpha a_i,$$

such as  $\|\underline{b}\| = |\alpha| \|\underline{a}\|$  and if  $\alpha > 0$ , then vectors  $\underline{a}$  and  $\underline{b}$  are similarly directed. If  $\alpha < 0$  – oppositely directed.

$$\alpha(\underline{a} + \underline{b}) = \alpha \underline{a} + \alpha \underline{b}; \quad \alpha(a_i + b_i) = \alpha a_i + \alpha b_i$$

$$(\alpha + \beta)\underline{a} = \alpha \underline{a} + \beta \underline{a}; \quad (\alpha + \beta)a_i = \alpha a_i + \beta a_i$$

- **Scalar product or dot product of vectors**

The dot product of two vectors is a scalar

$$\{\cdot\} : \mathbb{T}_1 \times \mathbb{T}_1 \rightarrow \mathbb{R}$$

$$\alpha = \underline{a} \cdot \underline{b} = \|\underline{a}\| \|\underline{b}\| \cos(\phi); \quad \alpha = a_i b^i = a^i b_i,$$

where  $\phi$  is the angle between vectors.

- Commutativity:  $\underline{a} \cdot \underline{b} = \underline{b} \cdot \underline{a}$ ;  $a_i b^i = b^i a_i$
- Distributivity:  $\underline{a} \cdot (\underline{b} + \underline{c}) = \underline{a} \cdot \underline{b} + \underline{a} \cdot \underline{c}$ ;  $a_i(b^i + c^i) = a_i b^i + a_i c^i$

- **Orthogonality**

Vectors are orthogonal if their dot product is zero

$$\underline{a} \perp \underline{b} \Leftrightarrow \underline{a} \cdot \underline{b} = 0; \quad a_i b^i = 0$$

- **Unit vector**

The unit vector of a nonzero vector  $\underline{a}$  is the vector  $\underline{\hat{a}}$  such that

$$\underline{\hat{a}} = \frac{\underline{a}}{\|\underline{a}\|}; \quad \hat{a}_i = \frac{a_i}{\|\underline{a}\|}$$

- **Projections**

Projection of the vector  $\underline{a}$  on the vector  $\underline{b}$  is a vector  $\underline{p}$  such as

$$\underline{p} = (\underline{a} \cdot \underline{\hat{b}}) \underline{\hat{b}}; \quad p_i = a_j \hat{b}^j \hat{b}_i.$$

By projection we often mean a scalar  $p$

$$p = \underline{a} \cdot \underline{\hat{b}}; \quad p = a_i \hat{b}^i$$

The projection of the vector  $\underline{a}$  on the plane with normal  $\underline{b}$  is a vector  $\underline{p}$  such that

$$\underline{p} = (\underline{I} - \underline{\hat{b}} \otimes \underline{\hat{b}}) \cdot \underline{a}; \quad p_i = (\delta_i^j - \hat{b}_i \hat{b}^j) a_j.$$

- **Vector product or cross product**

The vector product can be introduced only in oriented reference frames, contrary to previously defined operations which are valid also for non-orientation reference frames. Moreover the vector product is meaningful only in three dimensional spaces  $\mathbb{T}_1 = \mathbb{T}_1^3$ . So if there is a vector product in equation, spaces are implicitly supposed to be  $\mathbb{T}_n^3$ . The vector product of two vectors  $\underline{a}$  and  $\underline{b}$  is a vector

$$\underline{c} = \underline{a} \times \underline{b}; \quad c_i = \det \begin{bmatrix} \delta_1^i & \delta_2^i & \delta_3^i \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

such that

$$\begin{aligned} \underline{c} \cdot \underline{a} &= 0; \quad \underline{c} \cdot \underline{b} = 0; \quad \|\underline{c}\| = \|\underline{a}\| \|\underline{b}\| \sin(\phi) \\ c_i a^i &= 0; \quad c_i b^i = 0; \quad \|\underline{c}\| = \|\underline{a}\| \|\underline{b}\| \sin(\phi) \end{aligned}$$

Two vectors fulfill these conditions:  $\underline{c}$  and  $-\underline{c}$ . To determine the direction of the vector  $\underline{c}$  we employ a spin-vector  $\underline{c}_*$ , whose axis is parallel to the vector  $\underline{c}$  and whose rotation is oriented from vector  $\underline{a}$  to  $\underline{b}$  through the minimal angle. If the minimal angle is 0 or  $2\pi$  then by definition  $\|\underline{c}\| = 0$  so the orientation is meaningless. Finally the vector  $\underline{c}$  or  $-\underline{c}$  is associated with the spin-vector  $\underline{c}_*$  respectively to the orientation of the reference frame. For two polar vectors  $\underline{a}$  and  $\underline{b}$ , vector  $\underline{c}$  is axial. As can be demonstrated [Zhilin 01] vector product of two polar vectors  $\underline{a}$ ,  $\underline{b}$  is more meaningful for the spin-vector  $\underline{c}_*$  than for the axial-vector  $\underline{c}$ , since the spin-vector remains valid for mirror symmetries.

$$\underline{c} = \underline{a} \times \underline{b} = -\underline{b} \times \underline{a}; \quad c^i = \det \begin{bmatrix} \delta_1^i & \delta_2^i & \delta_3^i \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = -\det \begin{bmatrix} \delta_1^i & \delta_2^i & \delta_3^i \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{bmatrix}$$

$$\underline{a} \times (\underline{b} + \underline{c}) = \underline{a} \times \underline{b} + \underline{a} \times \underline{c};$$

$$\det \begin{bmatrix} \delta_1^i & \delta_2^i & \delta_3^i \\ a_1 & a_2 & a_3 \\ b_1 + c_1 & b_2 + c_2 & b_3 + c_3 \end{bmatrix} = \det \begin{bmatrix} \delta_1^i & \delta_2^i & \delta_3^i \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} + \det \begin{bmatrix} \delta_1^i & \delta_2^i & \delta_3^i \\ a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

- **Three vector products**

The mixed product of vectors

$$\beta = \underline{a} \cdot (\underline{b} \times \underline{c}) = \underline{c} \cdot (\underline{a} \times \underline{b}) = \underline{b} \cdot (\underline{c} \times \underline{a})$$

$$\beta = \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} = \dots = \dots$$

The double vector product of vectors

$$\underline{d} = \underline{a} \times (\underline{b} \times \underline{c}) = \underline{b}(\underline{a} \cdot \underline{c}) - \underline{c}(\underline{a} \cdot \underline{b})$$

$$d_i = \det \begin{bmatrix} \delta_{1i} & \delta_{2i} & \delta_{3i} \\ a_1 & a_2 & a_3 \\ \det \begin{bmatrix} b_2 & b_3 \\ c_2 & c_3 \end{bmatrix} & -\det \begin{bmatrix} b_1 & b_3 \\ c_1 & c_3 \end{bmatrix} & \det \begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \end{bmatrix} = b_i a_j c_j - c_i a_j b_j.$$

Note that brackets which points the order of operations, i.e. which vector product should be evaluated first are mandatory since

$$\underline{a} \times (\underline{b} \times \underline{c}) \neq (\underline{a} \times \underline{b}) \times \underline{c}$$

## A.2 Vector space basis

A basis is any set of linearly independent vectors  $\underline{e}_i$ ,  $i = 1, n$ . Vectors  $\underline{e}_i$  are linearly independent if and only if

$$\underline{e}_i \alpha^i = 0 \Leftrightarrow \alpha^i = 0$$

It can be shown that in 3D this condition is equivalent to the condition

$$\underline{e}_1 \cdot (\underline{e}_2 \times \underline{e}_3) \neq 0$$

Any vector  $\underline{a} \in \mathbb{T}_1^n$  can be presented as

$$\underline{a} = \alpha^i \underline{e}_i,$$

where  $\alpha^i$  are the coordinates of the vector  $\underline{a}$  in the basis  $\underline{e}_i$ . For two arbitrary bases

$$\underline{a} \cdot \underline{b} = \alpha^i \underline{e}_i \cdot \beta^j \underline{e}_j = \alpha^i \beta^j \underline{e}_i \cdot \underline{e}_j.$$

The dual basis  $\underline{e}^i$  is constructed such that

$$\underline{e}_i \cdot \underline{e}^j = \delta_i^j$$

Naturally vectors of the dual basis are linearly independent. If vectors of the basis  $\underline{e}_i$  are normalized, i.e.  $\|\underline{e}_i\| = 1$ , then the norms of the vectors of the dual basis are not smaller than one

$$\|\underline{e}^i\| \geq 0, \text{ since } \underline{e}_i \cdot \underline{e}^i = \underbrace{\|\underline{e}_i\|}_{=1} \|\underline{e}^i\| \cos(\phi) = 1 \Rightarrow \|\underline{e}^i\| = \frac{1}{\cos(\phi)},$$

where  $\phi$  is the angle between vectors  $\underline{e}_i$  and  $\underline{e}^i$ .

Now the summation by upper and lower repeated indices becomes more clear. Coordinates  $\alpha^i$  of the vector  $\underline{a}$  in the basis  $\underline{e}_i$  are determined by the dot product of  $\underline{a}$  with the corresponding coordinate of the dual basis and vice versa

$$\alpha^i = \underline{a} \cdot \underline{e}^i; \quad a_i = \underline{a} \cdot \underline{e}_i,$$



where  $a_i$  are the coordinates of the vector  $\underline{a}$  in the basis  $\underline{e}^i$ . If the basis  $\underline{e}_i$  is orthonormal, then the dual basis is identical to it

$$\underline{e}_i \cdot \underline{e}_j = \delta_i^j \Leftrightarrow \underline{e}_i = \underline{e}^i.$$

and the coordinates of vectors in both bases are equal  $a^i = a_i$ . Objects with upper and lower index will be called respectively contravariant and covariant objects. A more detailed explanation and definition is given below.

To read the end of this section, it is recommended to check briefly the S-structure notations introduced in Section A.6. S-structures are special constructions under the space of tensors of all ranks, which have been introduced in the dissertation to simplify the formalism and to avoid the indices. In the following, where it will be possible, instead of index notations, s-structure notations will be employed

$$\begin{aligned} \underline{e}_i &\longrightarrow \underline{\underline{e}}, & \underline{e}^i &\longrightarrow \underline{\underline{\bar{e}}} \\ a_i &\longrightarrow \underline{\underline{a}}, & a^i &\longrightarrow \underline{\underline{\bar{a}}} \end{aligned}$$

Using this notations we introduce a special t-scalar  $\underline{\underline{A}}$  called the first fundamental matrix or metric matrix, which is symmetric due to commutativity of the scalar product

$$\underline{\underline{A}} = \underline{\underline{e}} \cdot \underline{\underline{e}} = \underline{\underline{A}}^T; \quad A_{ij} = \underline{e}_i \cdot \underline{e}_j = A_{ji}$$

The term “matrix” is used to follow the tradition. It is important to remark that this object formally is not an object of tensor space  $\underline{\underline{A}} \in \mathbb{T}_2^n$ , since elements of the tensor space are invariant objects and this object is double covariant. However in coordinate form the metric matrix seems very similar to tensor and is called the metric tensor. From the previous definition it follows directly that

$$\underline{\underline{e}} = \underline{\underline{A}} \underline{\underline{\bar{e}}}; \quad \underline{e}_i = (\underline{e}_i \cdot \underline{e}_j) \underline{e}^j$$

In the same manner the t-scalar  $\underline{\underline{\bar{A}}}$  – metric matrix for the dual basis is determined; it is also symmetric

$$\underline{\underline{\bar{A}}} = \underline{\underline{\bar{e}}} \cdot \underline{\underline{\bar{e}}} = \underline{\underline{\bar{A}}}^T; \quad A^{ij} = \underline{e}^i \cdot \underline{e}^j = A^{ji}$$

$$\underline{\underline{\bar{e}}} = \underline{\underline{\bar{A}}} \underline{\underline{e}}; \quad \underline{e}^i = (\underline{e}^i \cdot \underline{e}^j) \underline{e}_j$$

Considering the dot product of two bases.

$$\underline{\underline{I}} = \underline{\underline{\bar{e}}} \cdot \underline{\underline{e}} = \underline{\underline{\bar{A}}} \underline{\underline{e}} \cdot \underline{\underline{e}} = \underline{\underline{\bar{A}}} \underline{\underline{I}} \underline{\underline{A}}; \quad \delta_j^i = \underline{e}^i \cdot \underline{e}_j = A^{ij} \underline{e}_j \cdot \underline{e}^k A_{lk} = A^{ij} \delta_j^k A_{lk}$$

Directly due to the symmetry of fundamental t-scalars

$$\underline{\underline{\bar{A}}} \underline{\underline{A}} = \underline{\underline{I}}, \quad \underline{\underline{\bar{A}}} = \underline{\underline{A}}^{-1}; \quad A^{ij} A_{jk} = \delta_k^i, \quad A^{ij} = \frac{\text{cofactor}(A_{ij})}{\det A_{ij}},$$

## A.2 Vector space basis

---

where by cofactor  $(A_{ij})$  we mean the determinant of matrix  $\tilde{A}_{ij}$  with suppressed  $i$ -th line and  $j$ -th column and multiplied by  $(-1)^{i+j}$ . It follows directly

$$\det \tilde{\tilde{A}} = \frac{1}{\det A}$$

Following the derived formulae, any vector  $\underline{a}$  can be easily written in both bases

$$\underline{a} = \tilde{\tilde{a}} \tilde{\tilde{e}} = \tilde{\tilde{a}} \tilde{\tilde{e}}; \quad \underline{a} = a^i \underline{e}_i = a_i \underline{e}^i$$

where  $\tilde{\tilde{a}}$  and  $\tilde{\tilde{a}}$  are the v-scalars of coordinates of the vector  $\underline{a}$  in the basis  $\tilde{\tilde{e}}$  and  $\underline{e}$  respectively

$$\tilde{\tilde{a}} = \tilde{\tilde{A}} \tilde{\tilde{a}}, \quad \tilde{\tilde{a}} = \tilde{\tilde{A}} \tilde{\tilde{a}}; \quad a_i = A_{ij} a^j, \quad a^i = A^{ij} a_j$$

### A.2.1 Transformation matrices, covariant and contravariant objects

The choice of the basis is arbitrary. Vectors does not depend on this choice, so for any basis the following equality holds:

$$\underline{a} = \tilde{\tilde{a}} \tilde{\tilde{e}} = \tilde{\tilde{a}} \tilde{\tilde{e}} = \tilde{\tilde{a}} \tilde{\tilde{e}} = \tilde{\tilde{a}} \tilde{\tilde{e}}$$

where  $\tilde{\tilde{e}}$  and  $\tilde{\tilde{e}}$  are the new basis and its dual basis respectively. As done for the dual basis, the transformation t-scalar can be constructed to determine coordinates of vectors in the new basis from known coordinates in the initial basis and vice versa. Vectors of the new basis are expressed by the vectors of the initial basis as follows

$$\tilde{\tilde{e}} = (\tilde{\tilde{e}} \cdot \tilde{\tilde{e}}) \tilde{\tilde{e}}; \quad \underline{e}_{i'} = (\underline{e}_{i'} \cdot \underline{e}^j) \underline{e}_j$$

the transition t-scalar  $\tilde{\tilde{P}}$ , traditionally referred as the pushforward transformation matrix, is then constructed as follows

$$\tilde{\tilde{P}} = \tilde{\tilde{e}} \cdot \tilde{\tilde{e}}; \quad P_{i'}^j = \underline{e}_{i'} \cdot \underline{e}^j$$

For backward transition, the t-scalar  $\tilde{\tilde{P}}'$  – pullback transformation matrix – is given by

$$\tilde{\tilde{P}}' = \tilde{\tilde{e}}' \cdot \underline{e}; \quad P_{j'}^{i'} = \underline{e}^{i'} \cdot \underline{e}_{j'}$$

Note that these transformation matrices in general are not symmetric, so indices are ordered

$$\tilde{\tilde{P}} \neq \tilde{\tilde{P}}^T; \quad \tilde{\tilde{P}}' \neq \tilde{\tilde{P}}'^T$$

It is easy to show that the t-scalars  $\tilde{\tilde{P}}^T$  and  $\tilde{\tilde{P}}'$  are inversely proportional:

$$\tilde{\tilde{P}}^T \tilde{\tilde{P}}' = \tilde{\tilde{I}}, \quad \tilde{\tilde{P}}' = \tilde{\tilde{P}}^{-T}; \quad P_{i'}^j P_{j'}^{i'} = \delta_{i'}^{j'}, \quad P_{i'}^j = \frac{\text{cofactor}(P_{j'}^{i'})}{\det P_{j'}^{i'}}$$

We clearly see that the initial basis and the new basis are related by the forward transition  $\underset{\sim}{P}$ ; The corresponding dual bases are related by the backward transition  $\underset{\sim}{P}'$ :

$$\begin{aligned}\underset{\sim}{e}' &= \underset{\sim}{P} \underset{\sim}{e} \\ \underset{\sim}{e}' &= \underset{\sim}{P}' \underset{\sim}{e}\end{aligned}$$

One of the key notions of the theory is the definition of the covariant and contravariant objects.

If an object  $\underset{\sim}{c}$  of order  $n$  depends on the basis choice, then this object is called

- *n-order covariant* if it follows the same transformation as the initial basis:

$$\underset{\sim}{c}' = \overbrace{\underset{\sim}{P} \underset{\sim}{P} \dots \underset{\sim}{P}}^{n \text{ times}} \underset{\sim}{c}; \quad c_{\alpha' \beta' \dots \gamma'} = \overbrace{P_{\alpha'}^i P_{\beta'}^j \dots P_{\gamma'}^n}^{n \text{ times}} c_{ij \dots n}$$

- *n-order contravariant* if it follows the same transformation as the dual basis:

$$\underset{\sim}{c}' = \overbrace{\underset{\sim}{P}' \underset{\sim}{P}' \dots \underset{\sim}{P}'}^{n \text{ times}} \underset{\sim}{c}; \quad c^{\alpha' \beta' \dots \gamma'} = \overbrace{P_i^{\alpha'} P_j^{\beta'} \dots P_n^{\gamma'}}^{n \text{ times}} c^{ij \dots n}$$

- *p-order covariant, q-order contravariant,  $p + q = n$*  if it changes  $p$  times according to the transformation of the initial basis and  $q$  times according to the transformation of the dual basis:

$$\underset{\sim}{c}' = \overbrace{\underset{\sim}{P} \underset{\sim}{P} \dots \underset{\sim}{P}}^{p \text{ times}} \overbrace{\underset{\sim}{P}' \underset{\sim}{P}' \dots \underset{\sim}{P}'}^{q \text{ times}} \underset{\sim}{c}; \quad c^{\varepsilon' \xi' \dots \zeta'}_{\alpha' \beta' \dots \gamma'} = \overbrace{P_{\alpha'}^i P_{\beta'}^j \dots P_{\gamma'}^k}^{p \text{ times}} \overbrace{P_l^{\varepsilon'} P_m^{\xi'} \dots P_n^{\zeta'}}^{q \text{ times}} c_{ij \dots k}^{lm \dots n}$$

An object of the first order  $c$  which does not follow covariant or contravariant transformation rules are called non tensorial object and the position of the index is not important. In such cases, the index is often put in brackets  $c_{(i)}$ . If an object is  $m$ -order covariant and  $m$ -order contravariant, then the object is invariant with respect to the change of basis. This is the case for a vector, which consists of covariant coordinates and contravariant basis vector or contravariant coordinates and covariant basis vectors. Since tensors of higher ranks are constructed by means of tensor products of vectors, a tensor of rank  $m$  is  $m$ -order covariant and  $m$ -order contravariant, so it is also invariant in any basis change. The Kronecker delta  $\delta_i^j = \underset{\sim}{e}_i \cdot \underset{\sim}{e}^j$  is a first-order covariant and first-order contravariant object, so is also invariant, and can be considered as the coordinates of the second order unity tensor. Another example is the metric matrix  $\underset{\sim}{A}$  which is second order covariant; the metric matrix of the dual basis  $\underset{\sim}{\bar{A}}$  is second order contravariant and naturally their product  $\underset{\sim}{A} \underset{\sim}{\bar{A}} = \underset{\sim}{I}$  is an invariant object.

### A.2.2 Gradient operator or Hamilton's operator

Let  $\underline{\underline{e}}$  be a moving coordinate system in the reference frame, then the vector  $\underline{r}$  of each point, fixed in the frame of reference, is determined as

$$\underline{r} = \underline{r}(\underline{\underline{y}}, t),$$

where  $\underline{\underline{y}} = \underline{\underline{y}}(t)$  are the coordinates of  $\underline{r}$  in the basis  $\underline{\underline{e}}$  at time  $t$ , i.e.

$$\underline{r} = \underline{\underline{y}} \underline{\underline{e}}$$

and

$$\dot{\underline{r}} = \frac{\partial \underline{r}}{\partial \underline{\underline{y}}} \dot{\underline{\underline{y}}} + \frac{\partial \underline{r}}{\partial t} = 0.$$

For an observer in the moving coordinate system  $\underline{\underline{e}}$ , a change of point  $\underline{r}$  leads to a change of coordinates  $\underline{\underline{y}}$ . Logically there exists changes of  $\underline{r}$  in which only one coordinate changes, for example  $y^i$ . So for a smooth function  $\underline{r}(\underline{\underline{y}})$  at a given time, there exists one curve in space determined by the vector  $\underline{r}(y^1, \dots, y^i, \dots, y^n)$  where all  $y^j = \text{const}$  and only  $y^i$  changes. The tangent vector for such a curve is determined by

$$\underline{r}_i = \frac{\partial \underline{r}}{\partial y^i}$$

A set of such tangent vectors is a covariant v-vector

$$\underline{\underline{D}} \underline{r} = \frac{\partial}{\partial \underline{\underline{y}}} \underline{r},$$

where  $\underline{\underline{D}}$  is a covariant operator

$$\underline{\underline{D}} = \frac{\partial}{\partial \underline{\underline{y}}}$$

The dual local basis can be constructed in a standard manner

$$\overline{\frac{\partial \underline{r}}{\partial \underline{\underline{y}}}} = \underline{\underline{A}} \frac{\partial \underline{r}}{\partial \underline{\underline{y}}},$$

where  $\underline{\underline{A}} = \underline{\underline{A}}^{-1}$  and

$$\underline{\underline{A}} = \frac{\partial \underline{r}}{\partial \underline{\underline{y}}} \cdot \frac{\partial \underline{r}}{\partial \underline{\underline{y}}}$$

Covariant or contravariant sets of such tangent vectors can be chosen as a basis. Since the basis depends on the points in the reference frame, such a basis is called local. It can be shown that the basis constructed in such a way is

covariant. It is worth mentioning, that in [Kagan 47], these sets of vectors were interpreted as first rank tensors with vector components – an interpretation of v-vectors.

Finally the well known gradient operator (invariant differential operator) is written as

$$\nabla = \overline{\frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}}} \circ \mathbb{D}$$

The gradient vector is an objective operator, which does not depend on the choice of the coordinate system, since it is first order covariant due to covariant operator  $\mathbb{D}$  and first order contravariant due to  $\overline{\frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}}}$ . Another example of an objective operator is the full time derivative  $\frac{d}{dt}$ . An example of a non-objective operator is the partial time derivative  $\frac{\partial}{\partial t}$ . The two following statements can be easily proven

$$\nabla \frac{d}{dt} = \frac{d}{dt} \nabla, \quad \nabla \frac{\partial}{\partial t} \neq \frac{\partial}{\partial t} \nabla$$

The gradient operator is very important for mechanical and physical theories. For a scalar field  $s(\mathbf{r}) = s(\mathbf{r}(\tilde{\mathbf{y}}))$ , the influence of the gradient operator is uniquely defined by

$$\text{grad}(s) = \nabla s = \overline{\frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}}} \circ \frac{\partial s}{\partial \tilde{\mathbf{y}}}$$

The gradient notation “grad” is equivalent to the tensor product of the gradient operator and the field; in the case of a scalar field, the tensor product is meaningless and so omitted. The gradient vector can act differently on tensor fields of first and higher orders  $\underline{\underline{T}}(\mathbf{r}) = \underline{\underline{T}}(\mathbf{r}(\tilde{\mathbf{y}}))$ . The side (left-hand or right-hand) and intermediate operators are important. Gradient of a tensor field:

$$\text{grad}(\underline{\underline{T}}) = \nabla \otimes \underline{\underline{T}} = \overline{\frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}}} \circ \otimes \frac{\partial \underline{\underline{T}}}{\partial \tilde{\mathbf{y}}} \neq \underline{\underline{T}} \otimes \nabla$$

The tensor product notation is often skipped. Divergence of a tensor field:

$$\text{div}(\underline{\underline{T}}) = \nabla \cdot \underline{\underline{T}} = \overline{\frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}}} \circ \cdot \frac{\partial \underline{\underline{T}}}{\partial \tilde{\mathbf{y}}}$$

Remark that the following equality is correct only for vectors and symmetric second-order tensors:

$$\nabla \cdot \underline{\underline{T}} = \underline{\underline{T}} \cdot \nabla$$

The rotor of a tensor field:

$$\text{rot}(\underline{\underline{T}}) = \nabla \times \underline{\underline{T}} = \overline{\frac{\partial \mathbf{r}}{\partial \tilde{\mathbf{y}}}} \circ \times \frac{\partial \underline{\underline{T}}}{\partial \tilde{\mathbf{y}}}$$

### A.3 Sub-basis, vector function of v-scalar argument

---

$$\nabla \times \underline{t} = -\underline{t} \times \nabla$$

Div and rot are alternative notations for the scalar and vector products of the gradient operator with a tensor field.

Another important operator in physics and mechanics is the Laplace operator or Laplacian  $\Delta$ , which is defined as

$$\Delta = \nabla \cdot \nabla = \nabla^2$$

It should not be confused with  $\nabla \otimes \nabla \neq \Delta$ .

Some examples:

- Balance of momentum in the deformable body

$$\nabla \cdot \underline{\underline{\sigma}} + \underline{f}_v = 0$$

- Cauchy's strain tensor

$$\underline{\underline{\varepsilon}} = \frac{1}{2} (\nabla \otimes \underline{u} + \underline{u} \otimes \nabla)$$

- Compatibility of (small) deformations

$$\nabla \times (\nabla \times \underline{\underline{\varepsilon}}) = 0$$

- Compatibility of (finite) deformations

$$\nabla \times \underline{F} = 0$$

- Maxwell's equations

$$\nabla \cdot \underline{E} = \frac{\rho}{\varepsilon_0}$$

$$\nabla \cdot \underline{B} = 0$$

$$\nabla \times \underline{E} = -\frac{\partial \underline{B}}{\partial t}$$

$$\nabla \times \underline{B} = \mu_0 \underline{J} + \mu_0 \varepsilon_0 \frac{\partial \underline{E}}{\partial t}$$

### A.3 Sub-basis, vector function of v-scalar argument

Let us consider a vector function  $\underline{\rho}$  of v-scalar argument  $\underline{\xi}$  such as

$$\underline{\rho}(\underline{\xi}) : \mathbb{R}^m \rightarrow \mathbb{T}_1^n, \quad m \leq n$$

If for each  $\underline{\xi}$  there exists only one element  $\underline{\rho}(\underline{\xi}) \in \mathbb{T}_1^n$ , where  $\mathbb{T}_1^n$  is a vector space with a reference frame, then the image of the space  $\mathbb{R}^m$  in  $\mathbb{T}_1^n$  is a set of points. If  $\underline{\rho}(\underline{\xi}) = \underline{\rho}(t, \underline{\xi})$ , then this set of points changes in time. For instance, in the 3D space at given time, the image  $\mathbf{Im} \mathbb{R}^m$  in the space is

- a point, if  $m = 0$ ;
- a curve, if  $m = 1$ ;
- a surface, if  $m = 2$ ;
- a volume if  $m = 3$ .

At each time at point  $\underline{\rho}(t, \underline{\xi})$  a local sub-basis  $\underline{e}$  can be established if the function  $\underline{\rho}(t, \underline{\xi}) \in C^1(\underline{\xi})$  is sufficiently smooth versus its second argument  $\underline{\xi}$

$$\underline{e} = \frac{\partial \underline{\rho}}{\partial \underline{\xi}}$$

For the given time  $t^0$  and point  $\underline{\xi}^0$  this basis determines a local subspace  $\mathbb{T}_1^m$  of the full space  $\mathbb{T}_1^n \subset \mathbb{T}_1^n$ :

- a straight line, tangential to the curve at  $\underline{\rho}(t^0, \underline{\xi}^0)$ , if  $m = 1$ ;
- a tangential plane of the surface at  $\underline{\rho}(t^0, \underline{\xi}^0)$  if  $m = 2$ ;
- a full space, if  $m = 3$ .

If needed, the basis  $\underline{e}$  can be complemented by vectors orthogonal to the subspace  $\mathbb{T}_1^m \ni \underline{e}_i \perp \mathbb{T}_1^m$ . Moreover in the particular case of three dimensional space  $n = 3$  and  $m = 2$ , the third basis vector can be constructed as the vector product of the basis vectors  $\underline{e}$

$$\underline{e}_3 = \underline{e}_1 \times \underline{e}_2.$$

However if we are interested only in the object described by the function  $\underline{\rho}(t, \underline{\xi})$  there is no need for such a completion.

The metric matrix associated with the object  $\underline{\rho}(t, \underline{\xi})$  is nothing but:

$$\underline{A} = \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = \underline{A}^T,$$

and consequently the dual basis is:

$$\bar{\underline{e}} = \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} = \bar{\underline{A}} \frac{\partial \underline{\rho}}{\partial \underline{\xi}},$$

where  $\bar{\underline{A}} = \underline{A}^{-1} = \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} \cdot \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} = \bar{\underline{A}}^T$ .

A corresponding sub-gradient operator is constructed like in the previous paragraph, (to avoid any confusions the order of the operator is marked):

$$\nabla^m = \bar{\underline{e}} \circ \frac{\partial}{\partial \underline{\xi}} = \overline{\frac{\partial \underline{\rho}}{\partial \underline{\xi}}} \circ \frac{\partial}{\partial \underline{\xi}}$$

Such a sub-gradient operator is used in theory of deformable rods ( $m = 1$ ) and shells ( $m = 2$ ).

As one can note, the metric matrix associated with an object in a  $n$  dimensional space has a dimension  $m \times m$  instead of  $n \times n$ . It is a normal situation in differential geometry. But the classical rule of summations by indices can not be used anymore, since the classical summation by indices is defined according to the dimension of the full space, i.e. from 1 to  $n$ . One possible solution employed in the literature is to introduce a reduced summation from 1 to  $m$ , this summation will be performed only within, for example small Greek indices or capital letter indices like

$$\underline{e}^\alpha = \frac{\partial \underline{\rho}}{\partial \xi^\alpha} = A^{\alpha\beta} \frac{\partial \underline{\rho}}{\partial \xi_\beta}; \quad \underline{e}^P = \frac{\partial \underline{\rho}}{\partial \xi^P} = A^{PQ} \frac{\partial \underline{\rho}}{\partial \xi_Q}$$

Let us demonstrate the difference between an intuitive form constructed with the direct tensor language and a half-index form based on Greek letters. For quantities related to the surface, both of them are equivalent, but as in practice the probability of error is lower if one uses the first formalism (this example is taken from Chapter 2:)

$$\begin{aligned} \Delta \delta \xi \approx & (g_n \underline{\mathbb{H}} - \underline{\mathbb{A}})^{-1} \left\{ \frac{\partial \underline{\rho}}{\partial \xi} \cdot \left( \delta \frac{\partial \underline{\rho}}{\partial \xi} \Delta \xi + \Delta \frac{\partial \underline{\rho}}{\partial \xi} \delta \xi \right) + \Delta \xi \left( \frac{\partial \underline{\rho}}{\partial \xi} \cdot \frac{\partial^2 \underline{\rho}}{\partial \xi^2} \right) \delta \xi - \right. \\ & - g_n \underline{n} \cdot \left( \delta \frac{\partial^2 \underline{\rho}}{\partial \xi^2} \Delta \xi + \Delta \frac{\partial^2 \underline{\rho}}{\partial \xi^2} \delta \xi \right) - g_n \Delta \xi \left( \underline{n} \cdot \frac{\partial^3 \underline{\rho}}{\partial \xi^3} \right) \delta \xi + \\ & + \left[ g_n \left( \delta \frac{\partial \underline{\rho}}{\partial \xi} + \frac{\partial^2 \underline{\rho}}{\partial \xi^2} \delta \xi \right) \cdot \frac{\partial \underline{\rho}}{\partial \xi} \underline{\mathbb{A}} - \delta g_n \underline{\mathbb{I}} \right] \left( \underline{n} \cdot \Delta \frac{\partial \underline{\rho}}{\partial \xi} + \underline{\mathbb{H}} \Delta \xi \right) + \\ & + \left[ g_n \left( \Delta \frac{\partial \underline{\rho}}{\partial \xi} + \frac{\partial^2 \underline{\rho}}{\partial \xi^2} \Delta \xi \right) \cdot \frac{\partial \underline{\rho}}{\partial \xi} \underline{\mathbb{A}} - \Delta g_n \underline{\mathbb{I}} \right] \left( \underline{n} \cdot \delta \frac{\partial \underline{\rho}}{\partial \xi} + \underline{\mathbb{H}} \delta \xi \right) \Big\} \end{aligned} \quad (\text{A.1})$$

$$\begin{aligned} \Delta \delta \xi^\alpha = & C^{\alpha\beta} \left\{ \frac{\partial \rho_i}{\partial \xi_\beta} \left( \delta \frac{\partial \rho^i}{\partial \xi_\gamma} \Delta \xi^\gamma + \Delta \frac{\partial \rho^i}{\partial \xi_\gamma} \delta \xi^\gamma \right) + \Delta \xi^\beta \frac{\partial \rho_i}{\partial \xi_\beta} \frac{\partial^2 \rho^i}{\partial \xi_\gamma \partial \xi_\eta} \delta \xi^\eta - \right. \\ & - g_n n_i \left( \delta \frac{\partial^2 \rho^i}{\partial \xi_\beta \partial \xi_\gamma} \Delta \xi^\gamma + \Delta \frac{\partial^2 \rho^i}{\partial \xi_\beta \partial \xi_\gamma} \delta \xi^\gamma \right) - g_n \Delta \xi^\gamma \left( n_i \frac{\partial^3 \rho^i}{\partial \xi_\gamma \partial \xi_\beta \partial \xi_\kappa} \right) \delta \xi^\kappa + \\ & + \left[ g_n \left( \delta \frac{\partial \rho_i}{\partial \xi_\beta} + \frac{\partial^2 \rho_i}{\partial \xi_\beta \partial \xi_\gamma} \delta \xi^\gamma \right) \frac{\partial \rho^i}{\partial \xi_\eta} A^{\eta\kappa} - \delta g_n \delta_\beta^\kappa \right] \left( n_i \Delta \frac{\partial \rho^i}{\partial \xi_\kappa} + H_{\kappa\tau} \Delta \xi^\tau \right) + \\ & + \left[ g_n \left( \Delta \frac{\partial \rho_i}{\partial \xi_\beta} + \frac{\partial^2 \rho_i}{\partial \xi_\beta \partial \xi_\gamma} \Delta \xi^\gamma \right) \frac{\partial \rho^i}{\partial \xi_\eta} A^{\eta\kappa} - \delta g_n \Delta_\beta^\kappa \right] \left( n_i \delta \frac{\partial \rho^i}{\partial \xi_\kappa} + H_{\kappa\tau} \delta \xi^\tau \right) \Big\}, \end{aligned} \quad (\text{A.2})$$

where  $C^{\alpha\beta} = (g_n H_{\alpha\beta} - A_{\alpha\beta})^{-1}$ , here  $\alpha, \beta, \dots, \kappa \in 1, 2$  and  $i \in 1, 2, 3$ .



## A.4 Tensors

Scalars and vectors are not sufficient for physical and mechanical theories. More complicated structures, higher order tensors, are required. Second-order tensors seem to be much more abstract and hardly imaginary objects than vectors and scalars. However, a second-order tensor is not a purely mathematical construction. In mechanics it appeared first as a mathematical formalization of intuitive associations not connected with coordinate systems.

The first second-order tensors (tensor of inertia of a rigid body and rotation tensor) have been introduced by Leonhard Euler in 1758. The term “tensor” has been proposed by W. Voigt only in 1900. In 1788 Joseph-Louis Lagrange introduced the second-order tensor of small deformations. In 1822 Augustin-Louis Cauchy introduced a second-order tensor to characterize the stress state – the stress tensor  $\underline{\underline{\sigma}}$  – and gave a consistent mathematical framework for the tensor space. Tensors became unavoidable objects for the description of deformable continua. Ever since, this stress tensor is called Cauchy’s stress tensor.

- **Definition**

The second-order tensor space is defined as *tensor* product of first order tensor spaces (vector spaces).

$$\mathbb{T}_2^n = \mathbb{T}_1^n \otimes \mathbb{T}_1^n$$

Note, that the direct product of such spaces  $\mathbb{T}_1^n \times \mathbb{T}_1^n$  does not lead to a tensor space; such a space is even not a linear space. Although, tensors are “physical” in 3D space ( $n = 3$ ), for the sake of generality, the formalism will be derived for an arbitrary  $n$  dimensional space, except some particularities related, for example, to vector product. Any element of the constructed space – second-order tensor  $\underline{\underline{A}} \in \mathbb{T}_2^n$  – can be written as a *formal* sum of tensor products of vectors

$$\underline{\underline{A}} = \underline{a} \otimes \underline{b} \oplus \underline{c} \otimes \underline{d} \oplus \dots \oplus \underline{e} \otimes \underline{f}$$

Note that a tensor product of vectors  $\underline{a} \otimes \underline{b}$  is an ordered combination of two vectors called *diads*

$$\underline{a} \otimes \underline{b} \neq \underline{b} \otimes \underline{a}.$$

The maximum number of independent diads forming the tensor is equal to the dimension of the space  $n$ . For diads we require that

$$\alpha(\underline{a} \otimes \underline{b}) = (\alpha \underline{a}) \otimes \underline{b} = \underline{a} \otimes (\alpha \underline{b}) = \alpha \underline{a} \otimes \underline{b}$$

Following this axiom, a *zero diad* is determined if  $\alpha$  is put to zero

$$0 \underline{a} \otimes \underline{b} = \underline{0} \otimes \underline{b} = \underline{a} \otimes \underline{0}$$

$$(\alpha + \beta) \underline{a} \otimes \underline{b} = \alpha \underline{a} \otimes \underline{b} + \beta \underline{a} \otimes \underline{b}$$

If  $\alpha = 1, \beta = 0$  then

$$(1 + 0)\underline{a} \otimes \underline{b} = \underline{a} \otimes \underline{b} + \underline{0} \otimes \underline{b} = \underline{a} \otimes \underline{b}$$

So a zero diad is a zero element of the diad space and can be noted as 0.

$$\underline{a} \otimes (\underline{b} + \underline{c}) = \underline{a} \otimes \underline{b} + \underline{a} \otimes \underline{c}$$

$$(\underline{a} + \underline{b}) \otimes \underline{c} = \underline{a} \otimes \underline{c} + \underline{b} \otimes \underline{c}$$

The formal sums of diads  $\oplus$  will be replaced by usual symbol  $+$ , which should be understood as a sign of union like in complex numbers  $x + iy$ .

- **Composition/union of tensors**

The sum of diads is not ordered and any their combination defines the same tensor

$$\begin{aligned} \underline{\underline{A}} &= \underline{a} \otimes \underline{b} + \underline{c} \otimes \underline{d} + \cdots + \underline{e} \otimes \underline{f} = \\ &= \underline{c} \otimes \underline{d} + \underline{a} \otimes \underline{b} + \cdots + \underline{e} \otimes \underline{f} = \\ &= \cdots = \underline{c} \otimes \underline{d} + \underline{e} \otimes \underline{f} + \cdots + \underline{a} \otimes \underline{b} \end{aligned} \quad (\text{A.3})$$

The product with a scalar results in

$$\begin{aligned} \alpha \underline{\underline{A}} &= (\alpha \underline{a}) \otimes \underline{b} + (\alpha \underline{c}) \otimes \underline{d} + \cdots + (\alpha \underline{e}) \otimes \underline{f} \\ (\alpha + \beta) \underline{\underline{A}} &= \alpha \underline{\underline{A}} + \beta \underline{\underline{A}} \end{aligned}$$

Composition, sum or union of tensors is a tensor from the same space

$$\underline{\underline{A}} + \underline{\underline{B}} = \underline{\underline{C}}$$

As it can be easily shown, tensors are invariant objects, i.e. they do not depend on the choice of basis. The second order tensor space is shown to be linear.

- **Coordinates of tensors**

In coordinate form, any second-order tensor for a given basis and a dual basis can be written in four ways

$$\underline{\underline{A}} = A_{i*}^{*j} \underline{e}^i \otimes \underline{e}_j = A_{*j}^{i*} \underline{e}_i \otimes \underline{e}^j = A_{ij} \underline{e}^i \otimes \underline{e}^j = A^{ij} \underline{e}_i \otimes \underline{e}_j$$

To avoid any confusions of components' order the use of stars is mandatory for mixed coordinates (co-covariant and co-contravariant), the same for higher order tensors

$${}^3 \underline{\underline{A}} = A_{i*k}^{*j*} \underline{e}^i \otimes \underline{e}_j \otimes \underline{e}^k$$

By analogy with the s-structure formalism, in the basis  $\underline{\underline{e}}$  and its dual  $\underline{\underline{\bar{e}}}$ , a tensor can be written also in four ways:

$$\underline{\underline{B}} = \underline{\underline{\bar{\bar{B}}}} \circ \circ \underline{\underline{\bar{\bar{E}}}} = \underline{\underline{\bar{\bar{B}}}} \circ \circ \underline{\underline{\bar{\bar{E}}}} = \underline{\underline{\bar{\bar{B}}}} \circ \circ \underline{\underline{\bar{\bar{E}}}} = \underline{\underline{\bar{\bar{B}}}} \circ \circ \underline{\underline{\bar{\bar{E}}}}$$

where t-tensor  $\underline{\underline{E}}$  contains tensor product of basis vectors and t-scalar  $\underline{\underline{B}}$  contains coordinates of tensor in this basis:

$$\underline{\underline{E}} = \underline{\underline{e}} \otimes \underline{\underline{e}}; \quad \underline{\underline{\bar{E}}} = \underline{\underline{\bar{e}}} \otimes \underline{\underline{\bar{e}}}; \quad \underline{\underline{\bar{E}}} = \underline{\underline{e}} \otimes \underline{\underline{\bar{e}}}; \quad \underline{\underline{\bar{E}}}^T = \underline{\underline{\bar{e}}} \otimes \underline{\underline{e}}$$

All possible coordinates of tensors are connected by the fundamental metric matrices  $\underline{\underline{A}}$  and  $\underline{\underline{\bar{A}}}$ . To avoid any confusion, in a component form, these matrices will be put in square bracket  $[A_{ij}]$ ,  $[A^{ij}]$ :

$$B^{ij} = [A^{ik}] B_{k*}^{*j} = [A^{jk}] B_{*k}^{i*} = [A^{ik}] [A^{jl}] B_{kl}$$

The change of basis is defined as the pushforward and the pullback transformations:

$$\underline{\underline{e'}} = \underline{\underline{P}} \underline{\underline{e}}; \quad \underline{\underline{\bar{e'}}} = \underline{\underline{P'}} \underline{\underline{\bar{e}}}$$

in component form:

$$\underline{e}_{i'} = [P_{i'}^j] \underline{e}_j; \quad \underline{e}^{i'} = [P^{i'}_j] \underline{e}^j,$$

square brackets imply that components  $[P_{i'}^j]$  do not represent tensors. Using the transformation matrices the coordinates of tensors can be defined in the new basis  $\underline{\underline{e'}}$  or its dual  $\underline{\underline{\bar{e'}}}$ , if coordinates in the basis  $\underline{\underline{e}}$  or  $\underline{\underline{\bar{e}}}$  are known.

$$\underline{\underline{B}} = B^{i'j'} \underline{\underline{e}}_{i'} \otimes \underline{\underline{e}}_{j'} = B^{kl} \underline{\underline{e}}_k \otimes \underline{\underline{e}}_l = B^{kl} [P_{i'}^{i'}] [P_{j'}^{j'}] \underline{\underline{e}}_{i'} \otimes \underline{\underline{e}}_{j'}$$

So the four pushforward transformation formulae for tensor coordinates are:

$$B^{i'j'} = [P_{i'}^{i'}] B^{kl} [P_{j'}^{j'}]; \quad B_{i'j'} = [P_{i'}^k] B_{kl} [P_{j'}^l]$$

$$B_{*j'}^{i'*} = [P_{i'}^{i'}] B_{k*}^{*l} [P_{j'}^l]; \quad B_{i'*}^{j'*} = [P_{i'}^k] B_{*l}^{k*} [P_{j'}^{j'}]$$

The pullback transformations can be obtained by simple substitution of indices without prime by indices with prime and vice versa. In s-structure notations, this set of transformations can be written as follows:

$$\underline{\underline{\bar{B'}}} = \underline{\underline{P'}} \circ \underline{\underline{\bar{B}}} \circ \underline{\underline{P'}}^T; \quad \underline{\underline{B'}} = \underline{\underline{P}} \circ \underline{\underline{B}} \circ \underline{\underline{P}}^T;$$

$$\underline{\underline{\bar{B'}}} = \underline{\underline{P}} \circ \underline{\underline{\bar{B}}} \circ \underline{\underline{P'}}^T; \quad \underline{\underline{\bar{B'}}}^T = \underline{\underline{P}} \circ \underline{\underline{\bar{B}}}^T \circ \underline{\underline{P'}}^T$$

- **Transposition**

Let  $\underline{\underline{A}} = \sum_i \underline{\underline{D}}_i$ , where  $\underline{\underline{D}}_i$  are diads. The transposition for diads  $\underline{\underline{D}} = \underline{\underline{a}} \otimes \underline{\underline{b}}$  is defined as follows:

$$\underline{\underline{D}}^T = (\underline{\underline{a}} \otimes \underline{\underline{b}})^T = \underline{\underline{b}} \otimes \underline{\underline{a}}$$

Transposition of tensor:

$$\underline{\underline{A}}^T = \sum_i \underline{\underline{D}}_i^T; \quad A_{i*}^{*jT} = A_{*i}^{j*}, \quad A_{ij}^T = A_{ji}, \dots$$

A tensor is called *symmetric* if:

$$\underline{\underline{A}}^T = \underline{\underline{A}}$$

A tensor is called *antisymmetric* if:

$$\underline{\underline{A}}^T = -\underline{\underline{A}}$$

Any tensor can be splitted into a sum of symmetric and antisymmetric tensors:

$$\underline{\underline{B}} = \underline{\underline{B}}^S + \underline{\underline{B}}^A, \quad \underline{\underline{B}}^T = \underline{\underline{B}}^S - \underline{\underline{B}}^A$$

so the symmetric and antisymmetric parts of a tensor are defined as follows:

$$\underline{\underline{B}}^S = \frac{1}{2}(\underline{\underline{B}} + \underline{\underline{B}}^T), \quad \underline{\underline{B}}^A = \frac{1}{2}(\underline{\underline{B}} - \underline{\underline{B}}^T)$$

The inverse and transposed second-order tensor often noted as

$$\underline{\underline{A}}^{-T}$$

#### • Dot product

- The inner product (contraction) of second-order tensors gives a second-order tensor

$$\{\cdot\} : \mathbb{T}_2^n \times \mathbb{T}_2^n \rightarrow \mathbb{T}_2^n$$

$$\underline{\underline{A}} = \underline{\underline{B}} \cdot \underline{\underline{C}} = (\underline{\underline{C}}^T \cdot \underline{\underline{B}}^T)^T; \quad A_{i*}^{*j} = B_{i*}^{*k} C_{k*}^{*j}$$

- The double inner product (double contraction) of second-order tensors gives a scalar

$$\{\cdot\} : \mathbb{T}_2^n \times \mathbb{T}_2^n \rightarrow \mathbb{R}$$

$$\underline{\underline{A}} : \underline{\underline{B}} = \underline{\underline{B}} : \underline{\underline{A}} = \alpha; \quad A_{i*}^{*j} B_{*j}^{i*} = \alpha$$

- The double scalar product of second-order tensors gives a scalar (in general different from the double inner product)

$$\{\cdot\} : \mathbb{T}_2^n \times \mathbb{T}_2^n \rightarrow \mathbb{R}$$

$$\underline{\underline{A}} \cdot \underline{\underline{B}} = \underline{\underline{A}} : \underline{\underline{B}}^T = \underline{\underline{B}} \cdot \underline{\underline{A}}^T = \alpha; \quad A_i^j B_i^j = B_j^i A_j^i = \alpha$$

The difference between the double inner product and the double scalar product consists in the following: if  $\underline{\underline{A}} \cdot \underline{\underline{A}} = 0$  then  $\underline{\underline{A}} = 0$ . For double inner product, the equality  $\underline{\underline{A}} : \underline{\underline{A}}^T = 0$  does not mean that  $\underline{\underline{A}} = 0$ . By definition, the double inner product and the double scalar product coincide if at least one of the tensors (in operation) is symmetric. A general rule for the scalar product of  $q$ -th order is:

$$\mathbb{T}_m^n \cdot \underbrace{\cdots}_{q \text{ times}} \cdot \mathbb{T}_l^n \rightarrow \mathbb{T}_{l+m-2q}^n.$$

It is obvious that  $l + m \geq 2q$  as well as  $l \geq q$  and  $m \geq q$ .

- **Operations with vectors**

- Left and right scalar products of vector and second-order tensor are defines as:

$$\{\cdot\} : \mathbb{T}_2^n \times \mathbb{T}_1^n \rightarrow \mathbb{T}_1^n; \quad \{\cdot\} : \mathbb{T}_1^n \times \mathbb{T}_2^n \rightarrow \mathbb{T}_1^n$$

$$\underline{\underline{a}} \cdot \underline{\underline{B}} = \underline{\underline{c}} \neq \underline{\underline{d}} = \underline{\underline{B}} \cdot \underline{\underline{a}}; \quad c^j = a^i B_{i*}^{j*} \neq d^j = B_{j*}^{*i} a_i$$

$$\underline{\underline{a}} \cdot \underline{\underline{B}} = \underline{\underline{B}}^T \cdot \underline{\underline{a}}$$

- Left and right vector products of vector and second-order tensor are defined in 3D space:

$$\{\times\} : \mathbb{T}_2 \times \mathbb{T}_1 \rightarrow \mathbb{T}_2; \quad \{\times\} : \mathbb{T}_1 \times \mathbb{T}_2 \rightarrow \mathbb{T}_2$$

$$\underline{\underline{a}} \times \underline{\underline{B}} = \underline{\underline{C}} \neq \underline{\underline{D}} = \underline{\underline{B}} \times \underline{\underline{a}}$$

$$\underline{\underline{a}} \times \underline{\underline{B}} = -(\underline{\underline{B}}^T \times \underline{\underline{a}})^T$$

- If the space of third order tensors is defined  $\mathbb{T}_3^n = \mathbb{T}_1^n \otimes \mathbb{T}_1^n \otimes \mathbb{T}_1^n$ , then left and right tensor products of a vector and a second-order tensor are defined by:

$$\{\otimes\} : \mathbb{T}_2^n \times \mathbb{T}_1^n \rightarrow \mathbb{T}_3^n; \quad \{\otimes\} : \mathbb{T}_1^n \times \mathbb{T}_2^n \rightarrow \mathbb{T}_3^n$$

$$\underline{\underline{B}} \otimes \underline{\underline{a}} = \underline{\underline{C}} \neq \underline{\underline{D}} = \underline{\underline{a}} \otimes \underline{\underline{B}}; \quad C_{i*k}^{*j*} = B_{i*}^{*j} a_k$$

- **Trace of a second-order tensor: scalar and vector invariants**

- Important for mechanics *trace* “tr”, or first invariant of a tensor, is also denoted as  $I$  or  $I_1$ . If a second-order tensor is written as a sum of diads  $\underline{\underline{A}} = \underline{a} \otimes \underline{b} + \cdots + \underline{c} \otimes \underline{d}$  then

$$I(\underline{\underline{A}}) = \text{tr} \underline{\underline{A}} = \underline{a} \cdot \underline{b} + \cdots + \underline{c} \cdot \underline{d}; \quad I(A_{*i}^{j*}) = A_{*i}^{i*}$$

$$\text{tr}(\underline{\underline{A}} + \underline{\underline{B}}) = \text{tr} \underline{\underline{A}} + \text{tr} \underline{\underline{B}}$$

$$\text{tr} \underline{\underline{A}} = \text{tr} \underline{\underline{A}}^T$$

$$\text{tr}(\underline{\underline{A}} \cdot \underline{\underline{B}}) = \text{tr}(\underline{\underline{B}} \cdot \underline{\underline{A}}) = \underline{\underline{A}} : \underline{\underline{B}}$$

$$\text{tr}(\underline{\underline{A}} \cdot \underline{\underline{B}}) = \text{tr}(\underline{\underline{A}}^T \cdot \underline{\underline{B}}^T)$$

$$\text{tr} \underline{\underline{B}} = \text{tr} \underline{\underline{B}}^S, \quad \text{tr} \underline{\underline{B}}^A = 0$$

$$\text{tr}(\underline{\underline{B}} \cdot \underline{\underline{C}}^S) = \text{tr}(\underline{\underline{B}}^S \cdot \underline{\underline{C}}^S)$$

- A vector invariant of any second-order tensor  $\underline{\underline{B}} = \underline{a} \otimes \underline{b} + \cdots + \underline{c} \otimes \underline{d} \in \mathbb{T}$  is a vector  $\underline{\underline{B}}_{\times}$  such that

$$\underline{\underline{B}}_{\times} = \underline{a} \times \underline{b} + \cdots + \underline{c} \times \underline{d}; \quad B_{\times}^i = B_{*j}^{i*} e_j \times e^j$$

$$(\underline{\underline{A}} + \underline{\underline{B}})_{\times} = \underline{\underline{A}}_{\times} + \underline{\underline{B}}_{\times}$$

$$\underline{\underline{B}}_{\times} = 0 \Leftrightarrow \underline{\underline{B}} = \underline{\underline{B}}^S$$

Any antisymmetric tensor can be written as:

$$\underline{\underline{B}}^A = \underline{\omega} \times \underline{\underline{I}} = \underline{\underline{I}} \times \underline{\omega}, \quad \underline{\omega} = -\frac{1}{2} \underline{\underline{B}}_{\times}^A$$

- Note the following useful formulae:

$$\underline{a} \times \underline{\underline{I}} \times \underline{b} = \underline{b} \times \underline{a} - (\underline{a} \cdot \underline{b}) \underline{\underline{I}}$$

$$(\underline{a} \times \underline{\underline{I}})_{\times} = -2\underline{a}$$

$$\underline{a} \times \underline{b} = (\underline{a} \times \underline{\underline{I}}) \cdot \underline{b}$$

$$\underline{a} \times \underline{\underline{I}} = \underline{\underline{I}} \times \underline{a}$$

- **Invariants of second-order tensors**

Besides the trace or first invariant  $I_1$ , two other invariants can be defined for the second-order tensors. These three invariants are called principle invariants of the tensor:

$$I_1 = \text{tr} \underline{\underline{A}}$$

$$I_2 = \frac{1}{2} \left[ (\text{tr} \underline{\underline{A}})^2 - \text{tr}[\underline{\underline{A}}^2] \right]$$

$$I_3 = \det \underline{\underline{A}} = \frac{1}{6} (\text{tr} \underline{\underline{A}})^3 - \frac{1}{2} \text{tr} \underline{\underline{A}} \text{tr}(\underline{\underline{A}}^2) + \frac{1}{3} \text{tr}(\underline{\underline{A}}^3)$$

The determinant of a tensor will be introduced later in Section A.5.

- **Unit second-order tensor**

The unit second order tensor (denoted  $\underline{\underline{I}}$ ) such that

$$\underline{\underline{a}} = \underline{\underline{I}} \cdot \underline{\underline{a}} = \underline{\underline{a}} \cdot \underline{\underline{I}}; \quad I_i^j = \delta_i^j; \quad a_i = \delta_i^j a_j$$

consequently

$$\underline{\underline{A}} = \underline{\underline{I}} \cdot \underline{\underline{A}} = \underline{\underline{A}} \cdot \underline{\underline{I}}$$

The unit second order tensor in any basis may be written as:

$$\underline{\underline{I}} = \underline{\underline{e}} \cdot \underline{\underline{e}} = \underline{\underline{e}} \cdot \underline{\underline{e}} = \underline{\underline{e}}_i \otimes \underline{\underline{e}}^i = \underline{\underline{e}}^i \otimes \underline{\underline{e}}_i$$

- **Projection tensors**

A tensor  $\underline{\underline{P}}$  is called a projection if

$$\underline{\underline{P}} = \underline{\underline{P}}^T, \quad \underline{\underline{P}} \cdot \underline{\underline{P}} = \underline{\underline{P}}$$

For example,

$$\underline{\underline{n}} \otimes \underline{\underline{n}}$$

projects any vector on the direction  $\underline{\underline{n}}$ ; and tensor

$$\underline{\underline{I}} - \underline{\underline{n}} \otimes \underline{\underline{n}}$$

projects any vector on the space orthogonal to the vector  $\underline{\underline{n}}$ .

- **Spherical and deviatoric tensors**

A tensor  $\underline{\underline{B}}$  is called *spherical* if it can be written as

$$\underline{\underline{B}} = \alpha \underline{\underline{I}}, \quad \alpha \in \mathbb{R}$$

The *spherical part* of  $n$ -dimensional tensor  $\underline{\underline{C}}$  is by definition:

$$\frac{1}{n} \text{tr} \underline{\underline{C}} \underline{\underline{I}} = \frac{1}{n} \text{tr} \underline{\underline{C}}^S \underline{\underline{I}}$$

Any symmetric tensor  $\underline{\underline{C}} = \underline{\underline{C}}^T$  for which  $\text{tr} \underline{\underline{C}} = 0$  is called *deviatoric tensor*.

Note that for  $\underline{\underline{C}}^T = -\underline{\underline{C}} \Rightarrow \text{tr} \underline{\underline{C}} = 0$  but the inverse statement is not true.

Any symmetric tensor  $\underline{\underline{B}} = \underline{\underline{B}}^T$  can be written as a sum of its spherical part and its deviatoric part

$$\underline{\underline{B}} = \frac{1}{n} \text{tr} \underline{\underline{B}} \underline{\underline{I}} + \underline{\underline{B}}^d$$

Any tensor  $\underline{\underline{B}} = \underline{\underline{B}}^S + \underline{\underline{B}}^A$  can be splitted into spherical and deviatoric part plus its antisymmetric part

$$\underline{\underline{B}} = \frac{1}{n} \text{tr} \underline{\underline{B}} \underline{\underline{I}} + \underline{\underline{B}}^d + \underline{\underline{B}}^A$$

In most mechanical and physical theories  $n = 3$ .

## A.5 Tensor as a linear operator on vector space

It can be shown that any linear vector operator on vector space, i.e. an operator which transforms one vector into another

$$\underline{f} : \mathbb{T}_1 \rightarrow \mathbb{T}_1 \quad \underline{y} = \underline{f}(\underline{x})$$

can be written as a right scalar product of a second-order tensor with the argument

$$\underline{y} = \underline{\underline{B}} \cdot \underline{x}$$

By definition tensors fulfill the requirements of linearity

$$\underline{\underline{B}} \cdot (\alpha \underline{a} + \beta \underline{b}) = \alpha \underline{\underline{B}} \cdot \underline{a} + \beta \underline{\underline{B}} \cdot \underline{b}$$

Let us write the conditions for an operator to be isomorphic, i.e. that it transforms uniquely the space  $\mathbb{T}_1$  in itself and that a unique inverse transformation exists. For an arbitrary tensor operator  $\underline{\underline{B}} \in \mathbb{T}_2$ , the kernel is all vector space, but the image is rigorously a subset of this space

$$\text{Im} \underline{\underline{B}} = \mathbb{T}_1, \quad \text{Ker} \underline{\underline{B}} \subset \mathbb{T}_1$$

Let us derive the condition  $\text{Ker} \underline{\underline{B}} = \mathbb{T}_1$ . If  $\underline{\underline{B}}$  transforms three linearly independent vectors into three independent vectors, then  $\text{Ker} \underline{\underline{B}} = \mathbb{T}_1$ . Let us consider three linearly independent vectors  $\underline{a}$ ,  $\underline{b}$ ,  $\underline{c}$ , they are independent if and only if

$$(\underline{a} \times \underline{b}) \cdot \underline{c} \neq 0$$

Their images

$$\underline{a}' = \underline{\underline{B}} \cdot \underline{a}, \quad \underline{b}' = \underline{\underline{B}} \cdot \underline{b}, \quad \underline{c}' = \underline{\underline{B}} \cdot \underline{c}$$

are linearly independent if and only if

$$(\underline{a}' \times \underline{b}') \cdot \underline{c}' \neq 0$$

The *Determinant of the tensor*  $\det \underline{\underline{B}}$  is introduced as follows

$$\det \underline{\underline{B}} = \frac{(\underline{a}' \times \underline{b}') \cdot \underline{c}'}{(\underline{a} \times \underline{b}) \cdot \underline{c}} \quad (\text{A.4})$$

Such a definition of the determinant makes it invariant, i.e. independent on the choice of the basis. It is easy to show that

$$\det \underline{\underline{B}} = \det[\underline{\underline{A}}] \det[\underline{\underline{\tilde{B}}}] = \frac{1}{\det[\underline{\underline{A}}]} \det[\underline{\underline{B}}] = \det[\underline{\underline{\tilde{B}}}] = \det[\underline{\underline{\tilde{B}}}^T]$$

$$\det \underline{\underline{B}} = \det[A_{ij}] \det B^{ij} = \frac{1}{\det[A_{ij}]} \det B_{ij} = \det B_{*j}^{i*} = \det B_{j*}^{*i}$$

where  $\underline{\underline{A}}$  or  $[A_{ij}]$  is the covariant fundamental metric matrix and

$$\underline{\underline{\tilde{B}}}, B^{ij}; \quad \underline{\underline{\tilde{B}}}, B_{*j}^{i*}; \quad \underline{\underline{\tilde{B}}}^T, B_{j*}^{*i}; \quad \underline{\underline{\tilde{B}}}, B_{ij}$$



are the contravariant, contra-covariant, co-contravariant and covariant coordinates of the tensor  $\underline{\underline{B}}$ .

It is worth noting that a tensor cannot be defined as the set of the matrix of its covariant or contravariant coordinates and the associated transformation rule:

$$B^{i'j'} = [P_k^{i'}] B^{kl} [P_l^{j'}]; \quad B_{i'j'} = [P_{i'}^k] B_{kl} [P_{j'}^l]$$

because the determinant of these matrices changes with a basis change:

$$\det B^{i'j'} = \det[P_k^{i'}]^2 \det B^{kl},$$

since  $\det[P_k^{i'}]^2 \neq 1$

$$\det B^{i'j'} \neq \det B^{kl}$$

A tensor is an invariant object and its determinant  $\det \underline{\underline{B}}$  should be also invariant. The determinant of a tensor is invariant if it is defined as in (A.4) or as the determinant of the matrix of co-contravariant or co-covariant coordinates.

It is easy to show that

$$\det \underline{\underline{I}} = 1; \quad \det \underline{\underline{A}} = \det \underline{\underline{A}}^T; \quad \det[\underline{\underline{A}} \cdot \underline{\underline{B}}] = \det \underline{\underline{A}} \det \underline{\underline{B}}$$

Now it is possible to introduce the *inverse tensor*  $\underline{\underline{B}}^{-1}$  such as

$$\underline{\underline{B}}^{-1} \cdot \underline{\underline{B}} = \underline{\underline{B}} \cdot \underline{\underline{B}}^{-1} = \underline{\underline{I}} \Rightarrow \det \underline{\underline{B}} \det \underline{\underline{B}}^{-1} = 1 \Rightarrow \det \underline{\underline{B}}^{-1} = \frac{1}{\det \underline{\underline{B}}}$$

The inverse tensor of  $\underline{\underline{B}}$  exists if and only if  $\det \underline{\underline{B}} \neq 0$ .

Similarly to the linear vector operator on the space of vectors, a linear tensor operator on the space of tensors can be introduced.

$$\underline{\underline{T}} : \mathbb{T} \rightarrow \mathbb{T}$$

Such a linear operator can be interpreted as a fourth order tensor  $\underline{\underline{C}}$ . Good examples are the Young-Cauchy elasticity tensor and its inverse are linear bijective operators which state that stress tensor is a linear function of strain tensor and vice versa.

## A.6 S-structures

In practice, in physical and mechanical theories, among widely employed “spaces on spaces” there are the following constructions: complex numbers, vectors, tensors of higher orders and matrices. Sometimes it is necessary to consider sets of such structures: basis (set of vectors), stress-strain state (set of stress and strain tensors), metric matrix (set of scalars), etc. To determine elements of such sets we make use of indices:  $\underline{e}_i$  –  $i$ -th basis vector,  $A_{ij}$  – element of the metric matrix and so on. As practice demonstrates, sometimes it is more convenient to consider such sets not as sets of elements but as “one-piece” elements of “higher” spaces and to work directly with them. To “work” means

perform some operations on these elements in order to obtain some results. For example, it seems very natural to determine bases  $\underline{e}_i, \underline{e}^i, i = 1, m$  as vectors of vectors  $\underline{\underline{e}}, \underline{\underline{e}}$  and to introduce operations to get the metric matrix not in component form  $A_{ij} = \underline{e}_i \cdot \underline{e}_j$ , but directly  $\underline{\underline{A}} = \underline{\underline{e}}^{\otimes} \cdot \underline{\underline{e}}$ .

From an abstract mathematical point of view, a space defined on the elements of another space – also called exterior algebra – has been employed since a long time. A good example, a similar to one type of s-structures, is a multivector – element of exterior algebra on a vector space, which consists in linear combinations of k-vectors also called blades: bivectors  $\underline{a} \wedge \underline{b}$ , trivectors  $(\underline{a} \wedge \underline{b}) \wedge \underline{c}$ . Operator  $\wedge$  is called the wedge operator and somehow is similar in our notations of a tensor product. A generalization of this algebra complemented by relations to exterior algebras is called the Clifford algebra [Lounesto 01], named after William Kingdon Clifford (1845-1879).

A similar idea of structure of structures appears in programming languages as “container” class which contain elements of another class. Such container can be used simply to keep some information. To make better use of such containers a set of operations should be defined: comparison of element, summation, subtraction, multiplication by scalar, etc. The operation result can be of the same class as container or of another class. In such a way the following structures are introduced in the programming code: complex numbers, vectors, tensors and matrices with related operations. It is also natural for programming languages to use sets of such structures: array of vectors, matrix of tensors, etc. However, to make use of it, a consistent set of operations has to be also defined as well as a suitable application.

A considerable part of the dissertation is devoted to the description of surface-point interaction in the framework of classical differential geometry. First, all computations have been performed using index notations. Both points and surfaces are determined in the three dimensional space  $n = 3$ , all points  $\underline{r}(x^i), i = 1, 2, 3$  are determined by coordinates  $x^i$  in three dimensional basis  $\underline{e}_i$ . However, a local coordinate system can be established on the surface  $\xi_\alpha$  and so it can be parametrized by a vector  $\underline{\rho}(\xi_\alpha)$ . To avoid any confusion we are forced to determine new indices  $\alpha = 1, 2$ . So we have to work with different dimensions: 3 dimensional space and 2 dimensional surface. In three dimensions we use established formalism of direct tensor language and in two dimensions Greek indices. Why not to say that  $\xi_1, \xi_2$  are covariant coordinates of a two dimensional vector? In order not to mix up 3D and 2D vector the last will be underlined by a wave  $\underline{\underline{\xi}}$ .

We are now facing the need to evaluate the full variation of the vector  $\underline{\rho}(\underline{\underline{\xi}})$ ; in index form, it is straightforward:

$$\delta \underline{\underline{\rho}} = \delta \underline{\underline{\rho}} + \frac{\partial \underline{\underline{\rho}}}{\partial \xi_\alpha} \delta \xi_\alpha$$

With our new notations of 2D vectors it is less trivial

$$\delta \underline{\underline{\rho}} = \delta \underline{\underline{\rho}} + \frac{\partial \underline{\underline{\rho}}}{\partial \underline{\underline{\xi}}} ? \delta \underline{\underline{\xi}}$$

What is the relation between the two last terms? The first component is nothing but a set of two 3D vectors and the second component is a 2D vector; their product should be a 3D vector. So we need to define a product operation between a 2D vector and a set of 3D vectors in a way that it gives a 3D vector. Moreover this operation should be commutative and the notation should differ from 3D notations. We propose:

$$\underline{a} = \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \circ \delta \underline{\xi} = \delta \underline{\xi} \circ \frac{\partial \underline{\rho}}{\partial \underline{\xi}}$$

It is not so hard to introduce such a product. The situation becomes more complicated when we need to evaluate the surface (induced) metric matrix

$$A_{\alpha\beta} = \frac{\partial \underline{\rho}}{\partial \xi_\alpha} \cdot \frac{\partial \underline{\rho}}{\partial \xi_\beta}$$

In new notations we get

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \boxtimes \frac{\partial \underline{\rho}}{\partial \underline{\xi}} = ?$$

On the left-hand side of the latter expression we have an operation between a pair of sets of two 3D vectors. The right-hand side should be a  $2 \times 2$  symmetric matrix  $A_{\alpha\beta}$ . In the vector space, there is only one operation which converts two vectors into a scalar: the dot product, so the left-hand part should contain a 3D dot product. On the other hand the two structures on the left-hand side can be considered as 2D vectors. In a vector space there is only one operation which converts two vectors into a higher order tensor – the tensor product. Finally it can be seen that between structures on the left-hand side one has to get two operations respectively to 2D and 3D vectors and in the right part a kind of 2D tensor. By analogy with 3D notations, 2D tensors will be underlined by two waves. The s-tensor product “ $\boxtimes$ ” for 2D structure is combined with the dot product “ $\cdot$ ” for 3D structures, so finally we obtain the following definition for the induced metric matrix:

$$\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \boxtimes = \underline{\underline{A}}$$

Here and further the upper operation (here  $\boxtimes$ ) relates to 2D structure and lower (here  $\cdot$ ) to 3D.

The initial attempt to introduce s-structure over 2D and 3D vector spaces was inspired from this example. After some trials to elaborate a more general form of s-structures, we realized that the use of such structures is much more rich both in mathematical and physical senses. Due to the complicated nature of mixed spaces the setting of a consistent s-structure “language” is not a trivial task. It is not yet fully elaborated, neither fully exploited. Below we make an attempt to derive a consistent framework for the space of generalized s-structures with related operations. The full formalism of diad-operations, like in  $\frac{\partial \underline{\rho}}{\partial \underline{\xi}} \cdot \frac{\partial \underline{\rho}}{\partial \underline{\xi}} \boxtimes$ , will be derived as well as its simplified form based on transpose operations, which is valid for several limited cases.

### A.6.1 Formal definition, notations and types

The s-structure formalism consists in the introduction of ordered sets of tensors. The “S” in the name refers to “set” or “super” (higher structure). S-structure can be interpreted as  $k$  order “tensor” on the  $m$  dimensional space of  $l$  order tensors of dimension  $n$ ; all s-structures are elements of S-spaces:

$$\underbrace{{}^m_k \mathbb{S}_l^n = {}^m_1 \mathbb{S}_l^n \otimes \dots \otimes {}^m_1 \mathbb{S}_l^n}_{k \text{ times}}$$

The sign  $\otimes$  means that elements are ordered, however contrary to objects of spaces  $\mathbb{T}_p^n$ , S-space elements are not mandatory invariant, i.e. they do not change according to the standard rules with change of the basis, that is why “tensor” is put in quotes. To avoid any confusion instead of the previously employed notation of tensor spaces a new one will be used for corresponding s-structures  $\tilde{\mathbb{T}}_p^n$ . By definition:

$${}_0^1 \mathbb{S}_0^1 = \mathbb{R}, \quad {}^m_1 \mathbb{S}_0^n = \tilde{\mathbb{T}}_1^m, \quad {}_0^1 \mathbb{S}_1^n = \mathbb{T}_1^n, \quad {}^m_2 \mathbb{S}_0^1 = \tilde{\mathbb{T}}_2^m, \quad {}_0^1 \mathbb{S}_2^n = \mathbb{T}_2^n$$

It is worth noting that s-structures are introduced in such a way, that:

$${}_k^n \mathbb{S}_k^n \neq \mathbb{T}_{k+k}^n, \quad {}_k^n \mathbb{S}_k^n \neq \tilde{\mathbb{T}}_{k+k}^n$$

For instance it is easy to show that  ${}_1^n \mathbb{S}_1^n \neq \mathbb{T}_2^n$ , if one considers the derivative of the simple vector  $\underline{r}(\underline{p})$  over v-scalar of its contravariant coordinates  $\underline{p}$  in basis  $\underline{\tilde{e}}, m = n$

$$\frac{\partial \underline{r}}{\partial \underline{p}} \in {}_1^n \mathbb{S}_1^n, \quad {}_1^n \mathbb{S}_1^n \neq \mathbb{T}_2^n, \quad {}_1^n \mathbb{S}_1^n \neq \tilde{\mathbb{T}}_2^n.$$

This expression is neither an ordinary tensor nor an s-tensor, but a v-vector – an element of  ${}_1^n \mathbb{S}_1^n$  space. Other definitions are possible but they may destroy the invariance of the vectors.

Let us suppose that a zero order s-structure is equal to a tensor space over which it is constructed:

$${}_0^m \mathbb{S}_l^n \equiv \mathbb{T}_l^n$$

Following this equivalence, one could consider an s-space is a super structure over s-space of zero order.

An s-structure of order  $k$  is a generalization of a first order s-structure. It would be convenient to define the space  ${}_1^m \mathbb{S}_l^n$  and the related operations. Elements of this space, are considered as generalized vectors. They are underlined by a single wave and will be called v-elements or v-tensors, since these “vectors” are defined over a tensor space. However, since the nature of these elements can be different, depending on the order  $l$  of the tensor space over which they are constructed, it would be more convenient to distinguish three types of v-tensors (see below). All the newly introduced operations will be similar to vector operations.

• **V-scalars**

V-scalars are elements of a linear space  ${}^m_1\mathbb{S}_0^n = \tilde{\mathbb{T}}_1^m$ ; note that the dimension of the internal space  $n$ , as will be shown later, makes sense and has to be retained. V-scalars and all s-structures are defined through a basis in s-space of dimension  $m$ . The sub-space or tensor space basis (of dimension  $n$ ) should not be mixed up with the basis in the s-space. Invariance of such s-elements will be discussed later. As the invariance of s-structures is not always ensured, we put a tilde over the notation of this vector space  $\tilde{\mathbb{T}}_1^m$ . So it is better to consider these structures as ordered set of scalars or simply as a matrix  $1 \times m$ . But contrary to matrix algebra the operation of transposition will not be introduced<sup>2</sup>. If needed, the dimension  $m$  will be mentioned as the left top index  ${}^m\tilde{\mathbf{b}}$ . Even if  $m = n$ , i.e. vectors and v-scalars are both defined in an  $n$  dimensional space, they remain elements of different spaces  $\tilde{\mathbf{a}} \in \tilde{\mathbb{T}}_1^n$ ,  $\mathbf{b} \in \mathbb{T}_1^n$  and no invariant operations can be defined between them. In such a case they are considered as vectors of completely different reference frames (see the difference between reference frame and coordinate system in Section A.1). Physically we would say that v-scalars and vectors are elements of physical spaces associated with different observers. These spaces are different, see illustration A.1. However, it does not make sense if there is no connection between those two spaces, further we will consider interdependent spaces. Some examples of v-scalar:

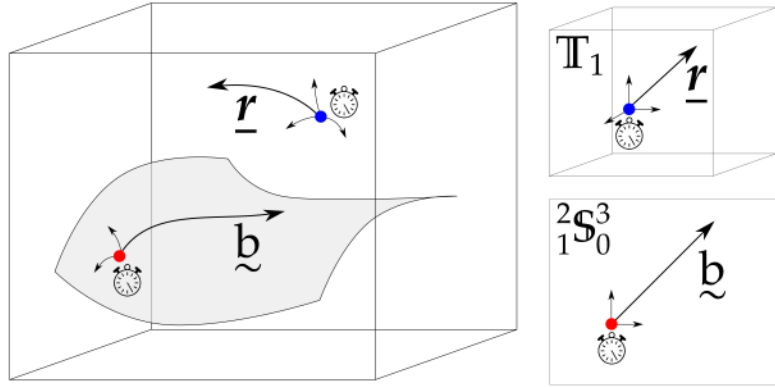


Figure A.1: A physical space for an observer  $A$  and two reference frames (two other observers) - blue and red corresponding to space  $\mathbb{T}_1$  and to two dimensional  $m = 2$  v-scalar space  ${}^2_0\mathbb{S}_0^3$  respectively. Corresponding vector and v-scalar are represented both in the physical space, as observer  $A$  can they see, and in their frames of reference on the right.

– a vector of dimension  $m$ ,  $\tilde{\mathbf{b}}$ , invariant for any choice of the s-space

<sup>2</sup>Transposition of v-scalars will appear in reduced form of s-structures, which can be employed as apparatus of a limited use. The aim is to avoid cumbersome two-level operations appearing in consistent description of s-structures as well as to represent not invariant structures. See Section A.7.

basis;

- a set of vector coordinates  $\underline{a}$  in any basis of dimension  $m$ ; in this sense, the dimension of s-space should be limited to  $m \leq n$ ; here this v-scalar is invariant respectively to the change of the s-space basis, but covariant or contravariant respectively to the  $n$  space basis;
- values of shape functions for 3D finite element as function of local coordinate  ${}^3\underline{a}$ , which is a set of three coordinates:  $\phi({}^3\underline{a})$ , here  $m$  is equal to the number of shape functions.

• **V-vectors**

V-vectors are elements of a linear space  ${}^m\underline{\mathcal{S}}_1^n = \tilde{\mathbb{T}}_1^m$  seen as ordered sets of invariant vectors. V-vectors are vectors of dimension  $m$ . In the given basis in s-space, components of such vectors are vectors. As this structure is invariant with respect to the vector space it cannot be invariant with respect to the s-space. Some examples of v-vectors:

- a set of vectors of covariant basis  ${}^n\underline{e}'$  is a v-vector of dimension  $n$  over vector space  $n$ ;
- a set of vectors of contravariant sub-basis  ${}^m\underline{e}'$  is a v-vector of dimension  $m$  over the vector space  $n$ ;
- a set of node vectors of a 3D finite element  ${}^m\underline{r}$ , where  $m$  is number of element, each vector  $\underline{r}$  is a three dimensional vector  $r \in \mathbb{T}_1$ .

• **V-tensors**

They are a generalization of v-scalars and v-vectors for higher order tensor ranks. Further elements of linear space  $\underline{\underline{A}} \in {}^m\underline{\mathcal{S}}_l^n$  will be called v-tensors of order  $l$ . By “v-tensors” we will simply understand v-tensors of order 2,  $\underline{\underline{B}} \in {}^m\underline{\mathcal{S}}_2^n$ .

The generalization of v-structures are t-structures (“t” for tensor). A t-structure of order  $k$  is defined as the tensor product of  $k$  v-spaces. Here we confine ourself to second order t-structures  $k = 2$

$${}^m\underline{\mathcal{S}}_l^n = {}^m\underline{\mathcal{S}}_l^n \otimes {}^m\underline{\mathcal{S}}_l^n$$

In this sense, this s-space is a generalization of the tensor concept in a space of tensors, consequently all operations to be introduced will be similar to tensor operations. By analogy with tensor theory notations, elements of such s-spaces will be underlined by two waves. We will distinguish

- **T-scalars** elements of space  ${}^m\underline{\mathcal{S}}_0^n$
- **T-vectors** elements of space  ${}^m\underline{\mathcal{S}}_1^n$
- **T-tensors** elements of space  ${}^m\underline{\mathcal{S}}_l^n$

Now we need to determine internal operations in all s-spaces and external operations connecting all s-spaces.

## A.6.2 Simple operations

### V-scalars

The most simple elements of s-spaces are v-scalars. All operations are similar to ordinary vector. Let  $\underline{a}, \underline{b}, \underline{c} \in \tilde{\mathbb{T}}_1^m$  and  $\alpha, \beta \in \mathbb{R}$ . A typical list is:

**Internal operations** of type  $\tilde{\mathbb{T}}_1^m \times \tilde{\mathbb{T}}_1^m \rightarrow \tilde{\mathbb{T}}_1^m, \tilde{\mathbb{T}}_1^m \times \mathbb{R} \rightarrow \tilde{\mathbb{T}}_1^m$

- $\underline{c} = \underline{a} + \underline{b} = \underline{b} + \underline{a}; \quad c^i = a^i + b^i = b^i + a^i$
- $(\underline{a} + \underline{b}) + \underline{c} = \underline{a} + (\underline{b} + \underline{c}); \quad (a^i + b^i) + c^i = a^i + (b^i + c^i)$
- $\alpha \underline{a} = \underline{a} \alpha = \underline{b}; \quad \alpha a^i = a^i \alpha = b^i$
- $(\alpha + \beta) \underline{a} = \alpha \underline{a} + \beta \underline{a}; \quad (\alpha + \beta) a^i = \alpha a^i + \beta a^i$
- For  $\alpha = 0, \beta = 1, (0 + 1) \underline{a} = 0 \underline{a} + \underline{a} = \underline{a}$ , so zero element  $0 = \underline{0} = 0 \underline{a}$

**External operations** of type  $\tilde{\mathbb{T}}_1^m \times \tilde{\mathbb{T}}_1^m \rightarrow \mathbb{R}$

- Scalar or dot product on s-space will be called s-dot product  $\underline{a} \circ \underline{b} = \underline{b} \circ \underline{a} = \alpha; \quad a_i \underline{e}^i \circ \underline{e}_j b^j = a^i b_j \delta_i^j = a^i b_i = \alpha$
- Orthogonal v-scalars  $\underline{a} \circ \underline{b} = \underline{b} \circ \underline{a} = 0; \quad a_i \underline{e}^i \cdot \underline{e}_j b^j = a^i b_i = 0$
- Norm  $\forall \underline{a} \exists \|\underline{a}\| \geq 0: \|\underline{a}\| = 0 \Leftrightarrow \underline{a} = 0,$   
 $\|\underline{a}\| = \sqrt{\underline{a} \circ \underline{a}} \geq 0, \quad \forall \underline{a} + \underline{b} = \underline{c} : \|\underline{a}\| + \|\underline{b}\| \geq \|\underline{c}\|$
- Unit v-scalar: if  $\|\underline{a}\| \neq 0 \Rightarrow \exists! \hat{\underline{a}} = \frac{\underline{a}}{\|\underline{a}\|}$ .

**External operations** of type  $\tilde{\mathbb{T}}_1^m \times \tilde{\mathbb{T}}_1^m \rightarrow \tilde{\mathbb{T}}_2^m$ .

Let  $\underline{C}, \underline{D} \in \tilde{\mathbb{T}}_2^m$

- S-tensor product  $\underline{a} \boxtimes \underline{b} = \underline{C} \neq \underline{D} = \underline{b} \boxtimes \underline{a} \quad a_i \underline{e}^i \boxtimes \underline{e}_j b^j \neq b_j \underline{e}^j \boxtimes \underline{e}_i a^i$

### V-vectors

All internal operations are formally similar to v-scalars. Let  $\underline{a}, \underline{b}, \underline{c} \in {}_1^m \mathbb{S}_1^n,$   
 $\underline{e}, \underline{f}, \underline{g} \in \tilde{\mathbb{T}}_1^m$  and  $\alpha, \beta \in \mathbb{R}$

**Internal operations** of type  ${}^m\mathbb{S}_1^n \times {}^m\mathbb{S}_1^n \rightarrow {}^m\mathbb{S}_1^n$  and  ${}^m\mathbb{S}_1^n \times \mathbb{R} \rightarrow {}^m\mathbb{S}_1^n$

- $\underline{\underline{c}} = \underline{\underline{a}} + \underline{\underline{b}} = \underline{\underline{b}} + \underline{\underline{a}}; \quad \underline{\underline{c}}^{(i)} = \underline{\underline{a}}^{(i)} + \underline{\underline{b}}^{(i)} = \underline{\underline{b}}^{(i)} + \underline{\underline{a}}^{(i)}$ , all indices are put in brackets as vectors are invariant in vector space  $\mathbb{T}_1^n$ ; but in s-space they are either covariant or contravariant coordinates of points in space  ${}^m\mathbb{S}_1^n$ , so the position of the index tells it, here, for example, that these vectors are considered to be contravariant coordinates of v-vector in s-space.
- $(\underline{\underline{a}} + \underline{\underline{b}}) + \underline{\underline{c}} = \underline{\underline{a}} + (\underline{\underline{b}} + \underline{\underline{c}}); \quad (\underline{\underline{a}}^{(i)} + \underline{\underline{b}}^{(i)}) + \underline{\underline{c}}^{(i)} = \underline{\underline{a}}^{(i)} + (\underline{\underline{b}}^{(i)} + \underline{\underline{c}}^{(i)})$
- $\alpha \underline{\underline{a}} = \underline{\underline{a}} \alpha = \underline{\underline{b}}; \quad \alpha \underline{\underline{a}}^{(i)} = \underline{\underline{a}}^{(i)} \alpha = \underline{\underline{b}}^{(i)}$
- $(\alpha + \beta) \underline{\underline{a}} = \alpha \underline{\underline{a}} + \beta \underline{\underline{a}}; \quad (\alpha + \beta) \underline{\underline{a}}^{(i)} = \alpha \underline{\underline{a}}^{(i)} + \beta \underline{\underline{a}}^{(i)}$
- For  $\alpha = 0, \beta = 1, (0 + 1) \underline{\underline{a}} = 0 \underline{\underline{a}} + \underline{\underline{a}} = \underline{\underline{a}}$ , so zero element  $0 = \underline{\underline{0}} = 0 \underline{\underline{a}}$

### A.6.3 Invariant s-structures

#### Invariant v-scalars

To make the s-structure language consistent, we need to prove that s-structures do not depend upon the change of basis. For v-scalars it is trivial, they are invariant by definition, as vectors. Let us derive the operations which allow to determine coordinates of a v-scalar in all bases.

Statement: *the v-scalar  $\underline{\underline{a}}$  does not depend on the change of basis.* Let us introduce a basis of  $m$  v-scalars  $\underline{\underline{e}}_i$ , such that all these v-scalars are linearly independent, i.e.

$$\underline{\underline{e}}_i \alpha^i = 0 \quad \Leftrightarrow \quad \alpha^i = 0$$

Then we construct a dual basis  $\underline{\underline{e}}^i$ , following the standard scheme

$$\underline{\underline{e}}_i \circ \underline{\underline{e}}^j = \delta_i^j$$

So now we can write a v-scalar  $\underline{\underline{a}}$  in these two bases

$$\underline{\underline{a}} = a^i \underline{\underline{e}}_i = a_i \underline{\underline{e}}^i \tag{A.5}$$

Then the s-dot product with  $\underline{\underline{e}}_i$  of the last equality gives:

$$a^i \underline{\underline{e}}_i \circ \underline{\underline{e}}_j = a_i \underline{\underline{e}}^i \circ \underline{\underline{e}}_j$$

By definition of the dual basis  $\underline{\underline{e}}^i \circ \underline{\underline{e}}_j = \delta_j^i$  and we get

$$a^i [S_{ij}] = a_j,$$

where  $[S_{ij}] = \underline{\underline{e}}_i \circ \underline{\underline{e}}_j$  is the fundamental metric matrix. The same operations are performed with the dual basis: we evaluate the s-dot product of (A.5) with vectors of the dual basis and get

$$a^i = \underline{\underline{e}}^i \circ \underline{\underline{e}}^j a_j = [S^{ij}] a_j$$



If  $\underline{e}_i$  are linearly independent, then  $\det[S^{ij}] \neq 0$  and

$$[S^{ij}][S_{jk}] = \delta_k^i$$

Following the process used for vectors, the pushforward and pullback transformation of coordinates can be derived. Introducing a new basis  $\underline{e}_{i'}$  and its dual basis  $\underline{e}^{i'}$ :

$$\underline{a} = a^{i'} \underline{e}_{i'} = a^i \underline{e}_i$$

The dot product of this expression with the new dual basis vectors gives

$$a^{j'} = a^i \underline{e}^{j'} \circ \underline{e}_i,$$

where

$$\underline{e}^{j'} \circ \underline{e}_i = [P_i^{j'}]$$

is the pushforward transformation matrix. In the same way

$$\underline{a} = a_{i'} \underline{e}^{i'} = a_i \underline{e}^i \Rightarrow a_{j'} = a_i \underline{e}_{j'} \circ \underline{e}^i$$

$$\underline{e}_{j'} \circ \underline{e}^i = [P_{j'}^i]$$

is the pullback transformation matrix. Finally

$$\underline{e}^{j'} = [P_i^{j'}] \underline{e}^i \quad \underline{e}_{j'} = [P_{j'}^i] \underline{e}_i$$

$$a^{j'} = [P_i^{j'}] a^i \quad a_{j'} = [P_{j'}^i] a_i$$

and

$$[P_{i'}^k][P_k^{j'}] = \delta_{i'}^{j'}$$

### V-vectors

The next question is: are v-vectors invariant or not? Yes and no, they are invariant in a certain sense, precisely each component of a v-vector is invariant with respect to the vector space, by definition of a vector, at the same time the set of these vectors is invariant in s-space. But a v-vector is not invariant respectively to both spaces simultaneously. Further we will call an element of a s-structure invariant if its components are invariant in tensor space (i.e. follow the transformation rules) and if the whole structure is invariant in the s-space, i.e. it follows the transformation rules defined in the s-space.

Let a v-vector  $\underline{a}$  be invariant in s-space, then

$$\underline{a} = \underline{a}^{(i)} \underline{e}_i = \underline{a}_{(i)} \underline{e}^i;$$

from this expression and the previous paragraph it follows:

$$\begin{aligned} \underline{a}^{(j)} &= [S^{ji}] \underline{a}_{(i)} & \underline{a}_{(j)} &= [S_{ji}] \underline{a}^{(i)} \\ \underline{a}^{(j')} &= [P_i^{j'}] \underline{a}^{(i)} & \underline{a}_{(j')} &= [P_{j'}^i] \underline{a}_{(i)}. \end{aligned} \tag{A.6}$$

As expected, vectors  $\underline{a}^{(i)}$  change in vector space with a changing basis in s-space.

### Remark on induced metric and local basis

To describe a surface in a three dimensional space, we use the induced metric t-scalar  $\underline{\underline{A}}$  and another t-scalar  $\underline{\underline{H}}$ . These structures are defined as follows

$$\underline{\underline{A}} = \frac{\partial \underline{\underline{p}}}{\partial \underline{\underline{\xi}}} \cdot \frac{\partial \underline{\underline{p}}}{\partial \underline{\underline{\xi}}}, \quad \underline{\underline{H}} = \underline{\underline{n}} \cdot \frac{\partial^2 \underline{\underline{p}}}{\partial \underline{\underline{\xi}}^2}$$

None of these structures are invariant in the aforementioned sense, because the components of v-vector  $\frac{\partial \underline{\underline{p}}}{\partial \underline{\underline{\xi}}}$  determining the local basis do not follow the rule

$$\frac{\partial \underline{\underline{p}}}{\partial \underline{\underline{\xi}}^i} = S^{ij} \frac{\partial \underline{\underline{p}}}{\partial \underline{\underline{\xi}}_j}$$

The latter statement is true only if  $A^{ij} = S^{ij}$ , which is not the case in general. That is why the whole developed theory of invariant s-structures remains for the moment a “thing-in-itself” and is applicable for considering problems only if the basis in s-space is fixed which is the case for the parent space in the Finite Element Method formalism. However, the range of possible application is not limited by the considered situation.

### A.6.4 Scalar products of V-vectors

#### Scalar product of v-vector and vector

The first product operation which will be introduced is the scalar product between a v-vector and a vector. This operation associates one v-scalar to each pair of a v-vector and a vector

$${}^m_1 \mathbf{S}_1^n \times \mathbb{T}_1^n \rightarrow {}^m_1 \mathbf{S}_0^n$$

Let  $\underline{\underline{c}} \in {}^m_1 \mathbf{S}_0^n, \underline{\underline{a}} \in {}^m_1 \mathbf{S}_1^n, \underline{\underline{b}} \in \mathbb{T}_1^n$  then:

$$\underline{\underline{c}} = \underline{\underline{a}} \cdot \underline{\underline{b}} = \underline{\underline{b}} \cdot \underline{\underline{a}}$$

This operation is easy to introduce:

$$c^i \underline{\underline{e}}_i = (\underline{\underline{a}}^{(i)} \cdot \underline{\underline{b}}) \underline{\underline{e}}_i = (\underline{\underline{b}} \cdot \underline{\underline{a}}^{(i)}) \underline{\underline{e}}_i$$

Let us demonstrate that the resulting s-scalar is bi-invariant, i.e. it does not depend on the choice of the basis, neither in vector space nor in s-space. Since vectors  $\underline{\underline{a}}^{(i)}$  and  $\underline{\underline{b}}$  do not depend on the choice of the basis in the vector space, it remains to show that:

$$c^i \underline{\underline{e}}_i = (\underline{\underline{a}}^{(i)} \cdot \underline{\underline{b}}) \underline{\underline{e}}_i = (\underline{\underline{a}}_{(i)} \cdot \underline{\underline{b}}) \underline{\underline{e}}^i$$

According to (A.6), the latter equalities can be transformed into:

$$([S^{ij}] \underline{\underline{a}}_{(j)} \cdot \underline{\underline{b}}) [S_{ik}] \underline{\underline{e}}^k = (\underline{\underline{a}}_{(k)} \cdot \underline{\underline{b}}) \underline{\underline{e}}^k,$$

since  $[S^{ij}][S_{ik}] = \delta_{kj}$ , the bi-invariance of such a v-scalar  $\underline{\underline{c}}$  has been demonstrated.

### Scalar product of v-vector and v-scalar

By analogy we define the scalar product operation between a v-vector and a v-scalar. For each such pair, this operation associates one vector:

$${}_1^m \mathbf{S}_1^n \times {}_1^m \mathbf{S}_0^n \rightarrow \mathbb{T}_1^n$$

Let  $\underline{\underline{a}} \in {}_1^m \mathbf{S}_1^n$ ,  $\underline{\underline{b}} \in {}_1^m \mathbf{S}_0^n$ ,  $\underline{\underline{c}} \in \mathbb{T}_1^n$ , then

$$\underline{\underline{c}} = \underline{\underline{a}} \circ \underline{\underline{b}} = \underline{\underline{b}} \circ \underline{\underline{a}}$$

In component form:

$$c^i \underline{\underline{e}}_i = \underline{\underline{a}}^{(i)} (\underline{\underline{e}}_i \circ \underline{\underline{b}}) = \underline{\underline{a}}^{(i)} (\underline{\underline{b}} \circ \underline{\underline{e}}_i)$$

To show that  $\underline{\underline{c}}$  is bi-invariant with respect to a basis change in s-space and vector space, it remains to show that

$$c^i \underline{\underline{e}}_i = \underline{\underline{a}}^{(i)} (\underline{\underline{e}}_i \circ \underline{\underline{b}}) = \underline{\underline{a}}_{(i)} (\underline{\underline{e}}^i \circ \underline{\underline{b}})$$

By analogy with the previous paragraph, according to (A.6), the last equalities can be transformed into:

$$[S^{ij}] \underline{\underline{a}}_{(j)} ([S_{ik}] \underline{\underline{e}}^k \circ \underline{\underline{b}}) = \underline{\underline{a}}_{(k)} (\underline{\underline{e}}^k \circ \underline{\underline{b}})$$

since  $[S^{ij}][S_{ik}] = \delta_k^j$ , the bi-invariance of this vector  $\underline{\underline{c}}$  is demonstrated.

### Double scalar product

Let us show that there is a scalar product operation such that for each pair of v-vector it associates a unique real number. This operation will be called full scalar product or double scalar product:

$${}_1^m \mathbf{S}_1^n \times {}_1^m \mathbf{S}_1^n \rightarrow \mathbb{R}$$

The requirement of obtaining a unique scalar implies that such an operation does not depend on basis change both in vector space and s-space. Let  $\underline{\underline{a}}, \underline{\underline{b}} \in {}_1^m \mathbf{S}_1^n$ ,  $\alpha \in \mathbb{R}$  for any value of  $m$  and  $n$

$$\langle \underline{\underline{a}}, \underline{\underline{b}} \rangle = \underline{\underline{a}} \circ \underline{\underline{b}} = \alpha$$

This operation can be introduced as follows:

$$\alpha = \underline{\underline{a}}^{(i)} \cdot \underline{\underline{b}}_{(i)} = \underline{\underline{a}}^{(i)} \cdot \underline{\underline{b}}_{(j)} \delta_i^j$$

which is equal to

$$\alpha = \underline{\underline{a}}^{(i)} \cdot \underline{\underline{b}}_{(j)} \underline{\underline{e}}_i \circ \underline{\underline{e}}^j = [\underline{\underline{a}}^{(i)} \underline{\underline{e}}_i] \circ [\underline{\underline{b}}_{(j)} \underline{\underline{e}}^j]$$

Let us demonstrate that the SS-product, introduced in such a way, is bi-invariant. For this purpose, let us write the v-vector in the s-space basis

$$\underline{\underline{a}} = \underline{a}^{(i)} \underline{e}_i = \underline{a}_{(i)} \underline{e}^i \quad \underline{\underline{b}} = \underline{b}^{(j)} \underline{e}_j = \underline{b}_{(j)} \underline{e}^j$$

Due to the invariance in the s-space, it is necessary and sufficient that:

$$\underline{a}^{(i)} \underline{e}_i \circ \underline{b}^{(j)} \underline{e}_j = \underline{a}^{(i)} \underline{e}_i \circ \underline{b}_{(j)} \underline{e}^j = \underline{a}_{(i)} \underline{e}^i \circ \underline{b}^{(j)} \underline{e}_j = \underline{a}_{(i)} \underline{e}^i \circ \underline{b}_{(j)} \underline{e}^j$$

Carrying  $\underline{e}_i \circ \underline{e}_j = [S_{ij}]$  and  $\underline{e}^i \circ \underline{e}^j = [S^{ij}]$  we get

$$\underline{a}^{(i)} \cdot \underline{b}^{(j)} [S_{ij}] = \underline{a}^{(i)} \cdot \underline{b}_{(i)} = \underline{a}_{(i)} \cdot \underline{b}^{(i)} = \underline{a}_{(i)} \cdot \underline{b}_{(j)} [S^{ij}].$$

According to (A.6) and the symmetry of the fundamental metric matrices:

$$[S_{ij}][S^{jk}]\underline{a}^{(j)} \cdot \underline{b}_{(k)} = \underline{a}^{(i)} \cdot \underline{b}_{(i)} = [S_{ij}][S^{ik}]\underline{a}^{(j)} \cdot \underline{b}_{(k)} = [S_{ik}][S^{ij}]\underline{a}^{(k)} \cdot \underline{b}_{(j)}$$

Since  $[S_{ij}][S^{jk}] = \delta_i^k$  and  $[S^{ij}] = [S^{ji}]$ , the equality is proven. So the SS-product is a bi-invariant operation.

### Scalar product or dot product

Let us introduce another important bi-invariant operation which to each pair of ordered v-vectors associates a unique t-scalar. T-scalars being invariant by definition in s-space, let us introduce the operation where the resulting t-scalar is invariant to the change of basis in vector space:

$$\{\cdot\} : {}^m\mathcal{S}_1^n \times {}^m\mathcal{S}_1^n \rightarrow {}^m\mathcal{S}_0^n$$

Let  $\underline{\underline{a}}, \underline{\underline{b}} \in {}^m\mathcal{S}_1^n, \underline{\underline{c}} \in {}^m\mathcal{S}_0^n$

$$[\underline{\underline{a}}, \underline{\underline{b}}] = \underline{\underline{a}} \cdot \underline{\underline{b}} = \underline{\underline{c}}$$

It is worth mentioning that  $n$  and  $m$  can take any value and that the resulting t-scalar may be not symmetric

$$\underline{\underline{a}} \cdot \underline{\underline{b}} \neq \underline{\underline{b}} \cdot \underline{\underline{a}}$$

This operation can be introduced in the following way:

$$\underline{\underline{c}} = C_{*j}^{i*} \underline{e}_i \boxtimes \underline{e}^j = (\underline{a}^{(i)} \cdot \underline{b}_{(j)}) \underline{e}_i \boxtimes \underline{e}^j = (\underline{a}^{(i)} \underline{e}_i) \cdot (\underline{b}^{(j)} \underline{e}_j)$$

The following condition is requested to ensure the invariance of t-scalar  $\underline{\underline{c}}$ :

$$C_{*j}^{i*} \underline{e}_i \boxtimes \underline{e}^j = C^{ij} \underline{e}_i \boxtimes \underline{e}_j = C_{ij} \underline{e}^i \boxtimes \underline{e}^j = C_{i*}^{*j} \underline{e}^i \boxtimes \underline{e}_j,$$

which, according to the relations between the basis and the dual basis, is equal to:

$$C_{*j}^{i*} \underline{e}_i \boxtimes \underline{e}^j = C^{ij} \underline{e}_i \boxtimes \underline{e}^k [S_{jk}] = C_{ij} [S^{ik}] \underline{e}_k \boxtimes \underline{e}^j = C_{i*}^{*j} [S^{ik}] \underline{e}_k \boxtimes \underline{e}^l [S_{jl}]$$

which, in turn, is equivalent to:

$$C_{*j}^{i*} = C^{ik}[S_{kj}] = C_{kj}[S^{ki}] = C_{k*}^{*l}[S^{ki}][S_{lj}], \quad (\text{A.7})$$

where the coefficients  $C$  are different scalar products of v-vector components. By definition, they are determined as follows:

$$C_{*j}^{i*} = \underline{a}^{(i)} \cdot \underline{b}_{(j)} \quad C^{ik} = \underline{a}^{(i)} \cdot \underline{b}^{(k)}$$

$$C_{kj} = \underline{a}_{(k)} \cdot \underline{b}_{(j)} \quad C_{k*}^{*l} = \underline{a}_{(k)} \cdot \underline{b}^{(l)}$$

which according to (A.6) can be rewritten as

$$C_{*j}^{i*} = \underline{a}^{(i)} \cdot \underline{b}_{(j)} \quad C^{ik} = \underline{a}^{(i)} \cdot \underline{b}_{(m)}[S^{km}]$$

$$C_{kj} = [S_{km}]\underline{a}^{(m)} \cdot \underline{b}_{(j)} \quad C_{k*}^{*l} = [S_{km}]\underline{a}^{(m)} \cdot \underline{b}_{(n)}[S^{ln}]$$

Substituting of these expressions in (A.7) gives

$$\underline{a}^{(i)} \cdot \underline{b}_{(j)} = \underline{a}^{(i)} \cdot \underline{b}_{(m)}[S^{km}][S_{kj}] = [S_{km}]\underline{a}^{(m)} \cdot \underline{b}_{(j)}[S^{ki}] = [S_{km}]\underline{a}^{(m)} \cdot \underline{b}_{(m)}[S^{lm}][S^{ki}][S_{lj}]$$

These three equalities are true as  $[S^{km}][S_{kj}] = \delta_j^m$ , so the bi-invariance of this scalar product has been proven. This operation associates only one t-scalar with each ordered pair of v-vectors. This t-scalar is invariant in the s-space.

### S-Scalar product or s-dot product

By analogy, another bi-variant scalar operation associates a unique tensor for each ordered pair of v-vectors:

$$\{\circ\} : {}^m\mathbb{S}_1^n \times {}^m\mathbb{S}_1^n \rightarrow {}^m\mathbb{S}_2^n$$

Let  $\underline{a}, \underline{b} \in {}^m\mathbb{S}_1^n, \underline{C} \in {}^m\mathbb{S}_2^n$

$$\{\underline{a}, \underline{b}\} = \underline{a} \circ \underline{b} = \underline{C}$$

It is defined in the following way:

$$\underline{a} \circ \underline{b} = \underline{a}^{(i)} \underline{e}_i \circ \underline{b}^{(j)} \underline{e}_j = \underline{a}^{(i)} \otimes \underline{b}^{(j)} \underline{e}_i \circ \underline{e}_j$$

As previously made this operation is introduced for any  $n, m$  and in general:

$$\underline{a} \circ \underline{b} \neq \underline{b} \circ \underline{a}$$

Again, it can be shown that this operation is bi-invariant and the resulting tensor is invariant in tensor space.

### A.6.5 Inverse v-vector

A v-vector  $\underline{\underline{B}} \in {}^m_1\mathbb{S}_1^n$  v-vector will be called inverse of a v-vector  $\underline{\underline{A}} \in {}^m_1\mathbb{S}_1^n$  respectively to the s-space, if the following conditions are fulfilled

$$\underline{\underline{A}} \cdot \underline{\underline{B}} = \underline{\underline{I}} \in {}^m_2\mathbb{S}_0^n$$

According to its definition, a unit t-scalar in any basis can be presented as  $\underline{\underline{I}} = \underline{\underline{e}}_i \boxtimes \underline{\underline{e}}^i$ . Then

$$\underline{\underline{e}}_i \boxtimes \underline{\underline{e}}^i = \underline{\underline{A}} \cdot \underline{\underline{B}} = \underline{\underline{A}}^{(i)} \underline{\underline{e}}_i \cdot \underline{\underline{B}}_{(j)} \underline{\underline{e}}^j = [\underline{\underline{A}}^{(i)} \cdot \underline{\underline{B}}_{(j)}] \underline{\underline{e}}_i \boxtimes \underline{\underline{e}}^j$$

In order to fulfill this equality we should require that

$$\underline{\underline{A}}^{(i)} \cdot \underline{\underline{B}}_{(j)} = \delta_j^i \quad \Leftrightarrow \quad \underline{\underline{A}}^{(i)} \cdot \underline{\underline{B}}_{(j)} = 1, i = j \quad \underline{\underline{A}}^{(i)} \cdot \underline{\underline{B}}_{(j)} = 0, i \neq j$$

if  $\underline{\underline{A}}^{(i)} = A_j^{(i)} \underline{\underline{e}}^j$  and  $\underline{\underline{B}}_{(i)} = B_{(i)}^j \underline{\underline{e}}_j$  the previous condition implies that

$$B_{(i)}^i = \frac{A_i^{(i)}}{(A_i^{(i)})^2}, \quad B_{(i)}^j = 0 \Rightarrow \quad (A.8)$$

$$\underline{\underline{A}}_j^{(i)} \cdot \underline{\underline{B}}_{(i)}^j = A_j^{(i)} (\underline{\underline{e}}^j \cdot \underline{\underline{e}}_j) \frac{A_j^{(i)}}{(A_j^{(i)})^2} = 1; \quad \underline{\underline{A}}_j^{(i)} \cdot \underline{\underline{B}}_{(i)}^k = A_j^{(i)} \underline{\underline{e}}^j \cdot \underline{\underline{e}}_k \frac{A_k^{(i)}}{(A_k^{(i)})^2} = 0$$

So the reasonable requirement for the inverse v-vector existence for  $\underline{\underline{A}}$  is

$$A_j^{(i)} \neq 0. \quad (A.9)$$

The inverse on the s-space v-vector to the v-vector  $\underline{\underline{A}}$  will be denoted by  $\underline{\underline{A}}^{-1^s}$ , where the small index  $s$  recalls that this is an inverse in the s-space. It can be shown that if  $\underline{\underline{A}}^{-1^s}$  is the inverse of a v-vector  $\underline{\underline{A}}$  on s-space, then  $\underline{\underline{A}}^{-1^s}$  is the inversed v-vector on tensor space and vice versa:

$$\underline{\underline{A}} \cdot \underline{\underline{A}}^{-1^s} = \underline{\underline{I}} \quad \Leftrightarrow \quad \underline{\underline{A}} \circ \underline{\underline{A}}^{-1^s} = \underline{\underline{I}}$$

Let us split the second equality:

$$\underline{\underline{A}} \circ \underline{\underline{A}}^{-1^s} = (\underline{\underline{A}}^{(i)} \underline{\underline{e}}_i) \circ (\underline{\underline{e}}^j \{ \underline{\underline{A}}^{-1^s} \}_{(j)}) = (A_k^{(i)} \underline{\underline{e}}^k \underline{\underline{e}}_i) \circ (\underline{\underline{e}}^j \{ A^{-1^s} \}_{(j)}^m \underline{\underline{e}}_m) = ((A_k^{(i)} \underline{\underline{e}}_i \circ \{ A^{-1^s} \}_{(j)}^m \underline{\underline{e}}^j) \underline{\underline{e}}^k \otimes \underline{\underline{e}}_m,$$

to get a unit tensor we should require the following:

$$((A_k^{(i)} \underline{\underline{e}}_i \circ \{ A^{-1^s} \}_{(j)}^m \underline{\underline{e}}^j) = \delta_m^k,$$

which is equivalent to:

$$((A_k^{(i)} \underline{\underline{e}}_i \circ \{ A^{-1^s} \}_{(j)}^k \underline{\underline{e}}^j) = 1; \quad ((A_k^{(i)} \underline{\underline{e}}_i \circ \{ A^{-1^s} \}_{(j)}^m \underline{\underline{e}}^j) = 0, m \neq k$$

$$A_k^{(i)} \{A^{-1s}\}_{(i)}^k = 1; \quad A_k^{(i)} \{A^{-1s}\}_{(i)}^m = 0, m \neq k$$

which according to (A.8) is for  $k = m$

$$A_k^{(i)} \frac{A_k^{(i)}}{(A_k^{(i)})^2} = 0 + \dots + 0 + A_k^{(k)} \frac{A_k^{(k)}}{(A_k^{(k)})^2} + 0 + \dots + 0 = 1;$$

and for  $k \neq m$

$$A_k^{(i)} \frac{A_m^{(i)}}{(A_m^{(i)})^2} = 0 + \dots + 0 + \overbrace{A_k^{(k)} \frac{A_m^{(k)}}{(A_m^{(k)})^2}}^{=0; m \neq i} + \dots + \overbrace{A_k^{(m)} \frac{A_m^{(m)}}{(A_m^{(m)})^2}}^{=0; m \neq i} + 0 + \dots + 0 = 0$$

Thus it is demonstrated that the inverse v-vector on s-space is the inverse on tensor space. Since all transitions are sufficient and necessary, the inverse statement is also true. From now on, we will note  $\underline{\underline{A}}^{-1}$  all inverse v-vectors; if such a v-vector exists, the following statements are true:

$$\underline{\underline{A}} \circ \underline{\underline{A}}^{-1} = \underline{\underline{I}}, \quad \underline{\underline{A}} \cdot \underline{\underline{A}}^{-1} = \underline{\underline{I}}$$

Since  $\underline{\underline{I}} = \underline{\underline{I}}^T$  and  $\underline{\underline{I}} = \underline{\underline{I}}^T$ :

$$\underline{\underline{A}} \circ \underline{\underline{A}}^{-1} = \underline{\underline{A}}^{-1} \circ \underline{\underline{A}} \Leftrightarrow \underline{\underline{A}} \cdot \underline{\underline{A}}^{-1} = \underline{\underline{A}}^{-1} \cdot \underline{\underline{A}}$$

### A.6.6 Isomorphism of s-space and tensor space

If the dimensions of the s-space and the tensor space are equivalent  $n = m$ , then they are isomorphic, i.e. for each element of the s-space of dimension  $n$  and order  $k$  there is only one element of the  $k$ -th order tensor space of dimension  $n$  and vice versa. All structures and operations are also preserved. Isomorphism between vectors and v-scalars is ensured by v-vector  $\underline{\underline{A}} \in {}^n\mathcal{S}_1^n$ , which can be considered as bijective linear function:

$$\underline{\underline{A}} : {}^n\mathcal{S}_0^n \rightarrow \mathbb{T}_1^n \quad \text{and} \quad \underline{\underline{A}} : \mathbb{T}_1^n \rightarrow {}^n\mathcal{S}_0^n$$

Not all v-vectors can establish an isomorphism.  $\underline{\underline{A}}$  should be defined such that

$$\text{if } \text{Ker} \underline{\underline{A}} \equiv {}^n\mathcal{S}_0^n, \text{ then } \text{Im} \underline{\underline{A}} \equiv \mathbb{T}_1^n,$$

$$\text{if } \text{Ker} \underline{\underline{A}} \equiv \mathbb{T}_1^n, \text{ then } \text{Im} \underline{\underline{A}} \equiv {}^n\mathcal{S}_0^n,$$

It means that  $n$  linear independent vectors (or v-scalars) are transformed into  $n$  linear independent v-scalars (or vectors).

**Theorem.** If  $\underline{\underline{A}} \in {}^n\mathcal{S}_1^n$  and  $\underline{\underline{A}} = \underline{\underline{A}}_{(i)} \underline{\underline{e}}^i$ , where  $\underline{\underline{A}}_{(i)}$  are  $n$  linearly independent vectors and  $\underline{\underline{e}}^i$  are  $n$  linearly independent v-scalars, then  $\underline{\underline{A}}$  is a bijection from the v-scalar

(vector) space to the vector (v-scalar space) space and the unique inverse bijection  $\underline{\underline{B}}$  from the vector (v-scalar space) space to the v-scalar (vector space) space exists, such that  $\underline{\underline{B}} = \underline{\underline{A}}^{-1}$ .

**Proof.** Let

$$\forall \underline{\underline{x}} \in {}^n\mathbb{S}_0^n \mid \underline{\underline{x}} = x^i \underline{\underline{e}}_i, \forall x^i \neq 0,$$

where the v-scalars  $\underline{\underline{e}}_i$  are linearly independent, then  $\underline{\underline{A}}$  is a bijection

$$\underline{\underline{A}} : {}^n\mathbb{S}_0^n \rightarrow \mathbb{T}_1^n,$$

if and only if

$$\exists \underline{\underline{y}} \in \mathbb{T}_1^n, \underline{\underline{y}} = y^k \underline{\underline{e}}_k, y^k \neq 0,$$

where  $\underline{\underline{e}}_k$  are linearly independent vectors. The bijection  $\underline{\underline{A}}$  is as follows

$$\underline{\underline{y}} = \underline{\underline{A}} \circ \underline{\underline{x}} \Leftrightarrow y^k \underline{\underline{e}}_k = \underline{\underline{A}}_{(k)} \underline{\underline{e}}^k \circ x^i \underline{\underline{e}}_i \Leftrightarrow y^k \underline{\underline{e}}_k = \underline{\underline{A}}_{(k)} x^k$$

if  $y^k = 0$ , then

$$\underline{\underline{A}}_{(k)} x^k = 0$$

since all  $x^k$  are non-zero by default, then vectors  $\underline{\underline{A}}_{(k)}$  are linearly dependent, which contradicts the condition of theorem, so  $y^k \neq 0$  and  $\underline{\underline{A}}$  is a bijection from  ${}^n\mathbb{S}_0^n$  to  $\mathbb{T}_1^n$ .

Let an inverse bijection  $\underline{\underline{B}}$  exists such that

$$\underline{\underline{B}} : {}^n\mathbb{S}_0^n \rightarrow \mathbb{T}_1^n,$$

then it transforms elements:

$$\underline{\underline{x}}' = \underline{\underline{B}} \cdot \underline{\underline{y}}$$

Substituting  $\underline{\underline{y}}$  as image of  $\underline{\underline{x}}$  gives:

$$\underline{\underline{x}}' = \underline{\underline{B}} \cdot (\underline{\underline{A}} \circ \underline{\underline{x}}),$$

to get  $\underline{\underline{x}}' = \underline{\underline{x}}$ , we require that  $\underline{\underline{B}} \cdot \underline{\underline{A}} = \underline{\underline{I}}$ , and by definition  $\underline{\underline{B}} = \underline{\underline{A}}^{-1}$ . The theorem is proven.

### Unit v-vector

At first glance it **seems** to be meaningful to introduce a special kind of bijective function – a unit v-vector  $\underline{\underline{I}}$  – which is its own inverse. The first form for this v-vector which crosses on mind can be defined only if  $n = m$  in the following way

$$\underline{\underline{I}} = \frac{1}{2} (\underline{\underline{e}}_i \underline{\underline{e}}^i + \underline{\underline{e}}^i \underline{\underline{e}}_i), \quad (\text{A.10})$$



where  $\underline{e}_i, \underline{e}^i$  are vectors and v-scalars of bases and  $\underline{e}^i, \underline{e}_i$  are vectors and v-scalars of dual bases in vector space and s-space respectively. Then we require that for any vector  $\underline{a} \in \mathbb{T}_1^n$  and v-scalar  $\underline{\alpha} \in {}^m_1\mathbb{S}_0^n$

$$\underline{a} = \underline{\underline{I}} \circ \underline{\alpha} \quad \underline{\alpha} = \underline{\underline{I}} \cdot \underline{a} \quad \underline{\underline{I}} = \underline{\underline{I}}^{-1}$$

Let us proof, that  $\underline{\underline{I}}$  defined like in (A.10) is its own inverse:

$$\underline{\underline{I}} \cdot \underline{\underline{I}} = \underline{\underline{I}}$$

$$\begin{aligned} \underline{\underline{I}} \cdot \underline{\underline{I}} &= \frac{1}{4} (\underline{e}_i \underline{e}^i + \underline{e}^i \underline{e}_i) \cdot (\underline{e}_j \underline{e}^j + \underline{e}^j \underline{e}_j) = \\ &= \frac{1}{4} (\underline{e}_i \underline{e}^i \cdot \underline{e}_j \underline{e}^j + \underline{e}_i \underline{e}^i \cdot \underline{e}^j \underline{e}_j + \underline{e}^i \underline{e}_i \cdot \underline{e}_j \underline{e}^j + \underline{e}^i \underline{e}_i \cdot \underline{e}^j \underline{e}_j) = \\ &= \frac{1}{4} ([A_{jk}][S^{jm}]\underline{e}_i \cdot \underline{e}^k \underline{e}^i \boxtimes \underline{e}_m + \delta_i^j \underline{e}^i \boxtimes \underline{e}_j + \delta_j^i \underline{e}^i \boxtimes \underline{e}_j + [A^{ik}][S_{im}]\underline{e}_k \cdot \underline{e}^j \underline{e}^m \boxtimes \underline{e}_j) \end{aligned} \quad (\text{A.11})$$

If  $[A^{ik}][S_{kj}] = \delta_j^i$  and  $[S^{ik}][A_{kj}] = \delta_j^i$  then

$$\underline{\underline{I}} \cdot \underline{\underline{I}} = \frac{1}{4} ([S_{ji}][S^{jm}]\underline{e}^i \boxtimes \underline{e}_m + [S^{ij}][S_{im}]\underline{e}^m \boxtimes \underline{e}_j) + \frac{1}{4} (\delta_i^j + \delta_j^i) \underline{e}^i \boxtimes \underline{e}_j$$

But since  $[A^{ij}]$  and  $[S^{ij}]$  are independent metric matrices in different spaces there is no connection between them, so the unit v-vector can be defined in spaces with equal metric in each point. Under this condition the following statement is true

$$\underline{\underline{I}} \cdot \underline{\underline{I}} = \frac{1}{4} 4 \underline{\underline{I}} = \underline{\underline{I}}.$$

For arbitrary tensor space and s-space, the question of the unit v-vector remains open.

### Isomorphism between sub-spaces of s-space and tensor space

A more interesting case arises when the tensor space and the s-space are of different dimensions, which is, by the way, the case considered in the dissertation: the s-space, which relates to the surface, is two dimensional and the tensor space is three dimensional.

Let us consider an isomorphism between a  $m$  dimensional s-space and a sub-space of vector space of dimension  $\mathbb{T}_1^m \subset \mathbb{T}_1^n$ ,  $m < n$ . Then it exists a v-vector  $\underline{\underline{S}} \in {}^m_1\mathbb{S}_1^n$  ensuring a bijective projection from one space into another:

$$\underline{\underline{S}} : {}^m_1\mathbb{S}_1^n \rightarrow \mathbb{T}_1^n$$

and its inverse  $\underline{\underline{S}}^{-1} \in {}^m_1\mathbb{S}_1^n$  ensuring the inverse projection:

$$\underline{\underline{S}}^{-1} : \mathbb{T}_1^n \rightarrow {}^m_1\mathbb{S}_1^n$$

such that:

$$\underline{\underline{\underline{S}}} \cdot \underline{\underline{\underline{S}}}^{-1} = \underline{\underline{\underline{I}}}$$

Let us consider a sub-basis  $\underline{e}_{i'} \in \mathbb{T}_1^m \subset \mathbb{T}_1^n$  allowing us to express any vector  $\underline{y}'$  in this subspace:

$$\underline{y}' = y^{i'} \underline{e}_{i'}$$

then there exists an isomorphism with a bijection function  $\underline{\underline{\underline{S}}}$  such that for any  $\underline{x} \in {}^m_1\mathbb{S}_0^n$

$$\underline{y}' = \underline{\underline{\underline{S}}} \circ \underline{x}$$

If such a bijection exists, then the inverse also exists, such that:

$$\underline{x} = \underline{\underline{\underline{S}}}^{-1} \cdot \underline{y}'$$

and:

$$\underline{\underline{\underline{S}}} \circ \underline{\underline{\underline{S}}}^{-1} = \underline{\underline{\underline{I}}}$$

but obviously:

$$\underline{\underline{\underline{S}}} \cdot \underline{\underline{\underline{S}}}^{-1} \neq \underline{\underline{\underline{I}}}$$

### A.6.7 Tensor product of v-vectors

The tensor product of two v-vectors or double tensor product (by analogy with double scalar product) implies that the tensor product is evaluated both between v-scalars and vectors. This product associates a t-tensor for each ordered pair of v-vectors:

$$\{\otimes\} : {}^m_1\mathbb{S}_1^n \times {}^m_1\mathbb{S}_1^n \rightarrow {}^m_2\mathbb{S}_2^n$$

Let  $\underline{a}, \underline{b} \in {}^m_1\mathbb{S}_1^n$  and  $\underline{c} \in {}^m_2\mathbb{S}_2^n$  then:

$$\underline{a} \otimes \underline{b} = \underline{c}$$

This operation can be introduced as:

$$\underline{a} \otimes \underline{b} = \underline{a}^{(i)} \underline{e}_i \otimes \underline{b}^{(j)} \underline{e}_j = \underline{a}^{(i)} \otimes \underline{b}^{(j)} \underline{e}_i \otimes \underline{e}_j = \underline{c}$$

To demonstrate that the resulting t-tensor is invariant in s-space, it is sufficient and necessary to show that:

$$\underline{a}^{(i)} \otimes \underline{b}^{(j)} \underline{e}_i \otimes \underline{e}_j = \underline{a}_{(i)} \otimes \underline{b}^{(j)} \underline{e}^i \otimes \underline{e}_j = \underline{a}^{(i)} \otimes \underline{b}_{(j)} \underline{e}_i \otimes \underline{e}^j = \underline{a}_{(i)} \otimes \underline{b}_{(j)} \underline{e}^i \otimes \underline{e}^j$$

which according to (A.6) rewrites:

$$\begin{aligned} \underline{a}^{(i)} \otimes \underline{b}^{(j)} \underline{e}_i \otimes \underline{e}_j &= [S_{ik}] \underline{a}^{(k)} \otimes \underline{b}^{(j)} \underline{e}_l \otimes \underline{e}_j [S^{il}] = \\ &= [S_{jk}] \underline{a}^{(i)} \otimes \underline{b}^{(k)} \underline{e}_i \otimes \underline{e}_l [S^{jl}] = [S_{ik}] [S_{jl}] \underline{a}_{(k)} \otimes \underline{b}_{(l)} \underline{e}_p \otimes \underline{e}_q [S^{ip}] [S^{jq}] \end{aligned} \quad (\text{A.12})$$

since  $[S_{ik}] [S^{kj}] = \delta_i^j$ , the invariance of the double tensor product is demonstrated.

## A.7 Reduced form of s-structures

In Chapter 2 we make use of the reduced form of s-structures. The reduced form of s-structure presents a matrix filled with scalars, vectors, tensors. To define operations between matrices the transposition operation has to be introduced. If the matrix is filled with tensors of second and higher order two transposition operations should be distinguished. Since the matrix is a two index construction, the use of the reduced form of s-structures is *limited* to second order structures. So formally this form is of limited use, see Remark 2.5. Here the main rules and forms of reduced s-structures are given.

For two v-vectors  $\underline{\underline{a}}, \underline{\underline{b}} \in {}^m\mathcal{S}_1^n$ , two matrices of vectors  $[\underline{\underline{a}}]_{\{1 \times m\}}, [\underline{\underline{b}}]_{\{1 \times m\}}$  are defined then two scalar products are defined as (here and further the  $\sim$  symbol means the equivalence between different representations

$$\begin{aligned} \boxed{\alpha = \underline{\underline{a}} \circ \underline{\underline{b}}} & \sim \\ \alpha = [\underline{\underline{a}}]^T \cdot [\underline{\underline{b}}] & \sim \alpha = \underline{\underline{a}}^{(i)} \cdot \underline{\underline{b}}^{(i)}, \quad i = 1, m \\ \boxed{\underline{\underline{A}} = \underline{\underline{a}} \boxtimes \underline{\underline{b}}} & \sim \\ [\underline{\underline{A}}] = [\underline{\underline{a}}] \cdot [\underline{\underline{b}}]^T & \sim A^{ij} = \underline{\underline{a}}^i \cdot \underline{\underline{b}}^j, \quad i, j = 1, m \end{aligned}$$

Note that an orthonormal basis in s-space is required. Remark also that

$$[\underline{\underline{a}}] \cdot [\underline{\underline{b}}]^T = \left[ [\underline{\underline{b}}] \cdot [\underline{\underline{a}}]^T \right]^T$$

For example, t-scalar, v-tensor and t-tensor,

$$\underline{\underline{D}} \in {}^m\mathcal{S}_0^n, \quad \underline{\underline{C}} \in {}^m\mathcal{S}_2^n, \quad \underline{\underline{E}} \in {}^m\mathcal{S}_2^n$$

for which the following matrices are defined

$$[\underline{\underline{D}}]_{\{m \times m\}}, [\underline{\underline{C}}]_{\{1 \times m\}}, [\underline{\underline{E}}]_{\{m \times m\}},$$

then the products with a v-vector can be presented in a reduced form as follows

$$\begin{aligned} \boxed{\underline{\underline{d}} = \underline{\underline{a}} \circ \underline{\underline{D}}} & \sim \\ [\underline{\underline{d}}] = [\underline{\underline{a}}]^T [\underline{\underline{D}}] & \sim \underline{\underline{d}}^j = \underline{\underline{a}}^i D^{ij}, \quad i, j = 1, m \\ [\underline{\underline{d}}] = [\underline{\underline{D}}]^T [\underline{\underline{a}}] & \sim \underline{\underline{d}}^j = D^{ji} \underline{\underline{a}}^i, \quad i, j = 1, m \\ \boxed{\underline{\underline{d}} = \underline{\underline{a}} \circ \underline{\underline{C}}} & \sim \\ \underline{\underline{d}} = [\underline{\underline{a}}]^T \cdot [\underline{\underline{C}}] & \sim \underline{\underline{d}} = \underline{\underline{a}}^i \cdot \underline{\underline{C}}^i, \quad i = 1, m \end{aligned}$$

$$\underline{\underline{d}} = [\underline{\underline{C}}^T] \cdot [\underline{\underline{a}}] \quad \sim \quad \underline{\underline{d}} = \underline{\underline{C}}^{iT} \cdot \underline{\underline{a}}^i, \quad i = 1, m$$

$$\boxed{\underline{\underline{d}} = \underline{\underline{a}} \circ \underline{\underline{E}}} \quad \sim$$

$$[\underline{\underline{d}}] = [\underline{\underline{a}}]^T \cdot [\underline{\underline{E}}] \quad \sim \quad \underline{\underline{d}}^j = \underline{\underline{a}}^i \cdot \underline{\underline{E}}^{ij}, \quad i, j = 1, m$$

$$[\underline{\underline{d}}] = [\underline{\underline{E}}^T]^T \cdot [\underline{\underline{a}}] \quad \sim \quad \underline{\underline{d}}^j = \underline{\underline{E}}^{Tji} \cdot \underline{\underline{a}}^i, \quad i, j = 1, m$$

For  $\underline{\underline{A}}, \underline{\underline{B}}, \underline{\underline{C}} \in {}^m_2\mathcal{S}_0^n$  the s-dot product in the reduced representation is defined as follows

$$\boxed{\underline{\underline{A}} = \underline{\underline{B}} \circ \underline{\underline{C}}} \quad \sim$$

$$[\underline{\underline{A}}] = [\underline{\underline{B}}][\underline{\underline{C}}] \quad \sim \quad A^{ij} = B^{ik}C^{kj}$$

For  $\underline{\underline{D}}, \underline{\underline{E}} \in {}^m_2\mathcal{S}_1^n$ , the double dot product is defined as

$$\boxed{\underline{\underline{A}} = \underline{\underline{B}} \circ \underline{\underline{C}}} \quad \sim$$

$$[\underline{\underline{A}}] = [\underline{\underline{B}}] \cdot [\underline{\underline{C}}] \quad \sim \quad A^{ij} = \underline{\underline{B}}^{ik} \cdot \underline{\underline{C}}^{kj}$$

And so on for other combinations, for example,

$$\underline{\underline{A}} \in {}^m_2\mathcal{S}_1^n, \underline{\underline{B}} \in {}^m_2\mathcal{S}_2^n, \quad \underline{\underline{A}} \circ \underline{\underline{B}} = \underline{\underline{C}} \in {}^m_2\mathcal{S}_1^n$$

Formally, the operations like the following are forbidden:

$$[\underline{\underline{a}}] \cdot [\underline{\underline{b}}], \quad [\underline{\underline{a}}] \cdot [\underline{\underline{B}}], \quad [\underline{\underline{B}}] \cdot [\underline{\underline{a}}]^T$$

but contrary to this restriction, we use such operations carrying in mind that there exist higher order s-structures behind these matrix notations, see Remark 2.5.

# Bibliography

- [ABA 07] *ABAQUS release notes, Abaqus release 6.7*, 2007.
- [Alart 88] P. Alart. *Multiplicateurs "augmentés" et méthode de Newton généralisée pour contact avec frottement*. Technical report, Document LMA-DME-EPFL, Lausanne, 1988.
- [Alart 91] P. Alart & A. Curnier. *A mixed formulation for frictional contact problem prone to Newton like solution methods*. Computer Method in Applied Mechanics and Engineering, vol. 92, pages 353–375, 1991.
- [Alart 97] P. Alart. *Méthode de Newton généralisée en mécanique du contact*. Journal de Mathématiques Pures et Appliquées, vol. 76, pages 83–108, 1997.
- [Alart 04] P. Alart, M. Barboteu & J. Gril. *A numerical modelling of non linear 2D-frictional multicontact problems: application to post-buckling in cellular media*. Computational Mechanics, vol. 34, pages 298–309, 2004.
- [Alexandrov 83] V.M. Alexandrov & S.M. Mhitaryan. *Contact problems for bodies with thin coatings and interlayers*. Nauka, Moscow, 1983.
- [Anand 93] L. Anand. *A constitutive model for interface friction*. Computational Mechanics, vol. 12, pages 197–213, 1993.
- [ANS 05] *ANSYS Contact technology guide, ANSYS release 10*, August 2005.
- [Arrow 58a] K. Arrow, L. Hurwicz & H. Uzawa. *Studies in nonlinear programming*. Stanford University Press, 1958.
- [Arrow 58b] K.J. Arrow & R.M. Solow. *Gradient methods for constrained maxima, with weakened assumptions*. In K.J. Arrow, L. Hurwicz & H. Uzawa, ed., *Studies in Linear and Nonlinear Programming*, pages 166–176. Stanford University Press, 1958.

## BIBLIOGRAPHY

---

- [Babuška 95] I. Babuška, F. Ihlenburg, E.T. Paik & S.A. Sauter. *A Generalized Finite Element Method for solving the Helmholtz equation in two dimensions with minimal pollution*. Computer Methods in Applied and Mechanical Engineering, vol. 128, pages 325–359, 1995.
- [Barboteu 99] M. Barboteu. *Contact, frottement et techniques de calcul parallèle*. PhD thesis, Université de Montpellier, France, 1999. Sc. advisor P. Alart.
- [Barboteu 02] M. Barboteu, P. Alart & S. Pagano. *Modélisation de problèmes non linéaires de grande taille : grandes déformations et autocontact dans un milieu cellulaire*. Revue européenne des éléments finis, vol. 11, pages 447–461, 2002.
- [Bathe 96] K.-J. Bathe. *Finite element procedures*. Prentice Hall, 1996.
- [Bay 87] N. Bay, T. Wanheim & B. Avitzur. *Models for friction in metal forming*. Manufacturing Technology review, vol. 2, pages 372–378, 1987.
- [Becker 03] R. Becker, P. Hansbo & R. Stenberg. *A finite element method for domain decomposition with non-matching grids*. Mathematical Modelling and Numerical Analysis, vol. 37, pages 209–225, 2003.
- [Belgacem 98] F.B. Belgacem, P. Hild & P. Laborde. *The mortar finite element method for contact problems*. Mathematical and Computer Modelling, vol. 28, pages 263–271, 1998.
- [Belytschko 08] T. Belytschko, W. Kam Liu & B. Moran. *Nonlinear finite elements for continua and structures*. John Wiley & Sons, 2008.
- [Benson 90] D.J. Benson & J.O. Hallquist. *A single surface contact algorithm for the post-buckling analysis of shell structures*. Computer Methods in Applied Mechanics and Engineering, vol. 78, pages 141–163, 1990.
- [Benson 07] D.J. Benson. *The History of LS-DYNA*. University Of California, San Diego, 2007.
- [Bernardi 90] C. Bernardi, N. Debit & Y. Maday. *Coupling finite element and spectral methods: First results*. Mathematics of Computation, vol. 54, pages 21–39, 1990.
- [Bertsekas 81] D.P. Bertsekas. *Augmented Lagrangian and differentiable exact penalty methods*. Technical report LIDS-P; 1113, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 1981.

- [Bertsekas 84] D.P. Bertsekas. Constrained optimization and Lagrange multiplier methods. Academic Press, 1984.
- [Bertsekas 03] D.P. Bertsekas, A. Nedic & A.E. Ozdaglar. Convex analysis and optimization. Athena Scientific, 2003.
- [Besson 97] J. Besson & R. Foerch. *Large scale object-oriented finite element code design*. Computer Methods in Applied Mechanics and Engineering, vol. 142, pages 165 – 187, 1997.
- [Bhashyam 02] G.R. Bhashyam. *ANSYS Mechanical - A powerful nonlinear simulation tool*. Technical report, 2002. <http://ansys.net/collection/824>.
- [Bonnans 06] J. F. Bonnans, J.Ch. Gilbert, C. Lemaréchal & C.A. Sagastizábal. Numerical optimization, theoretical and numerical aspect. Springer, second edition, 2006.
- [Bowden 50] F.P. Bowden & D. Tabor. The friction and lubrication of solids. Oxford University Press, 1950.
- [Brown 00] K. Brown, S. Attaway, S. Plimpton & B. Hendrickson. *Parallel strategies for crash and impact simulations*. Computer Methods in Applied Mechanics and Engineering, vol. 184, pages 375–390, 2000.
- [Bruneel 02] H.C.J. Bruneel & I. De Rycke. *QuickTrace: a fast algorithm to detect contact*. International Journal for Numerical Methods in Engineering, vol. 54, pages 299–316, 2002.
- [Bussetta 09] P. Bussetta, D. Marceau & J.-P. Ponthot. *Résolution du problème de contact mécanique frottant : méthode du lagrangien augmenté adapté*. In Proceedings of "9e Colloque National en Calcul des Structures", vol. 2, pages 623–628, Giens, France, 25-29 May 2009.
- [Campos 82] L.T. Campos, J.T. Oden & N. Kikuchi. *A numerical analysis of the class of contact problems with friction in elastostatics*. Computer Methods in Applied and Mechanical Engineering, vol. 34, pages 821–845, 1982.
- [Chamoret 04] D. Chamoret, P. Saillard, A. Rassineux & J.-M. Bergheau. *New smoothing procedures in contact mechanics*. Journal of Computational and Applied Mathematics, vol. 168, pages 107–116, 2004.
- [Chand 76] R. Chand, E.J. Haug & K. Rim. *Analysis of unbonded contact problems by means of quadratic programming*. Journal of Optimization Theory and Applications, vol. 20, pages 171–189, 1976.

## BIBLIOGRAPHY

---

- [Chen 07] Shaohua Chen & Huajian Gao. *Bio-inspired mechanics of reversible adhesion: Orientation-dependent adhesion strength for non-slipping adhesive contact with transversely isotropic elastic materials*. Journal of the Mechanics and Physics of Solids, vol. 55, pages 1001 – 1015, 2007.
- [Cocu 84] M. Cocu. *Existence of solutions of Signorini problems with friction*. International Journal of Engineering Science, vol. 22, pages 567–575, 1984.
- [COM 10] COMSOL Multiphysics 4.1 Release notes, 2010.
- [Courtney-Pratt 57] J.S. Courtney-Pratt & E. Eisner. *The effect of a tangential force on the contact of metallic bodies*. Proceedings of The Royal Society of London A, vol. 238, pages 529–550, 1957.
- [Crisfield 00a] M.A. Crisfield. *Non-linear finite element analysis of solids and structures, volume 1: Essentials*. John Wiley & Sons, 2000.
- [Crisfield 00b] M.A. Crisfield. *Non-linear finite element analysis of solids and structures, volume 2: Advanced topics*. John Wiley & Sons, 2000.
- [Crisfield 00c] M.A. Crisfield. *Re-visiting the contact patch test*. International Journal for Numerical Methods in Engineering, vol. 48, pages 435–449, 2000.
- [Curnier 84] A. Curnier. *A theory of friction*. International Journal of Solids and Structures, vol. 20, pages 637–647, 1984.
- [Curnier 88] A. Curnier & P. Alart. *A generalized Newton method for contact problem with friction*. Journal de mécanique théorique et appliquée, Special issue: Numerical methods in Mechanics of Contact Involving Friction, pages 67–82, 1988.
- [Curnier 95] A. Curnier, Q.-Ch. He & A. Klarbring. *Continuum Mechanics Modeling of Large Deformation Contact with Friction*. In M. Raous, M. Jean & J.J. Moreau, ed., Contact Mechanics, pages 145–158. Plenum, 1995.
- [de Saxce 98] G. de Saxce & L. Boussshine. *Limit analysis theorems for implicit standard materials: Application to the unilateral contact with dry friction and the non-associated flow rules in soils and rocks*. International Journal of Mechanical Sciences, vol. 40, pages 387 – 398, 1998.
- [Dick 06a] T. Dick & G. Cailletaud. *Analytic and FE based estimations of the coefficient of friction of composite surfaces*. Wear, vol. 260, pages 1305–1316, 2006.



- [Dick 06b] T. Dick & G. Cailletaud. *Fretting modelling with a crystal plasticity model of Ti6Al4V*. Computational Materials Science, vol. 38, pages 113–125, 2006.
- [Dick 08] T. Dick, S. Basseville & G. Cailletaud. *Fatigue modelling in fretting contact with a crystal plasticity model*. Computational Materials Science, vol. 43, pages 36–42, 2008.
- [Duvaut 71] G. Duvaut & J.L. Lions. *Elasticité avec frottement*. Journal de Mécanique, vol. 10, pages 409–420, 1971.
- [Duvaut 76] G. Duvaut & J.L. Lions. *Inequalities in mechanics and physics*. Springer-Verlag, 1976.
- [Farhat 94] C. Farhat & F.-X. Roux. *Implicit parallel processing in structural mechanics*. Computational Mechanics Advances, vol. 2, pages 1–24, 1994.
- [Fichera 63] G. Fichera. *Sul problema elastostatico di Signorini con ambigue condizioni al contorno*. Atti della Accademia Nazionale dei Lincei, Serie Ottava, Rendiconti, Classe di Scienze Fisiche, Matematiche e Naturali, vol. 34, pages 138–142, 1963.
- [Fichera 64] G. Fichera. *Problemi elastostatici con vincoli unilaterali: Il problema di Signorini con ambigue condizioni al contorno*. Atti della Accademia Nazionale dei Lincei, Mem., Cl. Sci. Fis. Mat. Nat., Sez. I, VIII. Ser., vol. 7, pages 91–140, 1964.
- [Fichera 72] G. Fichera. *Boundary Value Problems of Elasticity with Unilateral Constraints*. In C. Truesdell, ed., Mechanics of Solids, vol. 2, pages 391–424. Springer-Verlag, 1972.
- [Fischer 05] K.A. Fischer & P. Wriggers. *Frictionless 2d contact formulations for finite deformations based on the mortar method*. Computational Mechanics, vol. 36, pages 226–244, 2005.
- [Fischer 06] K.A. Fischer & P. Wriggers. *Mortar based frictional contact formulation for higher order interpolations using the moving friction cone*. Computer Methods in Applied Mechanics and Engineering, vol. 195, pages 5020 – 5036, 2006. John H. Argyris Memorial Issue. Part I.
- [Fletcher 70] R. Fletcher. *A class of methods in nonlinear programming with termination and convergence properties*. In J. Abadie, ed., Integer and Nonlinear Programming. North-Holland Publishing Company, 1970.

## BIBLIOGRAPHY

---

- [Fletcher 77] R. Fletcher & T. L. Freeman. *A modified Newton method for minimization*. Journal of Optimization Theory and Applications, vol. 23, pages 357–372, 1977.
- [Fortin 76] M. Fortin. *Minimization of some non-differentiable functionals by the augmented Lagrangian method of Hestenes and Powell*. Applied Mathematics and Optimization, vol. 2, pages 236–250, 1976.
- [Francavilla 75] A. Francavilla & O.C. Zienkiewicz. *A note on numerical computation of elastic contact problems*. International Journal for Numerical Methods in Engineering, vol. 9, pages 913–924, 1975.
- [Fritz 04] A. Fritz, S. Hübner & B.I. Wohlmuth. *A comparison of mortar and Nitsche techniques for linear elasticity*. CALCOLO, vol. 41, pages 115–137, 2004.
- [Fujun 00] Wang Fujun, Cheng Jiangang & Yao Zhenhan. *A contact searching algorithm for contact-impactor problems*. Acta Mechanica Sinica (English series), vol. 16, pages 374–382, 2000.
- [Fujun 01] Wang Fujun, Cheng Jiangang & Yao Zhenhan. *FFS contact searching algorithm for dynamic finite element analysis*. International Journal for Numerical Methods in Engineering, vol. 52, pages 655–672, 2001.
- [Galin 53] L.A. Galin. *Contact problems in elasticity*. GosTechTeorIzdat, Moscow, 1953.
- [Galin 76] L.A. Galin. *Development of the contact theory in USSR*. Nauka, Moscow, 1976.
- [Giannakopoulos 89] A.E. Giannakopoulos. *The return mapping method for the integration of friction constitutive relations*. Computers & Structures, vol. 32, pages 157 – 167, 1989.
- [Gibbs 84] J.W. Gibbs. *Elements of vector analysis*. New Haven, 1884.
- [Gibbs 60] J.W. Gibbs & E.B. Wilson. *Vector analysis: A textbook for the use of students of mathematics and physics. Founded Upon the Lectures of J. Willard Gibbs*, Dover, 1960.
- [Glowinski 89] R. Glowinski & P. Le Tallec. *Augmented lagrangians and operator-splitting methods in nonlinear mechanics*. SIAM, Philadelphia, 1989.
- [Goryacheva 98] I.G. Goryacheva. *Contact mechanics in tribology*. Springer, 1998.

- [Goryacheva 01] I.G. Goryacheva. *Mechanics of frictional interaction*. Nauka, Moscow, 2001.
- [Gosselet 06] P. Gosselet & Rey C. *Non-overlapping domain decomposition methods in structural mechanics*. Archives of Computational Methods in Engineering, vol. 13, pages 515–572, 2006.
- [Griebel 07] M. Griebel, S. Knapek & G. Zumbusch. *Numerical simulation in molecular dynamics*. Springer-Verlag, 2007.
- [Hansson 90] E. Hansson & A. Klarbring. *Rigid contact modelled by CAD surface*. Engineering Computations, vol. 7, pages 344–348, 1990.
- [Hartmann 09] S. Hartmann, J. Oliver, J.C. Cante, R. Weyler & J.A. Hernández. *A contact domain method for large deformation frictional contact problems. Part 2: Numerical aspects*. Computer Methods in Applied Mechanics and Engineering, vol. 198, pages 2607–2631, 2009.
- [Heegaard 93] J.-H. Heegaard & A. Curnier. *An augmented Lagrangian method for discrete large-slip contact problems*. International Journal for Numerical Methods in Engineering, vol. 36, pages 569–593, 1993.
- [Heegaard 96] J.-H. Heegaard & A. Curnier. *Geometric properties of 2D and 3D unilateral large slip contact operators*. Computer Methods in Applied Mechanics and Engineering, vol. 131, pages 263–286, 1996.
- [Heege 96] A. Heege & P. Alart. *A frictional contact element for strongly curved contact problems*. International Journal for Numerical Methods in Engineering, vol. 39, pages 165–184, 1996.
- [Hertz 82] H. Hertz. *Ueber die Berührung fester elastische Körper*. Journal für die reine und angewandte Mathematik, vol. 92, pages 156–171, 1882.
- [Hestenes 69] M.R. Hestenes. *Survey paper: Multipliers and gradient methods*. Journal of Optimization Theory and Applications, vol. 4, pages 303–320, 1969.
- [Heyliger 89] P.R. Heyliger & R.D. Kriz. *Stress intensity factors by enriched mixed finite elements*. International Journal for Numerical Methods in Engineering, vol. 28, pages 1461–1473, 1989.

## BIBLIOGRAPHY

---

- [Hjiaj 04a] M. Hjiaj, Z.-Q. Feng, G. de Saxcé & Z. Mroz. *On the modelling of complex anisotropic frictional contact laws*. International Journal of Engineering Science, vol. 42, pages 1013 – 1034, 2004.
- [Hjiaj 04b] M. Hjiaj, Z-Q Feng, G. de Saxcé & Z. Mróz. *Three-dimensional finite element computations for frictional contact problems with non-associated sliding rule*. International Journal for Numerical Methods in Engineering, vol. 60, pages 2045–2076, 2004.
- [Hughes 77] T.R.J. Hughes, R.L. Taylor & W. Kanoknukulchai. *A finite element method for large displacement contact and impact problems*. In K.J. Bathe, J.T. Oden, W. Wunderlich & E.L. Wilson, ed., Formulations and computational algorithms in FE analysis, pages 468–495. MIT Press, 1977.
- [Jarusek 99] J. Jarusek & C. Eck. *Dynamic contact problems with small Coulomb friction for viscoelastic bodies. existence of solutions*. Mathematical models and Methods in Applied Sciences, vol. 9, pages 11–34, 1999.
- [Jean 95] M. Jean. *Frictional contact in collections of rigid or deformable bodies: numerical simulation of geomaterial motions*. Studies in Applied Mechanics, vol. 42, pages 463–486, 1995.
- [Johnson 94] K.L. Johnson. *Contact mechanics*. Cambridge University press, 1994.
- [Kagan 47] V.F. Kagan. *Basics of surface theory in tensor form*, vol. 1 (in russian). M-L: OGIZ, 1947.
- [Kantorovich 48] L.V. Kantorovich. *Functional analysis and applied mathematics*. Uspehi Matematicheskikh Nauk, vol. 3, pages 89–185, 1948.
- [Karypis 95] G. Karypis & V. Kumar. *A fast and high quality multilevel scheme for partitioning irregular graphs*. In International Conference on Parallel Processing, pages 113–122, 1995.
- [Karypis 98] G. Karypis & V. Kumar. *A fast and high quality multilevel scheme for partitioning irregular graphs*. SIAM Journal on Scientific Computing, vol. 20, pages 359–392, 1998.
- [Kikuchi 88] N. Kikuchi & J.T. Oden. *Contact problems in elasticity: a study of variational inequalities and finite element methods*. SIAM, Philadelphia, 1988.

- [Klang 79] M. Klang. *On interior contact under friction between cylindrical elastic bodies*. PhD thesis, Linköping University, 1979.
- [Klarbring 86] A. Klarbring. *A mathematical programming approach to three-dimensional contact problems with friction*. Computer Methods in Applied Mechanics and Engineering, vol. 58, pages 175 – 200, 1986.
- [Klarbring 88] A. Klarbring. *On discrete and discretized nonlinear elastic structures in unilateral contact (stability, uniqueness and variational principles)*. International Journal of Solids and Structures, vol. 24, pages 459–479, 1988.
- [Klarbring 90] A. Klarbring. *Examples of non-uniqueness and non-existence of solutions to quasistatic contact problems with friction*. Ingenieur-Archiv (Archive of Applied Mechanics), vol. 60, pages 529–541, 1990.
- [Konyukhov 05] A. Konyukhov & K. Schweizerhof. *Covariant description for frictional contact problems*. Computational Mechanics, vol. 35, pages 190–213, 2005.
- [Konyukhov 06a] A. Konyukhov & K. Schweizerhof. *On a geometrical approach in contact mechanics*. In F. Pfeiffer & P. Wriggers, ed., Analysis and Simulation of Contact Problems, Lecture Notes in Applied and Computational Mechanics, vol. 27, pages 23–30. Springer, 2006.
- [Konyukhov 06b] A. Konyukhov & K. Schweizerhof. *A Special Focus on 2D Formulations for Contact Problems Using a Covariant Description*. International Journal for Numerical Methods in Engineering, vol. 66, pages 1432–1465, 2006.
- [Konyukhov 07a] A. Konyukhov & K. Schweizerhof. *On a Continuous Transfer of History Variables for Frictional Contact Problems Based on Interpretations of Covariant Derivatives as a Parallel Translation*. In Peter Eberhard, ed., IUTAM Symposium on Multiscale Problems in Multibody System Contacts, vol. 1, pages 95–101. Springer, 2007.
- [Konyukhov 07b] A. Konyukhov & K. Schweizerhof. *Symmetrization of Various Friction Models Based on an Augmented Lagrangian Approach*. In P. Wriggers & U. Nackenhorst, ed., IUTAM Symposium on Computational Methods in Contact Mechanics, vol. 3, pages 97–111. Springer, 2007.
- [Konyukhov 08] A. Konyukhov & K. Schweizerhof. *On the solvability of closest point projection procedures in contact analysis: Analysis and solution strategy for surfaces of arbitrary*

- geometry*. Computer Methods in Applied Mechanics and Engineering, vol. 197, pages 3045–3056, 2008.
- [Konyukhov 09] A. Konyukhov & K. Schweizerhof. *Incorporation of contact for high-order finite elements in covariant form*. Computer Methods in Applied Mechanics and Engineering, vol. 198, pages 1213–1223, 2009.
- [Krstulović-Opara 02] L. Krstulović-Opara, P. Wriggers & J. Korelc. *A C1-continuous formulation for 3D finite deformation frictional contact*. Computational Mechanics, vol. 29, pages 27–42, 2002.
- [Laursen 92] T.A. Laursen. *Formulation and Treatment of Frictional Contact Problems Using Finite Elements*. PhD thesis, Department of Mechanical Engineering, Stanford University, 1992. Sc. advisor J.C. Simo.
- [Laursen 93] T.A. Laursen & J.C. Simo. *A Continuum-Based Finite Element Formulation for the Implicit Solution of Multibody, Large Deformation Frictional Contact Problems*. International Journal for Numerical Methods in Engineering, vol. 36, pages 3451–3485, 1993.
- [Laursen 94] T.A. Laursen. *The Convected Description in Large Deformation Frictional Contact Problems*. International Journal of Solids and Structures, vol. 31, pages 669–681, 1994.
- [Laursen 02] T.A. Laursen. *Computational contact and impact mechanics: fundamentals of modeling interfacial phenomena in nonlinear finite element analysis*. Springer-Verlag, 2002.
- [Lounesto 01] P. Lounesto. *Clifford algebras and spinors*. Cambridge University Press, 2001.
- [Luenberger 03] D.G. Luenberger. *Linear and nonlinear programming*. Second Edition, Kluwer Academic Publishers, 2003.
- [Lurie 70] A.I. Lurie. *Theory of elasticity*. Nauka, Moscow, 1970.
- [Martins 87] J.A.C. Martins & J.T. Oden. *Existence and uniqueness results for dynamic contact problems with nonlinear normal and friction interface laws*. Nonlinear Analysis, Theory, Methods and Applications, vol. 11, pages 407–428, 1987.
- [Martins 94] J.A.C. Martins, M.D.P. Monteiro Marques & F. Gastaldi. *On an example of non-existence of solutions to a quasistatic frictional contact problem*. European Journal of Mechanics A/Solids, vol. 13, pages 113–133, 1994.

- [McDevitt 00] T.W. McDevitt & T.A. Laursen. *A mortar-finite element formulation for frictional contact problems*. International Journal for Numerical Methods in Engineering, vol. 48, pages 1525–1547, 2000.
- [Melenk 96] J.M. Melenk & I. Babuška. *The partition of unity finite element method: Basic theory and applications*. Computer Methods in Applied Mechanics and Engineering, vol. 139, pages 289 – 314, 1996.
- [Michalowski 78] R. Michalowski & Z. Mróz. *Associated and non-associated sliding rules in contact friction problems*. Archives of Mechanics, vol. 30, pages 259–276, 1978.
- [Middleton 85] J. Middleton & G.N. Pande, ed. *Penalty and augmented lagrangian formulation for contact problems*. Elsevier, 1985.
- [Mijar 00] A.R. Mijar & J.S. Arora. *Review of formulations for elastostatic frictional contact problems*. Structural and Multidisciplinary Optimization, vol. 20, pages 167–189, 2000.
- [Mijar 04a] A.R. Mijar & J.S. Arora. *An augmented Lagrangian optimization method for contact analysis problems, 1: formulation and algorithm*. Structural and Multidisciplinary Optimization, vol. 28, pages 99–112, 2004.
- [Mijar 04b] A.R. Mijar & J.S. Arora. *An augmented Lagrangian optimization method for contact analysis problems, 2: numerical evaluation*. Structural and Multidisciplinary Optimization, vol. 28, pages 113–126, 2004.
- [Moës 99] N. Moës, J. Dolbow & T. Belytschko. *A finite element method for crack growth without remeshing*. International Journal for Numerical Methods in Engineering, vol. 46, pages 131–150, 1999.
- [Mróz 94] Z. Mróz & S. Stupkiewicz. *An anisotropic friction and wear model*. International Journal of Solids and Structures, vol. 31, pages 1113 – 1131, 1994.
- [Murty 88] K.G. Murty. *Linear complementarity, linear and nonlinear programming*. Helderman-Verlag, 1988. Internet version: [http://ioe.engin.umich.edu/people/fac/books/murty/linear\\_complementarity\\_webbook/](http://ioe.engin.umich.edu/people/fac/books/murty/linear_complementarity_webbook/).
- [Muskhelishvili 66] N.I. Muskhelishvili. *Some basic problems in the mathematical theory of elasticity*. Nauka, Moscow, 5th edition, 1966.

## BIBLIOGRAPHY

---

- [Oatis 07] D. Oatis. *Getting in touch with ANSYS contact*. The Focus, Journal of Phoenix Analysis and Design Technologies, pages 1–3, 2007.
- [Oldenburg 94] M. Oldenburg & L. Nilsson. *The position code algorithm for contact searching*. International Journal for Numerical Methods in Engineering, vol. 37, pages 359–386, 1994.
- [Oliver 09] J. Oliver, S. Hartmann, J.C. Cante, R. Weyler & J.A. Hernández. *A contact domain method for large deformation frictional contact problems. Part 1: Theoretical basis*. Computer Methods in Applied Mechanics and Engineering, vol. 198, pages 2591 – 2606, 2009.
- [Oliver 10] J. Oliver, S. Hartmann, J.C. Cante, R. Weyler & J.A. Hernández. *On a new 3D contact domain method for large deformation contact problems*. In Plenary lecture at IV European Conference on Computational Mechanics, Palais des Congrès, Paris, France, May 16-21 2010. <http://www.eccm2010.org/cv/pdf/oliver.pdf>.
- [Padmanabhan 01] V. Padmanabhan & T.A. Laursen. *A framework for development of surface smoothing procedures in large deformation frictional contact analysis*. Finite Elements in Analysis and Design, vol. 37, pages 173 – 198, 2001.
- [Panagiotopoulos 85] P.D. Panagiotopoulos. *Inequality problems in mechanics, convex and nonconvex energy functions*. Birkhäuser, 1985.
- [Parisch 89] H. Parisch. *A consistent tangent stiffness matrix for three-dimensional non-linear contact analysis*. International Journal for Numerical Methods in Engineering, vol. 28, pages 1803–1812, 1989.
- [Pietrzak 97] G. Pietrzak. *Continuum mechanics modelling and augmented Lagrangian formulation of large deformation frictional contact problems*. PhD thesis, École Polytechnique Fédérale de Lausanne, 1997. Sc. advisor A. Curnier.
- [Pietrzak 99] G. Pietrzak & A. Curnier. *Large deformation frictional contact mechanics: continuum formulation and augmented Lagrangian treatment*. Computer Methods in Applied Mechanics and Engineering, vol. 177, pages 351–381, 1999.
- [Popov 10] V.L. Popov. *Contact mechanics and friction. physical principles and applications*. Springer-Verlag, 2010.



- [Powell 69] M.J.D. Powell. *A method for nonlinear constraints in minimization problems*. In R. Fletcher, ed., *Optimization*, pages 283–298. Academic Press, London, 1969.
- [Powell 78] M.J.D. Powell. *Algorithms for nonlinear constraints that use Lagrangian functions*. *Mathematical Programming*, vol. 14, pages 224–248, 1978.
- [Puso 02] M.A. Puso & T.A. Laursen. *A 3D Contact Smoothing Method Using Gregory Patches*. *International Journal for Numerical Methods in Engineering*, vol. 54, pages 1161–1194, 2002.
- [Puso 03] M.A. Puso & T.A. Laursen. *Mesh tying on curved interfaces in 3D*. *Engineering Computations*, vol. 20, pages 305–319, 2003.
- [Puso 04] M.A. Puso. *A 3D mortar method for solid mechanics*. *International Journal for Numerical Methods in Engineering*, vol. 59, pages 315–336, 2004.
- [Rabier 86] P.J. Rabier, J.A.C. Martins, J.T. Oden & L. Campos. *Existence and local uniqueness of solutions to contact problems in elasticity with nonlinear friction laws*. *International Journal of Engineering Science*, vol. 24, pages 1755–1768, 1986.
- [Rabinowicz 65] E. Rabinowicz. *Friction and wear of materials*. Wiley, 1965.
- [Rockafellar 70] R.T. Rockafellar. *Convex analysis*. Princeton University Press, 1970.
- [Rockafellar 73a] R.T. Rockafellar. *A dual approach to solving nonlinear programming problems by unconstrained optimization*. *Mathematical Programming*, vol. 5, pages 354–373, 1973.
- [Rockafellar 73b] R.T. Rockafellar. *The multiplier method of Hestenes and Powell applied to convex programming*. *Journal of Optimization Theory and Applications*, vol. 12, pages 555–562, 1973.
- [Rousselier 09] G. Rousselier, F. Barlat & J.W. Yoon. *A novel approach for anisotropic hardening modeling. Part I: Theory and its application to finite element analysis of deep drawing*. *International Journal of Plasticity*, vol. 25, pages 2383 – 2409, 2009.
- [Signorini 33] A. Signorini. *Sopra alcune questioni di elastostatica*. *Atti della Societa Italiana per Il Progresso delle Scienze*, 1933.

## BIBLIOGRAPHY

---

- [Signorini 59] A. Signorini. *Questioni di elasticità non linearizzata e semilinearizzata*. Rendiconti di Mathematica e dell sue Applicazioni, V. Ser., vol. 18, pages 95–139, 1959.
- [Simo 85] J.C. Simo, P. Wriggers & R.L. Taylor. *A perturbed Lagrangian formulation for the finite element solution of contact problems*. Computer Methods in Applied Mechanics and Engineering, vol. 50, pages 163–180, 1985.
- [Simo 92] J.C. Simo & T.A. Laursen. *An augmented Lagrangian treatment of contact problems involving friction*. Computers & Structures, vol. 42, pages 97–116, 1992.
- [Simo 98] J.C. Simo & T.J.R. Hughes. *Computational inelasticity*. Springer, 1998.
- [Spence 75] D.A. Spence. *The Hertz contact problem with finite friction*. Journal of Elasticity, vol. 5, pages 297–319, 1975.
- [Stupkiewicz 01] S. Stupkiewicz. *Extension of the node-to-segment contact element for surface-expansion-dependent contact laws*. International Journal for Numerical Methods in Engineering, vol. 50, pages 739–759, 2001.
- [Suh 81] N.P. Suh & H.C. Sin. *The genesis of friction*. Wear, vol. 69, pages 91–114, 1981.
- [Suh 86] N.P. Suh. *Tribophysics*. Prentice-Hall, 1986.
- [Taylor 91] R.L. Taylor & O. Papadopoulos. *On a patch test for contact problems in two dimensions*. In P. Wriggers & W. Wagner, ed., *Nonlinear Computational Mechanics*, pages 690–702. Springer, 1991.
- [Toselli 05] A. Toselli & O. Widlund. *Domain decomposition methods – Algorithms and theory*. Springer-Verlag, 2005.
- [Vallet 09] C. Vallet, D. Lasseux, P. Sainsot & H. Zahouani. *Real versus synthesized fractal surfaces: Contact mechanics and transport properties*. Tribology International, vol. 42, page 250–259, 2009.
- [Vorovich 01] I.I. Vorovich & V.M. Alexandrov, ed. *Mechanics of contact interactions*. PhysMatLit, Moscow, 2001.
- [Weber 90] G. Weber & L. Anand. *Finite deformation constitutive equations, and a time integration procedure for isotropic, hyperelastic viscoplastic solids*. Computer Methods in Applied Mechanics and Engineering, vol. 79, pages 173–202, 1990.

- [Wierzbicki 71] A.P. Wierzbicki. *A penalty function shifting method in constrained static optimization and its convergence properties*. *Archiwum Automatyki i Telemechaniki*, vol. 16, pages 395–416, 1971.
- [Wilkins 64] M.L. Wilkins. *Calculation of elastic–plastic flow*. In *Methods of computational physics 3*. Academic Press, New York, 1964.
- [Williams 99] J. R. Williams & R. O’Connor. *Discrete element simulation and the contact problem*. *Archives of Computational Methods in Engineering*, vol. 6, pages 279–304, 1999.
- [Wohlmuth 01] B.I. Wohlmuth. *Discretization methods and iterative solvers based on domain decomposition*. Springer, 2001.
- [Wriggers 90] P. Wriggers, T. Vu Van & E. Stein. *Finite element formulation of large deformation impact-contact problems with friction*. *Computers & Structures*, vol. 37, pages 319 – 331, 1990.
- [Wriggers 01] P. Wriggers, L. Krstulović-Opara & J. Korelc. *Smooth C1-interpolations for two-dimensional frictional contact problems*. *International Journal for Numerical Methods in Engineering*, vol. 51, pages 1469–1495, 2001.
- [Wriggers 06] P. Wriggers. *Computational contact mechanics*. second edition, Springer-Verlag, 2006.
- [Wriggers 08] P. Wriggers & G. Zavarise. *A formulation for frictionless contact problems using a weak form introduced by Nitsche*. *Computational Mechanics*, vol. 41, pages 407–420, 2008.
- [Wronski 94] M. Wronski. *Couplage du contact et du frottement avec la mécanique non linéaire des solides en grandes déformations*. PhD thesis, Université de Technologie de Compiègne, France, 1994. Sc. advisor G. Touzot.
- [Yang 05] B. Yang, T.A. Laursen & X. Meng. *Two dimensional mortar contact methods for large deformation frictional sliding*. *International Journal for Numerical Methods in Engineering*, vol. 62, pages 1183–1225, 2005.
- [Yang 08a] B. Yang & T.A. Laursen. *A contact searching algorithm including bounding volume trees applied to finite sliding mortar formulations*. *Computational Mechanics*, vol. 41, pages 189–205, 2008.
- [Yang 08b] B. Yang & T.A. Laursen. *A large deformation mortar formulation of self contact with finite sliding*. *Computer*

- Methods in Applied Mechanics and Engineering, vol. 197, pages 756 – 772, 2008.
- [Yastrebov 09] V.A. Yastrebov, G. Cailletaud & F. Feyel. *Traitement des problèmes de contact en calcul parallèle*. In Proceedings of "9e Colloque National en Calcul des Structures", vol. 2, pages 647–653, Giens, France, 25-29 May 2009.
- [Yastrebov 11a] V.A. Yastrebov, G. Cailletaud & F. Feyel. *A local contact detection technique for very large contact and self-contact problems: sequential and parallel implementations*. In G. Zavarise & P. Wriggers, ed., Trends in Computational Contact Mechanics, vol. 58, pages 227–251. Springer, 2011.
- [Yastrebov 11b] V.A. Yastrebov, J. Durand, H. Proudhon & G. Cailletaud. *Rough surface contact analysis by means of the Finite Element Method and of a new reduced model*. Comptes Rendus Mécanique, vol. 339, pages 473–490, 2011.
- [Zavarise 98] G. Zavarise & P. Wriggers. *A segment-to-segment contact strategy*. Mathematical and Computer Modelling, vol. 28, pages 497 – 515, 1998.
- [Zavarise 09a] G. Zavarise & L. De Lorenzis. *A modified node-to-segment algorithm passing the contact patch test*. International Journal for Numerical Methods in Engineering, vol. 79, pages 379–416, 2009.
- [Zavarise 09b] G. Zavarise & L. De Lorenzis. *The node-to-segment algorithm for 2D frictionless contact: Classical formulation and special cases*. Computer Methods in Applied Mechanics and Engineering, vol. 198, pages 3428–3451, 2009.
- [Zhilin 01] P.A. Zhilin. *Vectors and second-order tensors in three dimensional space (in russian)*. Saint-Petersburg, Nestor, 2001.
- [Zhong 88] W.X. Zhong & S.M. Sun. *A finite element method for elasto-plastic structure and contact problem by parametric quadratic programming*. International Journal for Numerical Methods in Engineering, vol. 26, pages 2723–2738, 1988.
- [Zienkiewicz 00a] O.C. Zienkiewicz & R.L. Taylor. *The finite element method, volume 1: The basis*, fifth edition. Butterworth-Heinemann, 2000.
- [Zienkiewicz 00b] O.C. Zienkiewicz & R.L. Taylor. *The finite element method, volume 2: Solid mechanics*, fifth edition. Butterworth-Heinemann, 2000.



## Mécanique numérique du contact : géométrie, détection et techniques de résolution

**Résumé :** Le but de ce travail était de fournir un cadre cohérent pour le traitement des problèmes de contact en utilisant une discrétisation de type nœud à segment. Trois aspects principaux de la mécanique numérique du contact ont été particulièrement considérés : la description de la géométrie, le problème de détection de contact et les techniques de résolution. Le manuscrit contient cependant une présentation complète de la mécanique du contact et des algorithmes numériques qui lui sont attachés. Un nouveau formalisme mathématique – les s-structures – est employé dans l'ensemble de la thèse. Il fournit un cadre de formulation intrinsèque qui permet d'exprimer de façon compacte un grand nombre de problèmes de mécanique et de physique. La thèse propose plusieurs idées originales et des extensions des techniques classiques, qui ont toutes été mises en œuvre dans le code de calcul par éléments finis ZéBuLoN (Z-set). Plusieurs études de cas, présentées dans la thèse, viennent démontrer les performances et la robustesse des méthodes numériques utilisées pour la détection et la résolution.

**Mots clés :** mécanique numérique du contact, méthode des éléments finis, technique de détection, parallélisation, géométrie précise du contact, méthode du Lagrangien augmenté, méthode de pénalisation, discrétisation nœud-à-segment.

## Computational contact mechanics: geometry, detection and numerical techniques

**Abstract:** The goal of this work is to derive a consistent framework for the treatment of contact problems within the Finite Element Method using the Node-to-Segment discretization. Three main components of the computational contact have been considered: geometry, detection and resolution techniques. For the sake of completeness, the mechanical aspects of contact as well as numerous numerical algorithms and methods have been discussed. A new mathematical formalism called "s-structures" has been employed through the entire dissertation. It results in a comprehensive coordinate-free notations and provides an elegant apparatus, available for other mechanical and physical applications. Several original ideas and extensions of standard techniques have been proposed and implemented in the finite element software ZéBuLoN (Z-set). Numerical case studies, presented in the dissertation, demonstrate the performance and robustness of the employed detection and resolution schemes.

**Keywords:** computational contact mechanics, finite element method, detection techniques, parallelization, geometrically precise theory of contact, augmented Lagrangian method, penalty method, node-to-segment discretization.